Investigating direct and cooperative microRNA regulation of *Pax6 in vivo* using a genome engineering approach

by

Bridget Ryan
BSc, University of Victoria, 2012

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Division of Medical Sciences

Supervisory Committee

Investigating direct and cooperative microRNA regulation of *PAX6 in vivo* using a genome engineering approach

by

Bridget Ryan
BSc, University of Victoria, 2019

Supervisory Committee

Dr. Robert L. Chow, Department of Biology
Supervisor

Dr. John S. Taylor, Department of Biology
Departmental Member

Dr. Perry L. Howard, Department of Biology
Departmental Member

Dr. Christopher J. Nelson, Department of Biochemistry and Microbiology
Outside Member

Abstract

Cells must employ a diversity of strategies to regulate the quantity and functionality of different proteins during development and adult homeostasis. Post-transcriptional regulation of gene transcripts by microRNAs (miRNAs) is recognized as an important mechanism by which the dosage of proteins is regulated. Despite this, the physiological relevance of direct regulation of an endogenous gene transcript by miRNAs *in vivo* is rarely investigated.

PAX6 is a useful model gene for studying miRNA regulation directly. PAX6 is highly dosage-sensitive transcription factor that is dynamically expressed during development of the eye, nose, central nervous system, gut and endocrine pancreas, and is mutated in the haploinsufficiency disease aniridia. Several miRNAs have been implicated in regulating PAX6 in different developmental contexts. Notably, miR-7 appears to regulate *Pax6* during specification of olfactory bulb interneurons in the ventricular-subventricular zone (V-SVZ) of the brain and during development of the endocrine pancreas.

Here, we produced a bioinformatics tool to enable selective mutation of candidate microRNA recognition elements (MREs) for specific miRNAs while ensuring that new MREs are not inadvertently generated in the process. We then performed the first comprehensive analysis of the mouse *Pax6* 3' untranslated region (3'UTR) to identify MREs that may mediate miRNA regulation of *Pax6* and to identify miRNAs capable of interacting with the 3'UTR of *Pax6*. Using *Pax6* 3'UTR genetic reporter assay,

we confirmed that two MREs for miR-7-5 located at 3'UTR positions 517 and 655 function together to regulate PAX6. We generated mice harbouring mutations in the *Pax6* 3'UTR that disrupt these miR-7-5p MREs, individually or in combination, to explore the biological relevance of miRNA regulation directly. PAX6 protein abundance was elevated in double miR-7-5p MRE mutants relative to wild type and single mutants in the ventral V-SVZ. However, this increase in PAX6 was not associated with an altered dopaminergic periglomerular neuron phenotype in the olfactory bulb.

Our findings suggest that, *in vivo*, microRNA regulation can be mediated through redundant MRE interactions. This work also reveals that directly mutating predicted MREs at the genomic level is necessary to fully characterize the specific phenotypic consequences of miRNA-target regulation.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

3’ RACE	3’ rapid amplification of cDNA ends
3’UTR	3’ untranslated region
5’UTR	5’ untranslated region
5aCON	Pax6(5a) paired domain consensus sequence
A	Adenosine
aa	Amino acid
AGO	Argonaute
AID	Activation-induced cytidine deaminase
AN	Aniridia locus
ANOVA	Analysis of variance
Antagomir	miRNA antisense oligonucleotides
AOB	Accessory olfactory bulb
ARE	AU-rich element
ban	*bantum*
bHLH	Basic-helix-loop-helix
BLAST	Basic local alignment search tool
bp	Base pair
BrdU	Bromodeoxyuridine
C	Cytosine
C. elegans	*Caenorhabditis elegans*
C-terminus	Carboxy terminus
CalB	Calbindin
CalR	Calretinin
Cas9	CRISPR-associated 9
C.I.	Confidence interval
CKO	Conditional knockout
CLASH	Crosslinking, ligation, and sequencing of hybrids
CldU	Chlorodeoxyuridine
CLIP-Seq	Crosslinking immunoprecipitation RNA sequencing
CNS	Central nervous system
Co-IP	Co-immunoprecipitation
CRISPR	Clustered regularly interspaced short palindromic repeats
Ct	Cycle threshold
D	Dorsal
DA	Dopamine
DAergic	Dopaminergic
DL	Dorsal-lateral
DMEM	Dulbecco’s modified eagle’s medium
E	Embryonic day
EdU	5-ethynyl-2’-deoxyuridine
Elavl1	ELAV-like protein 1
ELP4	*Elongation protein 4*
ena	*enabled*

ESC	Embryonic stem cells
ey	*eyeless*
FB	Forebrain
FBS	Fetal bovine syrum
G	Guanine
GABA	Gamma-aminobutyric acid
GC	Granule cell
GCL	Granule cell layer
GFAP	Glial fibrillary acidic protein
GFP	Green fluorescent protein
GL	Glomerular layer
Glut2	Glucose transporter 2
GRN	Gene regulatory network
HB	Hindbrain
HD	Homeodomain
hESC	Human embryonic stem cell
HITS-CLIP	High-throughput sequencing of RNA isolated by crosslinking immunoprecipitation
hMN	Hypoglossal motor neuron
Hprt	Hypoxanthine-phosphoribosyltransferase
IF	Immunofluorescence
INL	Inner nuclear layer
IPGTT	Intraperitoneal glucose tolerance test
kb	Kilobases
L	Lateral
lft2	*lefty2*
LOF	Loss of function
LV	Lens vesicle
MB	Midbrain
MBP	Maltose-binding protein
mESC	Mouse embryonic stem cell
mmu	*Mus musculus*
miRNA	MicroRNA
miR-RISC	MicroRNA-RISC complex
miR-SNP	MicroRNA small nucleotide polymorphism
miTRAP	miRNA trapping by in vitro affinity purification
MRE	MicroRNA recognition element
MRI	Magnetic resonance imaging
mRNA	Messenger RNA
ncRNA	Non-coding RNA
N-terminus	Amino terminus
NeuN	Neuronal nuclei
NLS	Nuclear localization signal
NSC	Neural stem cell
nt	Nucleotide
OB	Olfactory bulb

OE	Olfactory epithelium
Oligo	Oligonucleotide
OS-6mer	Offset 6mer
OSN	Olfactory sensory neuron
OV	Optic vesicle
P	Postnatal day
P3	Pax6 homeodomain consensus sequence
P6CON	Canonical Pax6 PD consensus sequence
PAI	N-terminal domain of the paired domain
Pax	*Paired box*
Pax6	*Paired homeobox-6*
Pax6ΔPD	Pax6 lacking the paired domain
PB	Phosphate buffer
PBS	Phosphate buffered saline
PC1/3	Prohormone convertase 1/3
PCR	Polymerase chain reaction
PD	Paired domain
PEST	Proline/glutamic acid/serine/ threonine
PFA	Paraformaldehyde
PGK	Phosphoglycerate kinase
PGN	Periglomerular neuron
Pri-miRNA	Primary microRNA
Poly(A)	Polyadenylation
prd	*paired*
PR	Photoreceptor
Pre-miRNA	Precursor microRNA
Pre-mRNA	Precursor messenger RNA
PST	Proline/serine/threonine-rich
qPCR	Quantitative polymerase chain reaction
RBP	RNA binding protein
RED	C-terminal domain of the paired domain
RISC	RNA-induced silencing complex
RMS	Rostral migratory stream
ROI	Region of interest
RPC	Retinal progenitor cell
RPE	Retinal pigmented epithelium
rRNA	Ribosomal RNA
RT	Reverse transcriptase
RT-qPCR	Reverse transcriptase quantitative polymerase chain reaction
S-phase	Synthesis phase
S.D.	Standard deviation
SE	Surface ectoderm
SEZ	Subependymal zone
Sey	*Small eye*
Shh	Sonic hedgehog
siRNA	Short interfering RNA

snoRNA	Small nucleolar RNA
SNP	Small nucleotide polymorphism
snRNA	Small nuclear RNA
SUMO	Small ubiquitin-like modifier
SVZ	Subventricular zone
TAD	Transactivation domain
TALEN	Transcription activator-like effector nuclease
Tbp	*TATA binding protein*
TBS	Tris buffered saline
TF	Transcription factor
TH	Tyrosine hydroxylase
TP	Target protector
tRNA	Transfer RNA
TuD	Tough decoy
U	Uracil
V	Ventral
V1/2	Ventral interneurons
VL	Ventral-lateral
V-SVZ	Ventricular-subventricular zone
VZ	Ventricular zone
WT1	*Wilm's Tumor 1*
WT	Wild type
YAC	Yeast artificial chromosome

Acknowledgments

First, I would like to extend a huge thank you to my parents. They inspired in me a desire to explore what lies beyond the horizon of human knowledge and the belief that I could contribute to humanity's understanding of the natural world. Without their contribution, I may never have considered pursuing a Ph.D.

I would also like to thank my supervisor, Dr. Robert Chow, for his support and mentorship. He gave me the opportunity to pursue a childhood dream and believed me capable of achieving something as challenging as a Ph.D. Thanks to his mentorship, I am more confident now, both as a scientist and more generally as a person, than I was when I began.

Many people have contributed to the success of this project, who deserve acknowledgement. I would like to thank my supervisory committee, Drs. Perry Howard, John Taylor and Chris Nelson. Their contributions and feedback on my project have both improved the quality of the project itself and my skills as a researcher. Drs. Yinhuai Chen, Spencer Alford, Kerry Delaney and Raad Nashmi have contributed valuable feedback and training on specific aspects of the project.

Finally, I would like to thank all the students who I have had the pleasure of working alongside and who have contributed work to this project: Emily Enns, Sam Story, Madison Wiebe, Kelly Hamilton, Kieran Lowe, Anneke Hylkema, Talveen Gil, Laura Hanson, and Lauren Braun. Additionally, I would like to acknowledge all the members of the Chow Lab, past and present, who have provided support: Dr. Lily Chen, Di Wu, Dr.

Oliver Krupke, Ana Litke, Peter Watson, Peter Socha, Chris Calvin, Alberto Ruiz, and Seb Gulka.

Dedication

To Torben,

Together, we have paddled against grueling currents

and summited challenging peaks.

Thank you for accompanying and supporting me on this adventure.

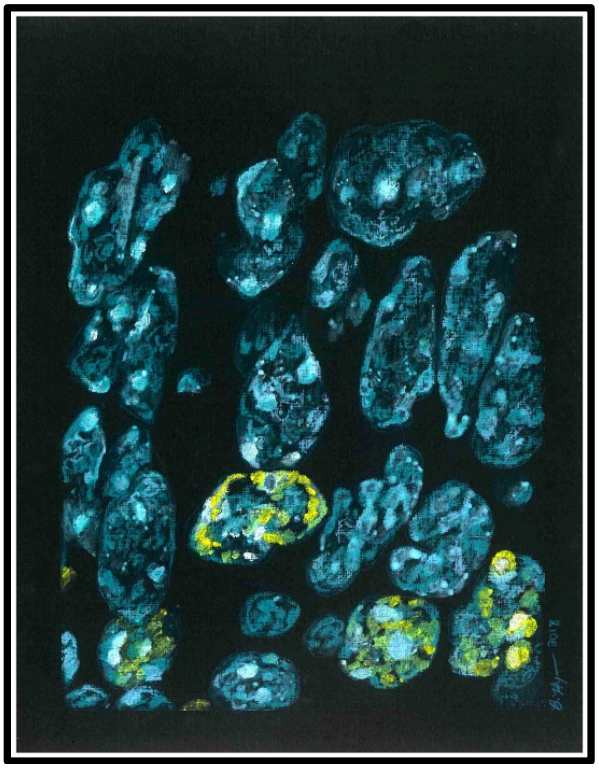
"Breach", pencil on paper (2016)

Decisions Left Unmade

Oh, to be a stem cell!
Pluripotent possibilities.
Committed choices and restricted fates?
I wish not to diff'rentiate!

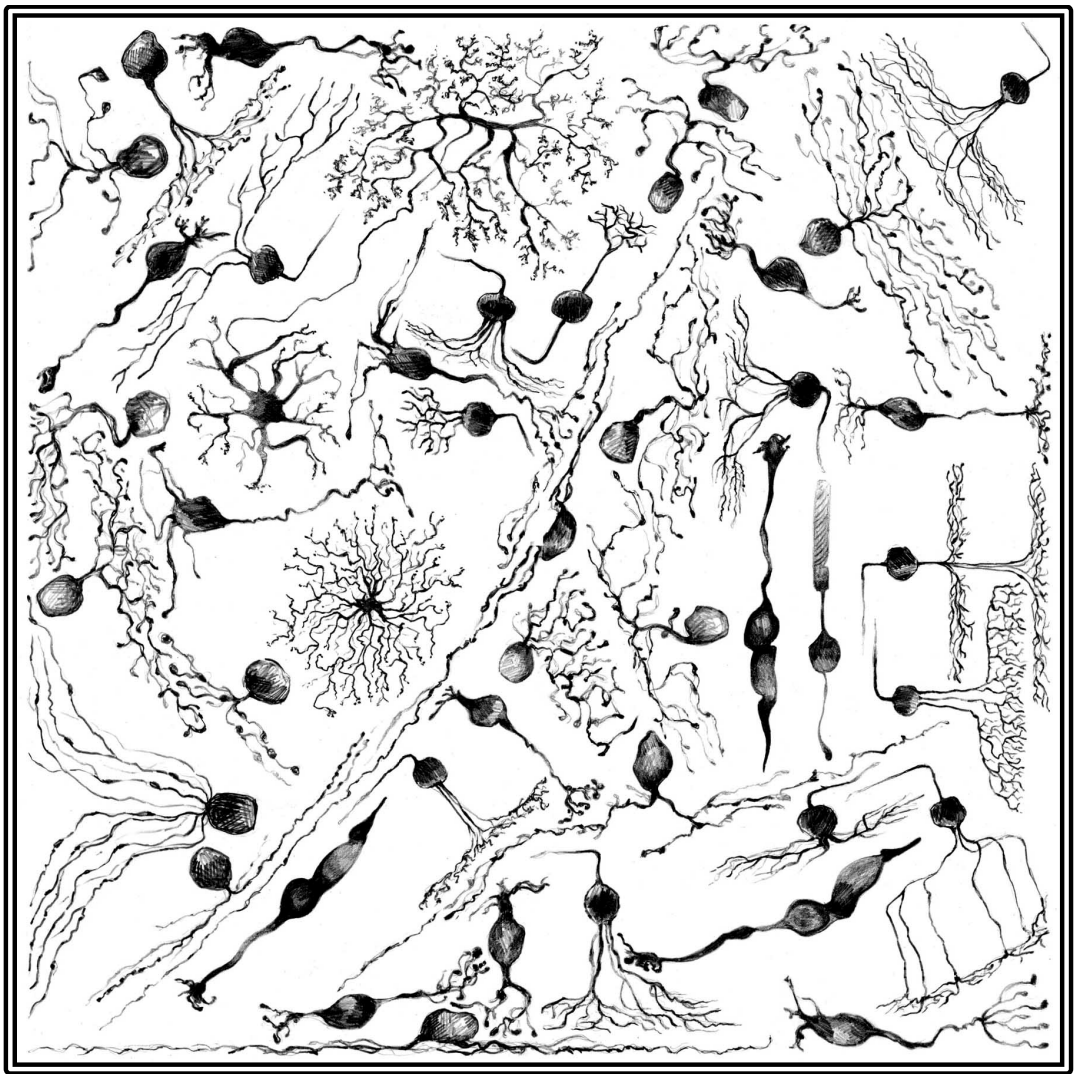
But if I do,
Factors may lead me to find
The other lives not left behind.

I can induce a change of state;
With a chance to explore a different fate.



"Synthesis", oil pastel on paper (2018)

Chapter 1: Introduction



"Regulated Decisions" Pencil on paper (2016)

1.1 From gene to protein: molecular mechanisms underlying cellular regulation

The processes by which transcription of a gene though to translation and protein function are regulated have important implications for cells. Since cells are constantly subjected to changing conditions, the abundance and activity of proteins must be dynamically regulated. It has become clear that all steps in the pathway from gene expression to the final protein are subject to regulation. Chromosome structure, *cis*-regulatory elements in the DNA and promoter usage can be used to control the "when" and "where" of transcription. From there, messenger RNA (mRNA) transcripts can be alternatively spliced, and their stability and use for translation subject to regulation. The activity of the final protein can then be further regulated through covalent attachment of various small molecules, interaction with other proteins, and ultimately degradation (Figure 1). Proper regulation of gene expression, protein stability and function are critical for correct development, response to stress and maintenance of homeostasis, and these processes are frequently dysregulated in disease.

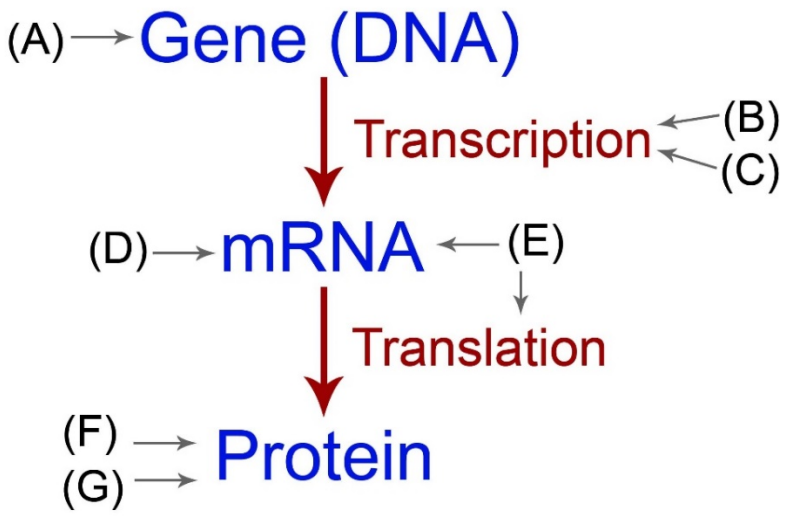


Figure 1. Mechanisms for regulating protein dosage and function

Many mechanisms can be used to regulate the quantity and functionality of a given protein in a cell. (A) Chromatin structure can be regulated to alter accessibility of the gene to transcription factors and RNA polymerase. This is accomplished by addition of various posttranslational modifications to histone proteins within nucleosomes [1]. Transcription of a gene at the level of the DNA can be regulated in several ways. (B) Cis-regulatory regions in the DNA, enhancers and silencers, can be used to control spatial and temporal aspects of transcription initiation [2]. (C) Additionally, alternative promoter usage can be employed to generate multiple different messenger RNAs (mRNAs) from the same genomic sequence, which can impact mRNA stability and produce different protein isoforms with varying functions [3]. mRNA can be regulated at the level of precursor-mRNA (pre-mRNA) processing: capping, splicing and polyadenylation, and though interaction with RNA binding proteins (RBPs). (D) Alternative splicing of the mRNA can be used to generate different protein isoforms [4] and alternative cleavage and polyadenylation of the mRNA 3' end can impact mRNA stability by altering the 3' untranslated region (3'UTR) length [5]. (E) Mature mRNAs can associate with a host of RNA binding proteins (RBPs) that regulate mRNA translation and decay [6]. RBPs can influence processes such as polyadenylation and deadenylation of mRNAs to regulate mRNA turnover [7](Zhang et al., 2010). An important example is regulation by microRNAs (miRNAs), which can interfere with initiation of translation and negatively affect mRNA stability by recruiting protein complexes to the mRNA [6]. Translation initiation is also highly regulated and can be affected by 5'UTR secondary structure [8]. Once a protein has been synthesized from a given mRNA, the stability and functionality of that protein can be regulated in many ways. (F) Degradation of proteins can be regulated by covalent attachment of the small protein ubiquitin [9] and other post translational modifications, such as phosphorylation, methylation, acetylation, hydroxylation and sumoylation, can be used to alter protein function of cellular localization [10]. (G) The function of a protein can also be modified though interaction with other proteins.

1.2 MicroRNAs as post-transcriptional regulators

1.2.1 Discovery of microRNAs

A large portion of the genome in complex organisms is transcribed into non-coding RNAs (ncRNA), RNA that is not translated into protein [11]. Functional ncRNAs were first identified in the form of infrastructural ncRNAs: transfer RNAs (tRNAs), ribosomal RNAs (rRNAs) and small nuclear RNAs (snRNAs), which play important roles in translation and splicing [11, 12]. More recently, trans-acting small regulatory RNAs have been discovered in plants and animals that play important roles in RNA editing, translation and mRNA stability: the small nucleolar RNAs (snoRNAs) and short interfering RNAs (siRNAs)/microRNAs (miRNAs) [11]. MicroRNAs (miRNAs) are a class of 21-25 nucleotide noncoding regulatory RNAs that are processed from stem loop precursors [13] and base-pair with complementary sequences in mRNAs to negatively regulate their translation and stability.

MicroRNAs were discovered through loss of function (LOF) mutations in *Caenorhabditis elegans* (*C. elegans*). *lin-4* was the first characterized miRNA [14]. It was identified by a LOF mutation in *C. elegans* that caused a defect in developmental timing. This miRNA negatively regulates the protein Lin-14 via a complementary antisense interaction with *lin-14* mRNA. Specifically, *lin-4* downregulates Lin-14 protein levels during the first larval stage, permitting developmental progression to the second larval stage [14]. Following this, the miRNA *let-7* was identified in *C. elegans* [15]. Like *lin-4*,

let-7 encodes a 22 nucleotide RNA that acts as a heterochronic gene switch. Specifically, it promotes a transition from the third to the fourth larval stage by temporally downregulating the protein Lin-41 via complementary base pairing to the 3' untranslated region (3'UTR) of the *Lin-41* mRNA [15].

Since their discovery in *C. elegans*, miRNAs were identified as a large class of regulatory molecules with many diverse targets. MicroRNAs are encoded in the genomes of most multicellular organisms studied [13]. Initial predictions estimated that the human genome encodes 200-250 miRNA genes, accounting for approximately 1% of the abundance of transcribed genes [16]. More recent work has produced significantly greater estimates of miRNA gene abundance, predicting that the human genome encodes approximately 1000 miRNA genes [17]. Short RNA deep-sequencing data has identified over 15000 miRNA gene loci and over 17000 mature miRNA sequences in 142 species. Specifically, over 2500 and 1900 distinct mature miRNA sequences have been identified in human and mouse, respectively [18]. If these sequences represent genuine mature miRNAs, it would mean that miRNA genes are one of the most abundant classes of regulatory genes in mammals.

The number of predicted miRNA targets is also very large. Computational approaches that consider evolutionary conservation of predicted microRNA recognition elements (MREs) in 3'UTR sequences suggest that 30-60% of human protein-coding genes are targeted by miRNAs [19, 20]. Other computational methods using pattern-based approaches for predicting miRNA-target heteroduplexes estimate that over 90%

of mammalian gene transcripts are directly regulated by miRNAs [21]. Taken together, these results suggest that miRNAs are a very abundant class of regulatory molecules with a huge number of target mRNAs.

1.2.2 MicroRNA biogenesis

Similar to mRNA transcription and processing, miRNA genes are transcribed as precursor RNAs by RNA polymerase II and are modified with both 5' cap structures and 3' polyadenylation (poly(A)) tails [22]. Though most miRNA genes are their own transcriptional units, some are located in the introns of precursor mRNAs (pre-mRNAs) and are processed from these introns [23]. Additionally, though most miRNA genes are isolated, many are arranged in clusters and are transcribed as multi-cistronic primary transcripts [24]. MiRNAs within such clusters are often related.

Once transcribed, the initial primary miRNA (pri-miRNA) transcript is processed in the nucleus by the enzyme Drosha into a 60-70 nucleotide intermediate RNA having hairpin secondary structure, the precursor miRNA (pre-miRNA) [25, 26]. Pre-miRNAs are then exported out of the nucleus where they are further processed to miRNA duplexes by the enzyme Dicer [27]. The double-stranded miRNA comprises the stem of the pre-miRNA hairpin. Imprecise processing by Drosha or Dicer can generate multiple distinct mature miRNAs from a single pri-miRNA, termed isomiRs [28]. One strand of the miRNA duplex, termed the guide strand, is retained as the mature miRNA and the other strand is degraded [29]. Guide strand selection is asymmetric, with either the 5' or 3' arm of the pre-miRNA being favoured (Figure 2) [18, 29]. Mature miRNAs are loaded into Argonaute (Ago) proteins [30–32] where they function as guides, directing the RNA-

induced silencing complex (RISC) to complementary sites in mRNAs to be silenced [33].

RISC is composed of the proteins Dicer, Ago and TRBP [26].

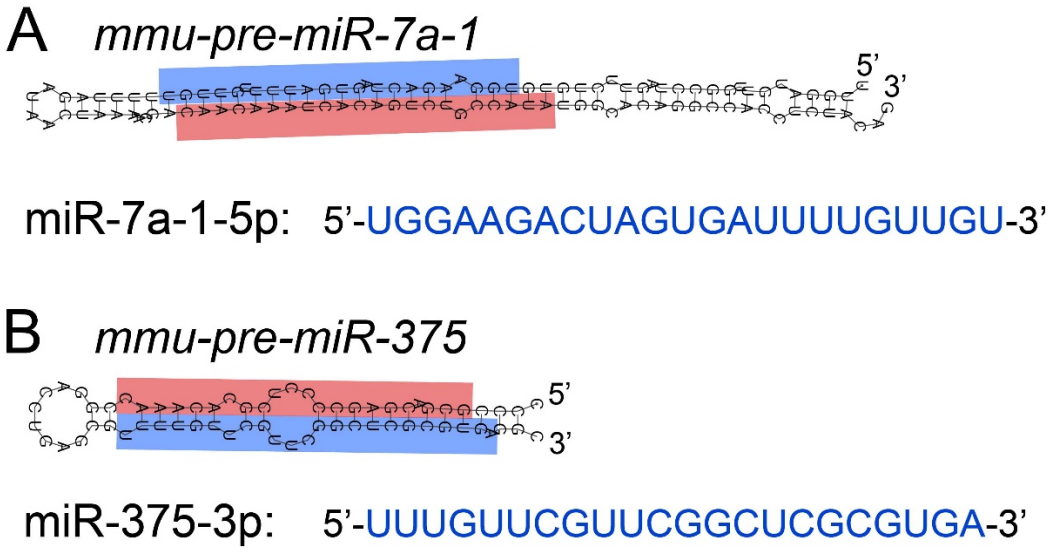


Figure 2. Pre-miRNA processing and synthesis of 5p versus 3p mature miRNA

Primary miRNA (pri-miRNA) transcripts are processed into approximately 60-70 nucleotide precursor miRNA (pre-miRNA) molecules by Drosha [25]. Pre-miRNAs have hairpin structure. (A-B) Example pre-miRNA sequences for *Mus musculus* (mmu)-miR-7a-1 and mmu-miR-375 from miRbase [18]. Further processing of the precursor miRNA by Dicer yields a 21-23 nucleotide mature miRNA (blue highlight in the pre-miRNA sequence), which is retained to serve as the guide strand in the RNA-induced silencing complex (RISC)[27]. The mature miRNA is derived from either the 5' or 3' arm of the pre-miRNA hairpin and the complementary passenger strand is degraded (red highlight). For most miRNAs, either the 5' or 3' arm of the pre-miRNA hairpin is favoured for synthesizing the mature miRNA (blue text) [29]. The mature miRNA nomenclature appends the miRNA name with either -5p or -3p to indicate which arm of the pre-miRNA hairpin the mature miRNA is derived from. (A) mmu-miR-7a-1: the 5' arm of the pre-miR-7a hairpin is preferentially retained and is designated miR-7a-5p. miR-7a-1-5p is 50X more abundant than miR-7a-1-3p based on deep sequencing read count [18]. (B) mmu-miR-375: the 3' arm of the pre-miR-375 hairpin is preferentially retained and is designated miR-375-3p. miR-375-3p is 100,000X more abundant than miR-375-5p based on deep sequencing read count [18].

1.2.3 Mechanism of miRNA target recognition

Much effort has been devoted toward investigating the mechanisms by which miRNAs recognize their targets, as this knowledge is valuable for predicting novel miRNA-target interactions. In *C. elegans*, the miRNAs *lin-4* and *let-7* were found to contain sequence complementarity to motifs within the 3'UTRs of their targeted transcripts [14, 15], setting a precedent for directing subsequent searches for miRNA targets to mRNA 3'UTRs [13]. Mechanistically, displacement of RISC by ribosomal complexes during translation may be the reason for this observed restriction of miRNA targeting to the 3'UTR of mRNAs [34]. This evidence is supported by observations that the number of predicted MREs conserved above chance is low in the first 15 nucleotides (nt) after the stop codon, and sites within 15 nt of the stop codon are less effective [35]. Though it is generally accepted that miRNAs target the 3'UTRs of mRNAs, functional MREs and Ago-occupied miRNA-MRE heteroduplexes have been identified in mRNA 5'UTRs and coding regions [36–40].

Animal miRNAs generally lack perfect or near-perfect sequence complementarity to their target mRNAs. Often, less than half of the miRNA sequence is complementary to the target [41]. This differs from plant miRNAs, which generally have perfect complementarity to their targets [42]. Consequently, identifying mRNA targets of animal miRNAs has presented a greater challenge. The miRNA 5' end, particularly nucleotides 2-8, referred to as the miRNA "seed" region, was suggested to be critical for mediating miRNA target recognition in animals [43]. This is supported by several observations: 5' segments of invertebrate miRNAs were perfectly complementary to their known 3'UTR

targets [44, 45], the 5’ ends of related animal miRNAs tend to be better conserved than the 3’ ends [19, 46], nucleotides upstream of most 3’UTR MREs are poorly conserved across homologous mRNAs [19], and mutations in the 5’ end of a miRNA that create mismatches between the miRNA and validated MREs abolish repression [47]. Additionally, the crystal structure of human Ago2 bound to miRNA reveals that binding to miRNA exposes the miRNA 5’ end to target recognition [31]. Overall, these results suggested that miRNA 5’ ends are most important for mediating target recognition, pairing to the miRNA 3’ end plays a limited role and that novel gene targets can be determined based on sequence complementarity to miRNA 5’ ends.

Several classes of functional MRE and the characteristics of more effective MREs have been identified in animals based on selective conservation of 3’UTR motifs to miRNA 5’ ends [43] (Figure 3A). These were referred to as canonical seed matches. In order of increasing selective conservation and efficacy, the canonical MREs are: offset-6mer (OS-6mer), 6mer, 7mer-A1, 7mer-m8 and 8mer [35, 48, 49]. OS-6mer MREs are complementary to miRNA positions 3-8 [20]. 6mer MREs are perfectly complementary to nucleotides 2-7 of the miRNA, starting the 5’ end. Two types of 7mer MREs exist: 7mer-m8 sites are complementary to nucleotides 2-8 of the miRNA, whereas 7mer-A1 sites are 6mer sites with an adenosine (A) across from position 1 of the miRNA. Finally, 8mer MREs are complementary to nucleotides 2-8 of the miRNA and have an A across from position 1 [19, 35]. Interestingly, it was observed that MREs targeted by miRNAs that do not begin with U usually have this conserved A, leading to the hypothesis that the RISC recognizes the conserved A and helps facilitate the miRNA-

mRNA interaction [19]. This hypothesis was validated by crystal structure of Ago bound to miRNA and target. The A across from the first miRNA nucleotide helps facilitate target recognition by binding Ago [31] and is not involved in Watson-Crick pairing with the miRNA [35, 48, 50].

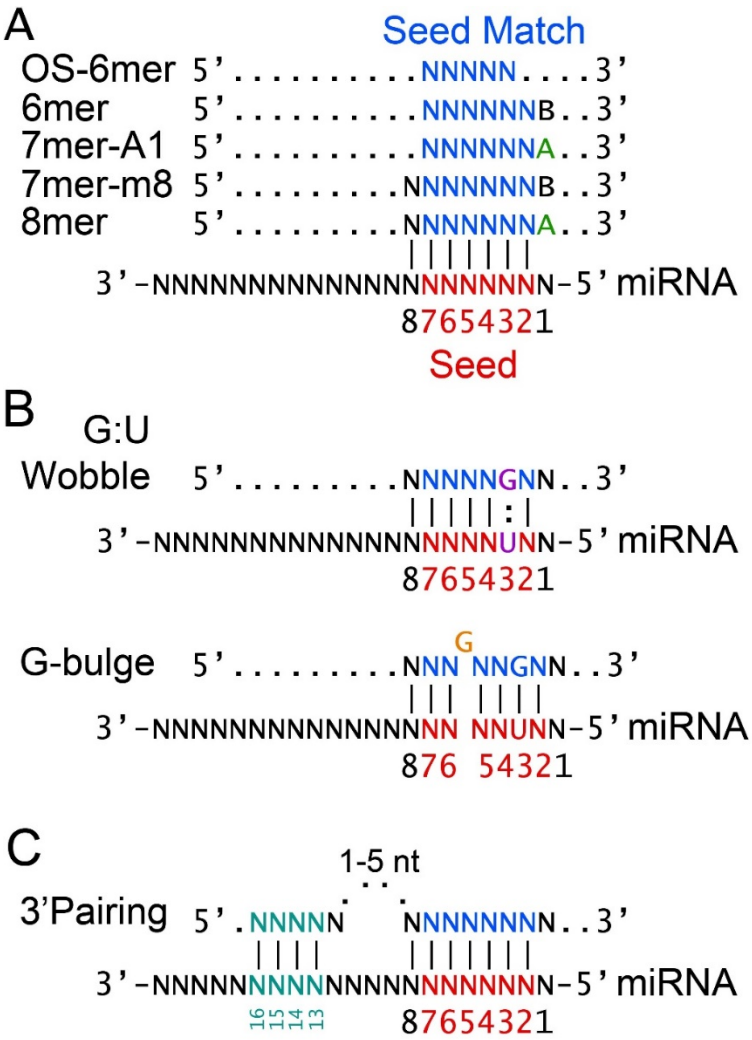


Figure 3. MicroRNA recognition element types (MREs)

(A) Canonical seed matched MREs. Canonical MREs in order from lowest to highest efficacy: offset 6mer (OS-6mer), 6mer, 7mer-A1, 7mer-m8, 8mer [20, 51]. The miRNA seed, nucleotides 2-7 starting from the miRNA 5' end, is shown in red. MRE in the mRNA 3'UTR is shown in blue. A across from miRNA position 1 (green) binds Ago and is not involved in Watson-Crick pairing with the miRNA [31]. (B) Non-canonical seed matched MREs. Functional analyses and Argonaute (Ago) crosslinking approaches identified several recurring non-canonical MREs: G:U wobble sites and G-bulge sites. Functional miRNA-MRE pairs can harbor G:U mismatches, "wobble pairs" (purple) [19, 47, 52, 53]. Ago crosslinking experiments identified many miRNA-MRE pairs harboring bulges in either the miRNA or mRNA [40]. An abundant bulged site is the G-bulge MRE [54] where a guanine (G) nucleotide in the mRNA is bulged between miRNA positions 5 and 6 (orange). (C) 3' pairing may function to supplement pairing to the 5' end or compensate for weak 5' pairing [19, 52]. Specifically, miRNA positions 13-16 (teal) appear to be most important for mediating 3' pairing [35]. Figure modified from [20] and [51].

Non-canonical MRE types have also been identified that are not selectively conserved but can bind miR-RISC and function to mediate target repression. Some miRNA-mRNA interactions have G:U mismatches, termed “wobble” pairs, or bulges between the miRNA seed and MRE (Figure 3B). Though these mismatches can function, they are generally considered to be detrimental [19, 47, 52]. Despite this observation, introduction of G:U wobbles into known functional MREs can still produce efficient target down-regulation, revealing that G:U wobbles may not always impair miRNA-target interactions [53]. Argonaute High-Throughput Sequencing of RNA isolated by crosslinking immunoprecipitation (Ago HITS-CLIP) has been used to validate that these non-canonical MREs can bind miR-RISC [40, 55–57]. An abundance of miRNA-MRE matches containing G:U mismatches and bulges were identified using this approach [40]. Though non-canonical MREs may bind miR-RISC, most of these MREs are unlikely to be functional [49].

Though the 3’ end of the miRNA is generally considered less critical for mediating miRNA-target recognition, it may function in the context of both canonical and non-canonical interactions (Figure 3C). Outside of the miRNA “seed”, nucleotides 13-16 are the best conserved between paralogous human miRNAs, leading to the hypothesis that these nucleotides may participate in supplementary or compensatory pairing [35]. In support of this, the crystal structure of miRNA bound to Ago reveals that nucleotides 13-16 are exposed for additional target recognition [31]. Functionally, 3’ pairing may enhance regulation [35], though mutations in the mRNA that disrupt 3’ pairing reveal that is generally does not play an important role in miRNA-mediated repression [47]. It

is important to note that extensive complementarity to the miRNA 3' end in the absence of a minimal 6mer MRE is not sufficient to facilitate targeting and optimizing pairing energy does not ensure identification of functional targets [52]. Interestingly, 3' compensatory pairing may provide target specificity between miRNA family members with identical 5' sequences [19, 52].

The position of target sites within a 3'UTR and the local 3'UTR environment can also influence miRNA targeting. Though complementarity to the miRNA seed is important, it may not be enough to confer repression. This was exemplified by experiments in *C. elegans* that moved functionally validated MREs from one 3'UTR into a 3'UTR for a different mRNA, or even to different locations within the same 3'UTR. From this, it was evident that the 3'UTR context impacts MRE functionality [53]. Additional observations suggested that MREs near the middle of the 3'UTR and within regions of high local guanine-cytosine (GC) content are less effective, and MREs that reside within local adenine-uracil (AU)-rich regions are more likely to be functional [35]. In contrast, experiments that artificially altered the AU content in the vicinity of validated MREs observed little impact on site efficacy [53]. mRNA secondary structure may impact miRNA regulation, with MREs located within regions of predicted secondary structure being associated with reduced miRNA-mediated repression [58]. This may explain the observations that MREs located within shorter 3'UTRs (<400 nt) tend to be associated with stronger repression than MREs located within longer 3'UTR (>800 nt) [59].

1.3 MicroRNA function

1.3.1 Mechanism of miRNA-mediated repression

As part of RISC, miRNAs act as sequence-specific guides that recruit RISC to mRNAs. The miRNA-RISC can downregulate gene expression by direct cleavage of target mRNAs [60], though direct cleavage of the mRNA is the mechanism employed primarily by plant miRNAs [61]. Animal miRNAs usually have a modest impact on target repression [50, 62] and can impact levels of both targeted mRNA and protein [50] through a combination of mRNA destabilization and translational repression [63]. However, if the miRNA is more abundant than its target, miRNAs can also function as switches [64]. mRNA destabilization is now thought to result from deadenylation and decapping of targeted transcripts, whereas translational repression is the consequence of inhibition of translation initiation [63, 65]. Some evidence suggests that reduction in protein levels following regulation by miRNAs is primarily the result of target mRNA destabilization [66], and translational inhibition is required first followed by mRNA degradation [67]. Though miRNAs are generally accepted to inhibit translation, they may be able to function to activate translation in quiescent cells by recruiting FXR1, a protein not normally part of the repressive miR-RISC [68].

1.3.2 Developmental importance of miRNAs

MicroRNAs play important roles during animal development and this importance is demonstrated by *Dicer*-null embryos, which are incapable of synthesizing mature miRNAs [27]. *Dicer*-null zebrafish embryos arrest at developmental day 10, once

maternal Dicer1 has been depleted [69]. Similarly, *Dicer*-null mouse embryos die early in embryonic development [70]. These results suggest that global miRNA function is essential for vertebrate development [27, 70].

In addition to the global function of miRNAs during the early stages of embryonic development, miRNAs are now known to be involved in many specific developmental processes. Conditional knockout (CKO) of *Dicer* using the Cre-loxP recombination system is used to interrogate the importance of global miRNA function during development of specific tissues. For example, conditional knockout of *Dicer* in the developing and adult endocrine pancreas revealed that miRNAs play important roles in development and survival of β-cells, and insulin biosynthesis [71–73]. Additionally, *Dicer* CKO demonstrates that global mRNA function is indispensable for normal central nervous system (CNS) development. Loss of *Dicer* in the developing cortex caused reduced cortical thickness due to apoptosis and disorganized cortical structure [74]. *Dicer* CKO in retinal progenitors produced a similar apoptotic phenotype in the retina [75, 76], along with reduced RPC competence [76, 77], improper boundary formation between the neural retina and neighbouring ciliary body [76] and defects in light responses [78]. Similarly, *Dicer* ablation from specific neuronal subpopulations causes impairments. For example, CKO of this enzyme in striatal dopaminergic neurons causes defects in motor behaviour [79] and CKO in excitatory forebrain neurons impairs neuronal differentiation, survival, and cell morphology [80]. In sum, all tissues likely require global miRNA function for correct development and maintenance.

It should be noted that, in addition to its functions in small RNA biogenesis, Dicer also has other cellular functions. For example, Dicer can translocate to the nucleus and is required for processing of pre-rRNA [81]. Consequently, phenotypes associated with *Dicer* knockout may not be solely due to defects in miRNA biogenesis.

Many miRNAs are expressed in specific spatial and temporal patterns during development [82–88], and it has been suggested miRNAs are primarily involved in differentiation and tissue maintenance in multicellular organisms [89, 90]. Many pieces of evidence support this hypothesis. First, with some exceptions [91], miRNA expression is largely absent from unicellular organisms, though components of the miRNA biogenesis pathway predate the evolution of multicellularity [92]. Second, more abundant and diversified miRNA expression is typically observed as development progresses [85, 88, 93, 94]. For example, miRNA abundance increases with differentiation in erythroid cells, skin and retina [95–97]. Third, cell lineage specification can be influenced by the complement of miRNAs expressed. Ectopic expression of specific combinations of miRNAs in hematopoietic stem cells can alter their cell fate choices [98] and though *Dicer*-null embryonic stem (ESCs) are viable in culture, they have differentiation defects [99]. Fourth, miRNAs generally have lower levels of expression in tumors relative to normal adult tissue [94, 100]. Finally, as embryonic development progresses mouse mRNA 3'UTR length tends to progressively increase [101] and mRNA 3'UTRs from the adult brain, a highly complex organ with many different cell types, tend to be longer than other tissues[102]. These findings suggest

that gene transcripts may be subject to increasing miRNA-mediated regulation at later developmental stages [101].

During development, transitions may occur temporally, as in differentiation, or during tissue patterning when spatial domains are established. MicroRNAs may function to sharpen these transitions by suppressing residual or unwanted transcripts [103] (Figure 4A-B). As evidence for this, anti-correlated expression patterns of miRNA and their predicted targets were observed in *Drosophila* [104]. Additionally, the first identified miRNAs, *lin-4* and *let-7*, negatively regulate their respective target transcripts and promote the transition from one stage of larval development to the next [14, 15]. Specifically, *let-7* promotes the temporal differentiation of hypodermal blast cells into cuticular alae at the end of the fourth larval stage [15]. Since these discoveries, the miRNA *let-7* has been identified across many different animal lineages where it is highly conserved in both sequence and onset of expression [105]. *let-7* continues to be expressed later in vertebrate development and into maturity. The lowest levels of *let-7* expression are seen in tissues that contain large proportions of immature cells, such as the bone marrow [105]. These results suggest that the miRNA *let-7* may play an important role in regulating the timing of tissue differentiation in vertebrates. MicroRNAs are also involved in defining the spatial boundaries of tissues. For example, during late embryonic development in zebrafish, miRNA-9 expression is required to define the boundary between the developing hindbrain and midbrain [106].

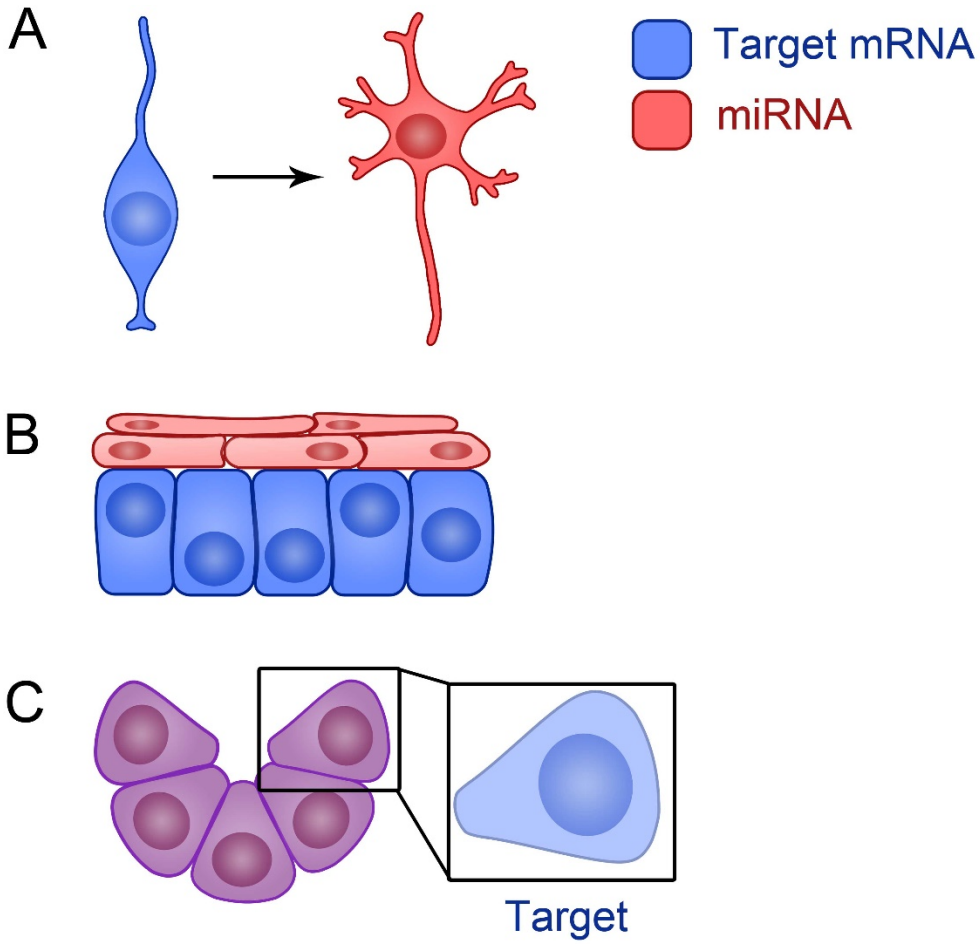


Figure 4. Developmental roles of miRNAs

(A) miRNAs can function as temporal switches to enhance state changes during progressive differentiation. Expression pattern of the target mRNA is shown in blue and the miRNA in red. (B) miRNA as a spatial switch to enhance boundaries during tissue morphogenesis. In both (A) and (B), the miRNA is expressed in distinct domains from the target. In these cases, microRNAs may play a role in sharpening transitions as cells switch states, to help to prevent systems from spontaneously changing states or to prevent ambiguous cell fate choices. (C) miRNAs can function as tuners, either dampening target to optimal levels or preventing unwanted fluctuations in target levels to provide stability. Here, the miRNA is coexpressed with target and target expression is maintained at low levels (see inset, light blue indicates low target level)[51, 107].

Though many miRNAs are highly conserved in vertebrates and animals, individual miRNA gene knockout animals are often viable and lack obvious

developmental phenotypes [87, 108, 109]. One explanation for this is functional redundancy. Many miRNAs are part of miRNA families that share the same seed sequence. Such miRNAs may function in combination to regulate the same targets [110], and deletion of some members of a miRNA family can be compensated for by remaining family members [111, 112]. However, most *C. elegans* mutants that lack multiple members of a miRNA family do not display overt abnormalities [113]. As an alternative explanation, though miRNA gene mutations may not typically be associated with gross abnormalities, these mutants are not actually normal. For example, systematic study in *Drosophila melanogaster* reveals that, despite having a normal appearance, over 80% of individual miRNA mutants show general defects in survival, lifespan, fertility or other developmental defects [114]. Interestingly, phenotypes associated with miRNA gene knockout may be exacerbated by physiological stress. For example, miR-7 deletion in flies alters expression of transcriptional regulators involved in photoreceptor and sensory organ development under conditions of temperature fluctuation [115]. Several mouse lines lacking specific miRNA genes or clusters are viable, fertile and lack overt abnormalities but show impaired responses to injury and tissue damage [116–119], mechanical stress [120, 121], synaptic transmission [122], aging [123] or glucose stress and obesogenic conditions [124–126]. Additionally, loss of individual miRNAs in worms generates mutant phenotypes in sensitized genetic backgrounds [127].

Observations from miRNA gene knockout animals have led to the hypothesis that the primary function of miRNAs is to provide stability and robustness to gene regulatory networks, particularly under conditions of physiological stress (Figure

4C)[103]. Developmental processes require the coordinated action of many transcription factors functioning in complex regulatory networks. An important feature of these networks is robustness, which results in decreased inter-individual variability while creating developmental stability in the face of environmental perturbations [128]. Computational methods provide evidence suggesting that regulatory networks containing miRNAs are recurrent in mammals [107]. For example, C-Myc positively regulates transcription of a transcription factor involved in cell cycle progression, E2F1, and the miR-17 cluster. Several miRNAs expressed as part of the miR-17 cluster negatively regulate E2F1, reducing positive feedback of E2F1 onto *c-Myc* [129]. This regulatory network containing miR-17p and miR-20a may provide tight regulation of proliferation in humans. Additionally, miR-7 is involved in regulatory networks for photoreceptor cell, proprioceptor organ, and olfactory organ development in *Drosophila* [130], where it may function to buffer developmental processes against environmental disturbances [115]. These networks are composed of feedback and feedforward network motifs, and though the mechanism of miRNAs is repressive, as part of networks, the ultimate result may not be repressive.

1.3.3 Cooperative and combinatorial regulation by miRNAs

Genes that encode for different functional classes of proteins are differentially represented as predicted targets of miRNAs. Of target transcripts predicted to be targeted by miRNAs in humans and flies, mRNAs encoding transcriptional regulators were found to be enriched [131–133]. Additionally, human genes involved in transcriptional regulation and developmental processes, or that encode for nuclear

proteins tend to contain many conserved predicted MREs in their 3'UTRs, suggesting that proteins involved in these processes are under strong regulation by miRNAs [59].

A single miRNA can target many different mRNAs (Figure 5B). In silico approaches relying on MRE conservation suggest that single miRNAs likely target many different mRNAs, and that regulation of a single mRNA by a single miRNA is rare [131, 134]. Bioinformatics predictions relying on evolutionary conservation of predicted MREs estimate that an individual miRNA will target, on average, 200 mRNA transcripts [135]. Overexpression and knockdown of individual miRNAs has been used to identify hundreds of putative targets [62] and Ago crosslinking immunoprecipitation RNA sequencing (CLIP-Seq) data reveals that a single miRNA may target hundreds of different mRNAs in a given cell type [40].

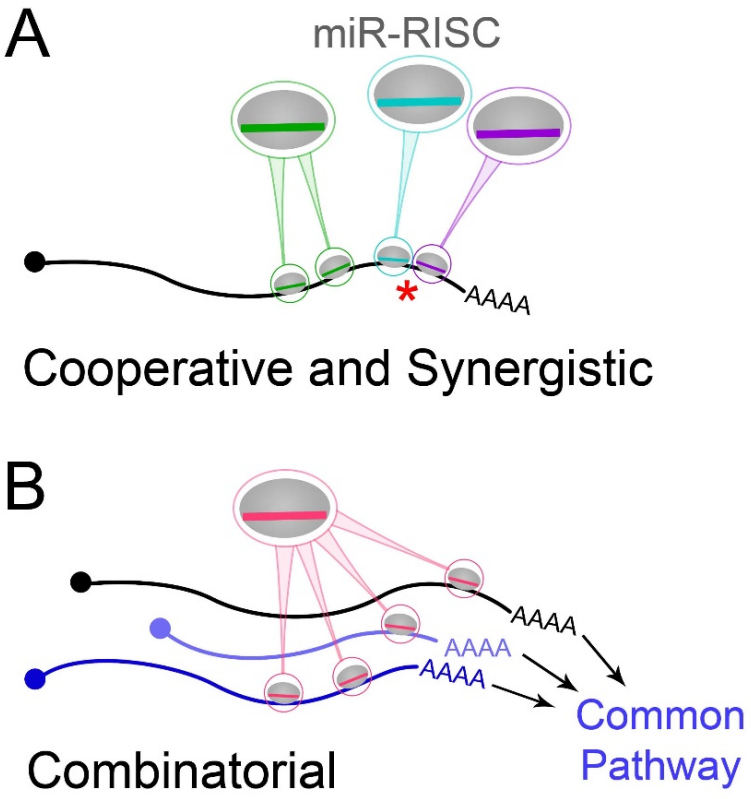


Figure 5. Many-to-many regulation to miRNAs

miRNA-RISC complexes (miR-RISC) are shown as grey ovals (RISC) with bright coloured lines (miRNA). Different miRNA species are represented by different colours. (A) An individual miRNA may be cooperatively regulated through multiple MREs, for the same miRNA or different miRNAs [47, 135–137]. Regulation may be synergistic if MREs are closely spaced (red asterisk)[35, 59, 138]. (B) An individual miRNA may target many different mRNAs in a combinatorial manner. Targeted mRNAs may encode proteins that participate in common pathways or functional modules [90, 131, 134]. Figure modified from [65].

The multiple gene transcripts predicted to be regulated by a single miRNA do not appear to be random, instead miRNAs may target multiple gene transcripts for proteins that participate in the same functional module [90]. A review of validated miRNA-target interactions in progenitor cell differentiation pathways for a variety of cell lineages highlights individual miRNAs that regulate multiple pathway components to produce

coherent outcomes. For example, miR-203 targets multiple regulatory proteins involved in promoting differentiation of epidermal stem cells [90]. Overexpression of miRNAs of interest combined with microarray or quantitative proteomics have also been used to identify common targets of single miRNAs or miRNA clusters that act coordinately as part of a single pathway [139–141].

One mRNA can also be targeted by multiple miRNAs (Figure 5A). The extent to which a target is repressed increases with increasing number of seed matches to a miRNA [47, 48, 59], and 3' UTRs with multiple MREs for a single miRNA are more likely to be regulated by that miRNA [19, 52]. Genetic reporter experiments have revealed that multiple miRNAs can regulate a single target through multiple MREs in the 3'UTR [47, 135–137]. Additionally, 3'UTRs containing multiple MREs recognized by the same or different miRNAs are associated with greater repression, particularly when the inter-site spacing is small. Specifically, functional and bioinformatics analyses reveal that MREs spaced approximately 10 to 40 nt apart mediate optimal target repression [35, 59, 138]. In summary, the regulatory relationship between miRNAs and their targets can be described as "many-to-many" [134].

1.3.4 Experimental approaches for studying miRNA regulation

See Table 1 for summary of key approaches used to study miRNA regulation.

Table 1. Experimental approaches for studying miRNA regulation

Approach	Purpose	Advantages and Limitations
Bioinformatics MRE prediction	Predict functional MREs in an mRNA, or mRNAs targeted by a given miRNA	**Advantage:** fast and inexpensive **Limitation:** high rate of false-positive and false-negative predictions
Ago-HITS-CLIP	Identify miRNA-mRNA binding events	**Advantage:** high throughput **Limitation:** identified miRNA-MRE interactions may not be functional
miTRAP	Identify miRNA-mRNA binding events	**Advantage:** exogenously expressed transcript contains MS2 hairpins and is easily purified **Limitation:** requires expression of an exogenous transcript. Identified miRNA-MRE interactions may not be functional
miR-CATCH	Identify miRNA-mRNA binding events	**Advantage:** identify miRNAs interacting with an endogenous mRNA **Limitation:** difficult to purify low abundance transcripts. Identified miRNA-MRE interactions may not be functional
miRNA LOF (gene knockouts, antagomirs, miRNA sponges, TuDs)	Identify miRNA targets and consequence of miRNA regulation	**Advantage:** identify many putative targets of a miRNA. Address the biological role of a miRNA **Limitation:** regulation of presumed targets may not be direct
miRNA overexpression	Identify miRNA targets	**Advantage:** identify many putative targets **Limitation:** regulation of presumed targets may not be direct. Can suggest interactions that do not occur normally. Can displace endogenous miRNAs by saturating RISC
Reporter systems: reporter gene fused to 3'UTR of interest	Identify functional MREs	**Advantage:** fast and easy validation of predicted MREs **Limitation:** results may not reflect regulation of endogenous gene; requires expression of exogenous reporter genes and frequently involves overexpression of miRNA
Target protectors	Identify and characterize functional MREs	**Advantage:** can target endogenous mRNA **Limitation:** not specific for an individual MRE; can block many MREs simultaneously
Mutation of MREs at the genomic level	Characterize functional MREs	**Advantage:** disrupts endogenous MRE, specific to MRE of interest **Limitation:** expensive and time consuming

Using knowledge of miRNA targeting, algorithms have been generated to predict mRNAs targeted by known miRNAs. Creating algorithms to predict miRNA targets in plants has been relatively easy since MREs have near perfect complementarity. In animals, functional duplexes are more variable in structure; consequently, predicting targets is much more difficult [142]. Many different software tools have been developed to predict potentially functional MREs in mRNAs [19–21, 45, 49, 58, 131, 135, 143–148]. These tools make use of several different parameters in their predictions, such as extent of complementarity to the miRNA 5' end, hybridization energy of the mRNA-miRNA heteroduplex, evolutionary conservation of predicted MREs within aligned orthologous 3'UTR sequences, mRNA secondary structure, and local 3'UTR context. However, establishing general rules for predicting functional MREs from 3'UTR sequences is difficult [149]. Consequently bioinformatics-based prediction of functional MREs suffers from a high rate of false positive predictions [150]. Additionally, approaches that rely on evolutionary conservation of an MRE within aligned orthologous 3'UTR sequences from many species may suffer from false negative results. Functional MREs may be conserved between orthologous sequences but may not be located within the same relative 3'UTR positions [151]. Ultimately, experimentation is required to validate predicted MREs.

Several high-throughput capture-based approaches have been developed to identify miRNA-mRNA binding events [150]. Immunoprecipitation methods have been developed to affinity purify components of the RISC, such as Ago-HITS-CLIP, and use high-throughput RNA sequencing, microarray or RT-qPCR to identify miRNA-target pairs. One limitation of HITS-CLIP approaches is that miRNA-mRNA heteroduplex components

must be sequenced separately, and putative binding maps are generated using bioinformatics. Crosslinking, ligation, and sequencing of hybrids (CLASH) has been used as a strategy to identify miRNA-mRNA interaction pairs [55]. The strength of these approaches is that they generate large-scale miRNA-mRNA interaction maps. However, they may generate also many false positive predictions. For example, the non-canonical miRNA-MRE interactions identified by these approaches may not generally be functional, despite binding miR-RISC [49].

RNA-bait approaches have been developed to identify miRNAs interacting with a mRNA of interest *in vitro* and *in vivo*. miRNA trapping by in vitro affinity purification (miTRAP) involves introduction of an exogenous reporter transcript fused to a 3'UTR of interest along with multiple MS2-loops into cells of interest. The MS2 RNA loops bind an MS2 protein. By fusing the MS2 protein to maltose-binding protein (MBP), the reporter transcript can be purified along with interacting miRNAs [152]. One limitation of miTRAP is that it relies on *in vitro* expression of an exogenous reporter transcript bearing the 3'UTR of interest. A different affinity purification strategy termed miR-CATCH was developed to enable affinity purification of endogenous mRNAs [153]. Here, a complementary biotin-tagged oligonucleotide is used to purify the transcript of interest along with associated miRNAs. miRNA-MRE interactions identified by these approaches need to be validated to address whether the interaction is associated with regulation.

Altering endogenous levels of a miRNA or interfering with miRNA activity can be used to address the biological role of a given miRNA. Several miRNA loss of function (LOF) approaches are available. The most reliable approach is to generate miRNA gene

knockouts; however, this approach is laborious and can be complicated by redundant miRNA genes. As alternatives to miRNA gene knockout, several miRNA competitive inhibition approaches have been generated for miRNA LOF: miRNA antisense oligonucleotides (antagomirs), miRNA sponges and tough decoys (TuD) [154]. Generally, these strategies function by binding a specific mature miRNA species and sequestering it away from its targets. miRNA overexpression has also been used to identify putative miRNA targets [155] and to assess the extent of the miRNA regulome in cultured cells [50, 62]. One limitation of miRNA LOF and overexpression strategies is that they do not demonstrate direct regulation. Additionally, overexpression of a miRNA can suggest interactions that do not occur *in vivo* and displace endogenous miRNAs by saturating RISC [156].

Several approaches have been developed to address whether identified MREs can function in miRNA-mediated repression *in vitro* and *in vivo*. Reporter systems involve expression of a reporter gene, such as luciferase or green fluorescent protein (GFP), fused to a 3'UTR sequence of interest. Levels of reporter protein are compared against a control reporter harboring a mutation in the MRE of interest. If the MRE is biologically functional, the presence of a targeting miRNA will direct RISC to the reporter mRNA, resulting in downregulation of reporter protein level. Reporter systems provide information about potentially functional miRNA-target interactions but are no guarantee that the endogenous transcript is regulated by the miRNA in question under normal physiological conditions [150]. Target protectors (TPs) have been developed as an alternative to exogenous reporter genes for the purpose of characterizing functional

MREs. TPs are antisense oligonucleotides designed to bind to the region of a 3'UTR sequence of interest containing the MRE, thus protecting the mRNA from miRNA-mediated repression [157]. The endogenous mRNA can be targeted using this approach, eliminating the need for exogenous reporters. One limitation of TPs is that they are not perfectly specific for the MRE of interest. Given that TPs are at least 25 nucleotides in length [157], they may block access to neighboring MREs for other miRNAs. Thus, changes in the level of target mRNA or protein, or physiological observations associated with TP use may attributable to regulation by additional miRNAs.

To achieve specificity in addressing the phenotypic consequence of endogenous regulatory loci, the gold standard involves mutation at the genomic level. A few cases of gene targeting approaches being used to disrupt endogenous MREs and assess their phenotypic consequences have been documented in the literature. Using a classical gene targeting approach, a miR-155 MRE was disrupted in the 3'UTR of the gene encoding the enzyme activation-induced cytidine deaminase (AID) in mice with the goal of addressing the role of miR-155 regulation of AID during B cell class switching directly. Mice heterozygous for the gene encoding AID have elevated mRNA and AID protein when the miR-155 MRE is mutated [158]. More recently, genome engineering using transcription activator-like effector nucleases (TALENs) and clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated-9 (Cas9) have been used to investigate the function of MREs *in vivo*. TALENs were used to delete an MRE for miR-430 in the 3'UTR of *lefty2* (*lft2*) in zebrafish embryos. *lft2* is upregulated in these mutant embryos, and embryos display cyclopia [159]. Additionally, CRISPR/Cas9 was

used to introduce an indel mutation into the *bantum* (*ban*) MRE in the 3'UTR of *enabled* (*ena*) in *Drosophila*. Level of *ena* in mutants overexpressing *ban* is unchanged. Expression of Ena in wing imaginal discs is important for tissue patterning. However, Ena is not upregulated in discs of mutants and wing development appears normal [159].

Despite the importance of performing genomic mutations of MREs to address the role of miRNA-target regulation directly, this approach is rarely employed. We sought to study the impact of miRNA regulation directly by using genome engineering to disrupt candidate MREs in the context of an endogenous 3'UTR. Particularly, we were interested in addressing whether an endogenous transcript can be cooperatively regulated through multiple MREs in the same 3'UTR. We chose the gene *Paired homeobox-6* (*Pax6*) for this investigation. *Pax6* encodes a transcription factor and developmental gene, it exhibits a dynamic and highly regulated pattern of expression, and proper development is very sensitive to the correct dosage of Pax6 protein, making *Pax6* an excellent model protein for studying miRNA regulation.

1.4 The transcription factor PAX6

1.4.1 Discovery of the paired box genes and *Pax6*

The *Paired box* (*Pax*) genes are part of a multigene family that encode transcription factors and were originally identified in vertebrates based on sequence homology to the *Drosophila* segmentation gene *paired* (*prd*) [160]. The *prd* gene contains a 384 base pair DNA sequence termed the paired box, which encodes a 128

amino acid paired domain (PD) [161]. This PD represented a novel DNA-binding domain, which is necessary and sufficient to mediate DNA binding [162]. A second DNA-binding domain, a helix-turn-helix homeodomain (HD), is also encoded by the *paired* gene [163]. The *prd* HD mediates DNA binding independent of the PD, and has different DNA sequence specificity [162].

Eight murine paired box-containing genes were originally isolated by genetic screening for paired box-containing genes in the *Mus musculus* (mouse) genome [164] and were named *Pax1-8*, as these genes encode transcription factors that all contain paired DNA binding domains [160]. Later, a ninth *Pax* gene was isolated from *Homo sapiens* (human) and mouse [165, 166]. *Pax* genes have spatially and temporally restricted expression patterns during development, suggesting an important role in cellular differentiation and tissue morphogenesis [160] and play indispensable roles during the development of many vertebrate organs and structures, particularly the CNS [166–176].

The sixth paired box-containing gene, *Pax6*, was originally isolated from mouse based on conservation of the paired box sequence motif with that of *Drosophila* [164]. The PAX6 amino acid sequence was deduced from its cDNA. The predicted protein is a 422 amino acid transcription factor that contains two DNA binding domains: a PD and a paired-like HD [177] (Figure 6A). *PAX6* was isolated in humans by positional cloning of a candidate cDNA at the *aniridia* (*AN*) locus [178]. Like its murine homologue, this gene is predicted to encode two DNA binding domains characteristic of *Pax* family members: a PD and a HD. Like other Pax proteins, the PD is located at the amino terminus (N-

terminus); however, the PD differs in sequence from other known PDs, suggesting differential DNA binding specificity [177, 179]. Mutations in the DNA sequence encoding the PD of *Pax6*, resulting in an amino acid substitutions, can reduce the DNA binding ability of this protein or alter its DNA targets, resulting in human disease [180]. The PD contains two independent DNA-binding subdomains [181, 182], the N-terminal subdomain PAI and the carboxy-terminal (C-terminal) subdomain RED [183]. Though PAI is most critical for binding DNA [181, 182], PAI and RED can function together to confer binding site specificity (Figure 6B)[183]. Additionally, though the PD and paired-like HD can recognize DNA motifs independently and the PD binds its consensus sequence more effectively than the HD binds its respective consensus sequence (Figure 6D)[184], they can also function cooperatively to expand the recognition repertoire of PAX6 [183].

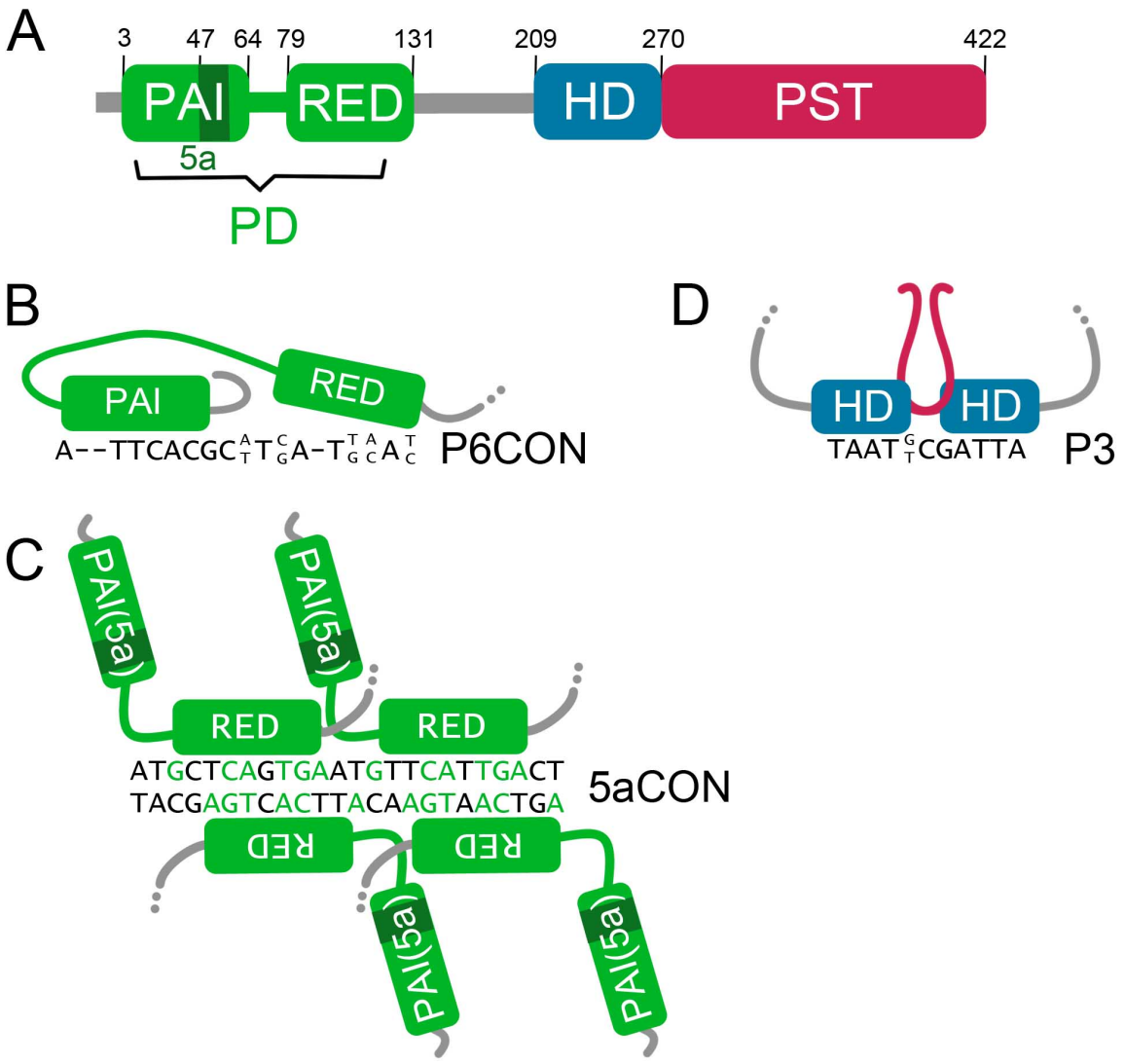


Figure 6. Structure and DNA binding of PAX6

(A) Schematic representation of the PAX6 protein with amino acid (aa) positions of the different functional subunits shown. The *Pax6* gene encodes two DNA binding domains, a paired domain (PD) and homeodomain (HD). *Pax6* also encodes a proline/serine/threonine (PST) rich transactivation domain [177]. The paired domain contains two independent DNA binding subdomains, PAI and RED [181, 182]. A 14 aa insertion into PAI encoded by an alternatively spliced exon, exon 5a, generates the isoform Pax6(5a). The PAX6 nuclear localization signal (NLS) spans the C-terminal of PAI to the N-terminal of RED [185]. (B) In canonical PAX6, PAI is primarily responsible for DNA binding and recognizes the consensus sequence P6CON [183, 184]. (C) Insertion of 5a into the PAI subdomain prevents PAI from participating in DNA binding. Consequently, PAX6(5a) recognizes a different DNA consensus sequence, 5aCON, using the RED subdomain as a dimer [186] or as a tetramer [187]. (D) The HD recognizes a unique DNA motif, P3, as a homodimer [184]. Figure modified from [188].

In addition to encoding DNA-binding domains, the *Pax6* gene encodes two additional functional domains. The carboxy terminus of the predicted PAX6 protein was found to be rich in proline, serine, and threonine (PST) [177]. Similarly, the human *PAX6* gene was found to encode a protein with a high proportion of serine and threonine residues at its C-terminus. This C-terminal domain of PAX6 was shown to transactivate transcription using reporter assays [184, 189, 190] and was referred to as the PST domain transactivation domain (TAD).[189] Additionally, *Gallus gallus* (chicken) PAX6 contains a nuclear localization signal (NLS) that includes the C-terminal region of the PAI subdomain, the linker between PAI and RED, and the N-terminus of RED [185] (Figure 6A).

1.4.2 Expression pattern of *Pax6* during development

Pax6 is expressed in a spatially and temporally restricted pattern during mammalian and more generally, vertebrate, development. The mRNA is present in whole mouse embryo starting at embryonic day 8 (E8) and is most abundant at E10.5 [177]. Additionally, *Pax6* mRNA is detectable in many developing mouse tissues: eyes, retinas, forebrain, thalamus, pituitary, cerebellum, spinal cord, olfactory bulb, and olfactory epithelium [177]. *Pax6* gene activity has also been documented in the developing and mature endocrine pancreas [191] and endocrine cells of the gut [192] using genetic reporters. More recently, PAX6 protein has been identified in osteocytes of developing calvaria and long bone [193] and in adult mouse testis [194] (Figure 7).

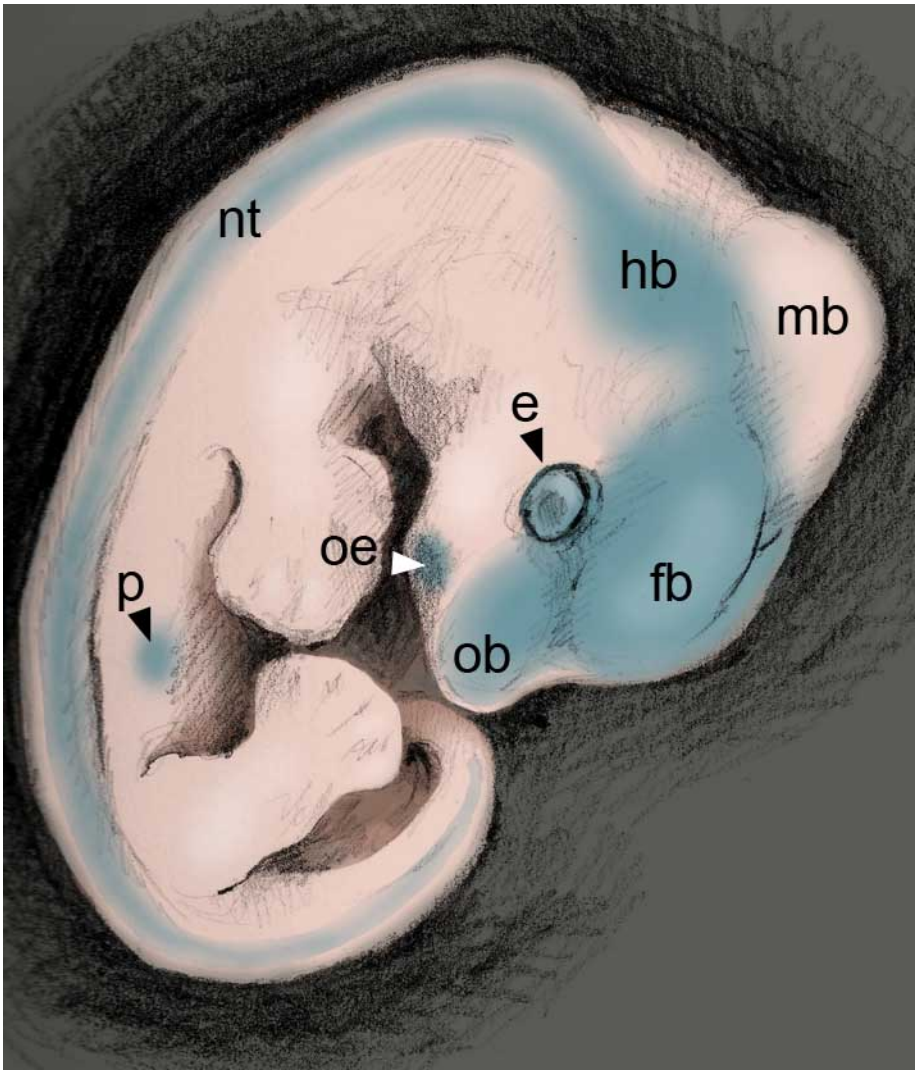


Figure 7. Summary of *Pax6* expression during mouse embryonic development

Illustration of an embryonic day (E) 12.5 mouse showing general pattern of *Pax6* expression in blue. *Pax6* is expressed in developing eye structures (e), in the developing olfactory epithelium (oe) and olfactory bulbs (ob). *Pax6* is also expressed in several regions of the developing brain: the forebrain (fb), hindbrain (hb) and neural tube (nt) of the presumptive spinal cord, but not in the roof of the developing midbrain (mb)[177] (Walther and Gruss). Pax6 is also present in the developing endocrine pancreas (p)[191, 195].

A series of coordinated morphological changes occur during vertebrate eye development and *Pax6* is dynamically expressed during this process. The eye develops

from facial surface ectoderm (SE) and underlying neurectoderm. *Pax6* expression is first detectable in the mouse following gastrulation, at E8 [196, 197], where it is present in a broad region of the head surface ectoderm [177]. By E8.5 the first signs of eye development are visible following neural tube closure as the head neurectoderm evaginates to form two protrusions of the lateral forebrain, the optic vesicles (OV), one on either side of the head [197, 198]. *Pax6* expression is visible in the closed neural tube and OVs of the developing eyes (Figure 8A)[177]. By approximately E9 the OVs then contact the lens-competent SE, inducing the SE to thicken and form lens placode [197]. *Pax6* expression is present in the vesicle epithelium, the optic stalk and the developing lens placode of the SE [177]. By E10.5, the lens placodes internalize to form lens vesicles and the OVs invaginate to form optic cups. The inner layer of the optic cup develops into the neuroretina and the outer portion forms the retinal pigment epithelium (RPE) (Figure 8B)[197, 198]. As embryogenesis progresses, *Pax6* expression continues in the developing retinal neurectoderm, lens and corneal SE. Expression is maintained in specific cell types of these adult tissues [177]. Analysis of *Pax6* expression during differentiation of lens, cornea, and retinal cells in *Danio rerio* (zebrafish) and mouse revealed that PAX6 is present in proliferating cells and that as cells differentiate, PAX6 levels decline [199]. Corneal and lens cells that maintain their proliferative capacity throughout life, and ganglion and amacrine cells of the mature retina continue to express *Pax6* (Figure 8C-D)[199]. Other ocular tissues such as the iris, ciliary body and lacrimal glands also express *Pax6* [200, 201]. Given that miRNAs can function to sharpen boundaries during differentiation, it is possible that negative regulation of *Pax6* during

differentiation of mature ocular cell types that do not express PAX6, such as photoreceptors and bipolar cells of the retina.

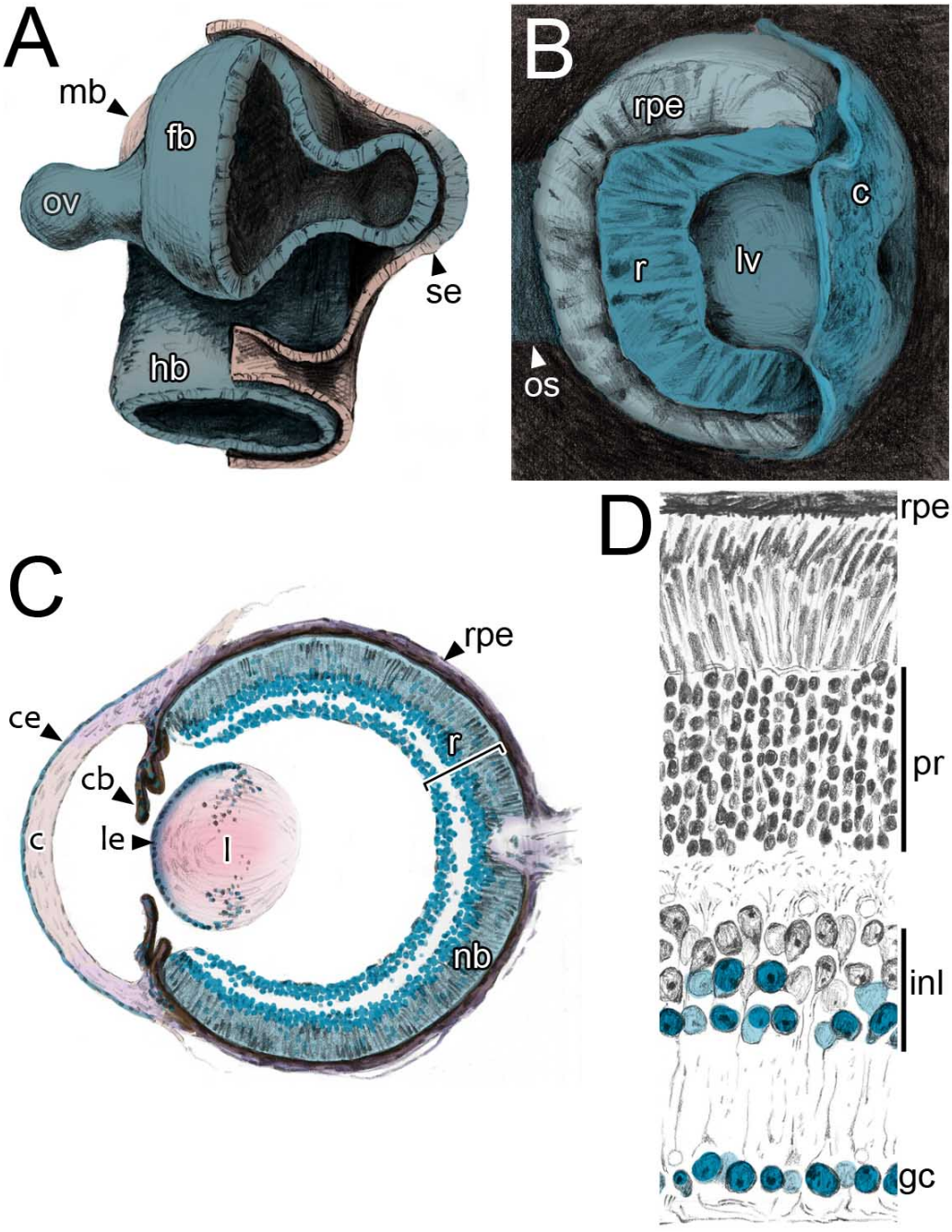


Figure 8. Summary of Pax6 expression during eye development

Illustrations of developing and adult mouse eye showing pattern of *Pax6* expression in blue. (A) Head end of the embryo at approximately embryonic day 9 (E9) showing head surface ectoderm (SE) and underlying neurectoderm. The dorsal surface of the head faces upward. The SE is removed to show optic vesicles, developing forebrain, midbrain and hindbrain. The neurectoderm is removed from the forebrain, optic vesicles and hindbrain to reveal neural tube structure. Note *Pax6* expression is visible in developing ocular structures, the optic vesicles and overlying SE, and that SE has thickened to form lens placode[177]. (B) Three-dimensional cross section through the E10.5 developing eye. *Pax6* is expressed in neurectoderm-derived structures: the optic stalk and optic cup, which consists of the inner developing retina and outer retina pigmented epithelium (RPE). *Pax6* is also expressed in surface ectoderm-derived structures, the lens vesicle and presumptive cornea. Note lower level of *Pax* expression in developing RPE [177]. (C) Cross section through the postnatal day 1 (P1) eye. *Pax6* is expressed in retinal progenitor cells of the neuroblastic layer and differentiated ganglion and amacrine neurons of the retina. Note low level of expression in RPCs. *Pax6* expression is also maintained in the ciliary body, corneal epithelium and lens epithelium [199, 202]. (D) Cross section through adult retina showing *Pax6* expression is retained in ganglion cells and amacrine cells of the inner nuclear layer. PAX6 is absent from bipolar cells of the inner nuclear layer, photoreceptor cells and RPE [199]. List abbreviations: c, cornea; cb, ciliary body; ce, corneal epithelium; fb, developing forebrain; gc, ganglion cell; hb, developing hindbrain; inl, inner nuclear layer; l, lens; le, lens epithelium; lv, lens vesicle mb, developing midbrain; nb, neuroblastic layer; os, optic stalk; ov, optic vesicle; pr, photoreceptors; r, retina; rpe, retina pigmented epithelium; se, surface ectoderm.

Following neural tube closure, *Pax6* is expressed in distinct patterns in the developing brain and spinal cord. *Pax6* mRNA and protein are present in the neuroepithelium of the developing forebrain, the telencephalon and diencephalon, and hindbrain, but is absent from the developing midbrain [177, 203, 204]. This expression pattern is maintained through cortical development. Specifically, *Pax6* is expressed in the dorsal and lateral walls of the telencephalon in neural progenitors of the ventricular zone (VZ). Here, PAX6 protein can be detected in a rostralateral-high to caudomedial-low gradient [205]. PAX6 is maintained in the proliferating neural progenitors of the

ventricular-subventricular zone (V-SVZ) postnatally, where it is expressed in a dorsal-high to ventral-low gradient [206]. Negative regulation of *Pax6* by miR-7 is proposed as a mechanism by which the spatial domain of PAX6 is established in the V-SVZ [206]. *Pax6* mRNA and protein are also expressed in the VZ of the developing spinal cord in a dorsal-high to ventral-low gradient [177, 207]. Given its role in the V-SVZ, miR-7 may also be playing a role in regulating *Pax6* in the developing spinal cord. Like the retina, most neurons in the brain and spinal cord stop producing PAX6 as they differentiate. However, some PAX6-positive cells in the adult forebrain and hindbrain co-express the neuron-specific protein Neuronal nuclei (NeuN) [208], and a sub-population of spinal cord interneurons continue to express PAX6 [209].

Two primary structures comprise the olfactory system, the olfactory epithelium and olfactory bulbs, and *Pax6* is expressed in both developing structures. Development of the nasal cavity and olfactory epithelium (OE) begins following closure of the neural tube, at approximately E9 in the mouse. The first indication of nasal development are thickenings of the head SE laterally on either side of the head, the nasal placodes [210]. The nasal placodes invaginate to form olfactory pits at E10.5 [210] at which point they begin to express *Pax6* [177]. *Pax6* continues to be expressed in the epithelium lining the nasal cavity through embryogenesis [177], and expression is retained in neurons, duct cells, support cells and stem cells of the adult OE [211]. Development of the olfactory bulbs (OBs) begins with specification of OB primordium from the rostral telencephalon. This is followed by evagination to form OB-like structures (olfactory lobes) and arrival of olfactory sensory neuron axons from the OE at E12.5 [210, 212]. *Pax6* is expressed in

olfactory lobes starting at approximately E12.5 [177]. Expression of *Pax6* continues in a subset of interneurons in the glomerular layer (GL) of the olfactory bulbs into adulthood [213], particularly the dopaminergic periglomerular neurons (PGNs) [214, 215]. Negative regulation of *Pax6* by miR-7 may be important for regulating specification of dopaminergic PGNs [206].

Most *Pax6*-expressing tissues are ectodermal in origin, with the most well-studied exception being the endocrine pancreas. The pancreas is a gland composed of both exocrine and endocrine cells and the endocrine pancreas plays an essential role in glucose metabolism and homeostasis throughout life [216]. The mature endocrine pancreas is composed of spheroid clusters of four cell types embedded in exocrine acinar cells: the α-, β-, δ- and PP-cells, expressing glucagon, insulin, somatostatin and pancreatic polypeptide, respectively [216]. During embryonic development, the endocrine pancreas also contains rare ghrelin producing cells, the ε-cells [216]. The presumptive pancreas is defined at E8.5, and pancreas organogenesis begins at E9.5 with evagination of dorsal and ventral buds from the foregut-midgut junction of the endoderm [216]. *Pax6* gene expression is first detected in the pancreas at E9 and by E10.5 colocalizes with cells expressing glucagon [191, 195]. Insulin-producing cells are detected at E15.5, and express PAX6 [195]. During pancreas organogenesis, miR-7 is co-expressed with *Pax6* and may function to refine the proportion of endocrine pancreas cell types formed [217]. *Pax6* continues to be expressed in the mature endocrine cells of the islets of Langerhans but is absent from the exocrine tissue [191].

1.4.3 Evolutionary conservation of Pax6 sequence and function

The sequence of the *Pax6* protein coding region is highly conserved across vertebrates and even between vertebrates and invertebrates. The proteins encoded by human and mouse *Pax6* genes are identical, and the amino acid sequence is 97% conserved between mice and zebrafish [218, 219]. Impressively, the *Pax6* homologue in *Drosophila, eyeless* (*ey*), is 94% identical in its amino acid sequence to mouse and human PAX6. Additionally, the PD and HD are 80-90% conserved between mammals and *Drosophila* [219–221].

In addition to a high degree of sequence conservation in its coding region, *Pax6* is highly conserved in expression pattern and function. *Pax6* expression in the retina, lens, and cornea is conserved among vertebrates [199]. These conserved patterns of expression are likely due to conserved cis-regulatory elements in the DNA. Several regulatory elements located within the mouse *Pax6* locus have high sequence identity between mouse, quail and pufferfish [222]. Additionally, transgenic mice expressing reporters driven by human or pufferfish *Pax6* regulatory sequences demonstrate conserved patterns of expression during mouse development [222, 223]. *Drosophila ey* is expressed in homologous structures to its vertebrate counterpart: it is present in the eye primordia and developing eye imaginal discs, and is also expressed in the nerve cord, optic lobes, and regions of the brain [220]. Misexpression of the mouse *Pax6* gene in various imaginal disc primordia of the *Drosophila* embryo causes the development of ectopic eyes on the wings, legs, or antennae [224], suggesting that PAX6 function is conserved across mammals and flies. The eye specification transcription factor network

[225, 226] and expression pattern of *Pax6* regulatory elements [227] are also highly conserved. Though there is no significant sequence identity between vertebrate and invertebrate *Pax6* non-coding regions, putative binding sites for PAX6 and basic-helix-loop-helix (bHLH) transcription factors are shared between vertebrates and invertebrates [228]. Despite the lack of sequence conservation between regulatory regions, insertion of a *Drosophila ey* enhancer upstream of a reporter in transgenic mice reproduces major features of the endogenous pattern of *Pax6* expression, directing protein expression to the developing eye and CNS [227].

The dynamic and specific expression pattern of the *Pax6* gene and protein during development and in adulthood suggests that PAX6 may be playing important roles in tissue specification, progenitor cell maintenance, differentiation and maintenance of differentiated cell identity. Due to the high degree of conservation in both sequence and function, many organisms can be used as models to dissect the function of PAX6 during human development and in human disease.

1.5 PAX6 function and dosage sensitivity

1.5.1 *Pax6* mutations provide a window into its developmental role

Mutations resulting in non-functional PAX6 proteins have provided a window into the function of Pax6 during development and adulthood. *Small eye* (*Sey*) is a semidominant mutation that has arisen independently in several mouse [229, 230] and rat [231] populations. Eyes and nasal cavities completely fail to develop in homozygous

mutants [230]. During the initial stages of eye development in homozygous mutants, the lens placode fails to develop and subsequently the optic cup degenerates. As a result of defective nasal cavity development, homozygotes die shortly after birth (Figure 9A)[230]. The *Sey* phenotype is the result of mutations in the *Pax6* gene, demonstrating that PAX6 is essential for eye and nasal development [230].

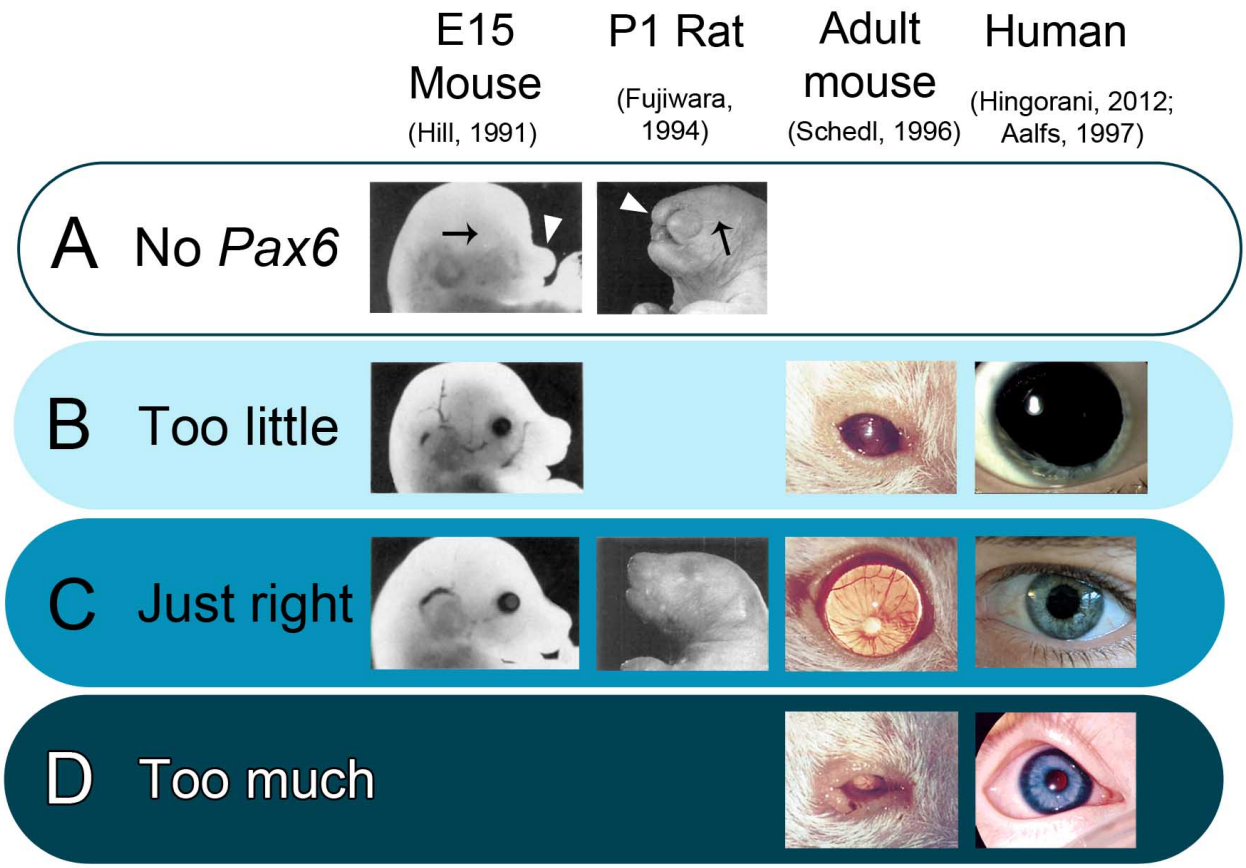


Figure 9. Sensitivity of eye development of *Pax6* dosage

(A) Eyes fail to develop in mice and rats completely lacking *Pax6*. The absence of ocular structures is clearly visible at embryonic day 15 (E15) in the mouse embryo and in postnatal day 1 (P1) rats (arrow). Additionally, nasal cavities fail to develop, which is visible as a shortened shout in the mouse embryo and clearly visible in the newborn rat (white arrowhead)[230, 231]. (B) Mutations in one copy of the *Pax6* gene cause haploinsufficiency. Decreased *Pax6* dosage produces a small eye phenotype in rodents, which is visible in the E15 mouse embryo and more dramatically in the adult [230, 232]. In humans, insufficient *PAX6* dose causes the disease aniridia, which is characterized primarily by complete or partial loss of the iris [233, 234]. (C) Proper *Pax6* dose is required for normal eye development. (D) Overexpression of *PAX6* can be deleterious to

eye development. Transgenic mice carrying multiple extra copies of *PAX6* display microphthalmia [232] and humans harboring duplications in the region of chromosome 11 containing the *PAX6* gene have mild ocular defects, such as iris translucency [235, 236]. Images of mouse embryos are modified from Hill *et al*., 1991; images of newborn rats are modified from Fujiwara *et al*., 1994; adult mouse images are modified from Schedl *et al*., 1996; image of the eye from a patient with aniridia is from Hingorani *et al*., 2012; image of the eye from a patient having a duplication of the *PAX6* gene is from Aalfs *et al*., 1997.

These naturally occurring mutations in the *Pax6* gene of rodents provide information about the role of this transcription factor during eye development. The progression of eye development can be compared between wild-type and mutant rodents to elucidate the role of Pax6 in this process. *Pax6* is expressed in both the OV and SE of wild-type mice [177, 237]. Though evagination of neurectoderm into OV occurs independent of *Pax6* expression, retinal development arrests at this stage [231, 237]. PAX6 is required for optic cup formation and differentiation of the neural retina and RPE. Interestingly, progenitor proliferation in *Pax6*-null OVs is reduced and accompanied by precocious differentiation into neurons between E9.5 and E10.5. These neurons do not express markers of differentiated retinal ganglion, amacrine or bipolar cells, and most do not survive past E13.5 [238]. Expression in the SE is also required for lens induction, as *Pax6*-null animals fail to form lens placode and develop lenses [231, 237].

Developmental events are sensitive to *Pax6* gene dose (Figure 9B). Mice carrying heterozygous mutations in the *Pax6* gene have defects in eye development and have small eyes [230]. Heterozygous *Sey* mice also have reduced body size and lens size compared to wild-type [229]. Normal eye development depends on PAX6 meeting a defined concentration threshold, and consequently lens placode formation is also

delayed in *Pax6* heterozygous mouse mutants [239]. These findings demonstrate that PAX6 functions in a semidominant manner and that eye development is very sensitive to levels of PAX6 protein.

PAX6 is also required for the development of the olfactory system and CNS. Similar to induction of lens placode, *Pax6* homozygous mutant mice and rats fail to develop nasal placode from presumptive nasal ectoderm. Additionally, the nasal pit, OE [237, 241] and OBs fail to develop [242]. Normally, the OBs develop as protrusions from the rostral telencephalon and initially consist of mitral cells, the output neurons that project from the olfactory bulb to the olfactory cortex. In *Pax6* homozygous mutants, OB-like structures form on the lateral telencephalon [241, 243]. Mitral cell genesis in the rostral telencephalon is normal between wild type and *Pax6* mutants, however cells migrate laterally in the mutants [244]. Transplantation of mitral cell progenitors into wild type (WT) and mutant embryos reveals that this mitral cell migratory defect is non-cell autonomous [244]. Though the olfactory system appears to develop normally in *Pax6* heterozygous mice, main OB mass and production of dopaminergic interneurons are reduced [245]. Gonadotropin-releasing hormone neurons are also specified from cells of the developing OE and migrate to the hypothalamus. These cells are completely absent in *Pax6*-null mice [246]. Defects in cortical development are also observed in *Pax6*-null mice, notably cortical thinning [242] and defective forebrain patterning [247, 248]. Finally, PAX6 plays a role in spinal cord development and is involved in specifying motor neuron and interneuron identity in the neural tube. PAX6 is expressed in dorsal neural tube progenitors and in a dorsal-high to ventral-low gradient in ventral neural

tube progenitors. Here, PAX6 is important for specifying dorsally derived hypoglossal motor neurons (hMNs), and ventral interneurons V1 and V2. *Pax6*-null mice have altered neural tube patterning and the generation of hMNs, V1 and V2 interneurons is impaired [207].

Development of the endocrine pancreas in also impacted by *Pax6* LOF. *Sey/Sey* mice thus have disorganized islets of Langerhans [191, 195]. Additionally, all endocrine cell types are reduced in number [191, 195], with glucagon-producing alpha-cells being completely absent [191]. Similar results were observed in *Pax6*-null rats [249]. Conditional inactivation of *Pax6* during endocrine pancreas development, starting at E9.5, results in hyperglycemia, hypoinsulinemia, reduction in α- and β-cells, loss of glucose transporter 2 (Glut2) expression in β-cells, a severe diabetic phenotype and death several days after birth [250]. PAX6 is also important for glucose homeostasis in adulthood. Induced conditional knockout of *Pax6* in the adult mouse endocrine pancreas produces a severe diabetic phenotype [251]. Additionally, though *Pax6* haploinsufficient mice do not present defects in pancreatic development, they display impaired proinsulin processing and glucose metabolism [252].

1.5.2 Regulation of *Pax6* during neural progenitor proliferation and differentiation

PAX6 is involved in neuronal proliferation, differentiation, cell cycle progression, and maintenance of progenitor multipotency. Small eyes and cortical defects observed in *Sey* mice suggest that decreased *Pax6* expression leads to disruptions in proliferation and differentiation [242]. Concomitant with differentiation, PAX6 protein levels are down-regulated in lens and most retinal cell types [199]. However, PAX6 is maintained

in lens, corneal and ciliary epithelial cells that retain proliferative capacity [199, 202], and expression is necessary for maintaining this proliferative capacity [202] further suggesting that *Pax6* plays a role in proliferation and differentiation. Conditional knockout of *Pax6* only in peripheral retinal progenitor cells (RPCs) of developing optic cup was used to address the role of PAX6 specifically in RPCs [253]. RPCs lacking PAX6 show reductions in proliferation and have limited proneural factor expression [253, 254]. In contrast to loss of *Pax6* at the optic vesicle stage, mutant retinal tissue lack ganglion, photoreceptor, bipolar, horizontal and Muller cells, and produce only amacrine interneurons [253, 254], demonstrating that PAX6 is important for proliferation and maintenance of RPC multipotency.

During CNS development, multipotent progenitors must undergo a transition from symmetric divisions that increase the progenitor cell pool to asymmetric divisions that give rise to cells that exit the cell cycle and initiate neuronal differentiation. Comparison of cell cycle length and neural progenitor differentiation between WT and *Sey* mouse cortices reveals that PAX6 plays a role in regulating cell cycle progression and the rate of transition from symmetric to asymmetric cell division. Cell cycle length was shortened in *Pax6*-null cortical progenitors accompanied by precocious differentiation [255]. Regulating the number of post-mitotic cells produced in the cortex and retina requires tight control of both cell cycle length and the number of cells that re-enter the cell cycle versus those that differentiate [256]. Depletion of PAX6 early in mammalian corticogenesis results in precocious neuronal differentiation, depletion of the progenitor pool, and a subsequent reduction in cortical size [256]. Similarly, the absence of PAX6 in

developing *Gallus gallus* (chick) neural tube leads to precocious differentiation of neural progenitors [257]. Additionally, chick retinal progenitor cells transfected with siRNA against *Pax6* reduces the number of cycling cells, reduces the number of differentiated ganglion cells, and enhances photoreceptor and amacrine cell production [258]. These results suggest that tightly controlled levels of *Pax6* expression are required to maintain the proliferative capacity of neural progenitors and for correct cell fate decisions.

1.5.3 *PAX6* and human disease

In addition to phenotypes associated with *Pax6* heterozygous mutations in rodents, *PAX6* mutations in humans are also associated with disease. Aniridia is a human congenital panocular disorder characterized by complete or partial absence of the iris, foveal and optic nerve hypoplasia, cataracts, early onset glaucoma, and vision impairments [259]. Congenital aniridia has an incidence between 1:64000 to 1:96000 and 67% of affected children have a parent with the disease The remaining 33% have sporadic mutations [233]. This disease is autosomal dominant and phenotypic expression can vary among individuals [233]. After its initial characterization, aniridia was found to be caused by heterozygous mutations in the *PAX6* gene, suggesting that aniridia may be the result of insufficient PAX6 [260] (Figure 9B).

A diversity of mutations is associated with aniridia. Classical aniridia is associated with nonsense mutations in the *PAX6* gene, leading to a truncated protein [179]. Many independent missense mutations in the gene sequence encoding the conserved PD have also been identified. Additionally, some mutations impact the HD sequence or affect alternative splicing [179]. These different missense mutations can differentially impact

PAX6 DNA binding and transactivation, and the impact of the mutation may also be dependent on cell type [261]. These observations lead to the suggestion that the more moderate phenotypes associated with these missense mutations are due to altered PAX6 function in a subset of PAX6-expressing cells [261]. Missense mutations have been identified in the TAD, which impact binding of the HD to its DNA consensus sequence without effecting transactivation [262]. The variable phenotypes seen in people with aniridia may be the result of these different mutations [179].

The phenotypic effects of aniridia are not confined to the eye. Structural magnetic resonance imaging (MRI) and tests of olfactory function demonstrate that many people with aniridia also have cerebral abnormalities and reduced olfaction [263]. People with aniridia display glucose intolerance due to impaired insulin processing and secretion [252, 264], leading to the suggestion that heterozygous *PAX6* mutations may be associated with risk for maturity onset diabetes of the young (MODY). Additionally, people with aniridia have higher rates of obesity compared to siblings that lack the disease [265].

1.5.4 Overexpression and ectopic expression of Pax6

Experiments involving overexpression or ectopic expression of *Pax6* during development provide additional clues as to its role in eye development. Misexpression of *Pax6* or the fly homologue *ey* in various imaginal disc primordia of the developing *Drosophila* embryo induce ectopic eye development on other structures [224]. Misexpression of *Pax6* in *Xenopus laevis* (frog) causes formation of fully differentiated ectopic eyes, typically on the dorsal side of the head [266]. Additionally, *Pax6*

misexpression induces ectopic expression of other genes involved in eye development, and ectopic retinas express molecular markers of the major differentiated retinal cell types [266]. These results indicate that PAX6 functions in a cell-autonomous manner and is necessary and sufficient to trigger eye development [266].

Given that *Pax6* and *ey* can induce ectopic eyes and ectopic expression of other proteins required for eye development they were named master control genes for eye morphogenesis in vertebrates and flies, respectively [221]. More accurately however, *Pax6* is part of a coordinated genetic hierarchy [226]. Mutations in other components of the eye-specification transcription factor network can disrupt eye development, and ectopic expression of these other circuit components can promote the development of eye structures. For example, ectopic expression of Six3, a transcription factor downstream of PAX6, induces the formation of ectopic OV-like structures [267], and *Xenopus* embryos misexpressing the transcription factor Rx develop ectopic retinal tissue [268]. The coordinated overexpression of all eye field transcription factors, those being: Otx2, ET, Pax6, Six3, Rx1, and Optx2, in *Xenopus* is sufficient to generate ectopic eyes outside the nervous system [226]. Misexpression of subsets of these transcription factors reveal that PAX6 is the most important circuit component with respect to generating ectopic eyes [226].

Not only is *Pax6* heterozygosity deleterious to normal development, increased PAX6 dosage can also be problematic (Figure 9D). Transgenic mice carrying five to seven copies of a human *PAX6*-bearing yeast artificial chromosome (YAC) display microphthalmia and lack photoreceptors [232, 269] and have cell autonomous defects

in late cortical progenitor proliferation, resulting in decreased thickness of superficial cortical layers [270]. Additionally, transgenic mice overexpressing *Pax6* during early pancreas development display perturbed development of the endocrine pancreas, β-cell apoptosis, and impaired glucose stimulated insulin secretion [271]. A few cases of *PAX6* gene duplication in humans have been reported, in which a band of chromosome 11, including *PAX6*, *Wilm's Tumor 1* (*WT1*) and *Elongation protein 4* (*ELP4*) genes, was duplicated causing mild ocular defects and mental retardation [235, 236]. These findings suggest that increased *PAX6* dosage in humans may be also deleterious. Taken together, these findings demonstrate that correct *Pax6* gene dose is critical for normal development.

1.6 Regulation of *Pax6* expression and function

1.6.1 Complex spatiotemporal control of *Pax6* expression

PAX6 is expressed in a complex spatial and temporal pattern, and correct maintenance of both expression pattern and protein dosage are very important for normal development and adult maintenance of many tissues. The pattern of expression, dosage and function of PAX6 has been shown to be regulated at every step along its biosynthesis pathway, from regulation of transcription, to mRNA processing and stability, to protein decay and post-translational modification.

The *PAX6* gene has a complex genomic organization. Initial characterization of the human *PAX6* gene found it to contain 14 exons, named 0-13 [260]. Translation of

Pax6 beigns in exon 4, the PD is encoded by exons 4-7 and the HD encoded by exons 8-10. The PST-rich TAD is encoded by exons 10-13 and the 3' end of exon 13 encodes the 3'UTR [177, 222]. Two additional exons were later identified, exon α and exon 5a, which are generated by alternative promoter use [222] and alternative splicing [260], bringing the total to 16 exons (Figure 10A).

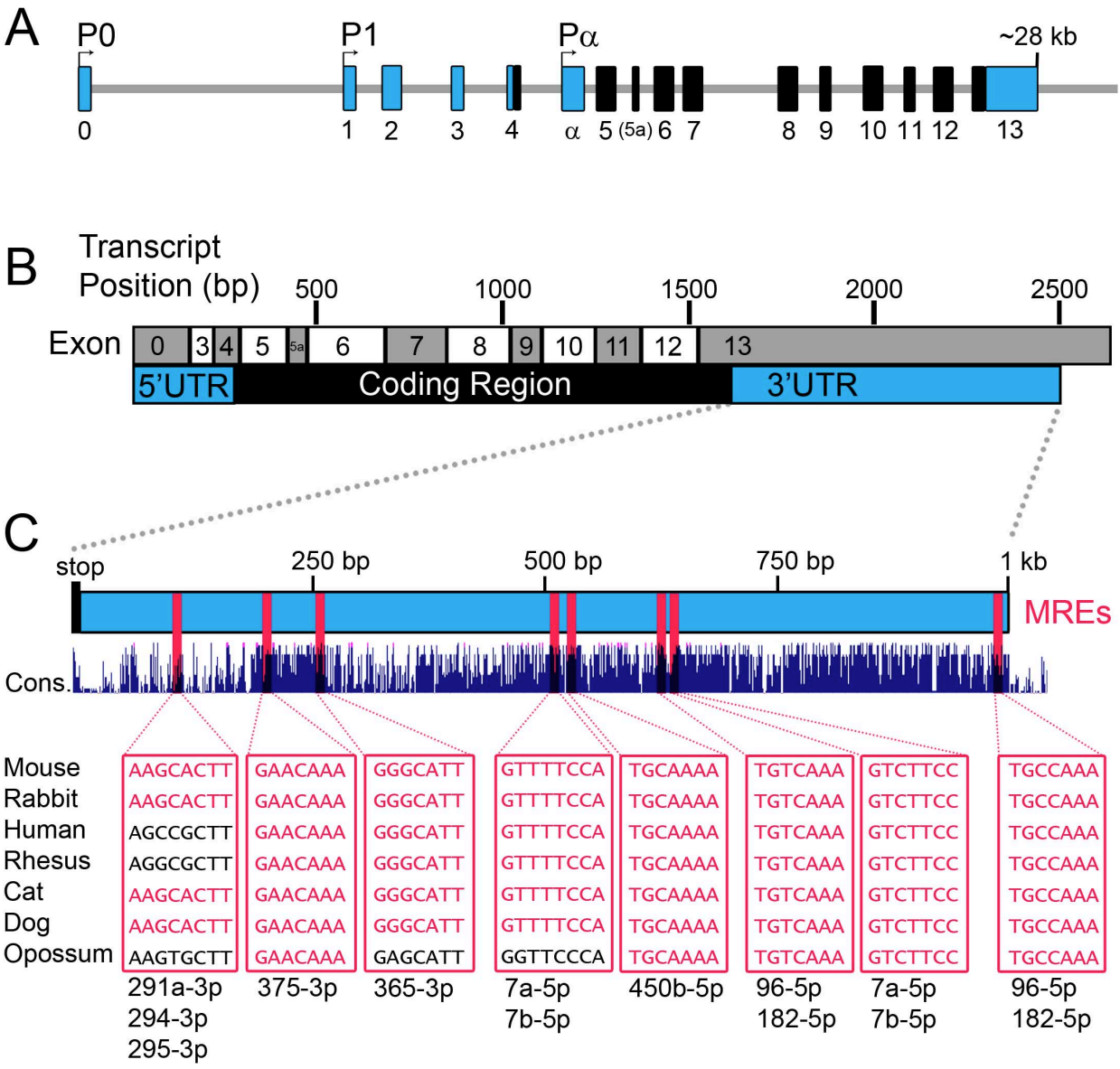


Figure 10. Gene and 3'UTR structure of *Pax6*

(A) Organization of mouse *Pax6* genomic locus. Non-coding exons are coloured blue and coding exons are black. Three transcription start sites, P0, P1 and Pα[222], and an alternatively spliced exon, 5a[177], are shown. *Pax6* contains 16 exons spread over approximately 28 kilobases (kb). Image modified from [188]. (B) Schematic of the mature *Pax6* mRNA (transcript 206, Ensemble, ENSMUST0000011108610) generated from P0 containing exon 5a. (C) Schematic of the mouse *Pax6* 3'UTR showing locations of microRNA recognition elements (MREs) in pink[206, 217, 272–279]. The placental mammal base-wise conservation is generated by the University California Santa Cruz (UCSC) genome browser[280]. Larger peaks indicate greater selective conservation. MRE sequences for miRNAs previously shown to regulate *Pax6* are shown for several species. Pink text shows conserved MRE sequences.

Pax6 transcription is regulated though multiple promoters. Two promoters located upstream of the *Pax6* translation start site, P0 and P1, were initially identified [222, 227, 281]. P0 and P1 derived transcripts are translated into the same protein but have different 5'UTR sequences [222, 281]. A third conserved transcription start site, Pα, was identified between exons 4 and 5 [222]. Transcription from Pα produces an alternative 5'UTR from exon α and results in a PAX6 isoform with a truncated paired domain called Pax6ΔPD (Figure 10A)[282, 283]. Overexpression of Pax6ΔPD causes microphthalmia in rodents and exacerbates rather than rescues the *Sey* phenotype [283], suggesting that the PAX6ΔPD isoform has distinct functions from canonical PAX6. Interestingly, PAX6ΔPD was shown to bind full-length PAX6 and enhance transcriptional activation through PAX6 [284]. Transcription from these three promoters is differentially regulated during development [227, 281, 285].

PAX6 is also regulated through alternative splicing. Previous characterization of the *Pax6* cDNA from mouse [177] and zebrafish [218] identified 42 base pair (bp) insertion in the paired domain encoding an additional 14 amino acids (aa). This insertion was later identified as an alternatively spliced exon, exon 5a (Figure 10A)[260]. 5a alters the DNA sequence recognition of the PD [286] and transcriptional activity of the PAX6 protein. Normally, the PAX6 PD binds DNA primarily using the PAI subdomain [182]. However, insertion of these 14 aa into PAI generates a PAX6 isoform that can recognize DNA exclusively using the RED subdomain (Figure 6C)[186]. The intact PD primarily recognizes the consensus sequence P6CON [182, 183] and with disruption of the RED subdomain, the PAX6 PD instead binds a different tetrameric consensus sequence,

5aCON, as a dimer [186] or as a tetramer [187]. Thus, PAX6 versus PAX6(5a) have distinct DNA-binding activities and altered transactivation potential [287]. Consequently, alternative splicing provides a mechanism to diversify the sequence motifs recognized by PAX6 and may enable PAX6 to switch its complement of regulated genes. The different isoforms are expressed in different relative proportions during chick retinal development [288] and mouse brain development [289]. Potential roles of PAX6 versus PAX6(5a) are revealed by missense mutations that disrupt either PAI or RED subdomains of the PD. Missense mutations in the PAI subdomain are associated with anterior segment defects in humans [290] whereas mutations in the RED subdomain are associated with foveal hypoplasia [291]. Cumulatively, these data suggest the two isoforms transactivate different target genes and have different developmental roles.

Appropriate gene function is dependent on more than just the coding sequence but also on non-coding regulatory elements. Numerous cis-regulatory elements were identified within the *Pax6* locus that are responsible for controlling the spatiotemporal expression pattern of *Pax6* [222, 227, 281, 292–295]. These cis-regulatory elements control transcription from P0, P1 and Pα and are located both upstream of the *Pax6* gene and within introns [188]. Transgene reporter assays were used to determine the distinct expression patterns of Pax6 cis-regulatory elements. Many of these enhancers drive *Pax6* expression in partially overlapping patterns, providing precision and robustness [188].

In addition to proximally located cis-regulatory elements, key developmental genes are often regulated by highly conserved elements located hundreds of kilobases

(kb) away from transcription start sites [296]. For example, other elements involved in regulating *Pax6* transcription were identified at a distance from the 3' end. Presence of these enhancers was first observed in individuals having either translocations in the *PAX6* gene that left the gene intact but still produced aniridia phenotypes [297–299] or small chromosomal deletions 3' to the *PAX6* gene [300]. Additionally, proximal enhancers for *Pax6* are insufficient to produce the complete pattern of Pax6 expression [283, 294] leading to the hypothesis that regulatory elements for *PAX6* also reside 100+ kb downstream of the *PAX6* gene. Location and function of these elements was confirmed by sequence conservation and transgenic reporter expression [283, 294, 295, 298, 301]. In sum, the *Pax6* gene contains a complex set of regulatory elements to control spatial and temporal aspects of its transcription.

1.6.2 Functional regulation of the PAX6 protein

In addition to alternative promoter use and splicing, which can generate different PAX6 isoforms having distinct functions, post-translational modification of PAX6 can influence transcriptional transactivation by PAX6. Reversable addition and removal of phosphate groups by protein kinases and phosphatases is commonly used at alter protein function [10]. Several sites of phosphorylation are present within the TAD that enhance transcriptional transactivation by PAX6 [190, 302]. Additionally, system-wide analysis of protein phosphorylation signatures during neurite growth [303] and mitosis [304] identified multiple sites of phosphorylation in the PAX6 PD and HD [305]. The specific role of phosphorylation of PAX6 at these locations remains unknown. PAX6 can also be dephosphorylated to dampen its transcriptional activity [306].

PAX6 can be modified by the addition of other small molecules, ubiquitin, the small ubiquitin-like modifier (SUMO) and carbohydrates, to alter function and stability. Large-scale proteomics analysis identified a site of ubiquitinylation in the PD of PAX6 [305, 307]. PAX6 can bind the E3 ubiquitin ligases Trim11 [308] and Mid1 [309], resulting in ubiquitination and degradation of PAX6. Ubiquitinylation of PAX6 may play an important role in regulating neurogenesis in mammals [308] and eye morphogenesis in *Xenopus* [309]. Binding of PAX6ΔPD to its HD consensus sequence, P3, is enhanced by sumoylation [310]. Vertebrate PAX6 also contains many predicted glycosylation sites [311] and quail PAX6 can be modified by addition of O-N-acetylglucosamine though the specific role of this modification remains unknown [312].

The transcriptional activity of PAX6 can be influenced by its interaction with other proteins. PAX6 can alter expression of target genes by binding proteins involved in histone modification and chromatin remodeling [313, 314]. *In vitro* co-immunoprecipitation (co-IP) and reporter assays reveal that PAX6 may interact with different homeodomain-containing proteins, such as Pbx1, HoxB1 and Chx10, and that these interactions may enhance transactivation by PAX6 [284]. PAX6 can function cooperatively though direct interactions with other transcription factors during dopaminergic fate specification in the olfactory bulb [315] lens development [316] and to control expression of genes encoding endocrine pancreas hormones [317, 318].

1.6.3 Regulation of *Pax6* at the level of mRNA turnover and protein synthesis

Given the importance of maintaining the correct expression pattern and dosage of PAX6, post-transcriptional regulation of *Pax6* by miRNAs may represent an important regulatory mechanism. This possibility was hinted at early on, with the discovery that the *Pax6* 3'UTR sequence is highly conserved. Glaser *et al.,* (1992) noted that the *Pax6* 3'UTR has 88% sequence identity between human and mouse, and positions 619-732 in the human *PAX6* 3'UTR are 99.1% identical to mouse and 89.4% identical to zebrafish [260]. The 3' end of the 3'UTR is better conserved between mouse and zebrafish than the coding sequence [218]. These observations suggest that the 3'UTR is playing an important role in regulating translation or stability of the *Pax6* transcript (Figure 10C).

Non-coding mutations located 3' to the *PAX6* coding sequence are recognized to be associated with aniridia. In addition to mutations in enhancer regions located over 100 kb downstream of the *PAX6* gene [319], some mutations the impact the 3'UTR of the *PAX6* transcript have been observed. Mutations that change the stop codon and cause the C-terminus to run into the 3'UTR have been identified in people with severe aniridia phenotypes [320]. For example, a 103 bp deletion was identified in a family with aniridia that produces a frameshift mutation and C-terminal extension of the *PAX6* coding region into the 3'UTR [321]. Though this mutation does alter the C-terminus of the PAX6 protein, extension of the stop codon into the 3'UTR could alter miRNA~~s~~ interaction with any MREs in the affected 3'UTR [34]. The region of 3'UTR converted into coding sequence contains predicted MREs for miR-375 and miR-365 [321]. An additional aniridia-associated mutation is present in the *PAX6* 3'UTR, though its impact

on miRNA regulation is unknown [322]. These observations further suggest that the *PAX6* 3'UTR is playing an important regulatory role.

Table 2. Summary of miRNAs predicted to regulate *Pax6*

MicroRNA	Position* (human)	Position* (mouse)	Proposed function
291a 294 295	NA	112	Promote early mesoderm and endoderm specification at the expense of ectoderm specification of mouse embryonic stem cells
96 182	660, 975	636, 958	Promote epidermal specification and prevent neural specification of human embryonic stem cells Promote proliferation and invasion in glioma
365	266	260	Inhibit astrocyte to neuron conversion following stroke Tumor suppressor, downregulated in retinoblastoma
7	537, 681	517, 655	Refine PAX6 gradient in the V-SVZ and negatively regulate dopaminergic periglomerular neuron fate in the olfactory bulb Promote oligodendrogenesis at the expense of neurogenesis during cortical development Promote the formation of ghrelin positive cells at the expense of insulin and glucagon positive cells during endocrine pancreas cell types
375	207	201	Inhibit specification of dopaminergic neurons in the frog accessory olfactory bulb in response to non-kin odorants Promote motor neuron differentiation of human embryonic stem cells
450b-5p	571	551	Prevent ocular commitment of eyelid epithelium
223	4854	NA	Promote proliferation and invasion in glioma
328	2500	NA	Increases risk for extreme myopia by negatively regulating *PAX6* Protects against Rolandic epilepsy by negatively regulating *PAX6*

* Position refers to the nucleotide starting position of the MRE in the 3'UTR. Position 1 is the first nucleotide following the stop codon.

Several miRNAs may play roles during temporal transitions in cell fate by directly targeting *Pax6* (Figure 10C, Table 2). The miR-290-295 cluster (miR-291a, 294 and 295) is highly expressed in ESCs [323] and *in vitro* work in mouse ESCs (mESCs) suggests that the miR-290 family of miRNAs facilitate early mesendoderm lineage specification by targeting the pro-ectodermal *Pax6.* miR-290-295-null mESCs have elevated levels of PAX6 protein and when induced to differentiate are biased toward expressing ectodermal markers [272]. During *in vitro* differentiation of human ESCs (hESCs) into neuroectodermal and epidermal cells, the miR-96 family of miRNAs (miR-96, 182, 183) may target *PAX6* in epidermal-fated ectodermal cells. In so doing, they may help sharpen the boundary between cells fated to become epidermis versus neuroectoderm. miR-96 and 182 are down-regulated in hESCs upon induction of neuroectodermal fate and overexpression of miR-96 in hESCs represses expression of neural markers [273]. Similarly, miR-7 was may promote oligodendrogenesis at the expense of neurogenesis *in vitro* by targeting *Pax6.* miR-7a is expressed at high levels in oligodendrocyte precursor cells relative to neural progenitor cells (NPCs). Overexpression and inhibition of miR-7 in NPCs promotes and represses oligodendrogenesis, respectively [274]. Following stroke, miR-365 repression of *Pax6* in astrocytes may impair conversion of astrocytes to neurons [275]. Following cerebral ischemia in rats, miR-365 expression is upregulated in ipsilateral striatum, and either miR-365 antagomir or PAX6 overexpression causes cells expressing an astrocyte-specific reporter to switch to expressing markers of mature neurons. miR-375 may influence specification of neurotransmitter cell fates during *Xenopus laevis* (frog) accessory olfactory bulb (AOB) development [276]. The proportion

of neurons expressing tyrosine hydroxylase (TH) or gamma-aminobutyric acid (GABA) in the frog AOB can be influenced by exposure to kin or non-kin odorants, with exposure to kin odorants resulting in a switch in neuronal phenotype from GABA-expressing to TH-expressing. PAX6 is associated with dopaminergic differentiation [324], and knocking down *Pax6* blocks adoption of TH fate by AOB neurons. miR-375 is upregulated in the AOB of larva exposed to non-kin odorants and blocking the miR-375 MRE in the *Pax6* 3'UTR using a TP results in an increase in the proportion of TH+ interneurons relative to GABA+. Finally, miR-375 regulation of *PAX6* may play an important role in reinforcing the developmental switch from neural progenitor to motor neurons in the spinal cord [277]. miR-375 is upregulated during motor neuron differentiation of hESCs and overexpression of miR-375 in cultured motor neuron progenitors results in a decrease in PAX6 protein. Taken together, miRNAs in the 290-295 cluster, in the miR-96 family, miR-7, miR-365 and miR-375 may play important roles in sharpening cell type transitions by negatively regulating *Pax6*.

MicroRNAs have also been identified that may play important roles in establishing spatial domains during embryogenesis by targeting *Pax6* (Figure 10C, Table 2). miR-450b-5p may regulate *Pax6* during eye development to prevent ocular commitment of presumptive eyelid epidermis and sharpen the boundary between PAX6-positive corneal epithelium and PAX6-negative epidermis [279]. miR-450b-5p and PAX6 are expressed in mutually exclusive domains during mouse eye development, with PAX6 being restricted to developing corneal epithelium and conjunctiva, and miR-450b-5p detected in corneal stroma and eyelid epithelium. In the early postnatal mouse V-

SVZ, *Pax6* is expressed in a dorsal-high ventral-low gradient. miR-7, which is expressed in an opposing gradient, may participate in defining the spatial expression pattern of Pax6 in this region [206]. Expression of PAX6 in V-SVZ progenitors is important for dopaminergic periglomerular neuron (PGN) specification in the olfactory bulb, and electroporation of miR-7 sponge into the ventral V-SVZ increases dopaminergic specification [206].

MicroRNAs may also function to tune *Pax6*, ensuring that the correct dosage of PAX6 is maintained during development and adulthood. During mouse pancreatic development, *Pax6* and miR-7 are coexpressed in developing islets and fine-tuning of *Pax6* levels by miR-7 may refine the proportion of endocrine cell types formed. Knockdown of miR-7 in embryonic pancreas explants increases insulin and glucagon positive cell formation at the expense of ghrelin positive cell formation in pancreatic islets [217]. TPs that shield the MREs for miR-7 and miR-375 in the *Pax6* 3'UTR are associated with increased PAX6 protein, both in cultured mouse β-cell line (βTC-6) and in isolated adult mouse pancreatic islets [278], providing support for direct regulation of *Pax6* through these MREs.

Deregulated fine-tuning of PAX6 levels by miRNAs may contribute to human cancers. The role of PAX6 in human cancers is complicated, with PAX6 acting as a tumor suppressor in some tissue contexts and promoting tumorigenesis in others [325]. For example, several human cancer cell lines and tumors have reduced levels of PAX6 with corresponding increases in levels of miRNAs predicted to target *PAX6*. miR- 223 and the miR-96 family, which are upregulated in glioma cell lines, may directly target *PAX6* to

promote proliferation and invasion [326, 327]. Conversely, PAX6 promotes tumorigenesis in human retinoblastoma cell lines [328]. miR-365-3p, which is downregulated in retinoblastoma, may function as a tumor suppressor by directly targeting *PAX6* [329].

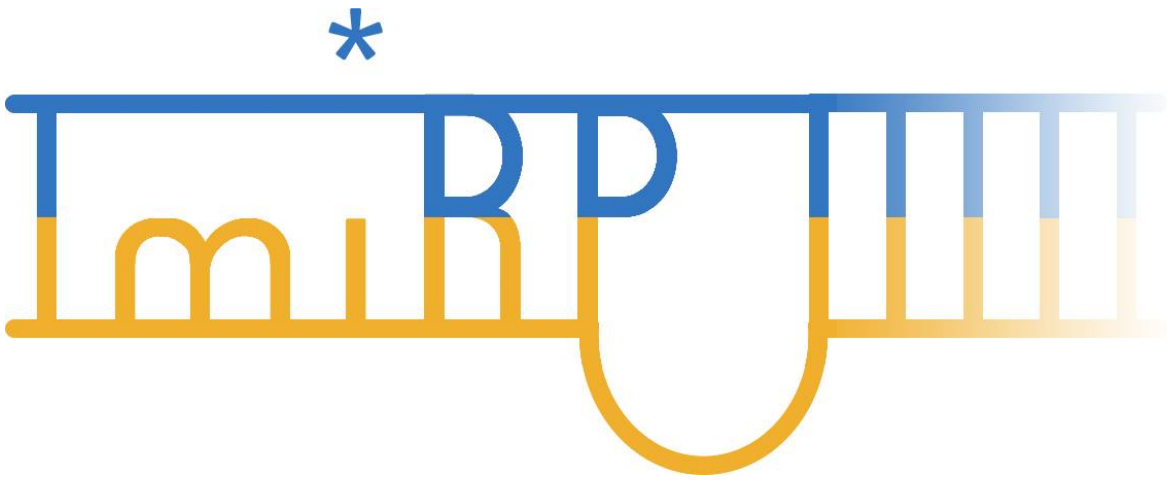
MicroRNA single nucleotide polymorphisms (miR-SNPs) are 3'UTR-located SNPs that alter MREs [330]. A miR-SNP was identified in the human *PAX6* 3'UTR that may alter regulation by miR-328. The allele predicted to create a miR-328 MRE, the C allele, is associated with risk for extreme myopia relative to the alternative T allele [331]. miR-328 is expressed in murine ocular tissues, and miR-328 overexpression decreases PAX6 mRNA and protein in a cultured human RPE cell line [332], suggesting that reduced PAX6 associated with the C allele increases the risk for extreme myopia. Conversely, the T allele lacking the predicted miR-328 MRE is associated with Rolandic epilepsy [333]. The authors suggest that increased expression of PAX6 due to disrupted miR-328 binding causes increased epilepsy susceptibility. Reporter assays demonstrate that miR-328 can potentially regulate through the *PAX6* 3'UTR [331–333]. Though these miR SNP studies provide information about potential functional consequences of disrupted and created MREs in an endogenous 3'UTR, they are correlative. Curiously, the miR SNP is not located in the 3'UTR sequence predicted to bind the miR-328 seed and overexpression of miR-328 can decrease reporter expression through the T allele, albeit to a lesser extent than the C allele [332], suggesting that this miR-SNP impacts miR-328 3' supplementary pairing.

1.7 Project objectives

Despite evidence that single miRNA-MRE interactions rarely mediate strong repression [50, 62] and regulation through multiple MREs in the same 3'UTR can enhance repression [19, 47, 52, 135–138], investigations into the functional relevance of miRNA regulation are often focused on *one* target regulated by *one* miRNA through *one* predicted MRE. Moreover, the physiological relevance of miRNA regulation of a single target is typically studied indirectly, using 3'UTR reporters, and miRNA overexpression or knockdown. Though these methods can suggest miRNA-target interactions, positive results do not guarantee that genuine regulation is occurring endogenously, under normal physiological conditions. Additionally, phenotypic effects associated with miRNA overexpression and knockdown are likely exaggerated and misrepresent the role one miRNA is playing in the regulation of one target. Target protectors provide useful alternatives, though they lack specificity for the MRE(s) of interest.

To address these limitations, we set out to accomplish three goals. First, we developed a software tool to facilitate design of MRE mutations. Successful mutations must disrupt the MRE of interest, preventing the intended miRNA from binding, while also avoiding introduction of new MREs for other miRNAs. Second, we used bioinformatics, miRNA expression analyses and assays in cultured cells to identify a cohort of candidate miRNA regulators of a single gene, *Pax6*. Finally, we generated mouse models to investigate the biological relevance of cooperative *Pax6* regulation by miRNAs *in vivo*.

Chapter 2: ImiRP, a computation approach to microRNA target site mutation



This chapter is adapted from the following publication:

Ryan BC, Werner TS, Howard PL, Chow RL. (2016) BMC Bioinformatics, 17:190.

Contributions:

BCR designed the project, tested the software, and prepared the manuscript. TSW designed, developed and tested the software, and edited the manuscript. PLH edited the manuscript. RLC edited the manuscript and supervised the entire project. All authors read and approved the entire manuscript.

2.1 Abstract

MicroRNAs (miRNAs) are small ~22 nucleotide non-coding RNAs that function as post-transcriptional regulators of messenger RNA (mRNA) through base-pairing to 6-8 nucleotide long target sites, usually located within the mRNA 3' untranslated region. A common approach to validate and probe microRNA-mRNA interactions is to mutate predicted target sites within the mRNA and determine whether it affects miRNA-mediated activity. The introduction of miRNA target site mutations, however, is potentially problematic as it may generate new, "illegitimate sites" target sites for other miRNAs, which may affect the experimental outcome. While it is possible to manually generate and check single miRNA target site mutations, this process can be time consuming, and becomes particularly onerous and error prone when multiple sites are to be mutated simultaneously. We have developed a modular Java-based system called ImiRP (Illegitimate miRNA Predictor) to solve this problem and to facilitate miRNA target

site mutagenesis.

The ImiRP interface allows users to input a sequence of interest, specify the locations of multiple predicted target sites to mutate, and set parameters such as species, mutation strategy, and disallowed illegitimate target site types. As mutant sequences are generated, ImiRP utilizes the miRBase high confidence miRNA dataset to identify illegitimate target sites in each mutant sequence by comparing target site predictions between input and mutant sequences. ImiRP then assembles a final mutant sequence in which all specified target sites have been mutated.

ImiRP is a mutation generator program that enables selective disruption of specified miRNA target sites while ensuring predicted target sites for other miRNAs are not inadvertently created. ImiRP supports mutagenesis of single and multiple miRNA target sites within a given sequence, including sites that overlap. This software will be particularly useful for studies looking at microRNA cooperativity, where mutagenesis of multiple microRNA target sites may be desired. The software is available at imirp.org and is available open source for download through GitHub (https://github.com/imirp).

2.2 Introduction

As described in chapter 1, miRNAs represent a class of abundant non-coding regulatory RNAs that are involved in many cellular processes. Given this, many computational tools have been developed for studying miRNAs. For example, several of programs are available to address challenges associated with miRNA discovery [16, 334] and numerous applications enable identification of miRNA recognition elements (MREs)

in mRNAs [19–21, 45, 58, 131, 135, 143–148, 335]. MRE prediction programs make use of a number of different parameters in their predictions, such as extent of complementarity to the miRNA 5' end, hybridization energy of the mRNA-miRNA heteroduplex, evolutionary conservation of predicted MREs, mRNA secondary structure, and local 3'UTR context. Additionally, other software tools are available for predicting the impact of 3'UTR SNPs on miRNA binding [146], for assessing the impact of mutations in miRNA genes [335], and for comparing the predicted set of target genes for two different miRNAs [335].

Currently, there are no software tools available for generating mutations to disrupt predicted MREs in a given mRNA sequence. Disrupting MREs by mutation is important for examining miRNA-mRNA interactions experimentally. Mutating a predicted MRE is commonly used in reporter-based experiments to determine whether a predicted MRE can be regulated by a candidate miRNA. In addition, mutations that disrupt one or more MREs in a given 3' UTR can be used to examine the biological relevance of the MRE(s) in question. One problem that can arise when an MRE is mutated is that the mutation itself can create a new or "illegitimate" MRE for a different miRNA (Figure 11). If the miRNA that targets the illegitimate site is present in the cell of tissue of interest, it could confound the data. Therefore, care must be taken when devising a mutation strategy to ensure that mutations do not inadvertently create new MREs. Small nucleotide polymorphisms (SNPs) that create functional MREs associated with phenotypes have been reported [331, 336–339].

Our primary goal is to study the direct role *Pax6* regulation by miRNAs

endogenously. Consequently, we want to ensure that, upon mutating MREs in the *Pax6* 3'UTR, functional MREs for other miRNAs are not created in the process.

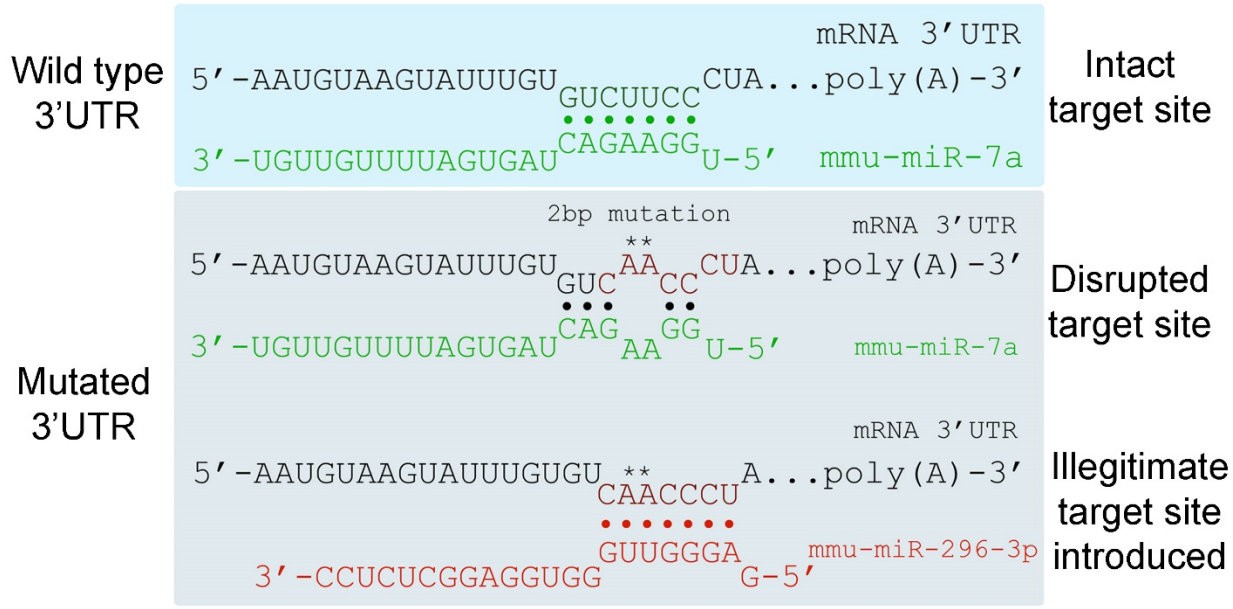


Figure 11. The problem associated with miRNA target site mutagenesis

A wild type 3'UTR sequence contains a predicted 7mer-m8 MRE for miR-7-5p (green). Mutation disrupts the interaction between miR-7 and its MRE but creates an illegitimate 7mer-m8 MRE for miR-296-3p in the process (red).

Our goal is to develop a modular, open source application that can run through a web interface or be downloaded and adapted to specialized projects. We created a Java-based system called ImiRP (Illegitimate miRNA Predictor) to generate MRE mutations in any sequence of interest and address the problem of illegitimate MRE creation. Users input a DNA or RNA sequence and specify the location of predicted MREs to be mutated. ImiRP then automates the processes of mutation generation, illegitimate MRE identification, and synthesis of a final mutant sequence lacking new

predicted MREs. ImiRP is particularly useful for investigating regulation of a single mRNA by multiple miRNAs, where situations such as target site overlap make mutagenesis design more challenging to perform manually. Currently, PITA is the only program that permits input of any mRNA sequence of interest. Using PITA, the MRE predictions for a wild type sequence can be compared to mutated sequences to identify a mutation strategy that does not generate illegitimate MREs. However, the process of generating mutations and analyzing PITA output manually is time consuming, prone to error, and becomes very challenging when multiple predicted MREs are mutated simultaneously.

2.3 Implementation

2.3.1 Input user interface

The ImiRP application is available at imirp.org. After creating a new project, the user enters a DNA or mRNA sequence of interest (i.e. 3'UTR) and specifies the organism of interest. All animal species with mature miRNA sequences available through miRBase version 21 (mirbase.org) are available for selection [18]. Users then define the position(s) of predicted MRE(s) to be mutated by selecting the nucleotide position of the input sequence that is complementary to position 7 of the targeting miRNA (Figure 12A). ImiRP will then highlight a stretch of six nucleotides starting from the nucleotide position that was entered. These six nucleotides define the mRNA region complementary to miRNA positions 2-7 starting from the 5' end, the miRNA "seed", and are collectively called a "mutation site" (Figure 12A). As an example, to introduce

mutations into a 7mer-m8 MRE, in which base pair complementarity occurs at miRNA positions 2-8, the input nucleotide position that is complementary to miRNA position 7 is entered in the “Mutation Site” input window. ImiRP introduces substitution mutations into the 3’UTR sequence specified as mutation sites.

We chose the design strategy of specifying six nucleotide “mutation sites” into which substitution mutations are introduced for several reasons. MREs vary in size and type (e.g. 6mer, 7mer-m8, 7mer-A1 or 8mer) and the miRNA seed region is used for mRNA target recognition for all major MRE types (Figure 3) [20, 51]. Consequently, disrupting base pairing between miRNA positions 2-7 and the MRE, by mutating multiple nucleotides in the MRE, can interfere with miRNA-MRE interaction [31] and miRNA mediated repression [47, 52, 53]. Additionally, this approach to defining mutation sites was chosen for practical reasons. Computational demands on the software are reduced by limiting the size of regions of the input sequence that can be mutated to six-nucleotide windows.

Upon defining the project parameters, the ImiRP user interface allows researchers to specify mutation parameters (Figure 12B). The user can select the types of MREs, termed “invalid sites”, that they wish to exclude from their mutant sequence. Any mutant sequence containing MRE types defined as “invalid” will be rejected.

ImiRP can identify five classes of MRE: 6mer, 7mer-m8, 7mer-A1, 8mer, and offset 6mer (OS-6mer) [20, 51]. 6mer MREs are complementary to positions 2-7 of the miRNA, 7mer-A1 MREs are complementary to positions 2-7 and have an A across from

miRNA position 1, 7mer-m8 MREs are complementary to miRNA positions 2-8 and 8mer sites are 7mer-m8 sites with an A across from position 1 [20]. OS-6mer MREs are complementary to positions 3-8 of the miRNA [20]. ImiRP also recognizes predicted MREs containing a single G:U mismatch, called “wobble” pairs [19, 47, 52, 53, 149]. By manually selecting and using arrow buttons in the user interface, invalid site types can be designated (Figure 12B). For instance, perfect 6-8mer MREs can be specified as invalid by moving them into the select list entitled “Invalid Site Types”. In this example, the resulting mutant sequence(s) would lack newly created 6-8mer sites that are perfect matches to the miRNA, but permit newly created 6-8mer sites containing G:U wobble base pairs between the miRNA and MRE.

The interface also enables users to specify a desired mutation strategy. The nucleotide bases to use for mutagenesis can be selected from a checklist. For example, by selecting only G (guanine), ImiRP will generate mutant sequences by substituting nucleotides within the defined MRE(s) to G only. In this example, if the unmutated MRE already contains guanine residues, these nucleotides will not be mutated. The number of nucleotide changes to introduce per specified MRE, ranging from two to six, can be defined from a drop-down menu. For example, selecting “2 changes per site” will ensure that two adjacent nucleotide changes are introduced into each specified MRE.

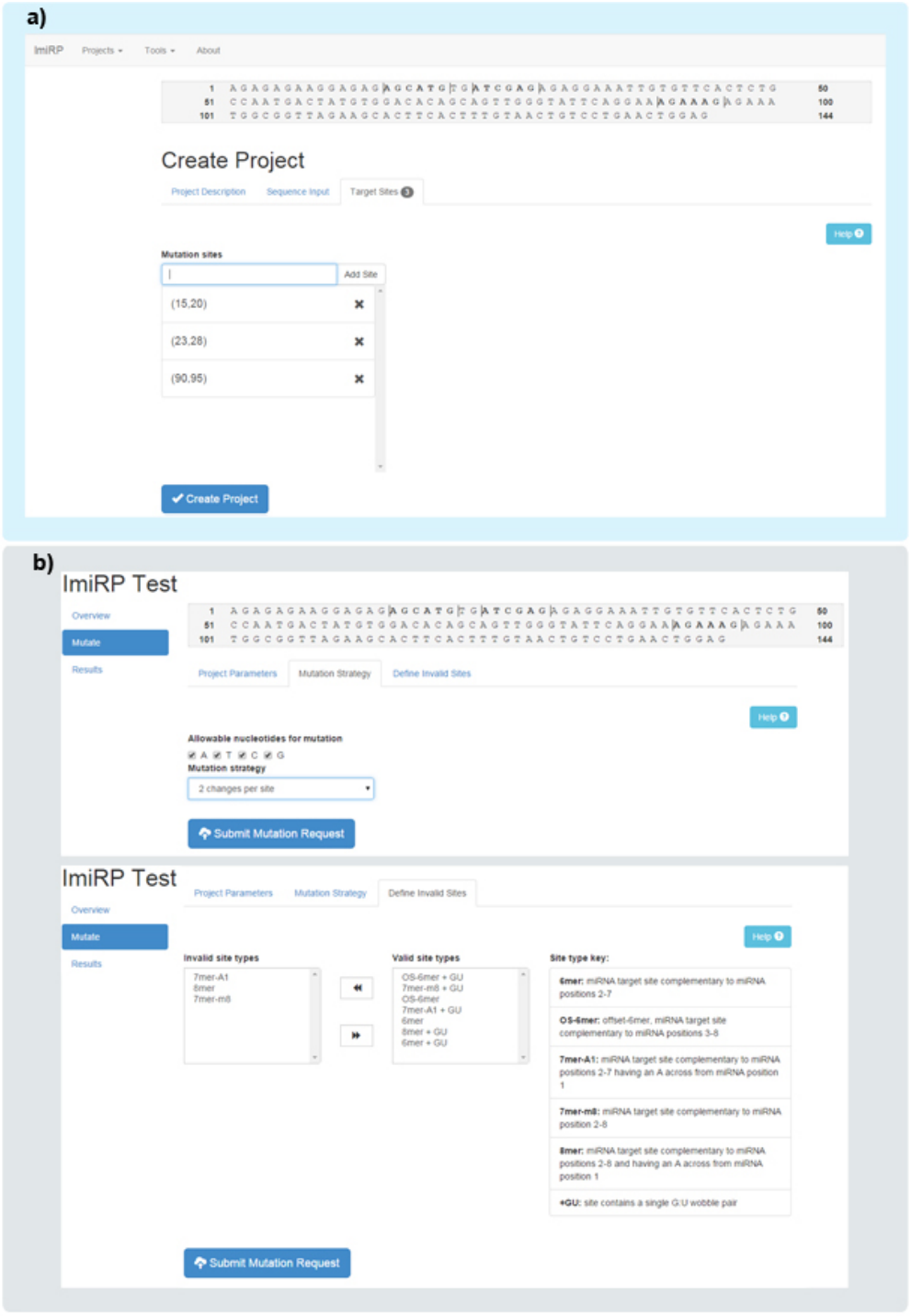


Figure 12. ImiRP user interface

(A) The user is first directed to create a project and specify project parameters. The project can be named, and in this example we have named the project “ImiRP Test”. A DNA or RNA sequence must be input along with information about the species of interest. The input sequence will be displayed for viewing. Finally, at least one “mutation site” (the region into which mutations will be introduced) is specified by typing the sequence position complementary to miRNA position 7 into the textbox. All selected mutation sites, complementary to miRNA positions 2-7, appear bolded in the displayed original input sequence for inspection. In this example, a segment of the mouse *Pax6* 3’UTR and three hypothetical predicted MREs are used as input. (B) Upon creation of a project, the user is asked to specify mutation parameters. In the “Mutation Strategy” tab, nucleotides to use for mutation and number of nucleotide changes per specified mutation site can be selected. In the “Define Invalid Sites” tab, users can specify the types of newly created MREs they do not want present in their final mutant sequence. In this case, any mutant sequences provided as output will contain two adjacent nucleotide changes per specified MRE using all four nucleotides for mutation, and no mutant sequences containing newly created 8mer, 7mer-m8, or 7mer-A1 predicted MREs will be provided as output.

2.3.2 ImiRP Workflow

Conceptually, ImiRP can be subdivided into three modules (Figure 13):

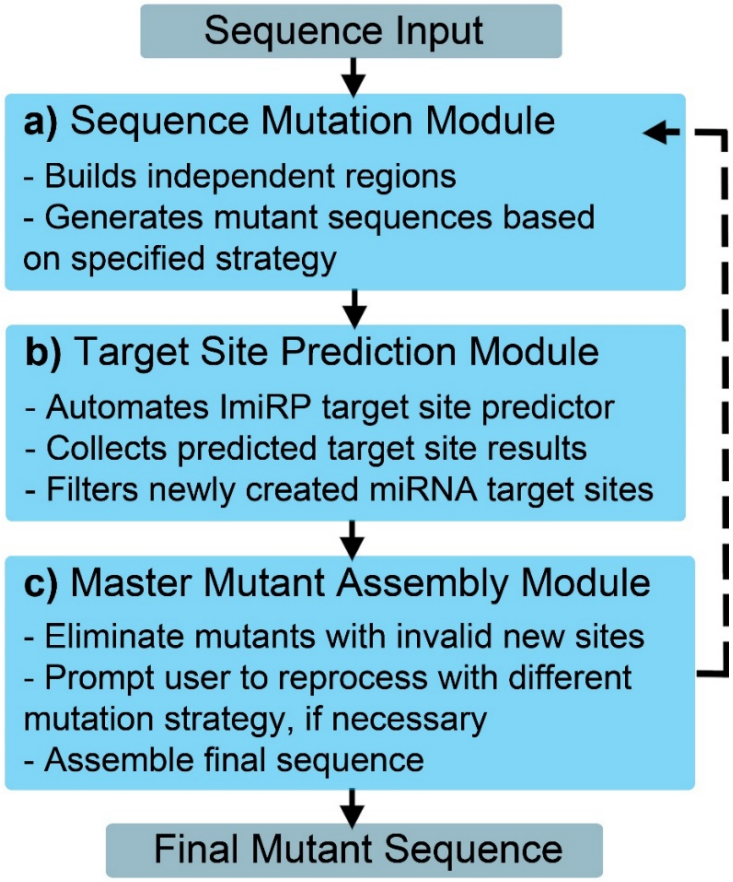


Figure 13. ImiRP workflow

ImiRP can be divided into three major modules. The Sequence Mutation module (A) feeds mutated sequences into the Target Site Prediction module (B), the results of which are supplied to the Master Mutant Assembly module (C) to optimize and output a final mutant sequence.

a) Sequence mutation module

When ImiRP is run, information about the input sequence, defined regions to mutate, and desired mutation strategy are sent to the Sequence Mutation module

(Figure 13A). This module first divides the input sequence into “mutationally independent” regions based on the spacing between MREs to be mutated. Since the largest recognized MRE is eight nucleotides [20], two predicted MREs spaced at least seven nucleotides apart can be mutated independent of one another without generating an illegitimate MRE that spans both mutations. Simultaneously mutating two predicted MREs spaced less than seven nucleotides apart could generate a new MRE for a different miRNA containing mutated nucleotides from each of the original target sites. Consequently, predicted MREs spaced less than seven nucleotides apart are grouped into a single independent region, and each independent region is annotated based on the positions of the first and last nucleotides of sites within the region relative to start of the input sequence (Figure 14A). All specified MREs within a given independent region are mutated as a unit, while MREs within other regions are left unchanged. This process is repeated for each independent region (Figure 14B).

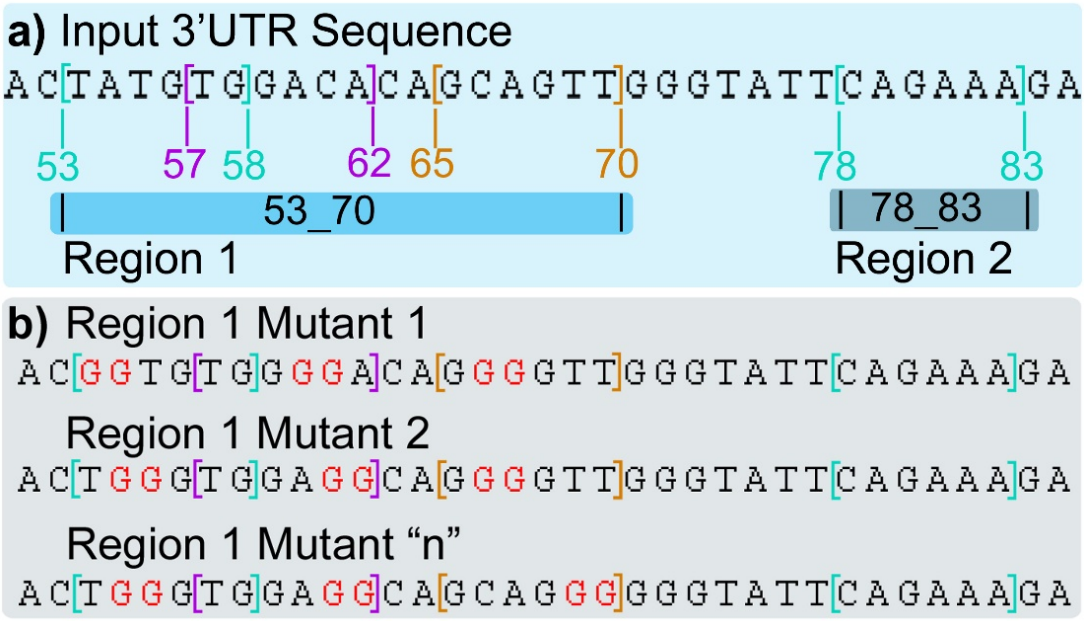


Figure 14. The Sequence Mutation module

(A) The input sequence is divided into "mutationally independent" regions. User-specified MREs spaced less than 7 nucleotides apart are grouped into a single independent region and are annotated based on the positions of the first and last nucleotides of sites within the given region. (B) To reduce computational overhead, MREs within a given independent region are mutated as a unit, while MREs in other regions are unchanged.

b) Target site prediction module

Mutant 3'UTR sequences generated by the Sequence Mutation module are sent to the Target Site Prediction module (Figure 13B). This module is responsible for identifying predicted miRNA MREs present in each mutant sequence that are absent from the original input sequence. The custom ImiRP miRNA prediction component uses pattern recognition to identify regions of each mutant 3'UTR sequence that are complementary to the 5' seed regions of known mature miRNAs. High confidence FASTA-format miRNA sequences were obtained from the most recent release of miRBase, version 21 [18]. High confidence miRNAs have at least ten small RNA

sequencing reads that map to each arm of the pre-miRNA hairpin sequence, or have at least five reads mapped to each arm and at least 100 reads in total [18].

The prediction module compares the MRE predictions for the input sequence to predictions for each mutant sequence. The predicted MREs that are present in each mutant sequence and absent from the input sequence (i.e. the illegitimate MREs) are stored in a database along with the respective mutant sequences (Figure 15). Information in the database is accessed by the Master Mutant Assembly module (Figure 13C) for synthesis of a final mutant sequence.

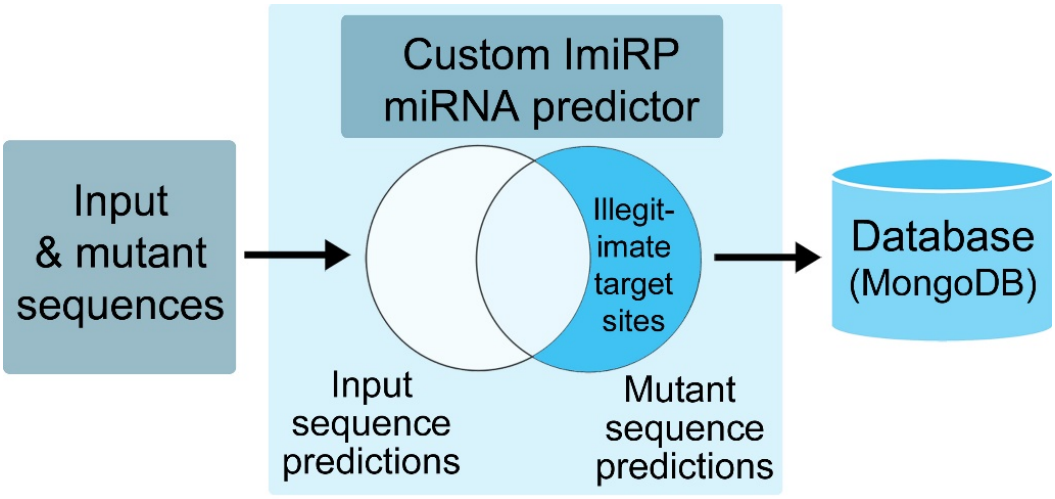


Figure 15. The Target Site Prediction module

Predicted miRNA target sites within input and mutant sequences are identified using a custom component that identifies predicted target sites based on complementarity to the miRNA 5' end. FASTA-format miRNA sequence data was collected from miRBase version 21 [11]. Information about predicted sites present in each mutant sequence that are absent from the input sequence are stored in a database for analysis.

c) Master mutant assembly module

The Master Mutant Assembly module first eliminates all mutant sequences containing predicted MREs that were specified as invalid through the user interface. This process is repeated for each independent region. As valid mutations for each region are generated, they are displayed in an output user interface (Figure 16). The output interface displays up to five valid mutants for each independent region. The specified six nucleotide "mutation sites" and mutated nucleotides are displayed in the output sequences. Here, the user can select a single desired mutant for each independent region and ImiRP generates an assembled mutant sequence in which all specified MREs have been mutated without creating any illegitimate predicted MREs of the types specified.

A zip folder containing four files is made available for download through the output user interface. The project information text file contains the assembled mutant DNA sequence, and three CSV files contain information about: i) MRE predictions for the input sequence, ii) MRE predictions for the assembled mutant sequence, and iii) new predicted MREs present in the mutant sequence. If any regions fail to generate valid results, the user will be alerted to reprocess those regions using a different mutation strategy. Accumulated results from multiple mutation runs are displayed in the results tab of the output user interface. In the future, we plan to annotate each valid mutant sequence displayed in the results with a "mutation run identifier" so that it is clear which mutation parameters were used to generate each sequence.

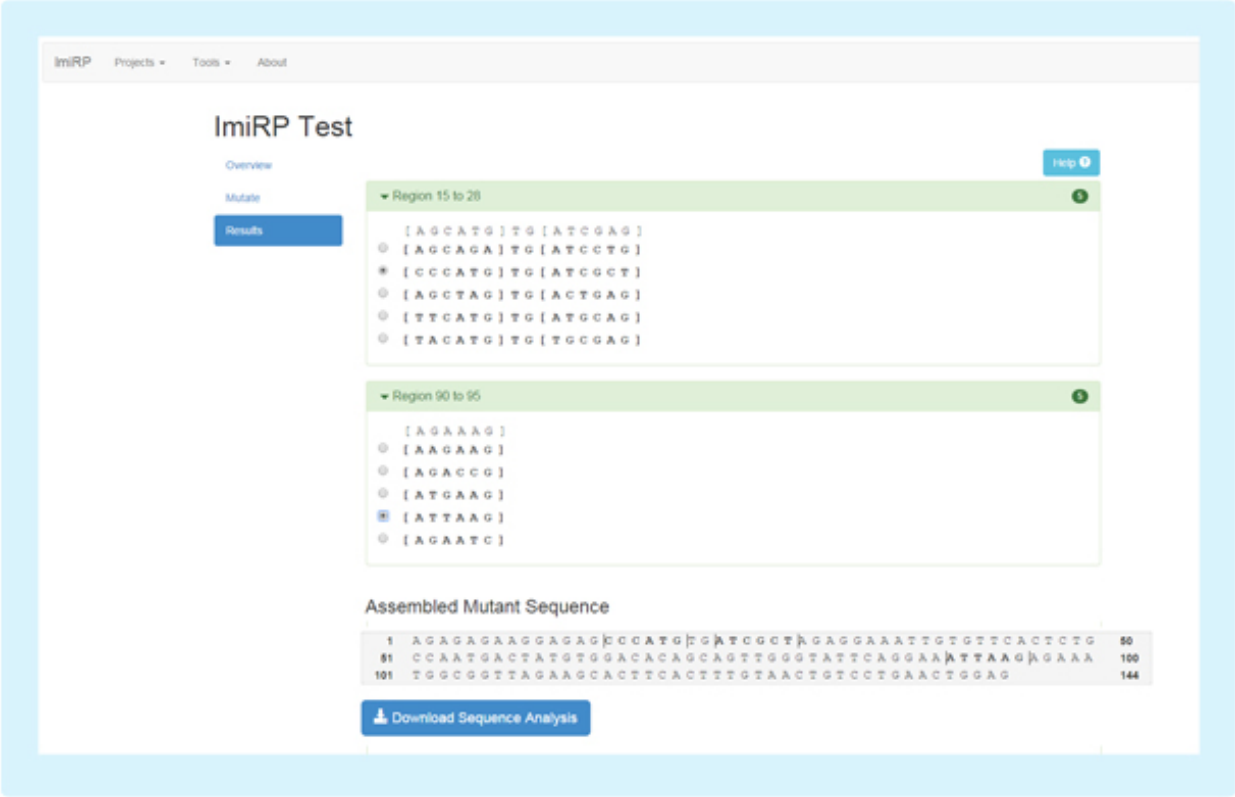


Figure 16. ImiRP Output User Interface

Input mutation sites spaced less than seven nucleotides apart are grouped into independent regions and each region is mutated independently of other regions. The output user interface displays up to five mutant sequences for each region that satisfy the specified mutation parameters. For example, the output mutations for ImiRP Test (Figure 12) will not have generated any new 8mer, 7mer-m8, or 7mer-A1 predicted MREs. The original input sequence for each region is displayed above the mutants for comparison, and brackets denote the bounds of each specified mutation site. The user must select one desired mutant for each region, and the changes are displayed as an assembled mutant sequence. Once the selection process is complete, the user may download a folder containing information about their assembled mutant sequence and predicted MREs in the input and mutant sequences.

2.3.3 Project organization

ImiRP's code is organized into two separate Scala Build Tool (SBT) projects: ImiRP Core and ImiRP Web. ImiRP Core contains the data model and the primary business logic but has no presentation layer. Core functionality is exposed through Java services and

Akka actors and must be programmatically interacted with. ImiRP Web depends on ImiRP Core and exposes the functionality of ImiRP Core to an end user via a Web interface. This multi-project organization scheme was chosen to keep the presentation logic separate from the core business logic, thus allowing ImiRP to support alternative core or presentation implementations in the future. Additionally, this organizational approach allows ImiRP to be run headless (without a presentation layer) or to be integrated with other projects or systems as a library.

2.3.4 Mutant sequence generation algorithm

ImiRP's mutation generation algorithm is essentially a random nucleotide sequence permutation generator. As input, it accepts a DNA or RNA nucleotide sequence, a list of mutation sites, a set of allowable nucleotide types for mutagenesis, number of allowable nucleotide changes per MRE, a mutation region, and a callback. The algorithm iterates over each mutation site in the given independent mutation region to apply nucleotide changes to each site. To accomplish this, the algorithm selects a random nucleotide from the list of allowable nucleotides and replaces the nucleotide in the input sequence at the start (5' end) of the mutation site. This process is repeated for subsequent nucleotides contained within the given mutation site. A mutant sequence is produced once all mutation sites have been processed in this way. Due to the simplicity of this algorithm, some mutant sequences will not satisfy all specified mutation criteria. As such, all mutant sequences must be passed through a filter that either accepts or rejects the mutant based on whether it satisfies all criteria. Finally, if a mutant sequence is accepted by the filter, the callback is activated and the

mutant sequence is passed on for further processing (see Appendix A Equation 2 for pseudocode).

2.3.5 MicroRNA recognition element prediction algorithm

The MRE predictor uses a pattern-matching algorithm to detect miRNA seed matches in an input sequence. To detect seed matches, the algorithm first iterates over a list of miRNA sequences from the miRBase high confidence database. The algorithm compares the first eight nucleotides starting at the miRNA 5' end (the miRNA "seed") of each miRNA entry against the input sequence to identify complementarity to the input sequence. An entry is compared to a sequence by scanning across the sequence using an eight-nucleotide sliding window. Before each advancement of the window, a hierarchy of conditions are checked to determine whether a seed match is present. If a seed match is found, the result is stored, the window advances, and the process repeats until the end of the sequence is reached.

The conditional hierarchy begins by comparing the first nucleotide (position 1) of the input sequence window to the last nucleotide of the miRNA 5' end (position 8). If those nucleotides are a match, the next five nucleotides (positions 2-6 of the sequence, positions 7-3 of the miRNA 5' end) are checked for matches. If any of these nucleotides are not a match, no seed match is possible and the window will advance. If they do match, then an 8mer, 7mer-m8, or OS-6mer seed match is known to be present. The condition hierarchy is then further evaluated to check for a match between sequence position 7 and miRNA position 2. If this match is not present, then an OS-6mer seed match is identified. If this match is present, the seed match is either an 8mer if the last

nucleotide of the sequence window (position 8) is an adenosine (A), or it is a 7mer-m8 if the last nucleotide is not adenosine. If the first nucleotide of the sequence window did not match miRNA position 8, then only a 7mer-A1 or a 6mer seed match are possible, given that sequence positions 2-7 and miRNA positions 7-2 have been found to match. The seed match is a 7mer-A1 if the last nucleotide of the sequence window is adenosine, or the seed match is a 6mer if the last nucleotide is not adenosine. Two nucleotides are considered a match if the following conditions are true: the sequence nucleotide is A and the miRNA 5' end nucleotide is T/U or vice versa, the sequence nucleotide is C and the miRNA nucleotide is G or vice versa. If G:U pairs are allowed, and if a G:U pair has not already been found, two nucleotides will be considered a match if the sequence nucleotide is T/U and the miRNA nucleotide is G or vice versa (see Appendix A Equation 3 for pseudocode).

2.3.6 System architecture

ImiRP was developed in the Java programming language. Java was chosen for its extensive catalogue of third party libraries and for its cross-platform support. ImiRP's architecture is based on a model-view-controller (MVC) architecture consisting of a view layer for presenting information, a controller layer for applying business logic, and a model layer for storing information. In its current form, the view layer is driven by the Play Framework (Play) and AngularJS. Play accepts incoming HTTP requests and maps those requests to view templates. These view templates are partially rendered by Play and then returned as HTML and Javascript for final rendering by AngularJS in a user's browser. The controller layer is a shared responsibility between ImiRP Web and

ImiRP Core. ImiRP Web's portion of the controller layer is responsible for mapping HTTP requests to ImiRP Core's services and returning results from those services as rendered templates. Currently, the model layer exists entirely within ImiRP Core, but this may change in the future as ImiRP Web may require some of its own models for features that are specific to presentation.

2.3.7 ImiRP data import

Currently, known miRNA data is imported from the high confidence dataset (mature.fa) that is provided by miRBase version 21. Mature.fa is a text file database of FASTA-format, mature miRNA sequences. ImiRP parses this database and imports it into its own internal format for use by ImiRP's target site predictor. To accomplish this, an import utility that can read the mature.fa dataset and import miRNAs was created. The import utility was designed to allow alternative miRNA datasets to be imported if they follow a similar FASTA format. This enables ImiRP to stay current as new miRNA datasets become available. Presently, an alternative miRNA dataset can only be used by altering ImiRP's global configuration file and restarting the application. In future versions of ImiRP, we may enable use of custom miRNA datasets that could be uploaded for specific projects. This way, each individual project could specify a miRNA dataset that best meets its requirements.

2.4 Results and Discussion

We designed ImiRP, a program that automates the entire process of generating mutant sequences, with the goal of facilitating MRE mutations in a mRNA 3' UTR. A key feature of ImiRP is its ability to suggest mutations that lack illegitimate MREs. Predicted MREs that are created following 3'UTR mutation have the potential to be functional and complicate experiments aimed at examining miRNA-target regulation.

2.4.1 Computational time optimization

The number of sequences generated by the Sequence Mutation module increases rapidly with increasing predicted MREs to mutate. Each of these mutant sequences then needs to be processed to identify MREs present in the mutant sequence that are absent from the input sequence. As a result, one problem that needed to be addressed in the implementation of ImiRP was the large computational time required to generate a final mutant sequence when mutating many predicted MREs simultaneously. We implemented several strategies to address this problem:

i) Limit miRNA target site mutations to six nucleotide motifs

Nucleotide changes introduced into the region complementary to miRNA positions 2-7 create the most significant reduction in regulation by the targeting miRNA [52]. Additionally, by limiting the problem space for MRE mutagenesis to six nucleotide regions, we reduce the number of possible mutant sequences and thus the computation time. However, some examples of "seedless" MREs have been identified. These rely more heavily on binding the 3' end of the miRNA or contain mismatches or bulges

between the miRNA seed and its target [340, 341]. Despite this, comparative analysis of orthologous mRNAs suggests that MREs complementary to miRNA 5’ ends selectively conserved [20]. Additionally, mRNA and protein expression analyses reveal that messages containing miRNA seed matches are preferentially regulated by miRNAs [50, 155] and the miRNA 5’ end is exposed for target recognition when bound in a complex with Ago2 [31]. As such, we decided to focus on an approach that restricts mutagenesis to the six-nucleotide region complementary to miRNA positions 2-7.

ii) Introduce adjacent nucleotide changes per specified seed site

Single nucleotide mismatches introduced into an MRE complementary to miRNA positions 2-7 have a large impact on target site efficacy, even in the presence of extensive pairing to the miRNA 3’ end [52]. Based on this, we reasoned that a minimum of two nucleotide changes introduced into an MRE would effectively abolish regulation by the targeting miRNA. Requiring that nucleotide substitutions be adjacent has several benefits. First, it reduces the number of possible mutants generated, improving processing speed. Additionally, this strategy reduces the number of illegitimate MREs created by narrowing the region of sequence that is altered.

To illustrate this, consider a situation where three closely spaced but non-overlapping MREs are mutated by disrupting at least two nucleotides using a permutation of the four nucleobases. If any combination of at least two nucleotide substitutions are used to generate mutant sequences, greater than 60 billion different mutants are possible (Appendix A Equation 4A). However, when only two adjacent

nucleotide changes are permitted per MRE, the number of possible mutant sequences for an independent region containing three non-overlapping MREs is reduced to approximately 90,000 (Appendix A Equation 4B)

iii) Divide the input sequence into independent regions based on inter-site spacing

Since 8mer MREs are the largest recognized MREs [20], predicted MREs spaced seven or more nucleotides apart can be mutated without generating a new predicted MRE that spans both mutations. As such, sites spaced less than seven nucleotides apart are grouped into independent regions and all MREs within an independent region are mutated as a unit. Mutating one region at a time, while keeping the remainder of the sequence unchanged, significantly reduces the number of permutations of redundant mutations that are generated (Appendix A Equation 4C) and dramatically improves processing speed. Expanding on the above example, if two independent regions containing three non-overlapping predicted MREs are mutated together, over 8 billion mutant sequences are possible. However, if these two regions containing three MREs each are mutated separately of one another, approximately 180,000 mutant sequences are possible (Appendix A Equation 4D).

iv) Custom ImiRP MRE prediction component

Initially, we used the PITA executable in the Target Site Prediction module. PITA assigns each predicted MRE a score based on the difference between the free energy of miRNA-target duplex formation and the cost of unpairing mRNA secondary structure in

the region of the predicted MRE [58]. Though evidence suggests that MRE accessibility [58] and local MRE environment [35, 53] can impact MRE efficacy, some evidence also suggests that mRNA secondary structure is an insufficient predictor of MRE functionality [53]. Since we are primarily concerned with creating new MREs that have the potential to be functional, this feature of the PITA target site prediction component was not needed. ImiRP's custom MRE predictor uses only pattern recognition to identify five types of MRE (8mer, 7mer-m8, 7mer-A1, 6mer, OS-6mer)[20] and is therefore less computationally demanding, thus improving processing time.

v) Stop generating mutant sequences once five valid mutants have been identified

The ImiRP web application is designed such that many users can run projects simultaneously. Consequently, we wanted a avoid situations where a single user could consume a large portion of the web server's resources. The ImiRP application stops generating mutant sequences once at least five valid mutants are identified for each independent region or once the Sequence Mutation module effectively exhausts all mutation possibilities. This ensures that a single mutation request will not continue to process unnecessary sequences and continue to consume resources. Due to the Sequence Mutation module generating mutations faster than the Target Site Prediction module can scan them; it is possible that more than five valid mutations may be identified. Only five valid mutant sequences are displayed for each independent region in the results tab of the output user interface. If a user is interested in having all valid

mutations displayed, they can download the source code and run the application on their own machine.

2.4.2 Generation of mutant sequences

We made several attempts to devise an algorithm that would not generate duplicate mutant sequences or mutants that did not satisfy all criteria. In an attempt at preventing duplicate mutations, a sequential permutation generation strategy was used. However, this approach was subject to spending long time periods generating undesirable mutants. For some problems, the sequential approach finds solutions very quickly while with others it does not find any solutions within a feasible time period. The random approach was found to more reliably and consistently produce results across a diverse set of mutation problems.

We also initially developed a mutation generation algorithm that only generates mutant sequences that satisfy all specified criteria without wasting computation on generating mutants that do not. This approach was abandoned because it requires development of new mutation generation algorithms for every set of criteria. As such, we decided it was preferable to keep separate the concepts of mutant sequence generation and criteria selection, thus enabling the user to specify diverse mutation criteria. For projects requiring disruption of many MREs and having relatively simple criteria, it may be preferable to devise a mutation generation algorithm that does not require the additional filtering step. However, if disruption of one or a few sites is desired, these inefficiencies do not appreciably impact processing speed.

2.4.3 Testing the ImiRP target site predictor and mutation generator

To determine whether ImiRP could accurately predict MREs, we input 150 nucleotide segments from human, mouse and fly 3'UTR sequences and compared output predictions for wild type sequences to predictions generated by TargetScan [49] and PITA [58] for the same sequences. All 3'UTR sequence data was collected using the University of California, Santa Cruz (UCSC) genome browser [280]. Care was taken when making comparisons to PITA's predictions. As of the first release of ImiRP, PITA's database of mature miRNA sequences, obtained from miRBase version 11.0, is out of date. Additionally, PITA's MRE prediction logic is not based on the currently accepted MRE classification [20]. Consequently, mature miRNA sequences available through miRBase were used to verify PITA's MRE type predictions. Our results show that ImiRP is capable of accurately predicting 8mer, 7mer-m8, 7mer-A1 and 6mer MREs. One limitation of these tests is that TargetScan and PITA do not recognize OS-6mer MREs. However, manual comparison between ImiRP's OS-6mer MRE predictions and mature miRNA sequences in miRBase reveals that ImiRP is capable of correctly predicting OS-6mer MREs (data not shown).

We also tested ImiRP's capacity to successfully introduce mutations into many MREs simultaneously while ensuring that new predicted MREs for other miRNAs are not created. To do this, we used the aforementioned 150 nucleotide 3'UTR sequences and MRE predictions as input into ImiRP. We compared ImiRP's MRE prediction output between the wild type and mutant sequences to ensure that specified MREs were successfully mutated and that new MREs were not created.

2.4.4 Software limitations

Though ImiRP can identify predicted illegitimate MREs created following mutation of existing MREs, this software does not identify mRNA regulatory elements required for non-miRNA processes. For example, the process of introducing mutations into a 3'UTR sequence could create an RNA-binding protein (RBP) motif or abolish a pre-existing polyadenylation signal. The web applications Transterm and RBP Database (RBPDB) can be used to address this shortcoming [342, 343]. These tools can identify RBP motifs in mRNA 3'UTRs. Predicted 3'UTR motifs can be compared between the original input sequence and the mutated sequences to ensure that no known motifs had been created or disrupted. In future versions of ImiRP, we may implement a tool that identifies polyadenylation signal motifs and ensures that polyadenylation signals present in the input sequence are not disrupted and new polyadenylation signals are not created following mutation.

One limitation associated with identifying only five classes of MRE is that other MRE types have been identified that may also be functional. *In vivo* reporter assays suggest that 4mer and 5mer MREs having extensive complementarity to the miRNA 3' end can be functional [52]. However other studies investigating miRNA-dependent repression *in vivo* demonstrate that 5mer seed matches, complementary to miRNA positions 2-6, cannot mediate repression [53]. Consequently, we designed the ImiRP MRE predictor such that it does not identify 4mer or 5mer MRE types. Genome-wide miRNA-mRNA interaction maps from mouse brain generated by Argonaute High-

Throughput Sequencing of RNA isolated by crosslinking immunoprecipitation (Ago HITS-CLIP) also suggest that two additional MRE types may exist. A third type of 6mer MRE, 6merα (complementary to miRNA position 1-6), is capable of binding Ago-miRNA [57]. Additionally, miRNA recognition elements containing G nucleotide bulges at positions 5-6 in the mRNA (G-bulge sites) are evolutionarily conserved and functional *in vivo* [54]. ImiRP is currently incapable of identifying 6merα and G-bulge MREs and, as a consequence, these MRE types may be created upon mutation of existing MREs. Based on user demand, we may choose to enable identification of 6merα and G-bulge sites though the ImiRP MRE predictor.

Experiments designed to disrupt a specific sequence motif frequently opt to remove the sequence by deletion as opposed to introducing substitution mutations. ImiRP could be designed to permit deletion mutations as well, however this would complicate illegitimate MRE identification. ImiRP compares MRE predictions between the input and mutant sequences. Any predicted MREs that are shared between the input and mutant sequences, i.e. MREs for the same miRNA, located at the same position within the sequence, are not considered illegitimate. Information about the position of a predicted MRE within the sequence is critical for identifying legitimate versus illegitimate MREs. To enable deletion mutations will require a more complicated illegitimate MRE identification algorithm. We plan to enable the use of insertion and deletion mutations as part of ImiRP's mutagenesis strategies in future versions. For the time being, users can download the ImiRP source code and modify it to enable insertion/deletion mutations.

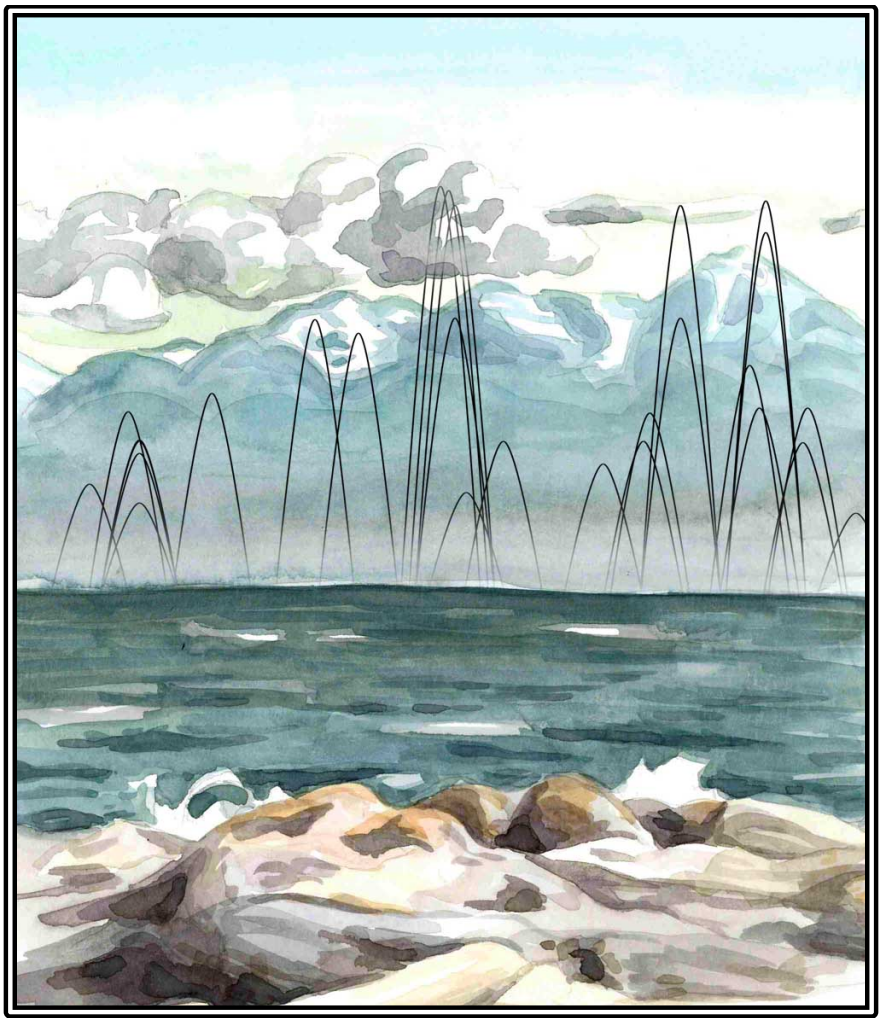
Unfortunately, it is challenging to predict with certainty whether a MRE has successfully been abolished by mutation without experimentation. Based on previous experimental evidence [52, 53, 149], introducing at least two nucleotide changes into the mRNA region complementary to miRNA positions 2-7 is likely to abolish regulation by the targeting miRNA. Similarly, predicting the functionality of a newly created MRE is also challenging without experimental validation. For example, a *lys-6* site from the *cog-1* 3'UTR in *C. elegans* is often nonfunctional when transplanted into different 3'UTR contexts [53]. ImiRP provides information about all predicted illegitimate MREs created in an output mutant sequence. Available databases documenting validated miRNA-target interactions [335, 344, 345] and known miRNA expression profiles [84, 346, 347] could be used to inform whether illegitimate MREs are likely to be functional.

2.5 Conclusions

In summary, we have described ImiRP, software that enables mutation of one or more predicted MREs within a sequence of interest while ensuring that new predicted MREs for other miRNAs are not created in the process. The ImiRP input interface allows the user to input a sequence of interest and specify the location of one or more MREs to mutate, species of interest, invalid MRE types, and mutation strategy. The output interface allows the user to select a desired mutant for each independent region and view a finalized mutant sequence. Upon completion, the assembled mutant sequence, and MRE predictions for the input and mutant sequences are made available for download. ImiRP source code is accessible open source through GitHub

(https://github.com/imirp) under the Apache Version 2 license and is available as a web application at imirp.org.

Chapter 3: Mapping the *Pax6* 3' untranslated region microRNA regulatory landscape



"Interactive landscape" Watercolor on paper (2016)

This chapter is adapted from the following publication:

Ryan BC, Lowe K, Hanson L, Gil T, Braun L, Howard PL, Chow RL. Mapping the *Pax6* 3' untranslated region landscape. (2018) BMC Genomics, 19:820.

Contributions:

BCR and RLC designed the project. **BCR** performed all the experiments and analyses for Figures 4-7. KL and TG performed 3'RACE experiments. LB contributed to RNA seq experiments. LH contributed to miTRAP experiments. **BCR** and RLC wrote the manuscript. PLH edited the manuscript provided feedback on project design. All authors read and approved the entire manuscript.

3.1 Abstract

PAX6 is a homeodomain-containing transcription factor that acts in a highly dosage-sensitive manner to regulate the development and function of the eyes, nose, central nervous system, gut, and endocrine pancreas. Several individual microRNAs (miRNA) have been implicated in regulating PAX6 in different cellular contexts, but a more general view of how they contribute to the fine-tuning and homeostasis of PAX6 is poorly understood. Here, a comprehensive analysis of the *Pax6* 3' untranslated region was performed to map potential miRNA recognition elements and served as a backdrop for miRNA expression profiling experiments to identify potential cell/tissue-specific miRNA codes. *Pax6* 3'UTR pull-down studies identified a cohort of miRNA interactors in pancreatic αTC1-6 cells that, based on the spacing of their recognition sites in the *Pax6*

3'UTR, revealed 3 clusters where cooperative miRNA regulation may occur. Some of these interacting miRNAs have been implicated in α cell function but have not previously been linked to Pax6 function and may therefore represent novel PAX6 regulators. These findings reveal a regulatory landscape upon which miRNAs may participate in the developmental control, fine-tuning and/or homeostasis of PAX6 levels.

3.2 Introduction

As described in chapter 1, PAX6 is a highly conserved transcription factor that is expressed in a specific spatio-temporal pattern in several tissues during vertebrate development, and plays an important role in cell fate determination and tissue differentiation [179]. *Pax6* is expressed in the developing retina, lens and cornea, and continues to be expressed in several mature ocular cell types [164, 199, 201, 237, 348]. *Pax6* is also expressed in the developing and mature endocrine pancreas [191, 195], central nervous system (CNS), and olfactory system [177, 237, 348], gut [192] and osteocytes [193]. In the complete absence of *Pax6*, eyes and nasal structures fail to develop [231, 349, 350], and patterning in the forebrain and specification of hormone-producing cells in the endocrine pancreas are severely perturbed [191, 195, 247]. *Pax6* is also required for maintenance of the progenitor cell pool in the cortex and spinal cord [256, 257] and in the retina for progenitor cell multipotency [253].

PAX6 function is particularly sensitive to dosage: too little or too much PAX6 can have profound effects on tissue development and maintenance. The requirement for precise PAX6 dose is exemplified by the semi-dominant phenotypes associated with

PAX6 haploinsufficiency and from overexpression phenotypes. Loss of a single copy of *Pax6/PAX6* results in a small eye phenotype in rodents [231, 349, 350] and is the primary cause of the poly-symptomatic and progressive disease aniridia in humans [179, 260, 351]. Though *Pax6/PAX6* haploinsufficiency is not associated with overt defects in pancreatic development, mice lacking one copy of *Pax6* have impaired proinsulin processing and glucose metabolism [252]. In humans, *PAX6* heterozygosity is associated with glucose intolerance [264]. However, the physiological mechanism(s) regulating precise PAX6 expression levels have not been elucidated.

Post-transcriptional regulation of *Pax6* by miRNAs may represent an important mechanism for maintaining the correct dosage of PAX6, and several miRNAs are suggested regulators of *Pax6*. Of interest to miRNA regulation of *Pax6*: regulation of an individual target transcript can be influenced by the cooperative activity of multiple miRNAs, acting through multiple target sites. For example, closely spaced miRNA target sites can act synergistically [35, 59, 138], multiple miRNAs can simultaneously bind [152] and cooperatively regulate a single target transcript [135, 137, 352], and transcription factors and developmental genes are enriched among genes predicted to be targeted by multiple miRNAs [59, 131].

Here, we performed the first comprehensive analysis of the *Pax6* 3' untranslated region to identify potential miRNA recognition elements. Using miRNA expression profiling experiments, we identified potential cell/tissue-specific miRNA codes. Finally, *Pax6* 3'UTR pull-down studies were used to identify a cohort of miRNA interactors in pancreatic αTC1-6 cells which, based on the spacing of their recognition sites in the

Pax6 3'UTR, revealed 3 clusters where cooperative miRNA regulation may be occurring. Our findings define the functional landscape of miRNA regulation of PAX6 expression.

3.3 Materials and methods

3.3.1 MRE prediction and selection

"Probability of Interaction by Target Site Accessibility" (PITA) was used to identify all miRNAs predicted to target 876bp of the mouse *Pax6* 3'UTR [58]. The miRNA target site prediction tools TargetScan [49], MicroCosm [143], Diana-microT [145] and miRanda [148] were also used. Only 7mer-m8, 7mer-A1 and 8mer MREs with or without a single G:U "wobble" pair, and 6mer MREs without a single G:U pair were considered for further analysis. The ImiRP target site prediction tool [353] was used to validate MRE type of PITA predictions. Multiz alignment of 60 vertebrate *Pax6* 3'UTR sequences available through the University of California Santa Cruz Genome Browser [280] was used to access MRE conservation, and miRBase was used to determine the organisms in which miRNAs of interested are expressed.

3.3.2 Animals

All research on mice was performed with approval of the University of Victoria Animal Care Committee in compliance with the Canadian Council on Animal Care (CCAC) guidelines for the ethical treatment of research animals. All experiments in this study were performed on 129S1 mice (strain #002448, The Jackson Laboratory, Bar Harbor, ME)

3.3.3 Tissue Harvesting and RNA Isolation

Adult tissues were harvested from male and female 2-month-old 129S1. For harvest of embryonic retina, 129S1 pregnant dams were euthanized 12 days post-coitus (E12.5). All tissues were dissected in ice-cold phosphate buffered saline (PBS).

To isolate embryonic day 12.5 retinal tissue from other ocular tissues, developing retinas were first separated from retina pigment epithelium (RPE) in PBS containing 500 units/ml DNase I (ThermoFisher, 18047019) and then incubated for 10 minutes in Hank's Balanced Salt Solution (ThermoFisher, 14025092) containing 0.8 units/ml Dispase II (Roche, 04942078001) with 5% carbon dioxide prior to dissection of the lens. All E12.5 retinas from a single litter were pooled, and three individual litters were collected.

All dissected tissues were put into 1 ml TRIzol (ThermoFisher, 15596-018) in tissue homogenizing tubes (Precellys, BER-KT0396110092) and homogenized for 1 minute at 3000 RPM using a digital disruptor genie (Scientific Industries, SI-DD38). Total RNA isolation was carried out according to manufacturer's protocol and RNA was resuspended in nuclease-free water and quantitated using a NanoDrop ND Spectrophotometer.

3.3.4 3'RACE (Rapid Amplification of cDNA ends)

3' RACE was performed based on a previously described protocol [354]. Briefly, Reverse transcription using the GeneRacer kit (Invitrogen) was carried out on 1 µg of total RNA using the Anchored PolyT Reverse primer to prime cDNA synthesis. Nested

amplification of cDNA ends was performed using Q5 High-Fidelity enzyme (New England BioLabs) using primer combinations F1+R1 followed by nested primers F2+R2 to amplify the *Pax6* 3'UTR, or F3+R1 followed by F4+R2 to amplify the reverse orientation transcript. The first amplification steps were carried out as follows: one 2-minute denaturation cycle at 94°C followed by 30 cycles consisting of 10 seconds at 98°C, 30 seconds at 70°C, 2 minutes at 72°C; and a final extension cycle for five minutes at 72°C. The nested amplification steps were carried out using the following conditions: one 2-minute denaturation cycle at 94°C followed by 30 cycles consisting of 10 seconds at 98°C followed by 2.5 minutes at 72°C; and a final extension cycle for five minutes at 72°C.

3' RACE Primers:

Anchored PolyT Reverse: 5'-

GCTCGCGAGCGCGTTTAAACGCGCACGCGTTTTTTTTTTTTTTTTTTVN-3'

R1: 5'-GCTCGCGAGCGCGTTTAAAC-3'

R2: 5'-GCGTTTAAACGCGCACGCGT-3'

F1: 5'-TGTCCTGAACTGGAGCCCGGGAATGGA-3'

F2: 5'-GGACCTTTGCGTACAGAAGGCACGGTAT-3'

F3: 5'- TAATCTAGGCCAGGACC-3'

F4: 5'-TTCCTAGTGAATCCCTTGTTGC-3'

3.3.5 RNA sequencing

Pre-enrichment of polyA RNA, standard sequencing library/sample preparation, and Illumina sequencing was performed by LC Sciences (Houston, TX, USA).

3.3.6 MS2-MBP and MS2 Binding Site Plasmids

The expression vector expressing the MS2-MBP fusion protein was a gift from Melissa Moore (University of Massachusetts, Worcester). The fusion protein was purified as described in Jurica *et al.* (2002) [355].

890bp of the mouse *Pax6* 3'UTR followed by the SV40 early polyadenylation signal and flanked by Spe1 and Afl2 restriction sites was synthesized by BioBasic Inc., and cloned downstream of the Turbo Green Fluorescent Protein (TurboGFP) gene in the pCMV-TurboGFP-dest1 plasmid vector (Evrogen, FP519) using Xba1 and Afl2 restriction sites. The Q5 site-directed mutagenesis kit (NEB, E0554S) was used to introduce a 2-nucleotide substitution into the miR-375 MRE located at *Pax6* 3'UTR position 201 in the TurboGFP-dest1-3'UTR plasmid. The primers used for miR-375 MRE mutation were F: TATCAGTTGGggCAAATCTTCATTTTGGTATCCAAAC and R: CCGTGCCTTCTGTACGCA, where lowercase letters indicate mutated sequences. We designed a sequence containing a tandem array of three MS2 binding sites, ACATGAGGATCACCCATGT, interspersed by 17 random nucleotides [356, 357]. The Q5 site-directed mutagenesis kit was used to introduce the MS2 binding sequence into TurboGFP-dest1 and TurboGFP-dest1-3'UTR plasmids using the following primers: Common forward: AGGATCACCCATGTCTCGGGAGTACCAGAGAACATGAGGATCACCCATGTAG-GTCCGTCATAATCAGCCATACCACA; TurboGFP-dest1-specific reverse:

CATGTCTTTATCATGACG-

AAGTACATGGGTGATCCTCATGTTTGACATGCTCTAGAGTCGCGGCCGATCC; TurboGFP-dest1-3'UTR-specific reverse:

CATGTCTTTATCATGACGAAGTACATGGGTGATCCTCATGTTTGACATGCAG-

GTTTAAAACTCTTGCAAG.

3.3.7 TurboGFP qPCR Primer Design and Efficiency

Primer-BLAST [358] was used to design quantitative PCR (qPCR) primers for quantification of TurboGFP transcript, with an amplicon size of 100-150 nucleotides. Selected primers had the sequences F: CCCGCATCGAGAAGTACGAG, R: GCGGATGATCTTGTCGGTGA. Primer pair efficiency was calculated using a ten-fold dilution series of cDNA prepared from TurboGFP-transfected αTC1-6 (ATCC CRL-2934) cell lysate. Cycle threshold (Ct) values were plotted versus dilution factors in a base-10 semi-logarithmic graph, the correlation coefficient was confirmed to be greater than 0.99, and amplification efficiency was calculated as $10^{(1/slope)}$. Primer efficiency was found to be 1.952.

3.3.8 Cell Culturing and Transfection

Mouse transgenic pancreatic alpha cells (αTC1-6, ATCC CRL-2934) were cultured in DMEM (low glucose, pyruvate, ThermoFisher, 31600-034) supplemented with 10% FBS (Life Technologies, 16000-044), 15 mM HEPES, 0.1 mM non-essential amino acids (ThermoFisher, 11140-050), 0.02% BSA (Sigma-Aldrich, A7906-50G), 1.5 g/L sodium bicarbonate, and 2.0 g/L glucose. Mouse transgenic pancreatic beta cells (βTC-6, ATCC

CRL-11506) were cultured in DMEM (ThermoFisher, 11960-044) supplemented with 4 mM L-glutamine (ThermoFisher, 25030), 1 mM sodium pyruvate (ThermoFisher, 11360), and 15% FBS. Both cell types were cultured at 37°C with 5% carbon dioxide. For total RNA harvest, cells were grown to approximately 80% confluence in 100mm culture dishes, washed once with PBS, and lysed with 1 ml TRIzol.

αTC1-6 cells were transfected in 6-well dishes with TurboGFP MS2 binding site plasmids 24 hours post-seeding, using jetPRIME transfection reagent (Polyplus, 114-07), following the manufacturer's protocol. Each well was transfected with 3ug of plasmid DNA and 6 μl jetPRIME reagent in 200 μl jetPRIME buffer. 48 hours post-transfection, cells were washed with PBS and lysed using a non-denaturing lysis buffer containing 20 mM Tris pH 7.5, 200 mM NaCl, 2.5 mM $MgCl_2$, 0.05% IGEPAL, 60 U/ml Superase-In (Ambion, AM2696), 1 mM DTT, and Complete protease inhibitor (Roche, 04693124001).

3.3.9 miTRAP

miTRAP protocol was performed as described in [152]. Cell lysates were briefly incubated on ice and supernatant was removed following centrifugation. 25 μl magnetic amylose beads (NEB, E8035S) per sample were blocked in lysis buffer containing 0.2 μg/μl yeast tRNA (Invitrogen, 15401-029) and 0.2 μg/μl BSA (Ambion, AM2616), and then bound with 2.5 μg MS2-MBP protein. Supernatants were incubated with MS2-MBP bound beads for 3 hours at 4°C. Beads were washed and resuspended in 50 μl lysis buffer and transferred to 1 ml TRIzol LS (ThermoFisher, 10296028) and RNA was isolated according to manufacturer's protocol. 1 μl RNA grade glycogen (ThermoFisher, R0551) was added to the aqueous phase of miTRAP products to improve RNA precipitation.

3.3.10 cDNA Preparation

15 μl reverse transcriptase (RT) reactions were prepared for miRNA RT-qPCRs. For miRNA profiling experiments and miTRAP experiments using TaqMan multiplex qPCR arrays, 750 ng and 60 ng RNA, respectively, was used per RT reaction. For miTRAP experiments using individual small RNA TaqMan assays for miR-375, 22.5 ng RNA was used per RT reaction. RT reactions were prepared using TaqMan MicroRNA Reverse Transcription kit (ThermoFisher, 4366596). A custom RT primer pool containing primers specific for the 95 selected miRNAs (ThermoFisher, 4449141) was used to prepare cDNA for TaqMan multiplex arrays, and an individual RT primer specific for miR-375 (ThermoFisher, 4427975, 000564) was used to prepare cDNA for individual miRNA qPCRs. The RT reactions were run following the manufacturer's protocols for custom TaqMan array miRNA cards with preamplification and for TaqMan small RNA assays.

TurboGFP cDNA synthesis was performed in 20 μl reactions using the Quantitect reverse transcription kit (Qiagen, 205311) following the manufacturer's protocols. 60 ng and 100 ng RNA per reaction were used to prepare TurboGFP cDNA for normalizing results from TaqMan multiplex arrays and individual miRNA qPCRs, respectively.

3.3.11 Quantitative PCR

miTRAP cDNA for use with TaqMan miRNA multiplex arrays was first preamplified using TaqMan PreAmp Master Mix (ThermoFisher, 4391128) and custom miRNA PreAmp primer pool. The qPCR reactions for miRNA profiling experiments and miTRAP experiments were prepared using TaqMan Universal Master Mix II with UNG (ThermoFisher, 4440038), and custom miRNA microfluidic cards were run on an Applied

Biosystems 7900 HT Fast Real Time PCR System fitted with the 384-well block. qPCR reactions for miR-375 were prepared using Universal Master Mix II with UNG and TaqMan small RNA assay for miR-375 (ThermoFisher, 4427975, 000564), and were run in MicroAmp fast 96-well reaction plates (0.1 ml, Applied Biosystems, 4346907) covered with optical adhesive covers (Applied Biosystems, 4360954) using the 7900 HT Fast Real Time PCR System fitted with the 96-well block. All protocols were performed following the manufacturer's instructions.

qPCR reactions for TurboGFP were performed using QuantiTect SYBR Green PCR kit (204143) and run using the Stratagene Mx300P qPCR system (Agilent Genomics). PCR reaction settings were as follows: hot start for 15 min at 95°C, and amplification 15 sec at 95°C, 30 sec at 60°C, 30 sec at 72°C repeated for 40 cycles with data recorded twice during the extension step.

3.3.12 Data Analysis

For miRNA tissue profiling analysis, three independent samples were collected and a miRNA was considered to be expressed only if it had a Ct value of less than 40 in all three samples. Levels of each miRNA were calculated relative to U6 snRNA by ΔCt = Ct(U6) – Ct(miRNA), and Relative miRNA level = $2^{\Delta Ct}$.

For miTRAP experiments, any samples for which Ct(no RT control) – Ct(+RT) was less than 10 were excluded from the analysis. Except for qPCR reactions performed using the TaqMan multiplex array cards, all qPCR reactions were performed in triplicate. If the standard deviation of the Ct value between triplicate technical replicates was greater than 0.5, the sample was discarded from the analysis. miTRAP experiments using

multiplex array cards were performed in quadruplicate, and miRNAs were only considered detected if the Ct values for purification with the wild type *Pax6* 3'UTR were less than 40 in all four replicates. If any Ct values for GFP-MS2 control purifications were undefined (i.e. greater than 40), but the miRNA was detected in all four purifications with the wild type *Pax6* 3'UTR, the undefined Ct was defined as 40.

Relative GFP expression for TurboGFP affinity purification analysis with and without the MS2 binding sites was calculated using Pfaffl's method [359] without a reference gene, where ΔCt = Ct[GFP] – Ct[GFP-MS2] and GFP Fold difference = $1.952^{\Delta Ct}$. Analysis of miR-375 purification with and without transfection of TurboGFP-MS2 was performed using the comparative Ct method without normalization to an internal control, where ΔCt = Ct[Untransfected] – Ct[GFP-MS2] and miR-375 Fold difference = $2^{\Delta Ct}$. Normalized relative miR-375 quantity (NRQ) with the wild type and miR-375 MRE mutant *Pax6* 3'UTRs was performed using Pfaffl's method with qBase+ software [360]. NRQ was calculated by ΔCt = Ct[GFP-MS2] – Ct[GFP-WT3'UTR-MS2] or ΔCt = Ct[WT3'UTR-MS2] - Ct[375MUT3'UTR-MS2], and NRQ = $2^{\Delta Ct[miRNA]}/1.952^{\Delta Ct[GFP]}$. Mann-Whitney U test was used for statistical analysis of qPCR data.

miTRAP ratio was calculated by dividing the relative miRNA abundance from the *Pax6* 3'UTR pull-down by the relative miRNA abundance in αTC1-6 cell lysate. Mean Ct for each miRNA from the four *Pax6* 3'UTR pull-down replicates was calculated. Relative miRNA abundance with the *Pax6* 3'UTR was calculated by $1.952^{\text{mean Ct[GFP]}}/2^{\text{mean Ct [miRNA}}$.

3.4 Results and discussion

3.4.1 Characterization of *Pax6* 3'UTR length

As a first step to determine how microRNAs regulate the expression of *Pax6*, we sought to identify a region that best represents the *Pax6* 3'UTR. The initial characterization of the mouse *Pax6* mRNA [177] identified a 3'UTR that was 1008 nucleotides long, however, the absence of a poly(A) signal at the end of the sequence made it unclear whether this represented the complete 3'UTR region. There are three putative polyadenylation (poly(A)) signals within the first 2000 nucleotides downstream of the mouse *Pax6* stop codon at positions 688, 861 and 1930 (Figure 17A). These three poly(A) signals are conserved at positions 714, 882 and 1955 in the human *PAX6* 3'UTR. Human *PAX6* encodes three additional poly(A) signals at nucleotide positions 488, 588 and 1661. Among the six human *PAX6* poly(A) signals, only those at positions 714 and 882 are conserved in all of the 23 amniote species examined (Figure 17B), making them good candidates for functional poly(A) signals *in vivo*.

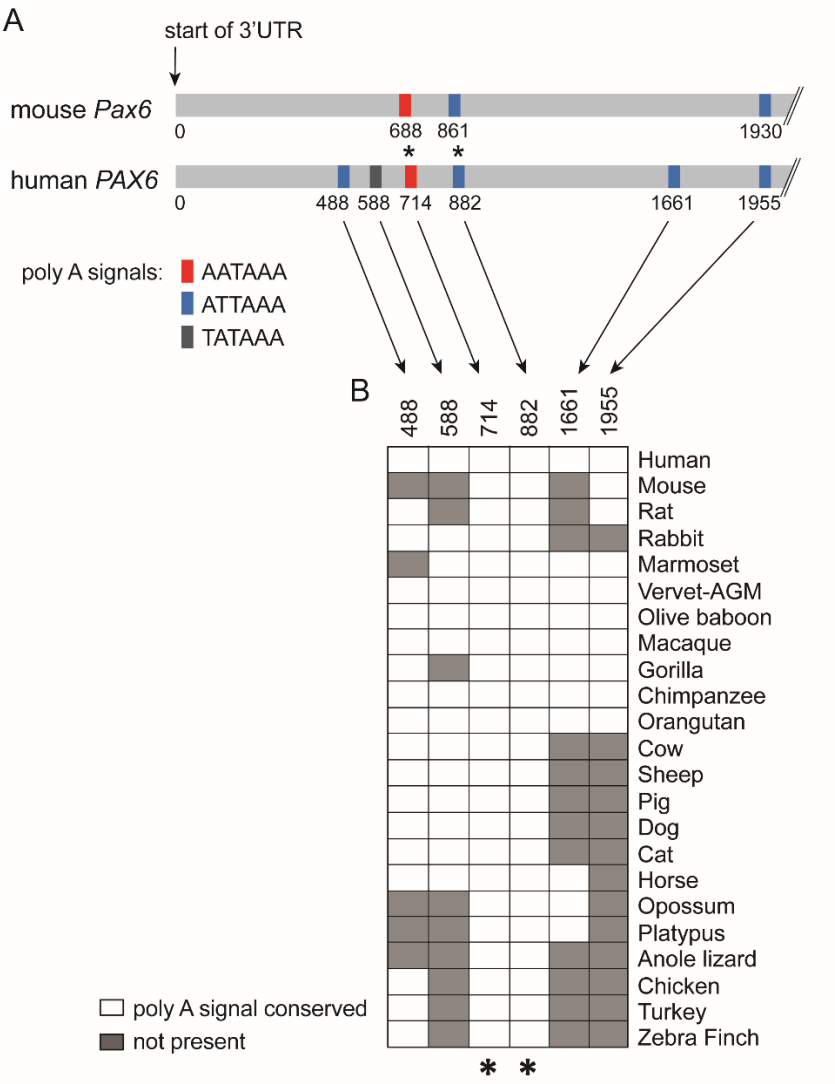


Figure 17. Predicted vertebrate Pax6 polyadenylation signals and conservation

(A) Location of putative polyadenylation signals (AATAAA, ATTAAA, TATAAA) in the mouse and human *Pax6*/*PAX6* genomic region downstream of the stop codon. Neither mouse nor human *Pax6*/*PAX6* have a AGTAAA polyadenylation signal within the region indicated. (B) Table showing conservation of the human *PAX6* polyadenylation signals across amniotes. Asterisks in (A) and (B) indicate two polyadenylation signals that are conserved in all of the species examined.

Two experimental approaches were next used to determine the length of *Pax6* 3'UTR. First, 3' rapid amplification of cDNA ends (3' RACE) was performed on adult mouse retinal total RNA (Figure 18A,B). This approach yielded a major band, which

based on our experimental design (Figure 18A), corresponds to a 3'UTR of 877 nucleotides. In addition to this major band, a number of weaker bands, both smaller and larger, were observed (Figure 18B). Subcloning and sequencing revealed that many of the 3'RACE products terminated after the poly(A) sequence at nucleotide position 861 of the 3'UTR (data not shown).

As a second approach, RNA-seq was performed on adult eye poly(A)-selected RNA (Figure 18C and D, top plot). Similar to the 3'RACE results, RNA-seq reads revealed a Pax6 3'UTR endpoint positioned in close proximity to the 861 poly(A) sequence (Figure 18, red dashed vertical line). A "CA" positioned at the end of this region at nucleotide 870 (Figure 18E) and 17 nucleotides upstream of a 16 nucleotide stretch that is 87.5% U/G is a good candidate for functioning in transcript cleavage. CA sequences facilitate transcript cleavage when positioned 15-30 nucleotides downstream of the poly(A) signal and ~20 nucleotides upstream of a U/G rich region [5]. Since 3'UTR length can be differentially regulated between developmental and adult stages [101, 361] RNA-seq was also performed on poly(A)-selected RNA from embryonic day 14 eye. Similar to the adult eye, RNA-seq reads from embryonic eye cDNA indicate that *Pax6* mRNA utilizes the poly(A) signal at nucleotide 861 (Figure 18D, lower panel).

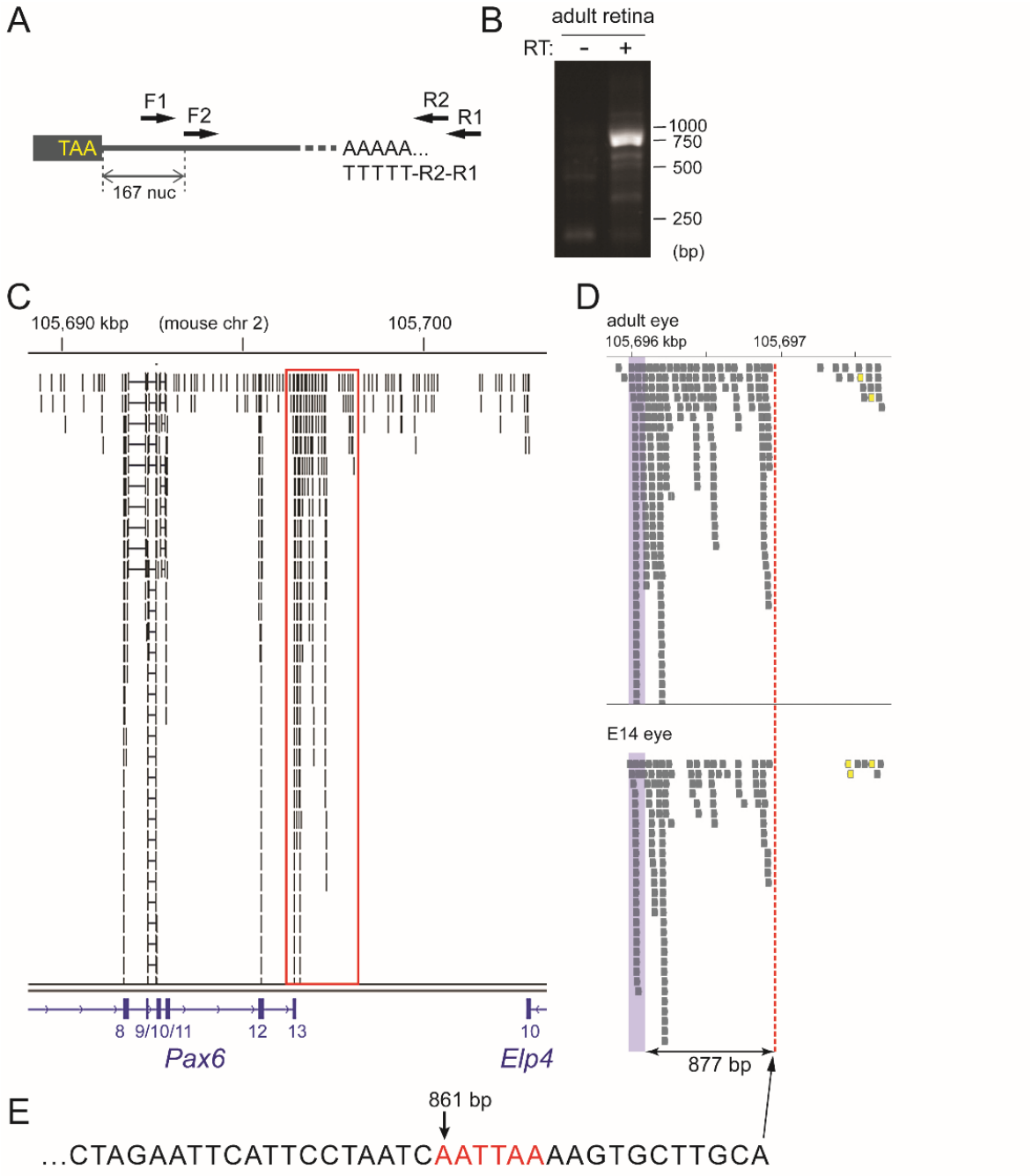

A
F1
F2
R2
R1
TAA
AAAAA...
TTTTT-R2-R1
167 nuc
B
adult retina
RT: - +
1000
750
500
250
(bp)
C
105,690 kbp
(mouse chr 2)
105,700
8
9/10/11
12
13
10
Pax6
Elp4
D
adult eye
105,696 kbp
105,697
E14 eye
877 bp
E
861 bp
...CTAGAATTCATTCCTAATCAATTAAAAGTGCTTGCA

Figure 18. Characterization of the mouse Pax6 mRNA 3' terminus

(A) Amplification strategy used in 3’ RACE approach to identify the *Pax6* mRNA 3’ end. “TAA” represents the Pax6 stop codon. Nested primers F1 and F2 were used in combination with reverse primers (R1, R2) built into the poly-T primer used to generate cDNA. (B) 3’ RACE performed on adult retina total RNA. A predominant band was observed at 700-750 bp in contrast to minus reverse transcriptase (- RT) negative control. Weaker bands above and below the predominant 700-750bp band were also observed. (C) Primary read data from RNA-seq experiment performed on polyadenylation-selected adult mouse eye mRNA. Reads were superimposed onto the 3’ end of the *Pax6* region containing exons 8-13 and the 3’ end of the adjacent gene *Elp4*. The box encompassing *Pax6* exon 13 is shown at a higher magnification in (D – top plot). The light blue shaded regions in (D) indicate the end of the *Pax6* coding region. The dotted line indicates the most 3’ end point for Pax6 in both adult eye (D, top plot) and embryonic day 14 (E14) retina (D, bottom plot). (E) Sequence read corresponding to the most 3’ read indicated by the dotted line in (D). The putative polyadenylation signal located at position 861 of the *Pax6* 3’ UTR is highlighted in red.

Interestingly, in both adult and embryonic eye tissues, sequence reads in reverse orientation to *Pax6* were observed immediately downstream of the 861 poly(A) signal (Figure 18D, yellow colored reads). 3’ RACE revealed a reverse transcript that is polyadenylated and utilizes a robust poly(A) signal (AATAAA) located 112 bp downstream of the *Pax6* 861 poly(A) signal (Figure 19A,B). Unlike the tissue-specific expression of *Pax6*, this reverse transcript appears to be expressed ubiquitously (Figure 19C). It is unclear whether this reverse orientation transcript is part of the 3'UTR for the adjacent *Elp4* gene which is transcribed in opposite orientation to *Pax6* and whose stop codon is ~5.5 kb away from the *Pax6* 861 poly(A) signal. Currently no non-coding RNAs map to this region.

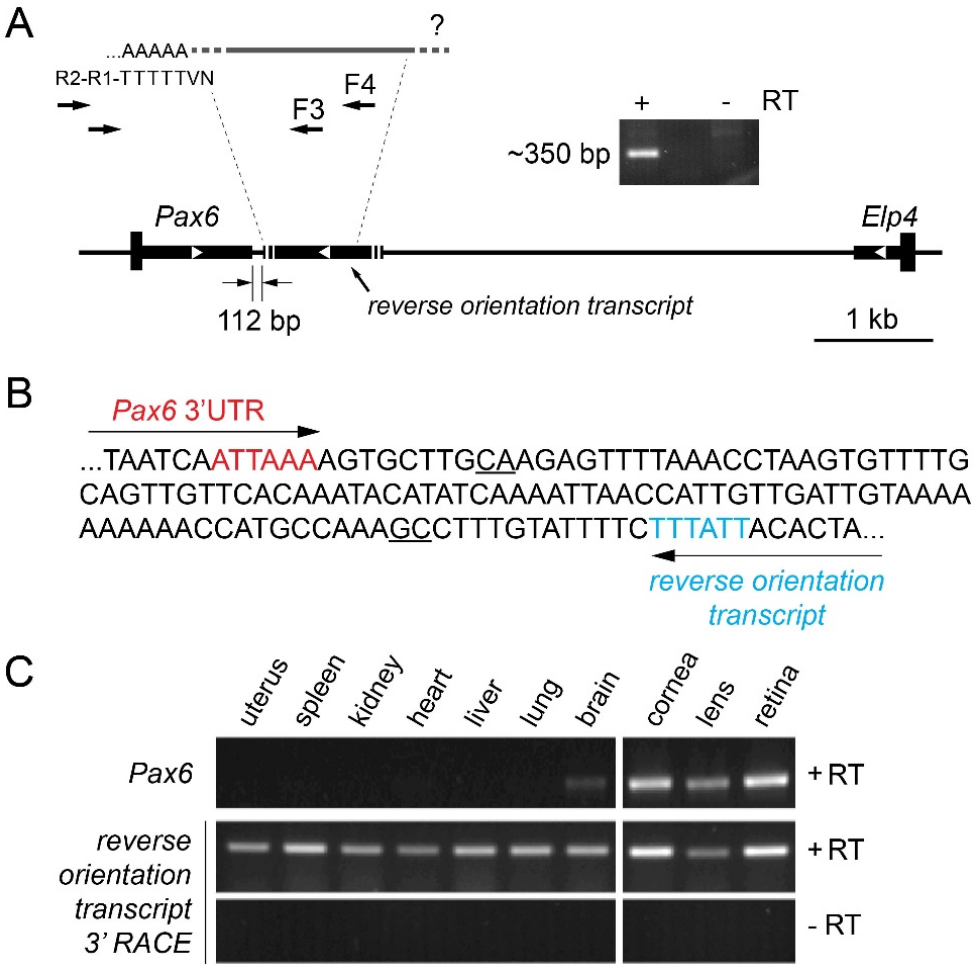


Figure 19. Characterization of a reverse orientation transcript terminating directly adjacent to the Pax6 mRNA 3' terminus

(A) 3' RACE amplification strategy used identify a reverse orientation transcript positioned close to the *Pax6* mRNA 3' end. The wider bars in the schematic diagram represent coding regions and the narrower bars represent 3'UTRs. Nested primers F3 and F4 were used in combination with reverse primers (R1, R2) built into the poly-T primer to amplify a robust ~350 base pair product. (B) Sequencing of the reverse orientation transcript shows that it terminates at a "GC" (underlined). The reverse transcript appears to utilize a polyA signal (blue font) that is 112 bp away from the end of the polyA signal at position 861 of the *Pax6* 3'UTR (red font). (C) 3'RACE of the reverse orientation transcript (bottom 2 panels) and reverse transcription of *Pax6* (top panel) showing widespread expression of the revere orientation transcript in contrast to the tissue specific expression of *Pax6*.

In summary, these data indicate that the majority of mouse *Pax6* transcripts predominantly utilize a highly conserved poly(A) signal at nucleotide 861 of the 3'UTR.

The remainder of this study examining the regulation of *Pax6* by microRNA will therefore focus on a 3'UTR terminating at this region. It should be noted, however, that the *Pax6* 3'UTR length is not rigidly fixed, as smaller and larger lengths (albeit less abundant) were observed. Indeed, the characterization of the original mouse *Pax6* mRNA clone [177], which focused on the largest clone obtained in that study, possessed a 3'UTR of 1008 nucleotides. Although we saw no apparent developmental differences in *Pax6* 3'UTR length we cannot rule out the possibility that *Pax6* 3'UTR length is differentially regulated in other cellular and developmental contexts.

3.4.2 Identification of predicted miRNA target sites within the mouse *Pax6* 3'UTR

To identify candidate miRNAs predicted to regulate *Pax6,* we ran the mouse *Pax6* 3'UTR through an unbiased bioinformatics screen. Although functional miRNA recognition elements (MREs) have been identified within the coding regions of several genes [36, 37, 39], and MRE clusters have been identified in 5'UTRs and coding regions by Ago HITS-CLIP [40], we chose to focus on the *Pax6* 3'UTR because this is the region most commonly involved in miRNA regulation [51]. We used the prediction tool Probability of Interaction by Target Accessibility (PITA) [58] because it screens for all MREs within a mRNA sequence of interest.

Our PITA screen included the four types of MREs that are matched to the miRNA 5' end and known to be selectively conserved: 6mer, 7mer-A1, 7mer-m8 and 8mer (Figure 3) [19]. The 6mer MRE is perfectly complementary to the "miRNA seed", miRNA nucleotide positions 2-7. The 7mer-A1 MRE consists of a seed match with an A across from miRNA nucleotide 1, 7mer-m8 consists of a seed match with a complementary

match to miRNA position 8, and the 8mer MRE consists of a seed match with both an A1 and m8. MRE types that were not considered in this analysis were offset 6mer (OS-6mer) sites, complementary to miRNA positions 3-8 [20, 51], and two site types identified by Argonaute High-Throughput Sequencing of RNA isolated by crosslinking immunoprecipitation (Ago HITS-CLIP), 6merα and G-bulge sites [54, 57]. Using PITA, 6665 unique hits, representing unique predicted miRNA-MRE interactions, were identified within 876 bp of the mouse *Pax6* 3'UTR (Fig 3A). Though this may seem like many predictions, this is not surprising. 4526 of these predictions are MREs harbouring both a mismatched pair and G:U pair between the miRNA and MRE. Additionally, 1187 of these predictions are 6mer MREs harbouring a single mismatch between the miRNA and MRE. Given that the PITA miRNA database contains 491 mouse miRNAs, the prediction of these types of MREs by chance is high.

In order to focus our analysis on those miRNAs most likely to regulate *Pax6*, we next screened the 6665 *Pax6* hits against a set of hierarchical criteria (Figure 20A). Mismatches between a miRNA and MRE are less likely to be associated with functional target sites [52] and are rarely under selection to be evolutionarily conserved [20]; these predicted MREs were excluded, narrowing the total number of predicted hits to 449. miRNA-target interactions containing single G:U "wobble" pairs can function effectively in downregulation endogenously [149] and in reporter screens [137], and simulations of Ago-miRNA:mRNA complexes and Ago HITS-CLIP data suggest that G:U wobbles can be tolerated [40, 362]. However, though 7 and 8mer MREs containing single G:U pairs functioned in down-regulation of reporters in *Drosophila in vivo*, 6mer MREs containing

single G:U pairs were non-functional [52]. For this reason, we chose to exclude 6mer+G:U MREs from consideration, which further narrowed the number of predicted hits to 191.

Target sites that show high conservation between orthologous 3'UTRs are more likely to be functional [19, 43]. Using UCSC Genome Browser [363] alignments, a total of 62 predicted MREs were found to be conserved in ≥85% of orthologous *Pax6* 3'UTR sequences from placental mammals. Finally, as MREs for miRNAs that are broadly conserved are more likely to be functionally conserved [20]. Using miRBase [18], we identified and excluded miRNAs that are rodent-specific and retained miRNAs that are found in humans. This reduced the number of unique hits to 55. Given that some MREs are predicted to be targeted by multiple miRNAs, and some miRNAs are predicted to target multiple MREs within the *Pax6* 3'UTR, we identified 47 candidate miRNAs of interest and 42 candidate MREs (Figure 20A, Appendix B Table 10). 12 of these MREs are predicted to be targeted by multiple miRNAs and 25 of the identified miRNAs are predicted to target multiple MREs within the *Pax6* 3'UTR. Of these 25 miRNAs, 8 have multiple MREs that are conserved in >85% of placental mammal alignments (Appendix B Table 10). In summary, we have identified 47 MREs within the *Pax6* 3'UTR that satisfy our selection criteria as strong candidates for miRNA interaction. Some of the MREs and the miRNAs that target them exhibit redundancy, which should be considered for any functional analyses. It should be recognized that although these 47 MREs represent strong candidates for miRNA interaction, it does not necessarily mean that they are functional *in vivo*, nor does it mean that other MREs that did not meet our criteria are

not functional. This analysis, however, does provide a starting point for studying *Pax6* regulation by miRNA.

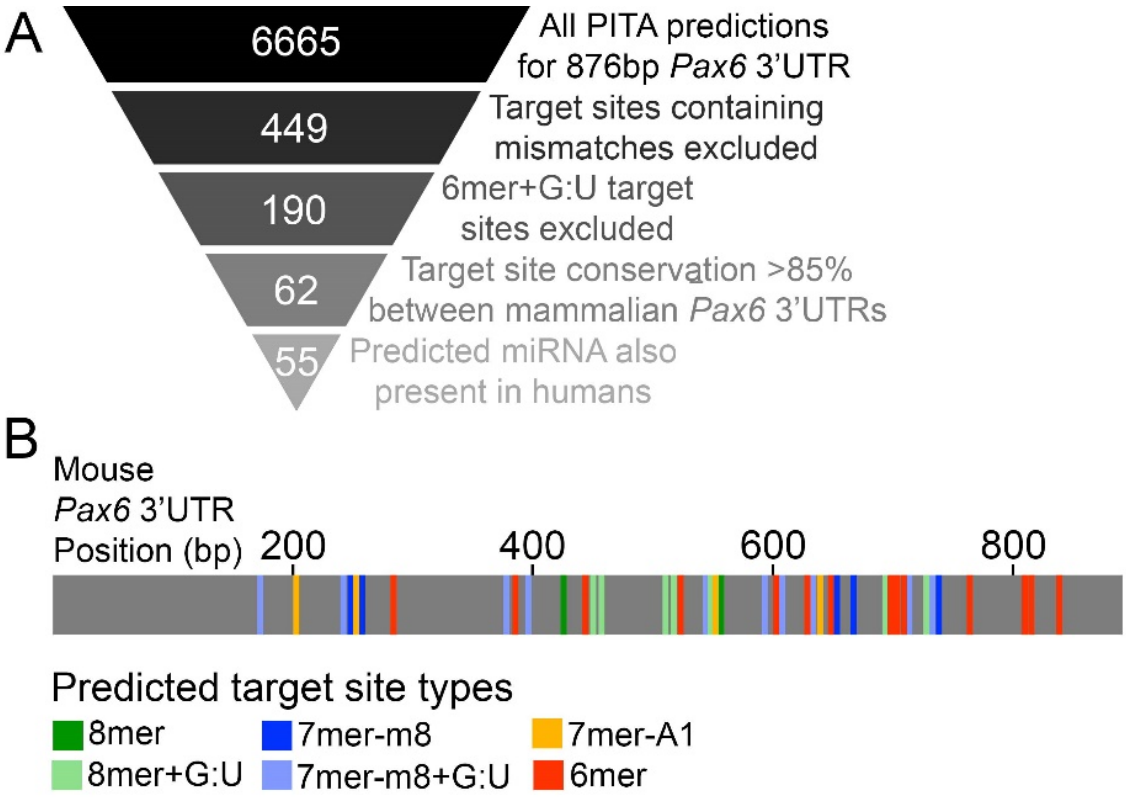


Figure 20. Predicted miRNA target sites in the mouse *Pax6* 3'UTR

(A) Strategy for miRNA target site selection. A total of 6665 unique hits were identified within 876bp of the *Pax6* 3'UTR using the miRNA target site prediction tool PITA. A unique hit encapsulates a single predicted MRE-miRNA interaction. Exclusion of predicted hits containing a single mismatch between the miRNA seed and the target site in the mRNA, and predicted 6mer target sites containing a G:U wobble base pair between the miRNA and the target reduced the number of hits to 449 and then 190, respectively. Requiring a high degree of conservation between orthologous mammalian Pax6 3'UTRs (≥85%), and requiring that targeting miRNAs also be present in humans reduced the number of predicted hits to 62 and 55, respectively. Since some miRNAs share common MREs, 42 unique MREs were identified. (B) Schematic of the mouse *Pax6* 3'UTR showing locations and target site types for the 42 unique MREs. These MREs are predicted to be targeted by 47 different miRNAs. Some miRNAs are predicted to target multiple sites within the *Pax6* 3'UTR.

3.4.3 Expression profiling of miRNAs predicted to target the mouse Pax6 3'UTR

We next sought to determine the expression profile of the miRNAs identified above that are predicted to target to *Pax6* 3'UTR. Our analysis focused on Pax6-expressing ocular tissues (embryonic day (E) 12.5 retina, adult retina and adult lens) and on the SV40-induced cell lines α-TC1-6 and β-TC6 which serve as in vitro models for endocrine pancreas α and β cells, respectively [364, 365]. Both cell lines endogenously express PAX6 [278, 366]. MiRNA expression was detected using TaqMan multiplex quantitative PCR (qPCR) arrays (Applied Biosystems). Extra wells available on the arrays allowed for the detection of additional miRNAs predicted to target less well conserved 7mer and 8mer MREs for expression profiling. Fifty-nine of the miRNAs examined were detected in at least one of the profiled cells or tissues, while 28 of the miRNAs did not have detectable expression in any of the tissues/cells examined (Figure 21A).

Our data revealed distinct tissue/cell-specific miRNA expression patterns arising from overlapping and non-overlapping miRNA expression (Figure 21A). All tissues/cells except for the adult retina had at least one uniquely expressed miRNA. These distinct expression patterns may allow for cell-type specific optimization of Pax6 expression through cooperative miRNA regulation.

The endocrine pancreas α and β cell lines for the most part had very similar miRNA expression patterns (Figure 21B). miR-375 was among the most highly expressed miRNAs in both α and β cells, consistent with previous expression studies which have also reported similar high miR-375 expression levels in endogenous human α and β cells [367]. miR-375, in addition to miR-16, 26a, 26b, 124 and 127, has been identified as an

abundant miRNA in MIN6 and αTC1 cells [368], and human islets [369]. Additionally, many miRNAs identified as islet-enriched relative to acinar cells (miR-7, 96, 127, 132, 183, 335) were identified in αTC1-6 or βTC6 cells by our analysis [370]. A smaller subset of α-specific and β-specific miRNAs was also observed. Some of these (miR-129, 145, 200b, 200c, 369-3p, 429), have previously been shown to be more abundantly expressed in β relative to α cells [367, 368], and serve to validate the veracity of the Taqman array cards. It should be noted that miR-96, a miRNA that we detected at low levels in αTC1-6 but not βTC6, was found to be expressed more abundantly in human β relative to α cells [367], possibly demonstrating a difference between the mouse and human endocrine pancreas cells.

Many of the miRNAs profiled displayed differential patterns of expression between E12.5 retina, adult retina, and adult lens (Figure 21C), similar to observations from microarray analyses finding differential expression of miRNAs in developing retina versus adult retina [371], and between adult ocular tissues [347]. For example, miR-124, 182, 183 and 96 were previously identified as adult retina-enriched miRNAs [372, 373] and these findings have been validated by our analysis. These results suggest that some of the differentially-expressed ocular miRNAs may have tissue-specific functions, potentially through the regulation of *Pax6*. Interestingly, the adult retina had an expression profile that was more similar to adult lens than to the E12.5 retina, which had the fewest miRNAs of the cells/tissue examined (Figure 21C). Additionally, of the miRNAs profiled, E12.5 showed overall lower miRNA expression than adult retina or lens (Figure 22C-E). This is not surprising given that miRNA abundance increases with cellular

differentiation [95, 96] and during development [88, 93, 94, 374], and fits with the proposal that miRNAs function primarily during differentiation and the maintenance of tissue identity [89, 90]. It should be noted that the expression level of some miRNAs in some tissues is highly variable, and this variability is particularly apparent for less abundant miRNAs. Notably, E12.5 retina-expressed miRNAs were especially variable in their level of expression (Figure 22D).

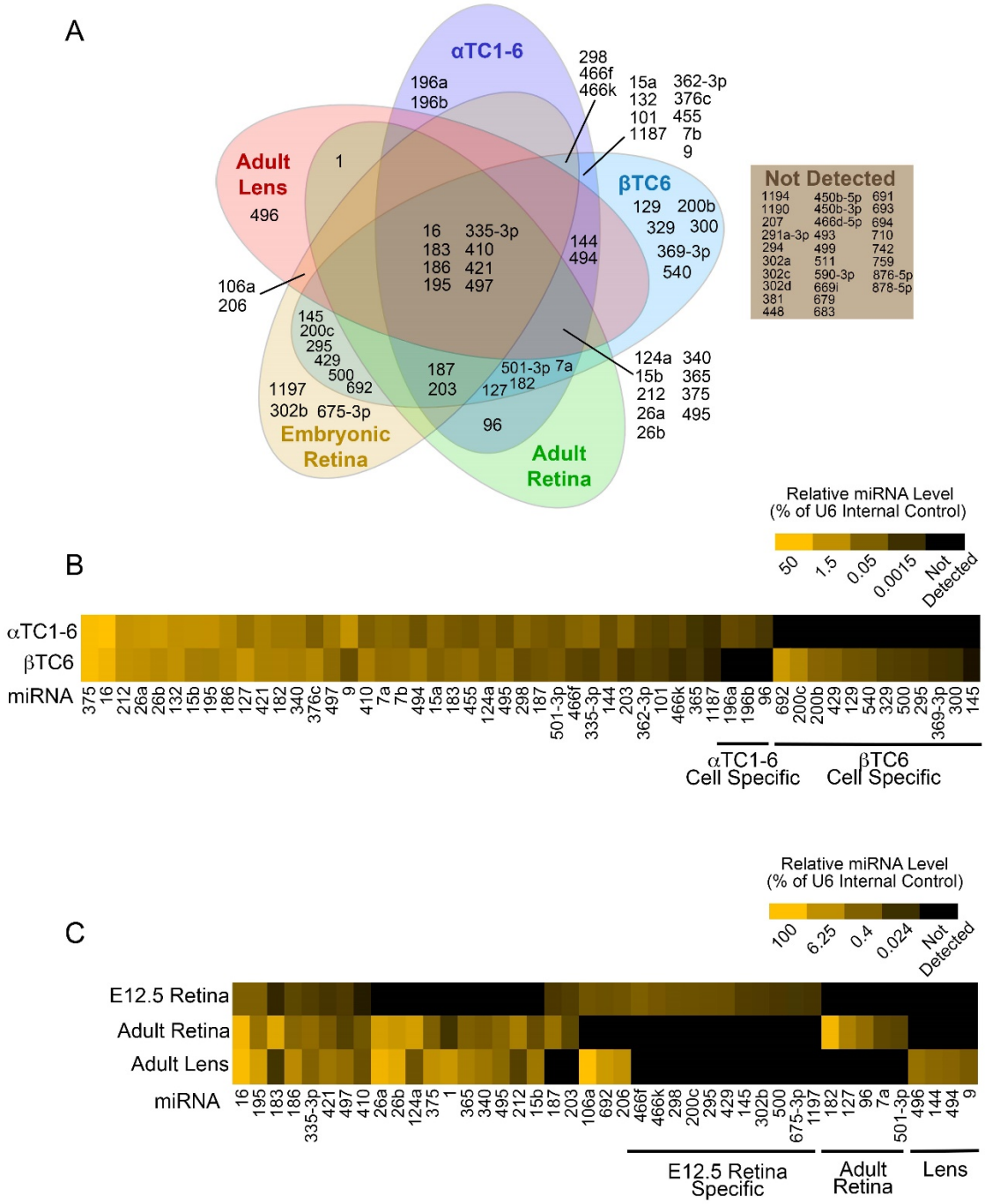


Figure 21. Expression profile of miRNAs predicted to target the mouse *Pax6* 3'UTR

(A) Expression profile of miRNAs having predicted target sites in the mouse *Pax6* 3'UTR in various *Pax6*-expressing cells and tissues. miRNAs assayed for include those identified by our prediction analysis (Fig. 3A) as well as others, as described in the results. miRNAs were assayed for using TaqMan multiplex qPCR array cards. Data represents a total of 3 replicates per tissue or cell type, with a miRNA considered to be expressed only if the cycle threshold was less than 40 for all three replicates. (B) Relative miRNA level in mouse cultured pancreatic α cell line, αTC1-6 and a cultured β cell line, βTC6. (C) Relative miRNA level in mouse E12.5 retina, adult retina, and adult lens. Heat map indicates relative miRNA expression as a percentage of an internal control, snRNA U6. Data represents the geometric mean of three independent samples. Note the scale differences between the two heat maps.

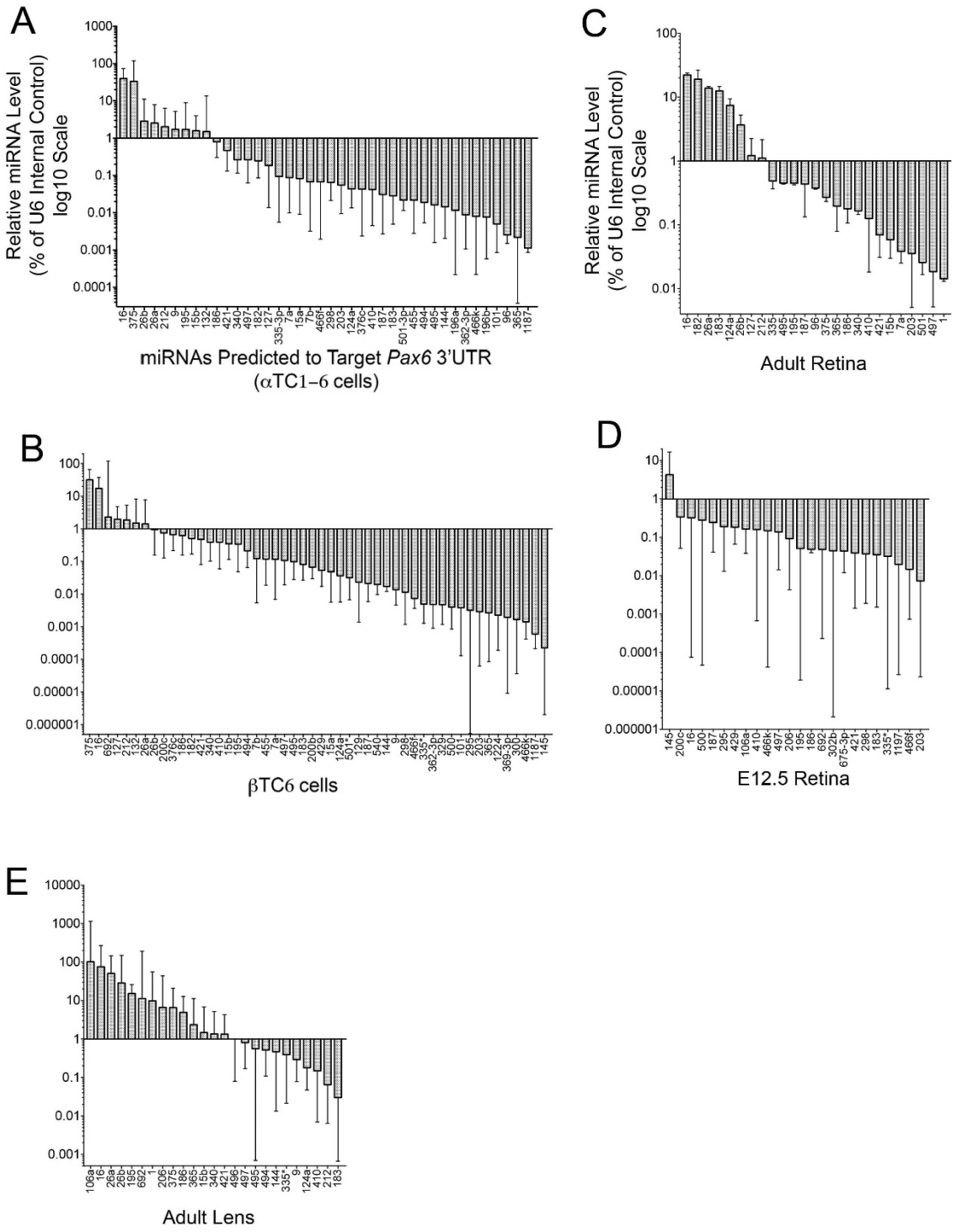


Figure 22. Relative levels of miRNAs predicted to target *Pax6* in *Pax6*-expressing cells and tissues

(A) Relative miRNA level in mouse cultured pancreatic α cell line, α-TC1-6, (B) a cultured β cell line, βTC6, (C) adult retina, (D) E12.5 retina and (E) adult lens expressed as a percent of snRNA U6 internal control. Data represents the geometric mean of three independent samples and error bars represent 95% confidence intervals. Data is represented using log10 scale; note scale differences between graphs.

3.4.4 Characterization of a *Pax6* miR-code in αTC1-6 cells

While the miRNA expression profiles generated above provide valuable information about potential cell and tissue-specific *Pax6* regulation, they do not provide information about whether any of the miRNAs actually interact with the *Pax6* 3'UTR. To address this, we utilized MS2 RNA sequence fused to exogenously-expressed mouse *Pax6* 3'UTR as "bait" RNA to co-purify bound miRNAs in an approach called miRNA Trapping by *in vitro* RNA Affinity Purification (miTRAP) [375]. MS2 is a single-stranded RNA bacteriophage, that binds a specific stem-loop RNA structure present in its genome using a coat protein to accomplish encapsidation [376]. This RNA stem loop and coat protein can be used for RNA affinity purification techniques [355]. miTRAP was originally used to identify miRNAs, miR-1 and miR-133, associated with the 3'UTR of *Hand2* in primary cardiomyocytes [152], and has since been used to identify miRNAs associated with the 3'UTRs of MYC and ZEB2 [375], and with the long intergenic non-coding RNA (lincRNA) p21 [377]. Our approach (summarized in Figure 23) utilized a reporter construct in which GFP was engineered to carry a modified *Pax6* 3'UTR containing three MS2 binding motifs positioned immediately 5' of the poly(A) signal. αTC1-6 cells were transiently transfected with MS2-tagged GFP constructs, and pull-down products were

amplified using TaqMan multiplex qPCR and normalized to GFP mRNA transcripts amplified using SYBR green qPCR.

Three steps were taken to demonstrate the specificity of our miTRAP approach. First, to show that MS2-containing mRNAs were selectively enriched, we compared the pull-down of GFP constructs with or without the MS2 binding motif. GFP reporter transcripts harbouring the 3xMS2 motif exhibited a 30-fold increase in enrichment compared to those lacking the MS2 binding motif (Figure 23B), demonstrating robustness of the pull-down process. Next, we examined the binding of the *Pax6* 3'UTR with a known interactor, miR-375 [276, 277]. Approximately five-fold more miR-375 was pulled-down with GFP constructs carrying the wild type *Pax6* 3'UTR compared to those with an SV40 3'UTR (Figure 23C), indicating *Pax6*-3'UTR mediated enrichment. To determine the specificity of this interaction, pull-down experiments were performed using constructs carrying mutations that disrupt a previously described 7mer-A1 MRE for miR-375 at position 201 of the *Pax6* 3'UTR [276, 277]. Mutation of this MRE resulted in 4-fold reduced binding of miR-375 compared to the wild type *Pax6* 3'UTR (Figure 23D), indicating its requirement for interaction with miR-375. A second miR-375 MRE (6mer site) is also present at position 288 of the *Pax6* 3'UTR, which may explain why binding is not completely lost. It is therefore possible that disruption of both MREs would further reduce miR-375 binding.

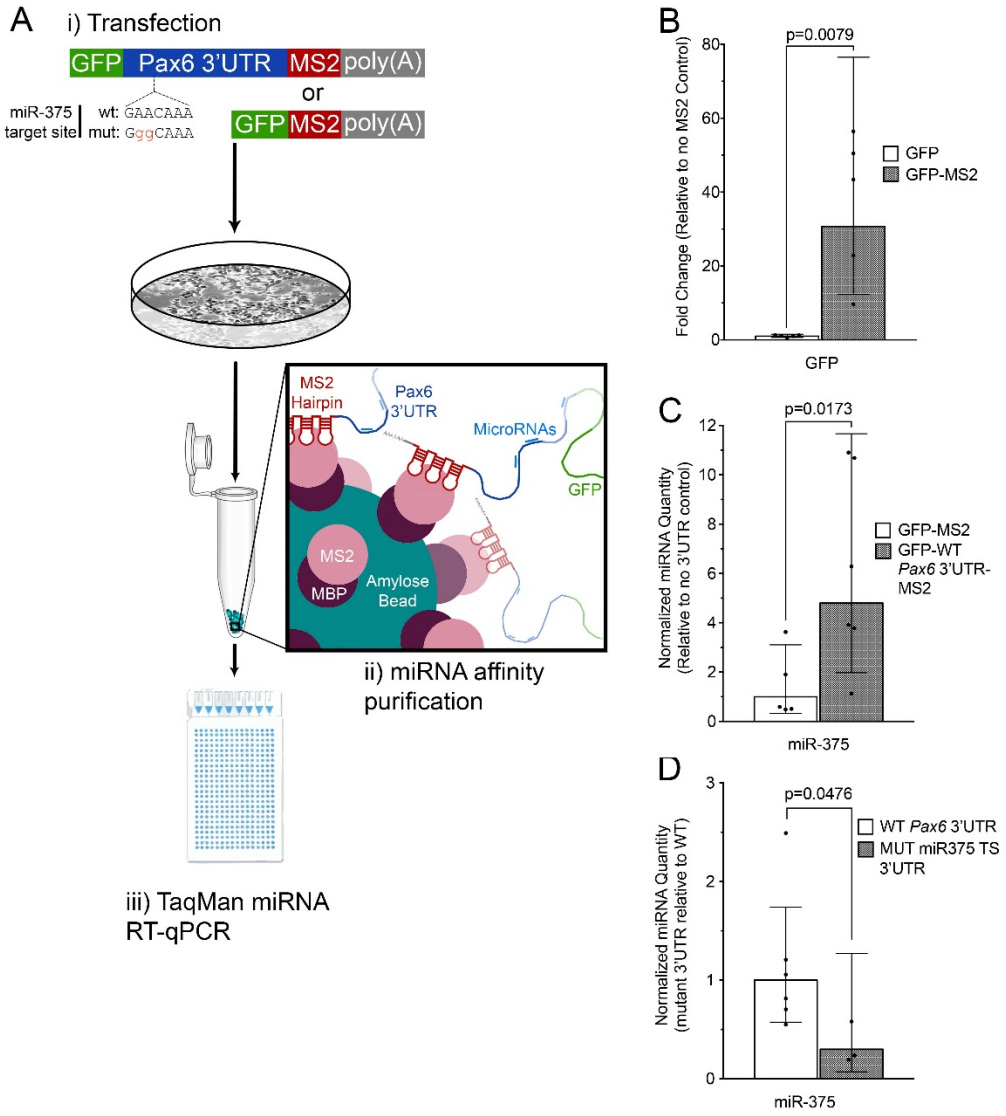


Figure 23. miTRAP as a strategy to purify *Pax6* 3'UTR-associated miRNAs

(A) Schematic of the *Pax6* 3'UTR affinity purification approach. (i) Plasmid vectors expressing GFP tagged with the MS2 RNA sequence motif followed by the SV40 polyadenylation signal are introduced into pancreatic αTC1-6 cells via transient transfection. (ii) MS2 coat protein fused to maltose binding (MS2-MBP) is used to purify GFP transcripts with bound miRNAs from αTC1-6 cell lysate. (iii) Real-time quantitative PCR (RT-qPCR) is used to detect GFP transcript and bound miRNAs. Schematic of the *Pax6* 3'UTR shows the location of the highly conserved miR-375 target site located at 3'UTR position 201 and miR-375 target site mutation. (B) Validation of the MS2-mediated affinity purification strategy by RT-qPCR quantification of GFP transcripts with and without the MS2 RNA motif. Fold change was calculated using Pfaffl's method [359]. qPCR results for GFP with the MS2 motif (grey bar) were expressed relative to data without the MS2 motif (unfilled bar). Data represents 5 independent samples, p=0.0079. (C) Affinity purification of miR-375 with the *Pax6* 3'UTR in αTC1-6 cells using TaqMan individual qPCR assays. Normalized relative quantity was calculated using Pfaffl's method, and a normalized relative quantity greater than 1 indicates that more target miRNA is purified with the *Pax6* 3'UTR (grey bar) than the control lacking a *Pax6* 3'UTR (unfilled bar). Data represents six independent samples, p=0.013. (D) Disruption of miR-375 binding to the *Pax6* 3'UTR following mutation of the miR-375 target site. Target miR-375 values were normalized to GFP as a reference gene, then normalized values for the mutant *Pax6* 3'UTR samples (grey bar) and are presented relative to the wt 3'UTR (unfilled bar). Data represents six independent wt 3'UTR samples and three miR-375 target site mutant 3'UTR samples, p=0.0476. Error bars represent 95% confidence intervals, and p-values were calculated using the Mann Whitney test. Note scale bar differences between the graphs.

The miTRAP approach was next used in conjunction with TaqMan multiplex qPCR arrays to identify the miRNAs that physically interact with the *Pax6* 3'UTR. Relative abundance of target miRNA transcripts associated with the *Pax6* 3'UTR was normalized to GFP and expressed relative to the SV40 3'UTR control using Pfaffl's method to produce a normalized relative quantity (NRQ) value [359]. An NRQ value greater than 1 indicates that more of a given miRNA was purified with the *Pax6* 3'UTR than with the control 3'UTR. Of the 40 miRNAs having MREs in the *Pax6* 3'UTR that were also expressed in cultured αTC1-6 cells (Figure 21), 25 were found to have NRQs greater than 1 (Figure 24A, summarized in Table 3). Eight of these 25 miRNAs, however, had 95%

confidence intervals extending below a NRQ of 1, indicating lower confidence that they bind the *Pax6* 3'UTR preferentially over the control SV40 3'UTR. miR-187 was pulled-down equally using MS2-tagged reporters with and without the *Pax6* 3'UTR suggesting non-specific interaction, while 14 of the 40 miRNAs assayed were undetectable in MS2-tagged *Pax6* 3'UTR reporter pull-downs. As such, we can be reasonably confident that 17 of the assayed miRNAs interact with the *Pax6* 3'UTR.

Table 3. Summary of miTRAP interactions in αTC1-6 cells

miTRAP Results		MRE Type	miRNA
Interaction (62.5%, 25/40)	High Confidence (42.5%, 17.40)	8mer	455
		7mer-m8	127*, 7a+, 7b+, 196a*, 298*, 421*+, 497*, 124a*, 16*, 26a*, 26b*
		7mer-A1	375+, 182*
		6mer	9+, 186, 183
	Low Confidence (20%, 8/40)	8mer	494*
		7mer-m8	132*+, 15b*, 195*, 376c*+++
		6mer	212+, 340+++, 203
Non-specific Interaction (2.5%, 1/40)		7mer-m8	187*
No Interaction (35%, 14/40)		8mer	335-3p*
		7mer-m8	1187*, 15a*, 196b*, 365+, 466k, 501-3p+
		7mer-A1	96*+
		6mer	101, 144, 362-3p, 410, 466f, 495

miRNAs identified in present study as well as by TargetScan, miRanda, MicroCosm, or DIANA-microT are depicted in red

* Predicted MRE-miRNA seed interaction contains a G:U pair

+ indicates miRNAs with additional, weaker MREs in the *Pax6* 3'UTR

miRNAs having MREs that are conserved in less than 85% of placental mammal *Pax6* 3'UTR sequence alignments are depicted in grey

A potential criticism of the miTRAP approach is that it could be biased toward pulling-down abundantly expressed miRNAs while low-expressed miRNA interactors might be missed [375]. We observed that this was not the case. Eleven of the 17 miRNAs pulled-down with the *Pax6* 3'UTR were expressed at levels less than 0.5% relative to the U6 internal control (miR-124a, 127, 182, 183, 196a, 298, 421, 455, 497, 7a and 7b; Supplementary Figure 2A). Notably, miR-196a, which was expressed at 0.01% relative to U6, was one of the more abundantly enriched miRNAs pulled-down with the *Pax6*

3'UTR, having a NRQ of 8.0 (Figure 24A). In contrast, several miRNAs with relative expression levels higher than miR-196a in αTC1-6 (e.g. miR-15a, 144, 335, 410, 466f, 495, 501), were not pulled-down at all (Figure 24C). Furthermore, miR-16 and miR-375, which were by far the most abundantly expressed miRNAs in αTC1-6 cells, had NRQ values (2.8 and 2.3, respectively) lower than that of miR-196a (Figure 24A). It is important to recognize that while our miTRAP data revealed miRNAs with the capacity to interact with the *Pax6* 3'UTR, it may not necessarily reflect miRNA binding to the endogenous *Pax6* 3'UTR. Bait mRNAs used for our experiments were driven from overexpression plasmids and harbour elements such as the MS2 binding domain, which could potentially influence miRNA binding. To address some of these issues and confirm our data, other miRNA:mRNA interaction approaches could be taken [150].

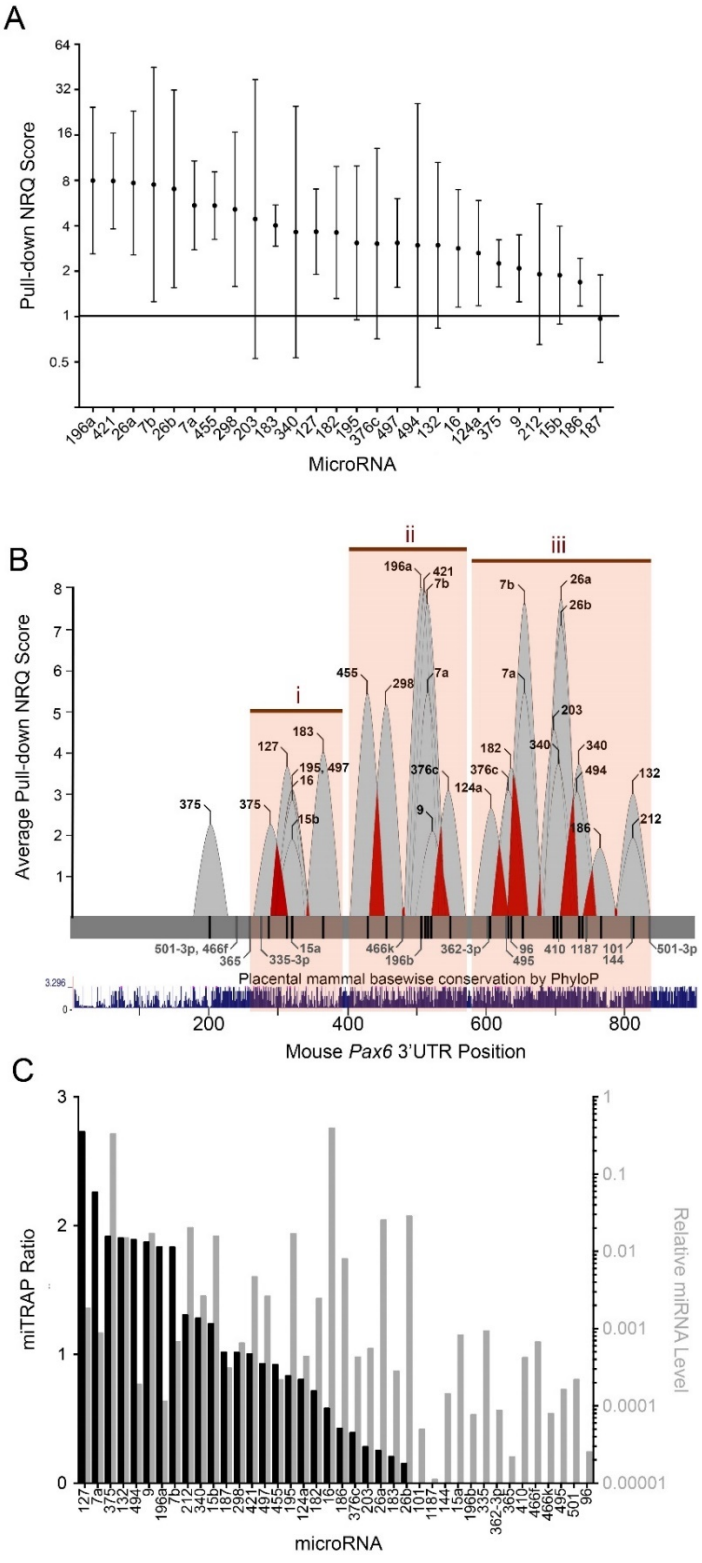


Figure 24. Characterization of miRNAs bound to the *Pax6* 3'UTR in pancreatic α cells

(A) Identification of miRNAs associated with the *Pax6* 3'UTR in αTC1-6 cells using TaqMan multiplex qPCR arrays. Normalized relative quantity (NRQ) greater than 1 indicates more target miRNA was purified with the *Pax6* 3'UTR than the control lacking a *Pax6* 3'UTR. Some miRNAs expressed in α cells that also have predicted MREs in the *Pax6* 3'UTR did not associate with the *Pax6* 3'UTR-containing transcript: miR-101, 1187, 144, 15a, 196b, 335-3p, 362-3p, 365, 410, 466f, 466k, 495, 501-3p, 96. Results represent four independent experiments. Geometric mean +/- 95% confidence intervals are shown. (B) Landscape of *Pax6* 3'UTR miRNA interaction in αTC1-6 cells. Average NRQ value for each interacting miRNA is indicated as a peak at the 3'UTR position(s) of the predicted MRE(s). Predicted MREs for interacting miRNAs are labeled in black, and non-interacting miRNAs in grey. Each peak has a 25-nucleotide width on either side of the MRE position. Overlap between interaction peaks spaced 8-50 nucleotides apart is indicated in red, and these interacting miRNAs may be capable of mediating cooperative regulation of *Pax6.* MREs for *Pax6* 3'UTR-interacting miRNAs were found to cluster into three regions, i-iii (orange boxes), located approximately at nucleotide positions 250-350, 420-550 and 600-810. Conservation of the *Pax6* 3'UTR sequence between orthologous placental mammal sequences is shown for the 876-nucleotide length. Secondary, poorly conserved MREs for interacting miRNAs having well conserved predicted MREs are not shown. (C) miRNA interaction with the *Pax6* 3'UTR is not directly related to miRNA abundance in αTC1-6 cells. miTRAP ratio and relative miRNA level for αTC1-6-interacting miRNAs is shown. miTRAP ratio was calculated by dividing the relative abundance of each miRNA with the *Pax6* 3'UTR by the relative abundance in αTC1-6 cells. Larger miTRAP values indicated greater enrichment of the miRNA with the *Pax6* 3'UTR relative to cellular abundance.

Individual miRNAs tend to have a modest impact on the protein output from targeted transcripts [50] and knock-out of individual miRNA genes or clusters are often not associated with overt phenotypes [108, 109, 111, 116, 120, 124, 125, 378, 379]. One explanation for these findings is that regulation of target transcripts by miRNAs may involve the collective action of multiple miRNAs. In support for this, multiple MREs can have a cooperative impact on target transcript destabilization [35] and genetic reporter protein synthesis [138] when MREs are approximately 8-50 nucleotides apart. We therefore probed the MRE positional code for the miRNAs identified by miTRAP to address the potential for cooperative miRNA regulation of *Pax6*. MREs for *Pax6* 3'UTR-

interacting miRNAs were found to cluster into three regions located approximately at nucleotide positions 250-350, 420-550 and 600-810 (Figure 24B) (note that some of these miRNAs share common or overlapping MREs, and would thus be incapable of acting cooperatively). Interestingly, miRNAs with the highest NRQ values clustered in the 3'-half of the 3'UTR with regions ii and iii. Additionally, the *Pax6* 3'UTR shows poor conservation of the 5' end between orthologous placental mammal *Pax6* 3'UTR sequences. Previous work has shown that miRNA repression is preferential at the ends of the 3'UTR and poor in the middle [35]. This study, however, was conducted using long (>1300 nt) 3'UTRs and assayed for repression of the target mRNA, so it is difficult to make any direct comparisons to our study.

Association of miRNAs to the *Pax6* 3'UTR does not provide direct evidence that multiple MREs are contributing to cooperative binding and regulation, or that bound miRNAs are acting to regulate transcript stability or protein synthesis. It is not clear from our miTRAP experiments whether multiple different miRNAs are binding a single transcript simultaneously and to what extent multiple MREs for individual miRNAs are contributing to binding. A sequential pulldown approach using biotinylated miRNA could be used to address simultaneous binding of multiple different miRNAs [152]. Additionally, performing miTRAP using the *Pax6* 3'UTR with mutations introduced into individual MREs, similar to the approach used in Figure 4E, could be used to determine whether binding of miRNAs to un-mutated closely spaced MREs are also impacted, and to address the relative contribution of multiple MREs for the same miRNA. Genetic reporter assays could be used to address the functionality of MREs that participate in

binding α cell-expressed miRNAs, and could also be used to address cooperative regulation through closely spaced MREs found to bind miRNAs. Ultimately, generation of mice harbouring mutations in multiple MREs found to bind α cell-expressed miRNAs could be used to address whether regulation of *Pax6* by these miRNAs plays an important role and α cell, and more generally endocrine pancreas, development and function.

Several of the miRNAs identified in our study by miTRAP have been implicated in processes relevant to Pax6 function in endocrine pancreas development. Two miRNAs that have previously been implicated in the regulation of *Pax6*, miR-7 and 375, may be critical for proper development of hormone-producing cells. Knockdown of miR-7 in embryonic mouse pancreas explants increased the proportion of insulin and glucagon-positive cells, and *in vivo* overexpression had an opposing effect, presumably through regulation of *Pax6* [217]. Knockdown of miR-375 during zebrafish embryonic development resulted in disorganized development of the pancreatic islet [380], and miR-375-null mice have reduced insulin positive cells per pancreas with a corresponding increase in the number of glucagon positive cells [124]. Three miRNAs identified in our miTRAP experiments that have not previously been implicated as *Pax6* regulators, miR-15, 16 and 195, result in reduced insulin and glucagon positive cells when overexpressed in developing pancreatic buds [381].

Several miRNAs identified using miTRAP are also associated with processes relevant to Pax6 function in glucose metabolism and endocrine pancreas gene transcription raising the possibility that they function, in part, through regulation of *Pax6*. For

instance, several *Pax6* 3'UTR interacting miRNAs (miR-124a, 132, 212, 494, and 9) are upregulated and miR-375 is downregulated in mouse β cells and pancreatic islets under hyperglycemic conditions [382–384]. Overexpression of *Pax6* 3'UTR interacting miRNAs miR-124a, 375 and 9 decreases the secretory capacity of β cell lines [368, 384, 385] and glucose tolerance improves in mice harboring genetic deletion of the miR-7a-2 gene [125]. Knockdown of miR-7 also increases insulin promoter activity in cultured β cell lines [72] and insulin mRNA from pancreas explants [217]. Together these findings support the idea that *Pax6* participates in a genetic network with the identified targets of these miRNAs to impact glucose-stimulated insulin secretion.

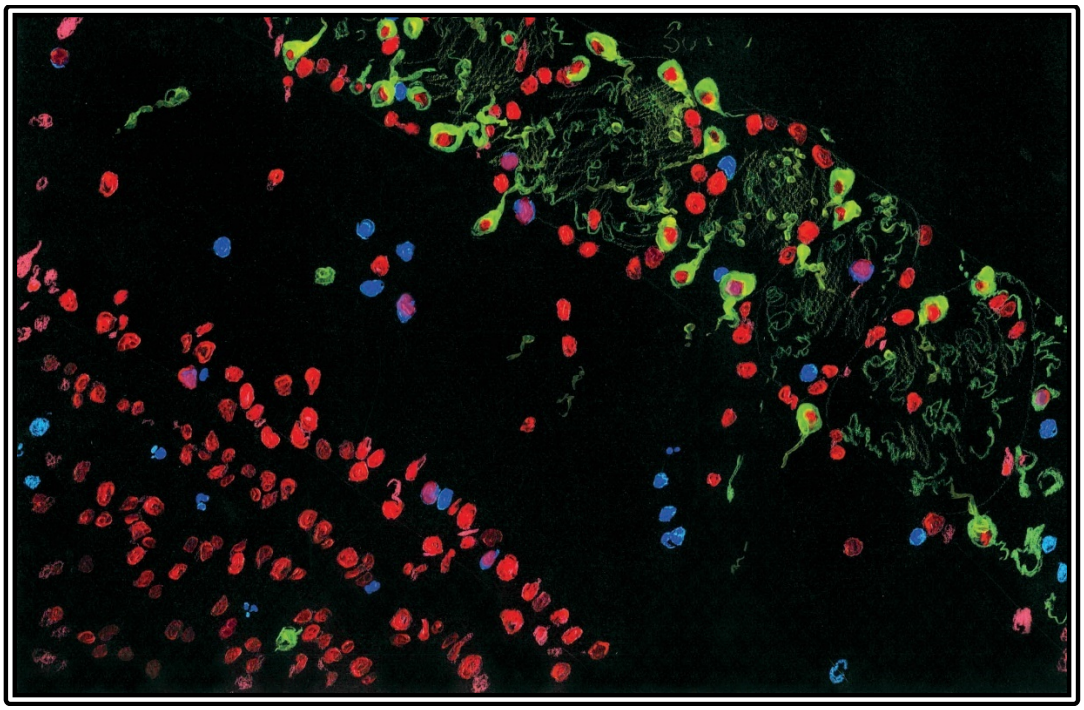
To date, only miR-375 and miR-7 have been investigated in the context of α cell hormone production and secretory function. *miR-375*-null mice had elevated plasma glucagon levels and glucagon secretion from isolated islets relative to controls [124] and miR-7 knockdown and overexpression during mouse pancreatic development elevates and represses glucagon mRNA expression, respectively [217]. How and whether the other miRNAs identified by miTRAP in α-TC1-6 cells contribute to α cell function remains to be determined.

3.5 Conclusions

Our work provides a comprehensive analysis of the *Pax6* 3'UTR and description of potential miRNA recognition elements. We have identified cell type-specific miRNA cohorts or "miR-codes" that form 3'UTR interaction clusters where cooperative regulation of *Pax6* may occur. As proper PAX6 function is highly sensitive to its dosage,

miRNA post-transcriptional regulation is predicted to play an important role in maintaining correct PAX6 levels. Our work reveals a regulatory landscape upon which the role of miRNAs on the developmental control, fine-tuning and/or homeostasis of PAX6 can be further investigated.

Chapter 4: Cooperative and direct regulation of *Pax6* by microRNA-7 through multiple recognition elements



"A long way from home" Acrylic on paper (2017)

This chapter is adapted from a manuscript in preparation for submission.

Ryan BC, Enns E, Story S, Wiebe M, Alford S, Hylkema A, Hamilton K, Lang RA, Howard PL, Chow RL.

Contributions: **BCR** and RLC designed the project. **BCR** performed luciferase assays, miR-CATCH pulldowns and RT-qPCR experiments, and all experiments. EE and WM contributed to staining, imaging and data analysis for PAX6 IF. SA cloned the pMIR-LUC construct and AH performed preliminary luciferase experiments. RL assisted with the miR-7 MRE mutant mouse design and generation. **BCR** performed EdU/BrdU injections and olfactory bulb dissections, and EE and **BCR** performed olfactory bulb sectioning and developed the olfactory bulb cell counting protocol. EE and SS performed olfactory bulb IF, imaging and cell counting. KH measured olfactory bulb glomerular layer areas. TW wrote the Fiji subdivision macro. **BCR** and RLC wrote the manuscript.

4.1 Abstract

PAX6 is a highly dosage-sensitive homeodomain transcription factor essential for development of the eye, nose, central nervous system, gut and endocrine pancreas, and is mutated in the haploinsufficiency disease aniridia. Though the cellular mechanisms regulating precise PAX6 expression levels have not been elucidated, post-transcriptional regulation of *Pax6* by miRNAs may represent an important mechanism for maintaining the correct dosage of PAX6. Several microRNAs (miRNA) have been implicated in

regulating PAX6 in different cellular contexts. Notably, miRNA-7 has been implicated in the formation of a dorsal-high, ventral-low gradient of PAX6 expression in the ventricular zone of the developing forebrain. The graded expression of PAX6, in part, underlies subtype diversity in neuroblasts that migrate to the olfactory bulb and become different types of interneurons. We show here that that two previously characterized miR-7 recognition elements located within the *Pax6* 3' untranslated region function in an additive manner to regulate PAX6 levels *in vitro*. However, the direct biological relevance of miRNAs in maintaining the correct dosage of PAX6 *in vivo* remains poorly understood. To address this, we generated mice harboring mutations that disrupt one or both miR-7 recognition elements within the *Pax6* 3'UTR. Deletion of either miR-7 site alone did not alter PAX6 expression in the postnatal ventricular zone. In contrast, and similar to our *in vitro* data, mutation of both sites resulted in the deregulation of PAX6 protein but not mRNA transcript. These findings provide direct evidence supporting a role for miR-7 in regulating *Pax6* during development. In addition, they indicate that *in vivo*, microRNA regulation can be mediated through redundant mechanisms.

4.2 Introduction

4.2.1 Important questions regarding miRNA regulation

Cells must dynamically regulate the quantity and functionality of many different proteins during development and adult homeostasis. Many mechanisms are available to

achieve this feat (Figure 1), one of which is regulation by microRNAs (miRNA). As mentioned in Chapter 1, microRNAs are a large class of small non-coding regulatory RNAs that target mRNA transcripts by complementary base pairing [13]. MicroRNAs function to interfere with protein synthesis and destabilize targeted mRNAs through recruitment of protein silencing complexes to microRNA recognition elements (MREs) located in mRNA 3' untranslated regions (3'UTRs)[34, 51, 63, 65].

The literature abounds with studies investigating the physiological relevance of *individual* miRNAs regulating *individual* target gene transcripts through *single* MREs. While a few studies have performed direct mutation of predicted MREs at the genomic level [158, 159] the majority do not fully demonstrate direct and specific regulation by miRNA(s) of interest *in vivo*. To the best of our knowledge, only one published study using direct mutation of a predicted MRE has been examined in mouse [158]. This is noteworthy as miRNAs typically exert a modest effect on the levels of targeted mRNAs and protein products [50, 62]. Additionally, multiple different miRNAs can bind a single mRNA [152] and regulation through multiple MREs in a single 3'UTR for the same or different miRNAs can enhance miRNA mediated repression [47, 48, 59, 135–138] .

Many studies employ miRNA overexpression or loss of function (LOF) to determine the impact of altered miRNA regulation on a target of interest and to determine the resulting phenotypic effects. However, single miRNAs can target many different gene transcripts [40, 62, 131, 134, 135] and produce complex regulatory effects as part of gene regulatory networks (GRNs)[90, 139–141]. Consequently, miRNA overexpression and LOF are insufficient to demonstrate that regulation of a single target

gene transcript is direct. Additionally, the phenotypes that result from these manipulations in miRNA quantity may be the consequence of miRNA regulation through many different gene transcripts. To complicate matters, overexpressed miRNAs may regulate mRNAs that are not normally targeted by the miRNA of interest under normal physiological conditions [156]. Given this, we wanted to address the impact and biological role of miRNA regulation directly, through specific genomic mutation of predicted MREs.

4.2.2 Pax6 as a model gene for studying miRNA regulation

We chose *Pax6* as a model gene for studying miRNA regulation directly for several reasons. First, *Pax6* encodes a transcription factor [177] and is critical for normal development of many tissues. Notably, PAX6 is essential for development of most ocular structures and many regions of the central nervous system, such as the olfactory system, forebrain and spinal cord [177]. Additionally, PAX6 is important for normal development of the endocrine pancreas and secretory cells of the gut (Figure 7)[191, 192, 195].

In addition to being dynamically regulated in space and time, PAX6 dosage must also be tightly regulated. This dosage requirement is exemplified by eye development (Figure 9). Mutations in both copies of the *Pax6* gene that produce non-functional PAX6 protein result in a failure for eyes to develop in rodents [230, 231]. Mutations in one copy of the *Pax6* gene results in haploinsufficiency. This produces a small eye phenotype in rodents [229, 230] and is the cause of the human ocular disease aniridia [233]. Overexpression of PAX6 can also be deleterious. Mice carrying multiple copies of the

human *PAX6* gene also develop small eyes [232] and, though rare in humans, chromosomal duplications that include the *PAX6* gene are associated with mild ocular defects [235]. Generally, eye development requires proper regulation of *Pax6* expression and gene dosage.

4.2.3 PAX6 in neural progenitors of the ventricular-subventricular zone

Other developing tissues also require correct PAX6 dosage and represent excellent developmental systems for studying miRNA regulation of *Pax6*. One such tissue is the early postnatal ventricular and subventricular zone (V-SVZ) lining the lateral ventricles (LVs)[206]. In this study, we chose to focus on the early postnatal V-SVZ for several reasons: 1.) Both Pax6 mRNA and protein are expressed in a gradient in the V-SVZ [386]. 2.) Neural progenitors lining the LVs continue to generate new neurons during the early postnatal period and into adulthood in rodents, providing a model system for studying neuronal subtype specification [387–389]. 3.) PAX6 influences the fate of neural progenitors from this region and changes in PAX6 level can alter cell fate choices[245, 324, 386]. 4.) miR-7 is implicated as a regulator of *Pax6* in the V-SVZ where it is expressed in an opposing gradient to PAX6 [206].

1.) PAX6 is expressed in a gradient in the early postnatal V-SVZ but not adult SEZ

In the first days after birth in mice, the VZ and SVZ are visible as distinct layers [388], both of which express PAX6 (Figure 25A)[206]. Like other regions of the developing CNS, such as the cortex and neural tube, PAX6 is present in a dorsal-high to ventral-low gradient in the V-SVZ. As development progresses, the VZ and SVZ become

much thinner and are no longer visible as distinct layers. In adulthood, the V-SVZ is called the subependymal zone (SEZ). Though PAX6 protein is detectable in the SEZ by IF, a clear dorsal-ventral gradient is no longer visible (Figure 25B-D), suggesting that PAX6 may be differentially regulated in the early postnatal versus adult V-SVZ. Given the absence of a clear PAX6 gradient in the adult SEZ, we decided to focus on PAX6 in the early postnatal V-SVZ.

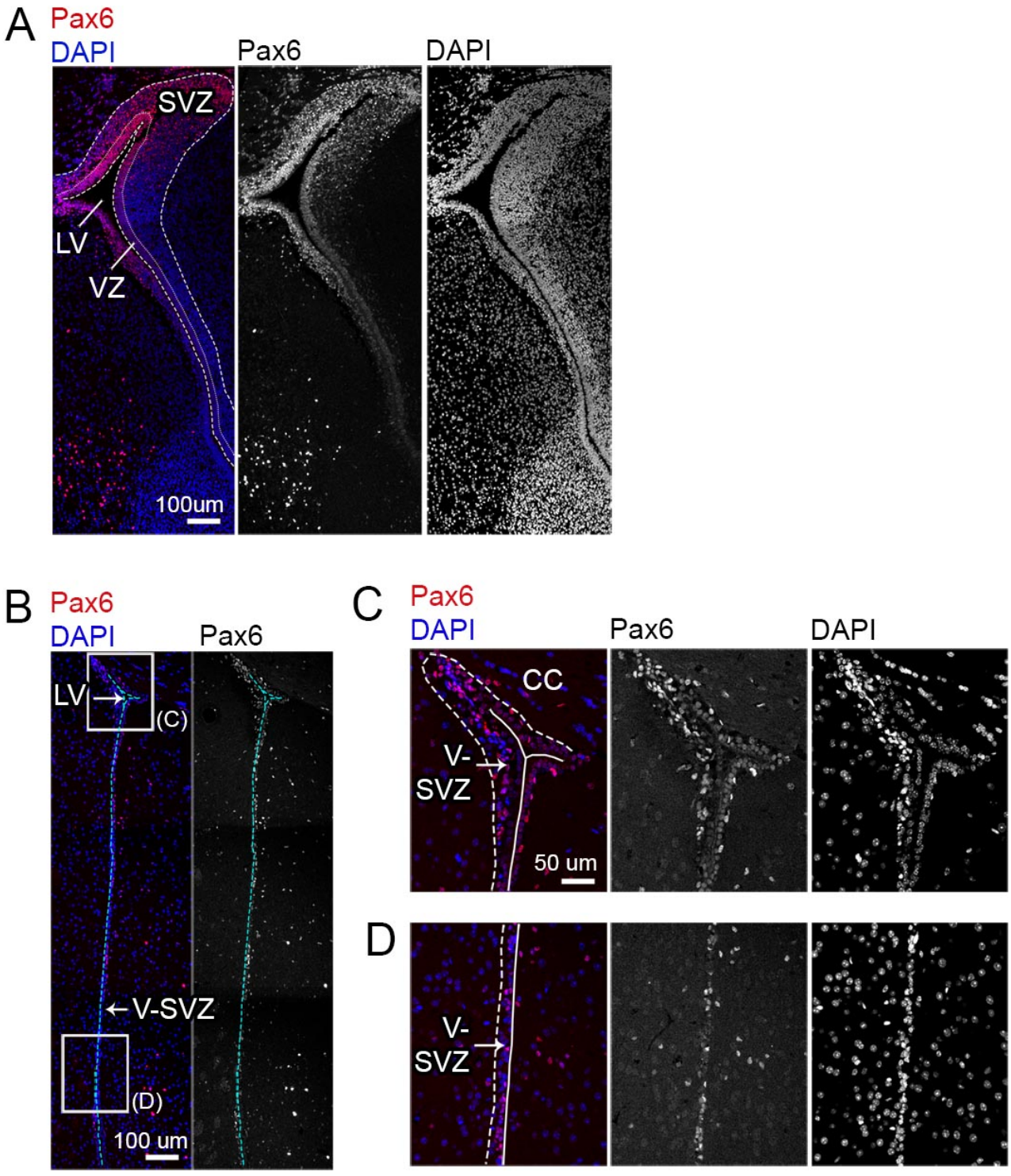


Figure 25. PAX6 immunofluorescence in the P1 and adult mouse V-SVZ

Immunofluorescence (IF) for PAX6 was performed using P1 (A) and P30 (B) paraffin sections, and confocal microscopy was used to collect images. DAPI labels all nuclei. (C) and (D) show high magnification of D and VL V-SVZ from adult. LV, lateral ventricle; VZ, ventricular zone; SVZ, subventricular zone; CC, corpus callosum.

2.) The early postnatal V-SVZ is composed of neural progenitors

The V-SVZ is composed of a mitotic population of progenitor cells that give rise to neurons and glia, and correct level of PAX6 protein in this region is important for regulating specification of some of these cell types. Thymidine analogues, such as 5-ethynyl-2'-deoxyuridine (EdU), can be used to reveal the proliferative capacity of this brain region and track the fate of differentiating cells. Dividing cells are located in both the VZ and SVZ (Figure 26A). Radial glia cells, the neural progenitors, make up the VZ and are organized into a pseudostratified epithelium [387, 388]. The radial glia give rise to transit amplifying cells and migrating neuroblasts [389] which comprise the SVZ [387, 388]. In adult mice, the transit amplifying cells divide approximately three times before giving rise to migrating neuroblasts, which divide once or twice [390]. The neuroblasts migrate along the rostral migratory stream (RMS) to the main olfactory bulb (OB), where they differentiate into interneurons of the granule and periglomerular layers (Figure 27B-C)[389]. Migration to the olfactory bulb and expression of interneuron subtype-specific markers takes approximately 15 days to 3 months and varies between different cell types [391, 392].

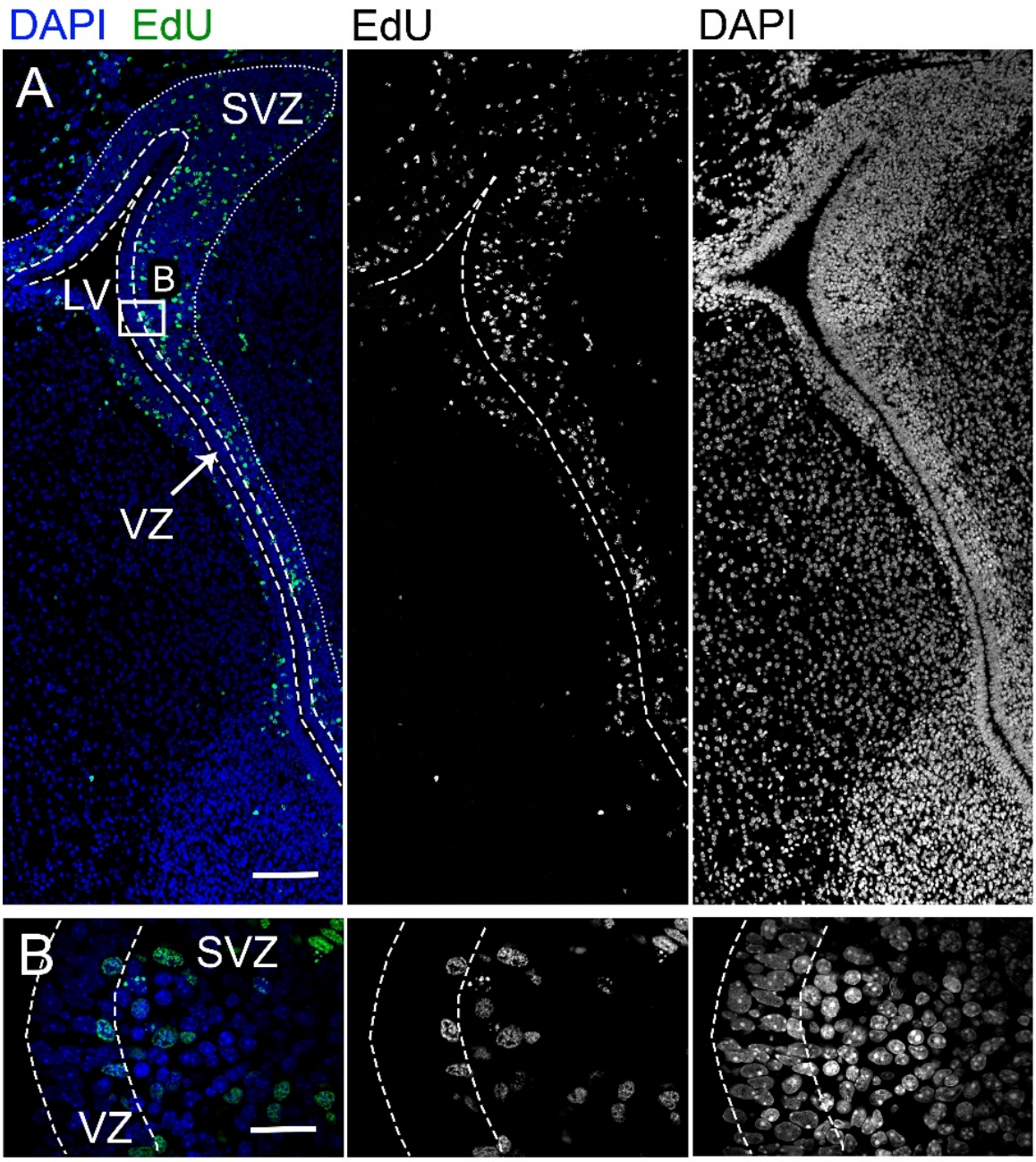


Figure 26. Region of proliferating cells in the P1 V-SVZ

(A-B) Fluorescent labeling for the thymidine analogue EdU (green) reveals proliferating cells in the ventricular-subventricular zone (V-SVZ). V-SVZ neural stem cells (NSCs) generate neuroblasts that migrate via the rostral migratory stream (RMS) to the main olfactory bulbs (OB) where they differentiate into OB interneurons. Tissue was harvested 1 hour post-EdU injection. DAPI labels all nuclei. Scale bar - 100 µm for A and 20 µm for B.

3.) PAX6 is important for specifying specific populations of OB interneurons

Different interneuron subtypes in the main rodent OB are specified by V-SVZ progenitors in spatially distinct subregions (Figure 27A)[388, 393, 394] and PAX6 plays an important role in this process. The OB interneurons can be divided into two general classes based on their locations within the OB: the periglomerular neurons (PGNs) and granule cells (GCs). The PGNs are interneurons that reside within the glomerular layer of the OB (Figure 27B) and generally, PAX6 is important for specification of PGNs [324]. We focused our analysis on three subtypes of PGN defined by the expression of non-overlapping markers: tyrosine-hydroxylase (TH), calretinin (CalR) and calbindin (CalB) [392, 395].

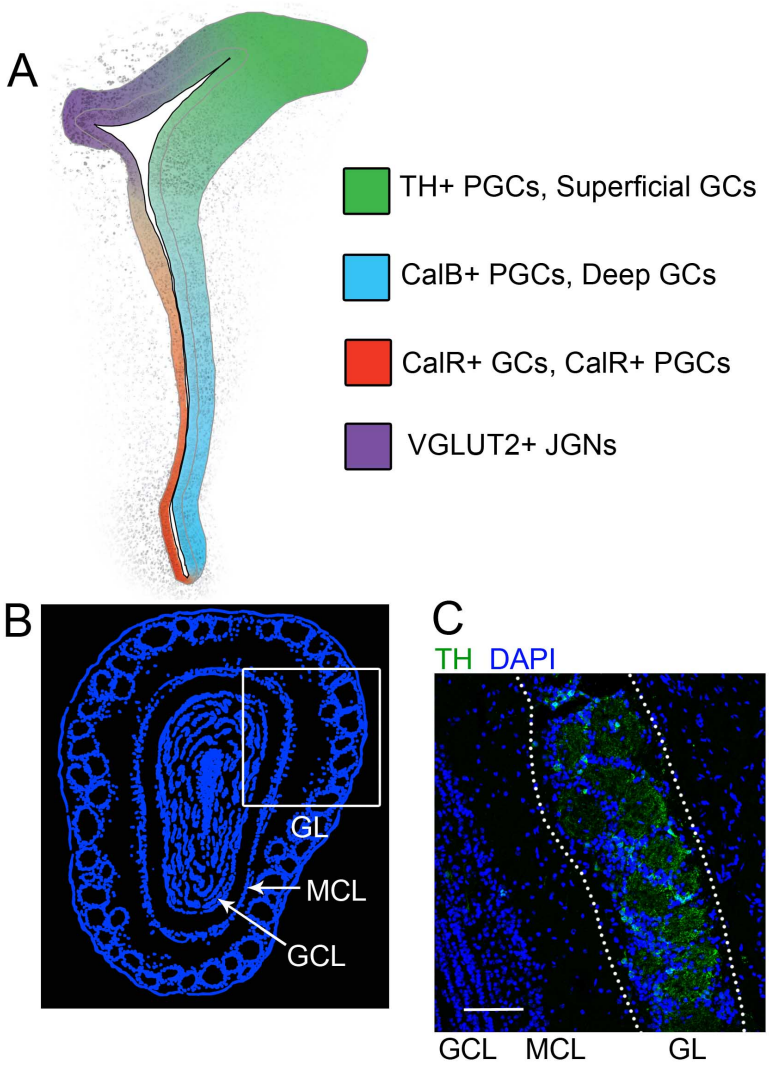


Figure 27. Spatial heterogeneity of V-SVZ progenitors and main olfactory bulb organization

(A) V-SVZ neural stem cells (NSCs) are spatially organized giving rise to distinct olfactory bulb (OB) interneuron subtypes [393]. Specifically, PAX6 is important for the specification of dopamine-producing tyrosine hydroxylase-positive (TH) OB interneurons [214]. (B) Schematic representation of a horizontal section through the mouse main OB. Cell nuclei are shown in blue. Olfactory sensory neurons (OSNs) project to the glomerular layer (GL) where they synapse onto mitral cells. The mitral cell bodies are in the mitral cell layer (MCL). The periglomerular interneurons (PGNs) make up the GL. An additional population of OB interneurons, the granule cells, comprise the granule cell layer (GCL). (C) Cell counting was restricted to the PGNs in the GL (white dotted line). The GL was defined as a region of interest (ROI) based on nuclear morphology and labeling for interneuron-specific markers. This image shows labeled dopaminergic PGNs (green), which express the enzyme tyrosine hydroxylase (TH). DAPI (blue) labels all nuclei. Scale bar represents 100 µm.

Specifically, PAX6 plays an important role in the specification and survival of TH-expressing cells in the olfactory bulb, the dopaminergic PGNs and the superficial granule cells [214, 215, 245], both of which are derived from DL V-SVZ progenitors (Figure 27A,C). Interestingly, though superficial granule cells produce *Th* mRNA, they do not appear to be immunoreactive for TH protein [214, 245, 396]. *Pax6*$^{Sey/+}$ mice have reduced TH positive PGNs relative to WT controls [245], and transplanted *Pax6*$^{Sey/Sey}$ progenitors fail to generate dopaminergic PGNs or superficial granule cells [214]. Additionally, electroporation of a *Pax6* overexpression construct into the lateral V-SVZ increases the proportion of TH-positive PGNs produced from electroporated cells [206] and *Pax6* overexpression in the SEZ and RMS of adult mice also increases the proportion of TH-positive PGNs formed [324]. PAX6 continues to be expressed in differentiated dopaminergic PGNs, where it is important for their survival [215]. Taken together, these results suggest that proper regulation of PAX6 dosage in neural progenitors of the V-SVZ is important for specifying the correct proportions of PGN subtypes.

4.) MicroRNA regulation of Pax6 in the V-SVZ

Pax6 is extensively regulated at the transcriptional level and post translationally [188]. For example, transcriptional regulation of *Pax6* is involved in establishing the V-SVZ gradient of PAX6 [206]. However, given the requirement for precise regulation of PAX6 during development, regulation by miRNAs may provide an important mechanism to maintain correct expression domains and dosage of PAX6. A previous study found that PAX6 protein is expressed in steeper dorsal-ventral gradient than *Pax6* mRNA or a

reporter under the control of *Pax6* regulatory elements, suggesting that miRNAs may be regulating *Pax6* post-transcriptionally (Figure 28A-B)[206]. Additionally, miR-7a is expressed in an opposing gradient to that of PAX6. Negative regulation of *Pax6* by miR-7 in the ventral V-SVZ may represent an important mechanism by which the steep dorsal-ventral gradient of PAX6 is maintained (Figure 28C)[206].

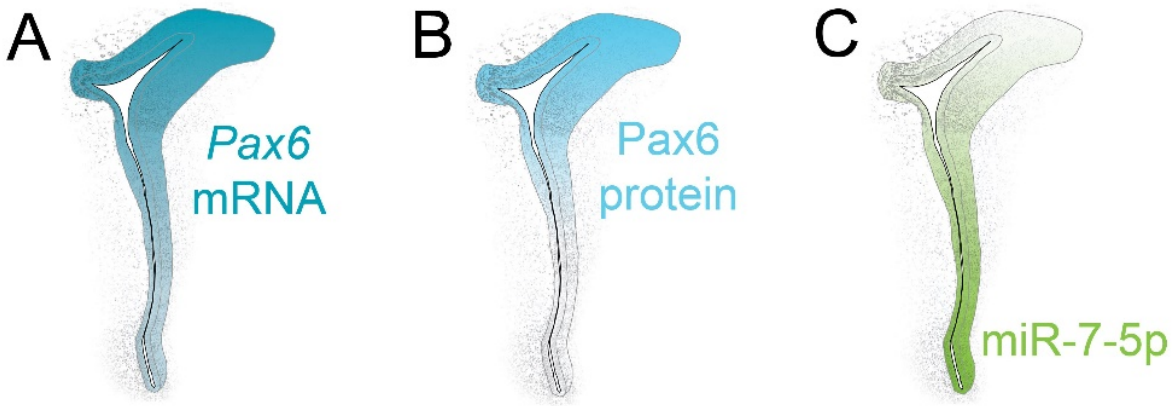


Figure 28. PAX6 and miR-7 expression in the P1 V-SVZ

Schematic representation of *Pax6* mRNA (A), PAX6 protein (B), and miR-7-5p RNA (C) expression domains in the V-SVZ. Note that PAX6 protein is more dorsally restricted than *Pax6* mRNA. This steep D-V gradient of PAX6 protein is predicted to be the consequence of negative regulation by miR-7-5p, which is expressed in an opposing gradient [206].

4.2.4 Choice to investigate miR-7

Though several miRNAs were previously identified as candidates to regulate *Pax6* endogenously [206, 217, 273–279, 326, 327, 329, 331–333], we chose to focus on miR-7. MicroRNA-7 represents a strong candidate to regulate *Pax6* for several reasons. In addition to its proposed role as a regulator of PGN cell type specification in the V-SVZ by targeting *Pax6* [206], miR-7 has been proposed to target *Pax6* in the developing

endocrine pancreas [217] and cortex [274]. Genes having multiple predicted MREs in their 3'UTR for a particular miRNA are more likely to be regulated by that miRNA [47, 48, 59]. Multiple predicted miR-7 MREs are present in the mouse [274, 278] and human [397] *Pax6* 3'UTRs, providing additional evidence that *Pax6* is a genuine target of miR-7. Additionally, MREs for broadly conserved miRNAs are more likely to be selectively maintained, suggesting that they are also more likely to be functional [20, 43]. The mature miR-7 sequence [18, 398] and general pattern of expression [69, 88, 370, 399] is highly conserved across animals. Specifically, the mature miR-7a-5p sequence is perfectly conserved from humans to flies [18]. Given this, we chose to focus on miR-7 as a regulator of *Pax6* in the early postnatal V-SVZ.

4.2.5 Aims, Predictions and Outcomes

Here, we performed a bioinformatics screen for predicted miR-7-5p and -3p MREs in the 3'UTR of *Pax6* and used *Pax6* 3'UTR genetic reporters to confirm that previously identified miR-7-5p MREs located at 3'UTR positions 517 and 655 are functional *in vitro.* Next, we generated mice harboring mutations in either the *Pax6* 3'UTR positions 517 or 655 miR-7-5p MREs individually or in combination using CRISPR-Cas9 genome editing. We predicted that disruption of the miR-7-5p MREs in the 3'UTR of *Pax6* would alter the gradient of PAX6 in the early postnatal V-SVZ. Specifically we predicted that *Pax6* mRNA and PAX6 protein would be elevated in the ventral V-SVZ in miR-7-5p mutants, with double mutants displaying the largest effect. To address this, we used reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) to assay for *Pax6* mRNA and quantitative immunofluorescence (IF) to assay for PAX6 protein. Using these

approaches, we observed little change in *Pax6* mRNA across the dorsal-ventral V-SVZ in mutants relative to wild type (WT) but observed an approximately two-fold increase in the intensity of PAX6 IF in double mutants relative to WT in the ventral V-SVZ.

Next, we wanted to address whether the observed increase in PAX6 is associated with altered specification of PGN subtypes, and we predicted that the proportion of TH positive dopaminergic PGNs would be elevated in miR-7-5p MRE mutants relative to WT. We did not observe any large changes in dopaminergic specification in *Pax6* 3'UTR miR-7-5p MRE mutants. Taken together, these results suggest that miR-7-5p is involved in refining the gradient of PAX6 in the early postnatal V-SVZ, though other regulatory mechanisms are likely also required for establishing the PAX6 gradient. Also, we provide evidence that miRNAs can function cooperatively though multiple MREs in an endogenous gene transcript *in vivo*.

4.3 Materials and methods

4.3.1 Cell culture

Mouse transgenic pancreatic alpha cells (αTC1-6, ATCC CRL-2934) were cultured in DMEM (low glucose, pyruvate, ThermoFisher, 31600-034) supplemented with 10% FBS (Life Technologies, 16000-044), 15 mM HEPES, 0.1 mM non-essential amino acids (ThermoFisher, 11140-050), 0.02% BSA (Sigma-Aldrich, A7906-50G), 1.5 g/L sodium bicarbonate, and 2.0 g/L glucose. Cells were cultured at 37°C with 5% carbon dioxide and were passaged once every 4 days when approximately 80% confluent. Cells were

dissociated using a 1:5 dilution enzyme-free cell dissociation buffer (ThermoFisher, 13151014) in HBSS for 15 min at 37*C and 2 million cells were plated per 100 mm culture dish. Culture media was refreshed every 2 days between passages.

4.3.2 Identification of miR-7 MREs in the *Pax6* 3'UTR

Previously, we identified a mouse *Pax6* 3'UTR length of 876 bp [400]. TargetScan [49], PITA [58] and ImiRP [353] were used to scan for predicted miR-7 MREs in the mouse *Pax6* 3'UTR. The UCSC Genome browser [280] multiz alignments were used to determine the conservation of predicted MREs.

4.3.3 Luciferase assay

Plasmid design

The luciferase reporter plasmids, which express firefly luciferase, were generated using the pMIR-REPORT vector (Invitrogen, AM5795) as a backbone. The *Pax6* 3'UTR upstream of the SV40 poly(A) sequence flanked by SpeI and NotI restriction sites was synthesized using BioBasic gene synthesis service (ON Canada), and the sequence was supplied in the pUC57 vector. SpeI and NotI restriction sites were used to clone 1009 bp of the *Pax6* 3'UTR downstream of the luciferase gene by digesting both pUC57-Pax6-3UTR-SV40poly(A) and pMIR-REPORT with SpeI (NEB, R3133S) and NotI (NEB, R3189S) for 4 hours at 37°C. Digest products were run on a 1.1% agarose gel for 45 min at 110 V, and desired bands were extracted from the gel using QIAquick gel extraction kit (Qiagen, 28704) following the manufacturer's protocol. Equimolar quantities of pMIR-REPORT backbone and Pax6-3'UTR-SV40poly(A) were ligated using T4 DNA ligase

(NEB, M0202) for 10 min at room temperature following the manufacturer's protocol, and ligated plasmid was used to transform subcloning efficiency DH5α competent *E. coli* (Invitrogen, 18265-017). The completed plasmid was named pMIR-LUC-WT3'UTR and was sequenced using Eurofins Genomics SimpleSeq DNA sequencing service with the primers pMIR-LUC1: AAGCATGAATTCAAGGTACC, pMIR-LUC2: TTTGGCACCAAAATCAACGG, pMIR-LUC3: AACGACATTTATAATGAACG, pMIR-LUC4: GATATGTGGATTTCGAGTCG, pMIR-LUC5: AATCCGGAAGCGACCAACGC, pMIR-LUC6: AAGGAGAGAGCATGTGATCG, pMIR-LUC7: TTGCATATAATTGAACCTGG.

Mutations were introduced into the pMIR-LUC-WT3'UTR to disrupt miR-7 MREs in the *Pax6* 3'UTR sequence. miR-7-5p positions 517 and 655, and miR-7-3p positions 626 and 758 MRE mutations were designed using ImiRP [353] to ensure that new predicted MREs were not created upon disruption of the sites of interest. The mutations were as follows: 517MUT GTTTTCCA to GTTTTggA, 655MUT GTCTTCC to GTCcgaa, 626MUT ATTTGTT to ATTTtgc, and 758MUT ATTTGTT to ATTgtTT, where lower case letters represent mutated nucleotides (Figure 29). Mutations were introduced into the pMIR-LUC-WT3'UTR vector using the Q5 site-directed mutagenesis kit (NEB, E0554S) following the manufacturer's protocols. The following primers were used for site-directed mutagenesis: 517MUT forward primer TGTTGGTTTTggAAAGGTTGTTAACAG, reverse primer GATAGTTTTTACATATATCTAGTGTG; 655MUT forward primer AGTATTTGTCcgaaCTAGAAATCCTCAGAATGATTTC, reverse primer TACATTTTGACATAAAACAAATTGTATTATATC; 626MUT forward primer AATACAATTTtgcTTATGTCAAAATGTAAGTATTTG, reverse primer

ATATCGAAGACACACTCTAC; 758MUT forward primer

TGCAATCATTgtTTTCTTTCTTGGCCAG, reverse primer

TGAAAATGTGTATAAAACATCTATAG. Mutants were sequenced using Eurofins Genomics SimpleSeq DNA sequencing service with the primers pMIR-LUC6 (see above),

F_Pax6Insert AAATCTTCATTTTGGTATCC, and R_Pax6Insert ATCTTGTCGATCATGGTTTCC.

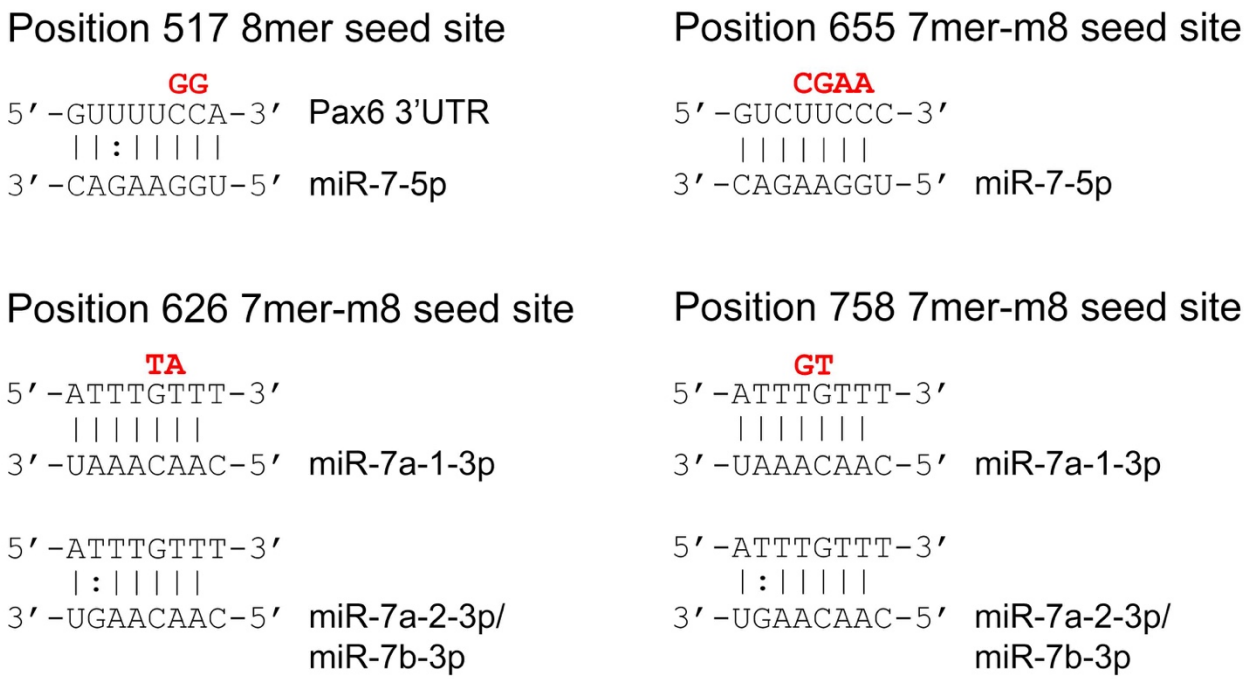


Figure 29. Mutagenesis strategy for miR-7 MREs in the *Pax6* 3'UTR

Top sequences show 8 nucleotides of the *Pax6* 3'UTR, representing the respective MREs. The bottom sequences show the 8 nucleotides at the mature miRNA 5' end, the miRNA seed. Dashes represent Watson-Crick base pairs and colons represent G-U pairs. *Pax6* 3'UTR miR-7 MRE mutations are displayed in red. Mutant sequences were generated using ImiRP [353].

Transfection and data analysis

αTC1-6 cells were transfected with the pMIR-LUC plasmids by first seeding 25,000 cells per well of a 96-well plate. 24 hours post-seeding, cells were transfected with 100 ng of plasmid DNA per well using JetPrime reagent (Polyplus Transfection, 114-07) following the manufacturer's protocol. The 100 ng of DNA was composed of 50 ng pCAG-GFP reporter [401], 10 ng pGL4.70 renilla luciferase internal control plasmid (Promega, E688A), and 40 ng pMIR-LUC firefly luciferase plasmid. Each 3'UTR condition was transfected as 6 technical replicate wells per plate. Every plate contained cells transfected with pMIR-REPORT without the *Pax6* 3'UTR for inter-plate normalization, and cells transfected with pMIR-LUC-WT3'UTR for comparing to mutant *Pax6* 3'UTR vectors.

24 hours post-transfection, firefly and renilla luciferase levels were measured using the dual-glo luciferase assay system (Promega, E2920) following the manufacturer's protocol. Luminescence was measured using the VICTOR multilabel plate reader (PerkinElmer, Massachusetts USA) and PerkinElmer2030 software. Luciferase luminescence (Fluc) was normalized to renilla luminescence (Rluc) for each well by Fluc/Rluc. Grubb's test was used to identify significant Fluc/Rluc outliers between the six technical replicate wells for each condition, and outlier wells were removed from analysis. Average normalized luciferase luminescence was calculated for each 3'UTR condition within a single plate. Since results can vary between different plate runs, inter-plate normalization was performed by dividing average normalized luciferase luminescence for each *Pax6* 3'UTR condition by the plate standard pMIR-

REPORT. Final results were scaled relative to the pMIR-LUC-WT3'UTR, such that all mutant *Pax6* 3'UTR conditions were expressed as a percentage of the WT *Pax6* 3'UTR condition. One-way ANOVA with Tukey's multiple comparisons test was used to compare normalized luciferase luminescence between different *Pax6* 3'UTR conditions.

4.3.4 Animals

All research on mice was performed with approval of the University of Victoria Animal Care Committee in compliance with the Canadian Council on Animal Care (CCAC) guidelines for the ethical treatment of research animals.

All experiments in this study were performed using mice on the 129S1 background (strain #002448, The Jackson Laboratory, Bar Harbor, ME) and both sexes were used. Animals were housed in a pathogen-free facility at the University of Victoria, Canada, in individually vented caging systems. A maximum of five adult mice were housed per cage, separated by sex. Cages were supplied with rodent bedding (Bed-O-Cobs) and rodent enrichment in the form of cotton nestlets (Ancare), paper fibers (Shepherd, Enviro-dri) and cardboard huts (WF Fisher, Bio-huts). Clean drinking water was available to animals at all times and made available through a reverse osmosis watering station. Animals were provided ad libitum access to dry animal diet (Lab-Diet) that has been CCAC approved.

Pax6 MRE mutant founder mice were generated at the Cincinnati Children's Hospital Medical Center and were backcrossed into 129S1 (Jackson Laboratory, 002448) for a minimum of 3 generations prior to being used for experimentation. For PAX6 protein quantification by immunofluorescence (IF) and olfactory bulb PGN fate tracking

experiments, heterozygous breeders were paired, and WT and homozygous mutant littermates were used for experiments. For *Pax6* mRNA quantification and mRNA affinity purification experiments, homozygous mutant breeders were paired to produce litters of exclusively homozygous pups, and WT 129S1 mice were crossed to generate WT pups.

For animal experiments, four different conditions were studied: *Pax6* wt, *517MUT*, *655MUT*, *517+655MUT*. Using the resource equation method, we aimed for a minimum sample size of 5-7 animals per genotype. Since heterozygous crosses were performed for the PAX6 IF and OB cell counting experiments, WT animals were collected from each of the three lines to ensure no differences were observed. Consequently, additional WT samples were harvested for these experiments.

4.3.5 Generation of miR-7-5p MRE mutant mice

Mutation design

miR-7-5p MRE mutants at *Pax6* 3'UTR positions 517 and 655 were designed using ImiRP [353]. Nucleotide substitutions were performed using any of the 4 nucleobases, and any mutants that generated new predicted 8mer +/- G:U, 7mer-m8 +/- G:U, 7mer-A1 +/- G:U or 6mer MREs were excluded. Since the position 655 MRE overlapped a PAM sequence for CRISPR/Cas9 editing, we selected a mutation that also disrupted the PAM.

In addition to using ImiRP, several other bioinformatics tools were used to design miR-7-5p MRE mutations. Since ImiRP relies of the miRbase high confidence miRNA dataset [18] to perform predictions, and some miRNAs that are not considered high

confidence may still be functional, we also scanned mutant sequences using PITA to ensure new predicted MREs were not created. Many types of regulatory motifs exist in mRNA sequences that could be altered by MRE mutation. We used RegRNA 2.0 [402], RBPDB [403] and UTRScan [404] to identify other RNA motifs in WT *Pax6* 3'UTR and mutant sequences. A predicted intron splicing silencer site was disrupted by the position 655 miR-7-5p MRE mutation, however since this site completely overlaps the position 655 MRE, designing a mutation that did not disrupt it was not possible.

CRISPR/Cas9 genome editing

Pax6^{517MUT} mice were generated using CRISPR/Cas9 gene editing by injecting guide RNA (gRNA), Cas9 mRNA and donor oligo into 129S1 (Jackson Laboratory, 002448) mouse embryos. The gRNA sequence was TTGGAAAACCAACAGATAGT and the donor oligo sequence was CATATAATTGAACCTGGGACAACACACACTAGATATATGTAAAAACTATCTGTTggTTTTGGAAAGGTTGTTAACAGATGAAGTTTATGTGCAAAAAAGGGTAAGATATGAATTCAAGGAGA, where lowercase underlined letters indicate mutated bases.

Pax6^{655MUT} and *Pax6*$^{517+655MUT}$ mice were generated by crossing homozygous *517MUT* mice to WT 129S1 mice. CRISPR/Cas9 gene editing was performed by injecting gRNA, Cas9 mRNA and donor oligo into heterozygous *517MUT* mouse embryos to generate single *655MUT* and double *517+655MUT* animals. The gRNA sequence was AGAAATCCTCAGAATGATTT and the donor oligo sequence was TGTGCAAAAAAGGGTAAGATATGAATTCAAGGAGAAGTTGATAGCTAAAAGGTAGAGTGTGTCTTCGATATAATACAATTTtgcTTATGTCAAAATGTAAGTATTTGTCcgaaCTAGAAATCCTCAGAAT

GATTTCTATAATAAAGTTAATTTC, where lowercase underlined letters indicate mutated bases. The donor oligo also included a mutation for the miR-7-3p MRE at position 626. Given the large distance between the 626 and 655 mutations, recombination crossover happened in the middle allowing for the generation of position 655 single mutants and 626+655 double mutants. miR-7-3p MRE position 626 mutants were not used in this study.

4.3.6 Genotyping and mutant sequence verification

gDNA preparation

Genomic DNA (gDNA) for genotyping and sequencing mutants was harvested from small ear biopsies collected using ear punches (World Precision Instruments, 500075). gDNA was isolated from ear clips by digestion at 98°C in 75 µl 0.05 M NaOH and was neutralized using 25 µl 0.5 M Tris pH 8.0. All PCR reactions for genotyping and sequencing were performed using Phusion DNA polymerase (NEB, M0530S) following the manufacturer's protocols and performed with a T3 thermocycler (Biometra).

PCR Genotyping

Genotyping was performed using a common forward primer and reverse primers specific for the *WT* or *MUT* alleles from the 3'UTR positions 517 and 655 miR-7-5p MREs. The primers were as follows: common forward primer ACCATGATCGACAAGATTTG, WT position 517 MRE reverse primer TCTGTTAACAACCTTTgg, MUT position 517 MRE reverse primer TCTGTTAACAACCTTTcc, WT position 655 MRE reverse primer TTCTGAGGATTTCTAGggaa, MUT position 655 MRE

reverse primer TTCTGAGGATTTCTAGttcg. Thermocycling conditions were as follows: 98˚C for 2 min, cycle through 98˚C for 15 sec, 60˚C for 30 sec and 72˚C for 15 sec 27 times, final extension 72˚C for 2 min.

PCR amplification of Pax6 3'UTR

Sequencing of the *Pax6* 3'UTR was performed using ear biopsy gDNA. PCR reactions were performed to amplify two overlapping regions of 3'UTR using the following primer pairs: forward primer 1 TAAAGAGAGAAGGAGAGAGC, reverse primer 1 GCTGTATTACTGTTCACAGTCCAA, forward primer 2 ATGGACTAGAACCAAGGACC, reverse primer 2 TTTCTTTGAGGACCATCAGG (Figure 30). Thermocycling conditions were as follows: 98˚C for 2 min, cycle through 98˚C for 15 sec, 60˚C/56˚C for 30 sec and 72˚C for 30 sec 35 times, final extension 72˚C for 2 min. 60˚C annealing temperature was used for primer pair 1 and 56˚C annealing was used for pair 2.

PCR amplification of Pax6 CDS

The *Pax6* coding region was sequenced from *655MUT* and *517+655MUT* homozygotes. The mouse *Pax6* cDNA sequence containing 13 exons was obtained from Ensembl (transcript Pax6-206, ENSMUST 00000111086.10).

Sequencing of the *Pax6* coding region from mutant mice was performed using total RNA harvested from adult retina. Retinas were dissected in ice-cold 1X PBS and quickly transferred to 1 ml TRIzol (ThermoFisher 15596-026) in precellys tubes (Precellys, KT03961-1-009.2) on ice. Tissues were homogenized using a disruptor genie

(Scientific Industries, 0-SI-DD38) at 3000RMP for 1 minute. Homogenized tissues were stored at -80° until ready for RNA isolation. RNA was isolated by phenol chloroform extraction and precipitated using 100% isopropanol. RNA concentration and purity were measured using a Nanodrop spectrophotometer (Nanodrop Technologies, ND-1000) and Nanodrop 1000 V3.8.1 software. RNA samples were stored at -80°C.

Complementary DNA (cDNA) was prepared for the *Pax6* transcript using QuantiTect Reverse Transcriptase Kit (Qiagen, 205313) and following the manufacturer's protocol with several modifications. 1.5 µg RNA was used per RT reaction and gDNA digestion was performed for 7min at 42˚C. 10 µM *Pax6* 3'UTR-specific primer, ATCTAGTGTGTGTTGTCCCAGG, was used for cDNA synthesis instead of the supplied random primers. The RT reaction was performed for 15 minutes at 42˚C and heat inactivated for 5 minutes at 95˚C.

Three overlapping PCR reactions were used to amplify the *Pax6* cDNA from mutant mouse retina, covering most of the 5'UTR, the entire coding region, and most of the 3'UTR. Primers were designed to span exon boundaries, to avoid amplifying gDNA. The region spanning exons 0-6 was amplified using forward primer 1 TTGAGAAGTGTGGGAACCAGC and reverse primer 1 ATGACACACTGGGTATGTTATCG, the region spanning exons 9-13 was amplified using forward primer 2 ATGCAGAACAGTCACAGCG and reverse primer 2 TTAGAAAACCATACCTGTATTCTTGC, and the region spanning exons 9-13 was amplified using forward primer 3 TCTGGAGAAAGAGTTTGAGAGG and reverse primer 3 TTACAAAGTGAAGTGCTTCTAACC (Figure 30, Table 4). Thermocycling conditions were as follows: 98˚C for 2 min, cycle

through 98˚C for 15 sec, 60˚C for 30 sec and 72˚C for 30 sec 35 times, with a final extension at 72˚C for 2 min.

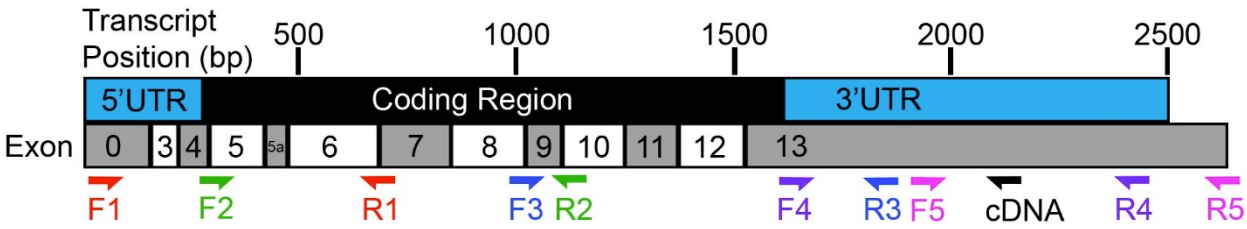


Figure 30. Schematic of the *Pax6* transcript with locations of PCR primers

Pax6-206 ENSMUST 00000111086.10, with relative locations of primers used for cDNA preparation and PCR amplification. The location of the primer used for cDNA synthesis is shown in black, and the primer pairs used to PCR amplify *Pax6* 5'UTR, coding region, and 3'UTR are shown in colour as forward (F) and reverse (R) 1-5. The PCR reactions using primer pairs 1-3 were performed using cDNA prepared from adult retina total RNA, and the PCR reactions using primer pairs 4 and 5 were performed using gDNA. Note that exons 1 and 2 are transcribed from promoter 1 (P1) and result in the production of several different 5'UTR variants, which are not shown here. PCR products were sequenced using the same primers used for amplification and different nested primers.

Sequencing PCR products

PCR products used for sequencing the *Pax6* 3'UTR and *Pax6* transcript from mutant mice were purified using QIAquick PCR purification kit (Qiagen, 28104) and sequenced by Eurofins Genomics SimpleSeq DNA sequencing service. Sequencing of the 3'UTR was performed using the same primers used for PCR amplification along with an additional nested primer for sequencing the product of primer pair 1: AAGAACAACACAGGCTGTTGG; and two nested primers for sequencing the product of primer pair 2: TTTGCGTACAGAAGGCACG and AAAGTAACCATTGATTTAGG. The *Pax6* transcript was sequenced using the same primers that were used for the PCR reactions with additional nested primers ATCTTCTATTCTAGAAGTGG and

ATTCGGGAAATGTCGCACGG used with the product of primer pair 1, and

TACGAGACTGGCTCCATCAG used with the product of primer pair 2.

Table 4. Primers used for sequencing *Pax6* transcript

Primer Name	Sequence (5'-3')	Location	Use
Pax6Seq_F1	TTGAGAAGTGTGGGAACCAGC	Exon 0 (5'UTR)	PCR/Sequencing
Pax6Seq_R1	ATGACACACTGGGTATGTTATCG	Exon 6-7 (CDS)	PCR/Sequencing
Pax6Seq_F2	ATGCAGAACAGTCACAGCG	Exon 4-5 (CDS)	PCR/Sequencing
Pax6Seq_R2	TTAGAAAACCATACCTGTATTCTTGC	Exon 9-10 (CDS)	PCR/Sequencing
Pax6Seq_F3	TCTGGAGAAAGAGTTTGAGAGG	Exon 8-9 (CDS)	PCR/Sequencing
Pax6Seq_R3	TTACAAAGTGAAGTGCTTCTAACC	Exon 13 (3'UTR)	PCR/Sequencing
Pax6Seq_F4	TAAAGAGAGAAGGAGAGAGC	Exon 13 (3'UTR)	PCR/Sequencing
Pax6Seq_R4	TTGGACTGTGAACAGTAATACAGC	Exon 13 (3'UTR)	PCR/Sequencing
Pax6Seq_F5	ATGGACTAGAACCAAGGACC	Exon 13 (3'UTR)	PCR/Sequencing
Pax6Seq_R5	TTTCTTTGAGGACCATCAGG	Exon 13 (3'UTR)	PCR/Sequencing
Pax6_cDNA	ATCTAGTGTGTGTTGTCCCAGG	3' of exon 13	cDNA
Pax6_3UTR_Seq	AAGAACAACACAGGCTGTTGG	Product of pair 4	Sequencing
Exon3_Seq	ATCTTCTATTCTAGAAGTGG	Product of pair 1	Sequencing
Exon5_Seq	ATTCGGGAAATGTCGCACGG	Product of pair 1	Sequencing
Exon6_Seq	TACGAGACTGGCTCCATCAG	Product of pair 2	Sequencing
3UTR_Seq_F5	TTTGCGTACAGAAGGCACG	Product of pair 5	Sequencing
3UTR_Seq_R5	AAAGTAACCATTGATTTAGG	Product of pair 5	Sequencing

Underline indicates exon junction.

4.3.7 Cell and tissue harvest

αTC1-6 cells were harvested for testing miRNA affinity purification with the *Pax6* transcript, using an approach called miR-CATCH [405], since they endogenously express *Pax6*. Unfixed cells were pelleted, flash frozen on liquid nitrogen, and stored at -80˚C.

miR-CATCH and RT-qPCR for miRNAs and *Pax6* were performed using P1 V-SVZ tissue. P1 mice were anesthetized with isoflurane and euthanized by decapitation. Brains were cut in half using a micro-scalpel (Fine Science Tools, 10316-14) and flat razor, and removed from the skull in ice-cold 1X PBS using a small spatula. Cerebellum was removed with a flat razor. Brain was stored in 1X PBS on ice until ready for sectioning. 3 animals were pooled per sample.

P1 V-SVZ was dissected from coronal brain sections prepared using a vibratome, following a general protocol provided by Dr. Ernie from the Cremer lab (Institut de Biologie de Développement de Marseille, France)[206]. A strip of 4% agar was glued to the platform of a Leica VT1000S vibratome (Leica Biosystems, Wetzlar Germany) using crazy glue, oriented parallel to the blade. Brain halves were crazy-glued caudal side down with the cut medial sides facing the strip of agar to support the tissue during cutting. Brain was cut into 400 µm coronal sections using carbon steel feather blades (TED Pella, Inc, 121-9) in ice-cold HEPES-buffered artificial cerebral spinal fluid (aCSF: 10 mM HEPES, 150 mM NaCl, 4 mM KCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 10 mM D-glucose). The region of V-SVZ harvested was visually identified by long open ventricles and visible olfactory limb of the anterior commissure. Each brain half yielded approximately 2 sections that contain the region of interest.

DEPC-treated (Sigma, D5758-50ML) or nuclease-free solutions were used for fixation and washes. Brain sections were fixed in aCSF with 1% paraformaldehyde (PFA, Electron Microscopy Sciences, 157-8). Fixation was performed for 10 min at room temperature on a plate rocker, and the cross-linking reaction was quenched in 250 mM Tris-EDTA (Sigma, T9285-100ML) for 10 min at room temperature. Cells and sections were washed three times in DEPC-treated PBS, pH 7.4 (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4).

V-SVZ regions were dissected in ice-cold 1X PBS in a rubber-coated dish under a dissecting microscope using fine forceps and a micro-scalpel (Fine Science Tools, 10316-14). For *Pax6* mRNA and miRNA quantification experiments, D, DL and VL V-SVZ was separated using Lumsden bioscissors (O.P.I.6-153). Whole D and L V-SVZ were used for *Pax6* pulldown experiments. Tissue was pelleted and flash frozen on liquid nitrogen. Tissue was stored at -80˚C for several weeks until all samples needed for a given experiment were collected (Figure 32i).

4.3.8 miR-CATCH

miR-CATCH oligo design

Oligonucleotides used to affinity purify *Pax6* transcript were designed according to Vencken *et al*. (2014)[405]. *Pax6* transcript (Pax6-206, ENSMUST 00000111086.10) secondary structure was predicted using RNAfold [406]. Several candidate 18-25 nucleotide oligonucleotide sequences were selected that were complementary to regions of the *Pax6* transcript not within the 3'UTR and not predicted to contain secondary structure. Mfold software [407] was used to identify candidate

oligonucleotides that are single stranded at 37˚C under 500 mM monovalent salt conditions. IDT's Oligo Analyzer (Integrated DNA Technologies) was used to ensure that the hybridization energy of candidate oligonucleotide dimers was less than 15% of the oligo-mRNA hybrid. Nucleotide BLAST was performed to ensure that candidate oligonucleotides share less than 15 nucleobases of complementarity to non-specific mouse endogenous RNAs. HPLC-pure DNA oligos 5' modified with biotin connected by a TEG spacer were ordered from Eurofins Genomics. Three candidate oligos were selected for testing affinity purification of *Pax6* mRNA: miR-CATCH_1 CTATTTTGCTTACAACTT, miR-CATCH_2 TGGAGTTGGTGTTCTCTC, miR-CATCH_3 TGGCTGGTAGACACTGGT. A negative control oligo antisense to miR-CATCH_2 was also used: GAGAGAACACCAACTCCA (Figure 31, Table 5).

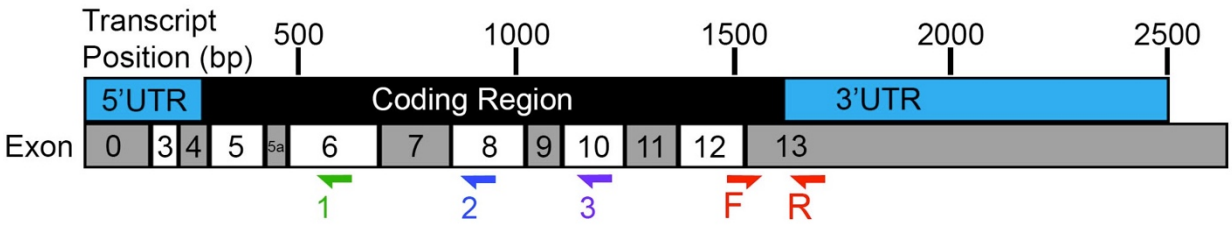


Figure 31. *Pax6* qPCR and pulldown primers

Pax6-206 ENSMUST 00000111086.10, with relative locations of primers used for Pax6 qPCR and pulldown. The forward (F) and reverse (R) primers used for qPCR are shown in red. Regions recognized by miR-CATCH oligos are shown as 1-3. miR-CATCH oligos 1-3 have a 5' Biotin-TEG modification.

Table 5. Primers used for *Pax6* qPCR and oligonucleotides used for *Pax6* pulldown

Oligo Name	Sequence (5'-3')	Use
Pax6_qPCR_F	TCAACAGGACTCATTTCACC	*Pax6* qPCR
Pax6_qPCR_R	TTCCTCTCTCGATCACATGC	*Pax6* qPCR
miR-CATCH_1	CTATTTTGCTTACAACTT	*Pax6* pulldown, 5' BIOTEG modification
miR-CATCH_2	TGGAGTTGGTGTTCTCTC	*Pax6* pulldown, 5' BIOTEG modification
miR-CATCH_3	TGGCTGGTAGACACTGGT	*Pax6* pulldown, 5' BIOTEG modification
miR-CATCH_negctrl	GAGAGAACACCAACTCCA	*Pax6* pulldown, antisense control

Underline indicates exon junction.

Lysate preparation

RNase-free solutions were prepared for performing *Pax6* miR-CATCH experiments. Solutions not containing tris were RNase inactivated using 0.1% DEPC (Sigma, D5758-50ML) at 37˚C overnight, followed by autoclaving. Tris solutions were made using nuclease-free 100X tris EDTA pH 8.0 (Sigma, T9285-100ML) and nuclease-free water (ThermoFisher, 10977015).

PFA-fixed cells and tissue were lysed as described by Vencken *et al*. (2014)[405] with several modifications. Lysis buffer was prepared by adding superase-in (ThermoFisher, AM2696, final concentration 87 U/ml), PMSF (Sigma, P7626-5G, final concentration 500 mM) and Complete protease inhibitor cocktail tablet (Roche, 4693116001, 1 tablet per 10 ml) fresh to a solution containing 140 mM NaCl, 50 mM HEPES, 1 mM EDTA, 1% triton, 0.1% sodium deoxycholate, pH 7.5. Unfixed αTC1-6 cells were resuspended in 1.4 ml lysis buffer per cell pellet, and V-SVZ tissue was resuspended in 0.25 ml lysis buffer. αTC1-6 cells were homogenized in 2 ml tubes with ceramic beads (Precellys, KT03961-1-009.2) for 8 min, and V-SVZ tissue was homogenized in 0.5 ml tubes with ceramic beads (Precellys Lysing Kit, Soft tissue

homogenizing CK14_0.5ml) for 10 min at 3000 RPM using a Digital Disruptor Genie (Scientific Industries, New York USA). Tissue was cooled on ice for 2 min between each 1 min of lysis. Following homogenization of V-SVZ tissue, lysate volume was brought to 0.5 ml with lysis buffer. EDTA was added to 20 mM and lysate was centrifuged at 16,000 g for 5 min at 4˚C. Supernatant was used for miR-CATCH (Figure 32ii).

Pax6 affinity purification

Magnetic streptavidin beads (ThermoFisher, Dynabeads MyOne Streptavidin C1, 65002) were prepared for miR-CATCH according to Venken *et al*. (2014)[405]. 200 ul of beads were used per capture. Beads were washed three times using B&W buffer (1 M NaCl, 0.5X Tris-EDTA pH 8.0), twice with solution A (0.1 M NaOH, 0.05 M NaCl) and once with solution B (0.1 M NaCl) as described. A 6-tube magnetic separation rack (NEB, S1506S) was used to immobilize the magnetic beads for 2 min between each wash. Beads were incubated with 200 µl of 5 µM capture oligonucleotide in 1X Tris-EDTA for 15 min at room temperature to facilitate oligonucleotide immobilization. Excess oligonucleotide was washed using hybridization buffer (2X Tris-EDTA pH 8.0, 1 M LiCl) as described.

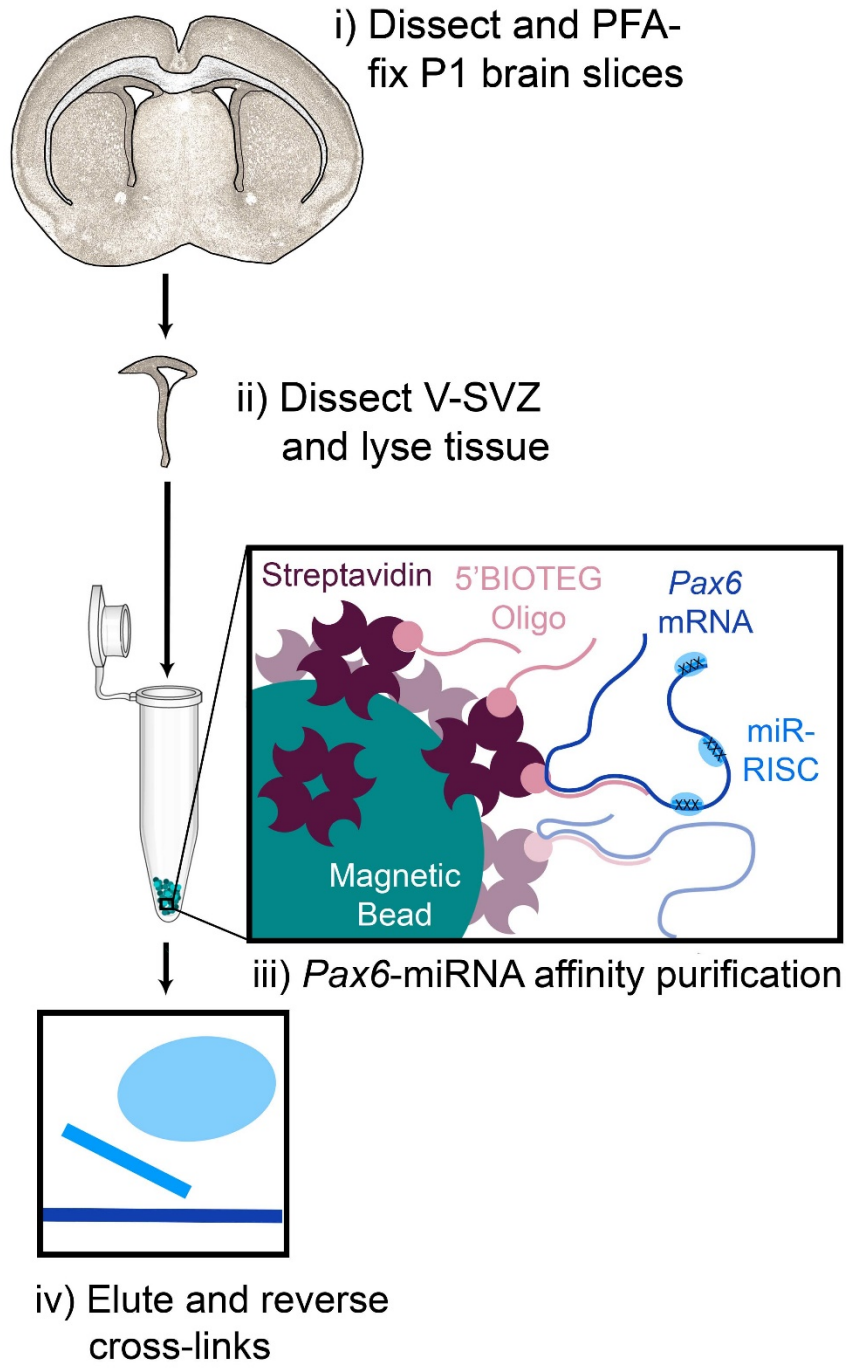


Figure 32. Summary of *Pax6* affinity purification by miR-CATCH

Pulldown of *Pax6* transcript was performed by resuspending the streptavidin beads in 0.7 ml hybridization buffer and 0.5 ml 1X Tris-EDTA at 37˚C. 200 µl lysate was used per capture (Figure 32iii). Beads were agitated with lysate at 37˚C in a hybridization

oven (VWR) for 1.5 hours. Following target immobilization, beads were washed twice with 37˚C wash buffer A (1X Tris-EDTA, 0.15 M LiCl, 0.5% SDS) and twice with 37˚C wash buffer B (1X Tris-EDTA, 0.15 M LiCl). Washes were performed for 5 min at 37˚C with agitation. To elute *Pax6* transcript, the bead pellet was resuspended in 30 µl of 60˚C 0.1X Tris-EDTA and incubated at 60˚C for 5 min. The supernatant was removed and PFA crosslinks were reversed by incubating at 70˚C for 45 min (Figure 32iv). Crosslink reversal was not performed for *Pax6* affinity purification from unfixed αTC1-6 cells. Purified RNA was stored at -80˚C and used for cDNA synthesis within 1 week.

4.3.9 RNA isolation from fixed V-SVZ

RNA was isolated from PFA-fixed D, DL and VL P1 V-SVZ tissue for miRNA and *Pax6* RT-qPCR. Lysate was prepared as described in section 4.3.8. 200 µl of lysate was prepared per V-SVZ sample and region. RNA was isolated using the miRNeasy FFPE kit (Qiagen, 217504) and a modified version of the kit's protocol. 190 µl buffer PKD and 10 µl proteinase K were added per 200 µl lysate. Lysate was incubated at 56°C for 15 min followed by 80˚C for 15 min. 40 µl DNase booster buffer and 10 µl DNase stock were added to the lysate, and DNase digestion was performed for 15 min at room temperature. Lysate was mixed with 800 µl of buffer RBC and 2.8 ml 100% ethanol, and was passed through the supplied column via centrifugation. RNA was extracted into 30 µl of 0.1X TE (Sigma, T9285-100ml) and was stored at -80°C for up to 1 week prior to use in cDNA synthesis for miRNA qPCR and up to two months for *Pax6* qPCR experiments.

4.3.10 *Pax6* RT-qPCR

Pax6 transcript quantified using RT-qPCR for the miR-CATCH experiments and for comparing mRNA levels between WT and mutant mouse V-SVZ. cDNA synthesis was performed using Quantitect reverse transcriptase kit (Qiagen, 205313) following the manufacturer's protocol. 600 ng RNA was used per cDNA synthesis reaction for whole V-SVZ tissue, and 250 ng RNA was used per cDNA reaction for D, DL and VL V-SVZ pieces. Digestion of gDNA was performed prior to cDNA synthesis for 7 min at 42˚C. Reactions were performed using a T3 thermocycler (Biometra). cDNA synthesis was performed at 42°C for 15 min followed by heat inactivation at 95°C for 3 min, and cDNA was stored at -20˚C and used within one week.

Primers used for qPCR amplification of *Pax6* cDNA were designed such that the forward primer spans the boundary between exons 12 and 13, to prevent amplification of gDNA (Figure 31, Table 5). The primers used for *Pax6* qPCR were: Pax6_qPCR_F TCAACAGGACTCATTTCACC and Pax6_qPCR_R TTCCTCTCTCGATCACATGC, producing an amplicon size of 128 bp. Primer efficiency was tested using a cDNA dilution series and was found to be 100%. TATA binding protein (*Tbp*) and Hypoxanthine-phosphoribosyltransferase (*Hprt*) were selected as reference genes, and primers were purchased from Qiagen (QT00198443, QT00166768).

qPCR was performed to quantify levels of *Pax6*, *Tbp* and *Hprt* using the Quantitect SYBR green PCR kit (Qiagen, 204145). 50 µl qPCR reactions were prepared following the manufacturer's protocol. qPCR reactions were run using the Mx3000P qPCR system (Stratagene, California USA) with the following thermocycling conditions:

initial heat activation at 95˚C for 15 min, cycle between 94˚C for 15 seconds, 60˚C for 30 seconds and 72˚C for 30 seconds forty times. Measurements were collected twice during the elongation step. Each sample was run in three technical replicate reactions. qPCR reactions being directly compared to one another were run on a single plate to avoid inter-run variability.

qPCR data was analyzed to determine whether the miR-CATCH strategy resulted in enrichment of *Pax6* transcript, and to determine whether *Pax6* transcript quantities were altered in the V-SVZ of miR-7-5p MRE mutants relative to WT mice. Grubbs test was used to detect outlier cycle threshold (Ct) values for technical replicates assays, and these were excluded. Mean Ct value of technical replicates was calculated for each sample and target. Standard deviation of technical replicate Ct values was calculated for each sample to ensure S.D. < 0.5 Ct. For miR-CATCH validation experiments using unfixed αTC1-6 cells and fixed V-SVZ tissue, relative *Pax6* quantity was calculated using Pfaffl's method without reference genes by the following equation: $2^{[\text{lysate Ct}]}/2^{[\text{Pax6 oligo Ct}]}$ or $2^{[\text{negctrl oligo Ct}]}/2^{[\text{Pax6 oligo Ct}]}$, where *Pax6* from cell lysate or from pulldown using the non-specific negative control oligonucleotide was used for normalization. For *Pax6* transcript quantification experiments, relative *Pax6* quantity was calculated for each sample using Pfaffl's method with multiple reference genes [408] by the following equation: $(2^{[\text{Tbp Ct}]} * 2^{[\text{Hprt Ct}]})^{0.5}/2^{[\text{Pax6 Ct}]}$. Relative *Pax6* quantity was normalized to D V-SVZ level to identify the gradient in WT animals and was normalized to WT levels to determine whether there were any changes with miR-7-5p MRE mutation.

4.3.11 miRNA RT-qPCR

Detection and quantification of mature miRNAs was performed using both the TaqMan RT-qPCR system (ThermoFisher) and the miScript SYBR green system (Qiagen). miR-7-5p RT-qPCR was performed to measure the interaction of miR-7-5p to the *Pax6* transcript by miR-CATCH, and was performed to compare miR-7-5p levels across the V-SVZ between WT and mutant mice. A TaqMan qPCR array card was also used to assay for many miRNAs simultaneously.

TaqMan RT-qPCR

TaqMan microRNA reverse transcription kit (ThermoFisher, 4366597) and TaqMan RT primers (miR-7a-5p, ThermoFisher, 4427975, ID 000268; U6 snRNA, ThermoFisher, 4427975, ID 001973) were used to prepare cDNA for miR-7-5p quantification. For miR-7-5p profiling experiments, 50 ng RNA was used per RT reaction, and cDNA preamplification was not performed prior to qPCR. For miR-CATCH and profiling experiments using the TaqMan array, cDNA synthesis was performed using a custom RT primer pool (ThermoFisher, 4459651) containing cDNA primers for all miRNA assays on the array card. cDNA synthesis reaction conditions were as follows: 16˚C for 30 min, 42˚C for 30 min followed by 85˚C for 5 min. cDNA was stored at -20˚C for a maximum of 1 week prior to use.

For miR-CATCH and profiling experiments using the TaqMan array, cDNA was preamplified using the TaqMan custom preamp pool (ThermoFisher, 4441856) and TaqMan preamp master mix (ThermoFisher, 4384266) following the manufacturer's protocol. cDNA was diluted 1:4 in nuclease-free water and 10 µl of cDNA were used per

preamplification reaction. No additional water was added to the reaction to maximize the volume of cDNA. Preamplification thermocycling conditions were as follows: 95˚C for 10 min, 55˚C for 2 min, 72˚C for 2 min, cycle between 95˚C for 15 sec and 60˚C for 4 min 12 times, heat inactivate at 99.9˚C for 10 min. Preamplification products were diluted 1:8 in nuclease free 0.1X Tris EDTA pH 8.0 (Sigma, T9285-100ML) and stored at -20˚C for a maximum of 1 week prior to use. cDNA synthesis and preamplification were performed using a T3 thermocycler (Biometra).

For expression and miR-CATCH analyses assaying for miR-7-5p individually, TaqMan 20X small RNA assays (miR-7a-5p, ThermoFisher, 4427975, ID 000268; U6 snRNA, ThermoFisher, 4427975, ID 001973) and TaqMan universal master mix II with UNG (ThermoFisher, 4440038) were used to prepare the qPCR master mix following the manufacturer's protocol with some adjustments. cDNA was diluted 1:4 and preamplification product was diluted an additional 1:4 in nuclease-free water prior to performing qPCR. 6 µl of diluted cDNA was used per reaction for profiling experiments, and 6 µl diluted preamplification product was used per reaction for miR-CATCH experiments. 20 µl qPCR reactions were performed in triplicate. Samples were loaded into 96-well microamp plates (ThermoFisher, 4346907) and sealed with an optical adhesive cover (ThermoFisher, 4360954). qPCR reactions were run using a 7900HT qPCR machine (Applied Biosystems) with the 96-well block and SDS software (ThermoFisher, V2.4) using the $\Delta\Delta$Ct (RQ) and standard run mode. qPCR thermocycling conditions were as follows: 50˚C for 2 min, 95˚C for 10 min, cycle between 95˚C for 15 sec followed by

60˚C for 60 sec repeated 40 times. Raw Ct data was exported into CSV format using the ExpressionSuite software (ThermoFisher, V1.0.3).

Custom TaqMan miRNA multiplex qPCR cards were designed to run 8 samples per card and assay for 48 small RNAs per sample. miRNAs predicted to target the *Pax6* 3'UTR were selected as described in chapter 3 [400]. Of those predicted to target *Pax6*, we selected miRNAs found to interact with the *Pax6* 3'UTR in αTC1-6 cells by miTRAP and miRNAs detected in rat SVZ neural stem cells [409]. Expression profiling results from αTC1-6 cells, βTC6 cells, E12.5 mouse retina, adult mouse retina and adult mouse lens (Figure 21) were also used to inform miRNA assay selection, with miRNAs expressed in most of these tissues being selected. miR-7-3p was also included because it has predicted MREs in the *Pax6* 3'UTR and is of interest as a possible regulator of *Pax6*. Two miRNAs, let-7a and miR-204-5p, were selected as negative controls. These miRNAs were detected in tissues of interest but lack predicted MREs in the *Pax6* 3'UTR. U6 was included as a normalizing control.

TaqMan miRNA array cards containing 8X48 small RNA qPCR assays (ThermoFisher, 4449139) were used for expression and miR-CATCH analyses to assay for many miRNAs simultaneously. qPCR master mix was prepared using TaqMan universal master mix II with UNG (ThermoFisher, 4440038) and 1.13 ul preamplification product according to the manufacturer's protocol. Cards were equilibrated to room temperature for 15 min before being removed from their packaging and loaded with 100 ul of qPCR master mix per port. TaqMan array cards were loaded following the manufacturer's instructions, were centrifuged using the Heraeus megafuge 40R (Thermo Scientific) and

sealed using the TaqMan micro fluidic card sealer (Applied Biosystems, model 7331770). qPCR reactions were run using a 7900HT qPCR machine with the 384-well block and SDS software (ThermoFisher, V2.4) using the ΔΔCt (RQ) and 384 wells TaqMan low density array run mode. qPCR thermocycling conditions were as follows: 50˚C for 2 min, 94.5˚C for 10 min, cycle through 97˚C for 30 sec and 59.7˚C for 1 min 40 times. Raw Ct data was exported into CSV format using the ExpressionSuite software (ThermoFisher, V1.0.3).

miScript RT-qPCR

cDNA synthesis for mature miRNAs was performed using the miScript II RT kit (Qiagen, 218161) using the HiSpec buffer following the manufacturer's protocol. 100 ng of RNA were used per 20 µl RT reaction and reaction conditions were as follows: 37˚C for 1 hr followed by heat inactivation at 95˚C for 5 min. Reactions were performed using a T3 thermocycler (Biometra). cDNA was diluted 1:10 in nuclease-free water and stored at -20˚C for up to 1 week.

Assays were performed for the target miRNA miR-7a-5p (Qiagen, 218300, MS00005880) and candidate controls U6 (RNU6-2, Qiagen 218300, MS00033740), miR-16-5p (Qiagen 218300, MS00037366), miR-26a-5p (Qiagen 218300, MS00032613), miR-106a-5p (Qiagen 218300, MS00011039), and let-7a-5p (Qiagen 218300, MS00032179). miR-145-5p (Qiagen 218300, MS00001631) and miR-375-3p (Qiagen 218300, MS00032774) were used to validate TaqMan card results. qPCR was performed using the miScript SYBR green PCR kit (Qiagen, 218075) and PCR reactions were set up following the manufacturer's protocol. qPCR reactions were run using the Mx3000P qPCR system (Stratagene, California USA) with the following thermocycling conditions:

initial heat activation at 95˚C for 15 min, cycle between 94˚C for 15 seconds, 55˚C for 30 seconds and 70˚C for 30 seconds forty times. Fluorescent data collection was performed twice during the elongation step. Each sample was run as three technical replicate reactions, with one –RT control reaction.

The stability of these candidates across the D, DL and VL V-SVZ, and between WT and 517+655MUT genotypes was compared by calculating the gene variability measure as described by Vandesompele *et al.* (2002)[408]. miR-26a-5p, miR-106a-5p and U6 were selected as normalizing controls for miR-7-5p expression analysis.

Data analysis

For V-SVZ profiling experiments using the TaqMan system to assay for many miRNAs simultaneously, relative miRNA quantity was calculated for each miRNA of interest using Pfaffl's method with multiple reference genes [408] by the following equation: $(2^{[\text{miR-26a}]} * 2^{[\text{miR-106a}]} * 2^{[\text{U6}]} * 2^{[\text{miR-16}]})^{1/4} / 2^{[\text{miRNA Ct}]}$. miRNAs were considered to be expressed only if they were detected in all replicate samples. Two-way ANOVA with Dunnett's multiple comparisons was used to compare relative miRNA levels in WT DL and VL V-SVZ, with the goal of identifying gradient-expressed miRNAs. To identify candidate differentially expressed miRNAs using TaqMan array cards, the relative miRNA quantity for 517+655MUT V-SVZ was normalized to WT. Fold change was calculated by a log2 transformation of the normalized relative miRNA quantities. D, DL and VL regions were analyzed in combination.

For V-SVZ profiling experiments using the miScript system to assay for miRNAs individually, mean cycle threshold (Ct) value of technical replicates was calculated for

each sample, and Grubbs test was used to detect outlier technical replicate Ct values which were excluded. Standard deviation of technical replicate Ct values was calculated for each sample to ensure S.D. < 0.5 Ct. Relative miRNA quantity was calculated by the following equation: $(2^{[26a]}*2^{[106a]}*2^{[U6]})^{1/3}/2^{[miR\text{-}7\ Ct]}$. Relative miR-7-5p quantity was also normalized to WT D V-SVZ level and one-way ANOVA with Tukey's multiple comparisons was used to compare miR-7-5p across genotypes within the D, DL and VL V-SVZ regions.

For miR-CATCH experiments comparing mutants to WT mice, relative miRNA quantity was calculated for each sample using Pfaffl's method with *Pax6* as the reference gene [359] by the following equation: $2^{[Pax6\ Ct]}/2^{[miRNA\ Ct]}$. Relative miRNA quantity for the mutants was normalized to WT to identify miRNAs having altered interaction to the *Pax6* 3'UTR with miR-7-5p MRE mutation. One-way ANOVA with Tukey's multiple comparisons was used to compare miR-7-5p interaction between WT and mutant *Pax6* transcript. To determine whether miR-7a-5p interacts with the WT *Pax6* 3'UTR, relative miR-7-5p quantity from WT V-SVZ samples was calculated using Pfaffl's method without reference genes by the following equation: $2^{[negctrl\ oligo\ Ct]}/2^{[Pax6\ oligo\ Ct]}$, where miR-7-5p from the negative control oligo pulldown condition was used for normalization. TaqMan array cards were used to identify other miRNAs with altered interaction to the 517+655MUT *Pax6* transcript compared to WT *Pax6*. The relative miRNA quantity for 517+655MUT pulldown was first normalized to WT and fold change was calculated by a log2 transformation of the normalized relative miRNA quantities.

4.3.12 PAX6 Immunofluorescence

Tissue harvest and section preparation

Brain tissue was harvested from P1 pups for PAX6 quantification. P1 pups were anesthetized via inhalant isoflurane (AVP, 8061652) and euthanized by decapitation. Skin and nasal region were removed to permit entry of fixative into the brain. Brains were transferred into 4% PFA in 0.1 M phosphate buffer (PB: 60 mM Na_sHPO_4, 20 mM NaH_2PO_4) and fixed for 2 hours at room temperature on a plate rocker. Brains were dissected from the skull in 0.1 M PB and fixed overnight in 4% PFA in 0.1 M PB at 4°C on a rotator. The cerebellum was removed from fixed brains with a straight-edge razor and brains were transferred into 70% ethanol.

Adult mouse brain tissue was harvested for PAX6 IF. Mice were anesthetized via inhalant isoflurane (AVP, 8061652) and given a 1 mg/kg intramuscular injection of a 1:1 v/v mixture of ketamine (100mg/ml) and dexmedetomidine hydrochloride (0.5mg/ml, Pfizer, cat#02333929). Animals were pinned in dorsal recumbency, fur was matted with 70% ethanol and the abdomen was cut open by an incision in the stomach. Two incisions made up either side of the ribcage, parallel to the sternum, to the apex of the chest and the diaphragm was cut away horizontally. The ribs and sternum were pinned back with a hemostat (Fine Science Tools, 13004-14) to expose the heart. A 23G butterfly needle (Becton Dickinson, cat# 367253) inserted into the left ventricle and the right atrium of the heart cut to allow blood and perfusate to escape. Animals were first perfused with 50 ml ice-cold 0.1 M phosphate buffer, followed by 30 ml of ice-cold 4% PFA in phosphate buffer. Solutions were pumped at 4 ml/min by peristaltic pump

(Masterflex Easy Load, ColeParmer, cat# EW-07518-00). Following perfusion, brains were dissected and fixed for 24 hours at 4°C in 0.1 M phosphate buffer with 4% PFA, then transferred into 70% ethanol for 2 days.

Paraffin embedding and sectioning were performed by Wax-it Histology Services (Vancouver, Canada). 4 µm coronal sections were prepared through the brain region visually identified by long open lateral ventricles, corresponding to approximately 1 mm to 1.6 mm from the rostral end of the P1 brain (Figure 33).

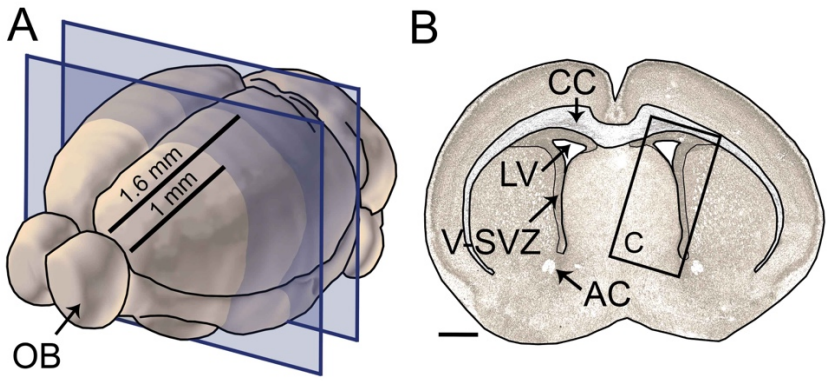


Figure 33. P1 brain sectioning for PAX6 immunofluorescence

(A) Postnatal day 1 (P1) mouse brain. Coronal sections were collected between 1-1.6 mm of the brain (blue sectioning planes) starting from the anterior end of the brain, not including olfactory bulbs (OB). (B) Coronal section through the region of interest revealing elongated lateral ventricles (LV), well-defined corpus callosum (CC) and anterior commissures (AC). Scale bar - 500 µm.

PAX6 immunofluorescence

Paraffin sections were deparaffinized via successive washes twice for 5 min in 100% xylenes, twice for 3 min in 100% ethanol, once for 1 min in 95% ethanol, once for 1 min in 80% ethanol and once for 1 min in 1X PBS. Antigen retrieval was performed

using the steamer method described by Tang *et al.* (2007)[410] and a 5.5-quart digital steamer (Hamilton Beach, 37530C). Slides were covered in sodium citrate buffer (10 mM Tri-sodium citrate, 0.5% tween-20, pH 6.0) and were incubated in the steam phase at approximately 100°C for 30 min. Slides were washed once with 1X PBS before applying antibody.

PAX6 IF was performed using rabbit anti- PAX6 (Covance, PRB-278P) diluted 1:500 in 1X PBS with 0.1% triton. Antibody dilutions were performed in PBS prepared from PBS tablets (Sigma, P4417-100TAB). Slides were incubated with primary antibody overnight at 4°C in a humidified chamber. Slides were washed three times in 1X PBS with 0.1% triton and 0.01% tween-20. Alexa 555 donkey anti-rabbit (ThermoFisher, A-31572) was diluted 1:500 and Draq5 nuclear stain (ThermoFisher, 62251) was diluted to 5 μM in 1X PBS with 0.1% triton, and slides were incubated with secondary antibody for one hour at room temperature. Slides were washed three times in 1X PBS with 0.1% triton and 0.01% tween-20, and mounted with a #1.5 glass coverslip using immu-mount (Shandon, 9990402).

Image acquisition

Images were collected using a Nikon Eclipse Ni microscope and C2+ confocal system (Nikon Instruments Inc.) and Nikon NIS-Elements software. Photomicrographs used for PAX6 IF intensity quantification were collected using the Nikon Plan APO 20X/0.75 objective at 1024 image size, 5.3 pixel dwell and with a laser power of 10. A uniformly autofluorescent plastic slide (Chroma, 92001) was imaged under the same conditions used for image collection every imaging session for flat field correction. High

magnification images for PAX6 nuclear localization analysis were collected using the Nikon APO TIRF 60X/1.49 oil objective with type A immersion oil (Cargille, 16482), and 22 pixel dwell. The dorsal V-SVZ was imaged first. The midpoint for the z-stacks was established using the Draq5 stain, the gain was quickly set for the PAX6 channel to ensure that the brightest V-SVZ pixels were just below saturation. Three z-stacks were collected centered around the midpoint, spaced 1 µm apart. All acquisition settings were kept constant for imaging the remainder of the V-SVZ. 3 separate images were required to capture one V-SVZ region. 3-4 V-SVZ regions were imaged per animal, and LVs were oriented either horizontally or vertically for imaging to permit image analysis.

Image processing and data analysis

Image processing was performed using ImageJ software [411]. A flat field reference image was prepared using the uniformly autofluorescent image by changing the image to 32-bit, selecting the channel of interest, smoothing the image using a median filter with a 20-pixel radius, and dividing the entire image by the brightest pixel. V-SVZ images were flat field corrected by changing to 32-bit, selecting the channel of interest, and dividing the image by the flat field reference image. A maximum intensity z-projection was performed, the three images comprising one V-SVZ region were stitched using the pairwise stitching function [412]. The background fluorescence in the brain slices was measured by defining an ROI in the striatum neighboring the V-SVZ, where PAX6 is absent. The average pixel intensity was measured in the background ROI and this value was subtracted from the entire image [413].

The dorsal and lateral VZ regions of interest (ROIs) were defined by morphology using the polygon selection tool in ImageJ. The ROI subdivision macro (Appendix A Equation 1) was used to divide the lateral VZ ROI into 16 equal area regions along the horizontal or vertical axis, depending on the orientation of the LV. Mean pixel intensity was measured for the dorsal VZ ROI and for each of the 16 lateral VZ regions (Figure 34). Immunofluorescence labeling, imaging and data analysis were performed blind to the genotype.

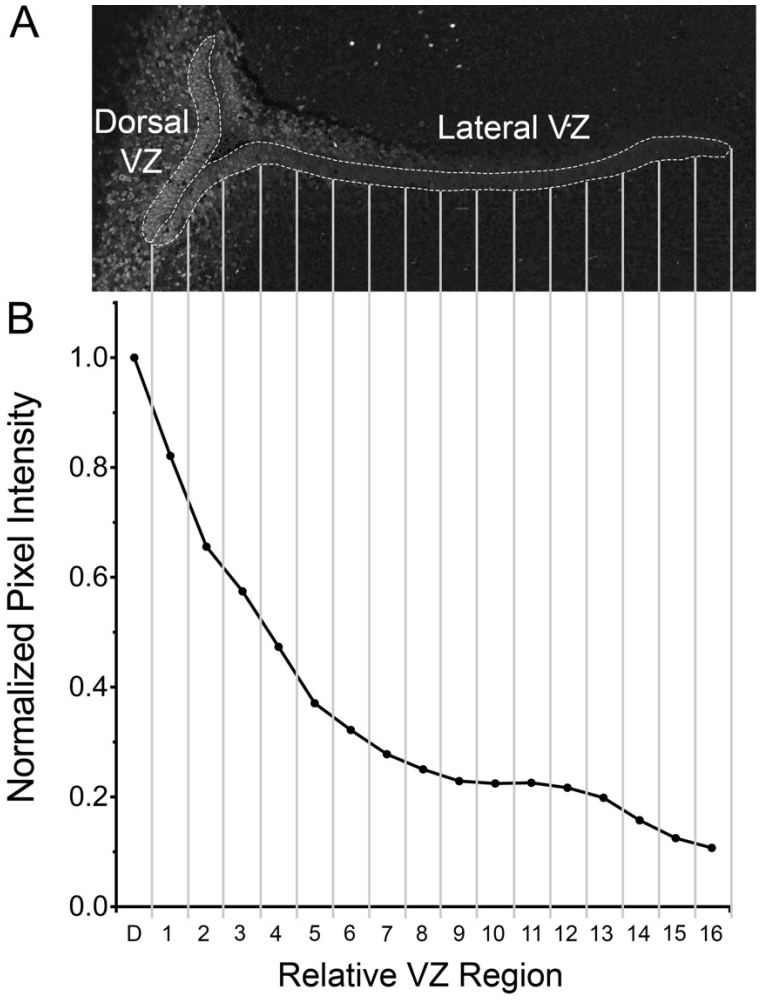


Figure 34. Image analysis strategy for P1 VZ PAX6 IF

(A) Dorsal and lateral VZ were defined as regions of interest (ROIs) by morphology. The lateral VZ ROI was subdivided into 16 equal-area pieces along the D-V axis using a subdivision macro. (B) Average pixel intensity was measured for each ROI and was normalized to the pixel intensity of the dorsal ROI.

Pixel intensity data was analyzed by averaging the pixel intensity for each ROI between the replicate VZ images collected for a single animal. Within each genotype, the average pixel intensity of each lateral region was normalized to the dorsal region intensity (Figure 34). Normalized PAX6 IF in the ventral-most region of the lateral VZ for WT and mutant genotypes were compared by 2-way ANOVA with Tukey's post-hoc test.

4.3.13 Olfactory bulb periglomerular neuron (PGN) fate tracking

Injections and tissue harvest

Heterozygous crosses were set up from the *517MUT*, *655MUT* and *517+655MUT* lines for EdU labeling and WT 129S1 (Jackson Laboratory, 002448) crosses were set up for BrdU labeling. 62.5 mg/ml EdU stock (AbCam, ab146186) and 250 mg/ml BrdU stock (Millipore Sigma, B5002-100MG) were prepared in a 3:1 DMSO: cell biology grade PBS (ThermoFisher, 10010023) and filter sterilized using a 0.2 um cellulose acetate syringe filter (VWR, 28145-477). EdU was diluted to 1 µg/µl and BrdU was diluted to 5 µg/µl in cell biology grade PBS for mouse injections. P1 pups were injected with 10 ug/g EdU or 50 ug/g BrdU twice, six hours apart, to label V-SVZ cells in S-phase. Injections were performed subcutaneously using a 1 ml insulin syringe with a 31-gauge needle (BD, 324912). All injections were performed by the same researcher. WT and homozygous mutant pups were used for olfactory bulb harvest.

Olfactory bulbs were harvested 60 day post-EdU or BrdU injection. Mice were anaesthetized via inhalant isoflurane (AVP, 8061652) and were euthanized by cervical dislocation. Mice were decapitated using surgical scissors (FST, 14001-12), cranial fur was matted using 70% ethanol and fine scissors (FST, 14061-09) were used to remove the skin and calvaria. The mandible was removed using dissector scissors (FST, 14081-09) and a straight-edge razor (VWR, 55411-050) was used to remove the caudal portion of the brain. Olfactory bulbs still embedded in nasal bone were transferred into ice-cold 0.1 M PB. Dorsal nasal bones were removed from around the olfactory bulbs using fine

scissors and forceps. Vannas spring scissors (FST, 15000-10) were used to cut the olfactory sensory nerves so olfactory bulbs could be gently removed from the skull while remaining connected to the rostral brain. Bulbs were fixed in 4% PFA with 0.1 M PB overnight at 4°C, and cryoprotected in 20% sucrose with 0.1 M PB for 48 hours at 4°C.

Prior to embedding, ventral brain tissue was removed to position the ventral olfactory bulb surface flat in 15 mm X 15 mm cryosection base molds (Leica Biosystems, 3803025). Molds were filled with O.C.T. Compound (Tissue-Tek, 4583) and flash frozen on liquid nitrogen. 14 µm transverse sections were prepared using a cryostat (Leica, CM850 UV), starting 1mm into the bulbs (Figure 35). Successive sections were collected 36 µm apart and mounted onto adhesive coated slides (Newcomer Supply, 5070). Sections were dried overnight at room temperature and stored long term at -20°C.

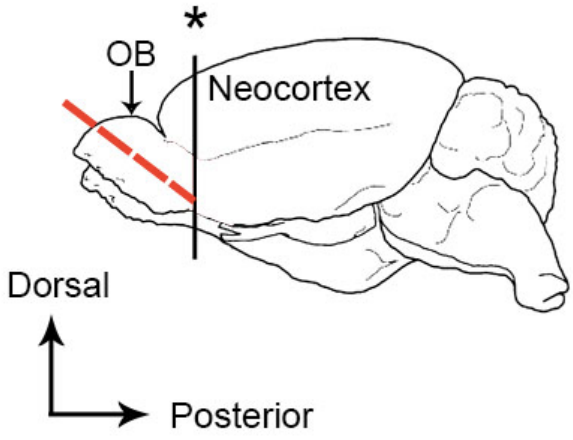


Figure 35. Olfactory bulb sectioning

Following dissection and fixation of whole brain, olfactory bulbs (OB) were dissected from the rest of the brain as indicated by the black line and asterisk (*) with the intention of preserving some of the neocortex to aid in sample orientation. Sections were prepared through the middle-most part of the olfactory bulb in the transverse plane (red dashed line), such that both medial and lateral sides of each bulb were included in a single section.

Immunofluorescence and click reaction

Prior to antibody labeling, slides were washed once for 15 min in 1X PBS to remove OCT. For tissue undergoing BrdU labeling, antigen retrieval in sodium citrate buffer was performed as described above. The following primary antibodies used: chicken anti-tyronsine hydroxylase (Aves Labs, TYH), rabbit anti-calretinin (Sigma, AB-5054), rabbit anti-calbindin (Sigma, C2724-2ML), mouse anti-NeuN (Millipore, MAB377), and sheep anti-BrdU (abcam, Ab1893-125). 100 µl of primary antibody was applied directly to slides. Secondary antibody was diluted with 2 µg/ml DAPI (ThermoFisher, D1306). The following secondary antibodies were used: donkey anti-chicken Alexa 488 (Jackson Immunoresearch, 703-545-155), donkey anti-rabbit Alexa 488 (ThermoFisher, A21206), and goat anti-mouse Alexa 555 (ThermoFisher, A21424), donkey anti-sheep Alexa 488 (ThermoFisher, A11015). All antibodies were diluted 1:500 in 1X PBS (Sigma, P4417-100TAB) with 0.1% triton X-100. All antibody incubations were performed overnight at 4°C in a humidified chamber, and slides were washed three time in 1X PBS following each round of labeling.

For tissue undergoing EdU labeling, a “click” reaction solution was prepared fresh and contained 100 mM TBS, 4 mM $CuSO_4$, 0.1 M sodium ascorbate and 4 mM Alexa 647 azide (ThermoFisher, C10340). Click solution was applied to the slides and slides were incubated for 30 min at room temperature in a humidified chamber. Following labeling, slides were mounted with a #1.5 glass coverslip using immu-mount (Shandon, 9990402).

Image acquisition and data collection

Images were collected using a Nikon Eclipse Ni microscope and C2+ confocal system (Nikon Instruments Inc.) and Nikon NIS-Elements software. Photomicrographs used for OB cell counting were collected using the Nikon Plan APO 20X/0.75 objective at 1024 image size, 5.3 pixel dwell and with a laser power of 15. The entire GL from transverse OB sections was imaged in a non-overlapping manner. 5 z-stacks were collected per field of view spaced 2.625 µm apart.

Cell counting was performed using ImageJ with the FIJI package [411]. For each image, the GL was defined as a region of interest (ROI) based on nuclear morphology. A threshold (IJ_IsoData) was taken for the EdU or BrdU labelling in the central-most z-stack and a selection was created of the 16-bit threshold image. The EdU/BrdU selection was overlaid onto the NeuN labelling, and all double positive cells were selected and counted. EdU+NeuN or BrdU+NeuN selections were overlaid onto the CalR, CalB or TH staining and triple positive cells were counted. Orthogonal view was used to confirm colocalization, and DAPI staining was used to confirm localization of EdU signal within nuclei.

Staining, image acquisition and cell counting were performed blind to the genotype. For EdU and BrdU labeling experiments, the number of triple positive cells was expressed as a proportion of the thymidine analog+NeuN cells. The total number of thymidine analog+NeuN cells and triple positive cells were also calculated per unit area. This was performed by calculating the GL ROI area for each image using ImageJ, and

then dividing the number of counted cells in each image by this number. Between 10 and 20 images were counted per staining condition (TH, CalR or CalB) per animal. Results were analyzed by 2-way ANOVA with Sidak's multiple comparisons.

4.4 Results

4.4.1 Identification and selection of miR-7 MREs for *in vivo* mutagenesis

Predicted miR-7 MREs in the Pax6 3'UTR

MicroRNA-7 is a promising candidate for regulating of *Pax6* in the brain and endocrine pancreas [206, 217, 274, 397]. These studies identified an MRE for miR-7 located at mouse *Pax6* 3'UTR position 655 as being involved in mediating miR-7 regulation of *Pax6*. Additionally, Zhao *et al*. (2012)[274] noted a second miR-7 MRE located at 3'UTR position 517. We wanted to perform a comprehensive scan of the mouse *Pax6* 3'UTR sequence to identify all candidate miR-7 regulatory elements in addition to those previously noted. We scanned 876 bp mouse *Pax6* 3'UTR sequence using the ImiRP MRE prediction tool [353] for predicted MREs for miR-7-5p and miR-7-3p. We chose to focus on 876 bp of the 3'UTR because our RNA sequencing (RNA-seq) and 3' rapid amplification of cDNA ends (3'RACE) data suggests that this is the primary location where the mouse *Pax6* transcript is cleaved and polyadenylated [400].

We identified nine predicted MREs for miR-7 in the 3'UTR of mouse *Pax6* (Table 6). Five of these were either 6mer MREs with a single G:U mismatch (6mer+G:U) or offset-6mer (OS-6mer) MREs, which are unlikely to be functional [35, 48, 50, 51, 62,

414] and were consequently not considered further. We chose to further investigate two predicted MREs for miR-7-5p at *Pax6* 3'UTR positions 517 and 655, and two predicted MREs for miR-7-3p at positions 626 and 758.

Table 6. Predicted MREs for miR-7 in the mouse *Pax6* 3'UTR

3'UTR position	microRNA	MRE type
250	miR-7-3p	OS-6mer
444	miR-7-3p	OS-6mer+G:U
517*	miR-7-5p	8mer+G:U
609	miR-7-5p	OS-6mer
626*	miR-7-3p	7mer-m8
651	miR-7-3p	OS-6mer
655*	miR-7-5p	7mer-m8
758*	miR-7-3p	7mer-m8
765	miR-7-5p	6mer+G:U

* Indicates MREs considered for further analysis

Next, we sought to determine the extent to which these four miR-7 MREs are conserved across orthologous vertebrate *Pax6* 3'UTRs. Three of these four predicted miR-7 MREs are conserved between orthologous mouse and human *PAX6* 3'UTR sequences and are well conserved across orthologous mammalian *Pax6* 3'UTRs (Figure 36A-B). Additionally, miR-7 MREs at mouse *Pax6* 3'UTR positions 626 and 655 were found to be well conserved across vertebrates. Particularly, the miR-7-5p MRE located at mouse *Pax6* 3'UTR position 655 is very highly conserved, suggesting that it may be the primary miR-7 MRE involved in regulation of *Pax6.* However, the miR-7 MREs located at 3'UTR positions 517 and 626 may be involved in mediating miR-7 regulation of *Pax6* in mammals and vertebrates, respectively.

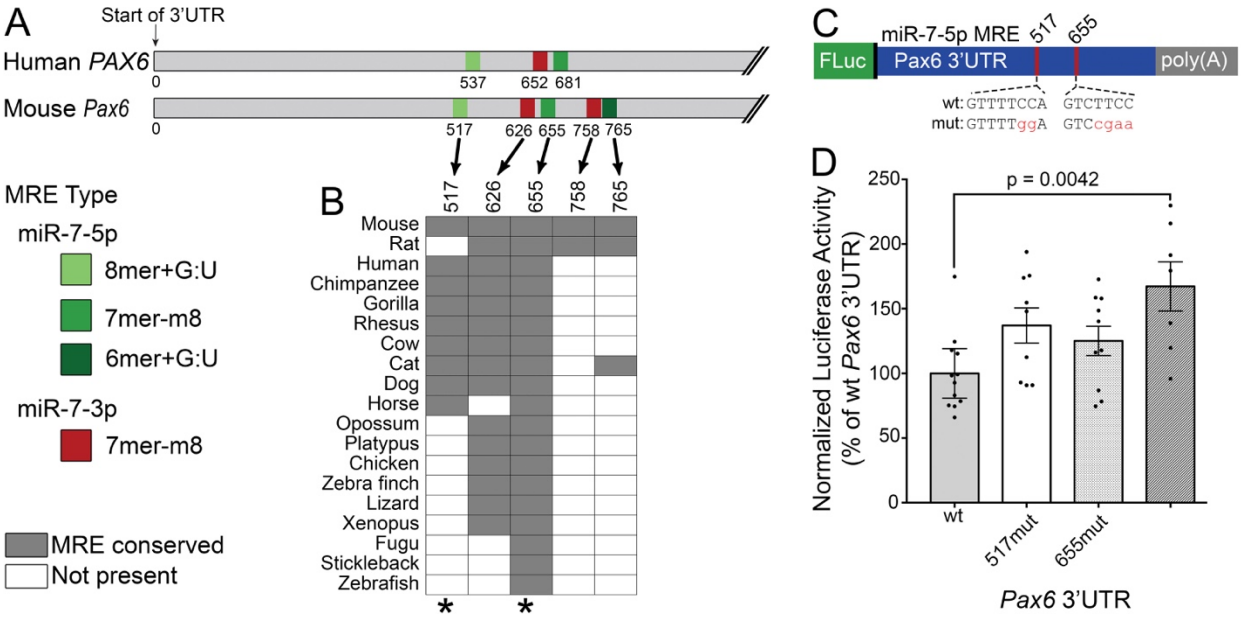


Figure 36. Identification and in vitro functional analysis of predicted *Pax6* 3'UTR miR-7 MREs

(A) Location of miR-7 MREs in the human and mouse *PAX6/Pax6* genomic region downstream of the stop codon predicted by ImiRP (4). Green and red indicate MREs for miR-7-5p and miR-7-3p, respectively. (B) Table showing conservation of mouse miR-7 MREs across vertebrates. Asterisks indicate MREs selected for *in vivo* mutagenesis. (C) Schematic of the firefly luciferase (FLuc) transcript containing the *Pax6* 3'UTR. Locations of MREs for miR-7-5p at 3'UTR positions 517 and 655 and mutations are shown. (D) Plasmid vectors were transiently transfected into αTC1-6 cells, and FLuc activity was normalized to an internal control (renilla luciferase) and expressed as a percentage of the wt 3'UTR control. +/- 95% CI. n=7-12. Statistics were performed using one-way ANOVA with Tukey's multiple comparisons test.

Expression and conservation of miR-7-5p relative to miR-7-3p

We next wanted to determine the relative abundance of miR-7-5p versus -3p with the goal of addressing their relative capacities to regulate *Pax6.* Both mature miR-7-5p and -3p are processed from different arms of the same pre-miR-7 RNA hairpin by Dicer (Figure 2). Dicer cleaves the pre-miRNA to produce a duplex consisting of the 5p and 3p mature miRNAs [65]. One strand of the duplex becomes the guide strand and is loaded into Argonaute (AGO). The other strand, termed the passenger strand, is

degraded. Given this, we suspected that processing of one arm of the pre-miR-7 hairpin may predominate.

To determine relative abundances of miR-7-5p versus -3p, we used miRbase, a database containing the results of many small RNA deep sequencing experiments from a variety of mouse tissue types [18]. miRbase data suggests that mouse miR-7-5p is more abundant and thus preferentially processed over miR-7-3p. Small RNA deep sequencing read counts show that mouse miR-7a-5p has 205117 reads per million from deep sequencing experiments, whereas miR-7a-1-3p has 881. These results suggest that miR-7-5p is approximately 200 times more abundant than miR-7-3p across several different mouse tissues. The extent to which a given miRNA represses its target mRNAs depends in part on relative the abundance of miRNA to target, with larger miRNA:target ratios being associated with greater repression [64]. Given this, the number of deep sequencing read counts for miR-7-5p relative to -3p suggests that miR-7-5p is playing a more important role in miRNA mediated repression of target gene transcripts.

In addition to its lower level of expression, miR-7-3p is also less well conserved relative to miR-7-5p. In humans and mice, miR-7 is encoded by three separate genes. Alignments of the pre-miR-7 sequences reveals that miR-7-5p is extensively conserved between miR-7 paralogs, whereas miR-7-3p shows less extensive conservation (Figure 37A-B). Additionally, the miR-7-5p sequence is perfectly conserved within animals, whereas orthologous mature miR-7-3p sequences are less well conserved (Figure 37C-D). Taken together, relative abundance and expression of miR-7-5p and -3p suggest that miR-7-5p plays a dominant regulatory role.

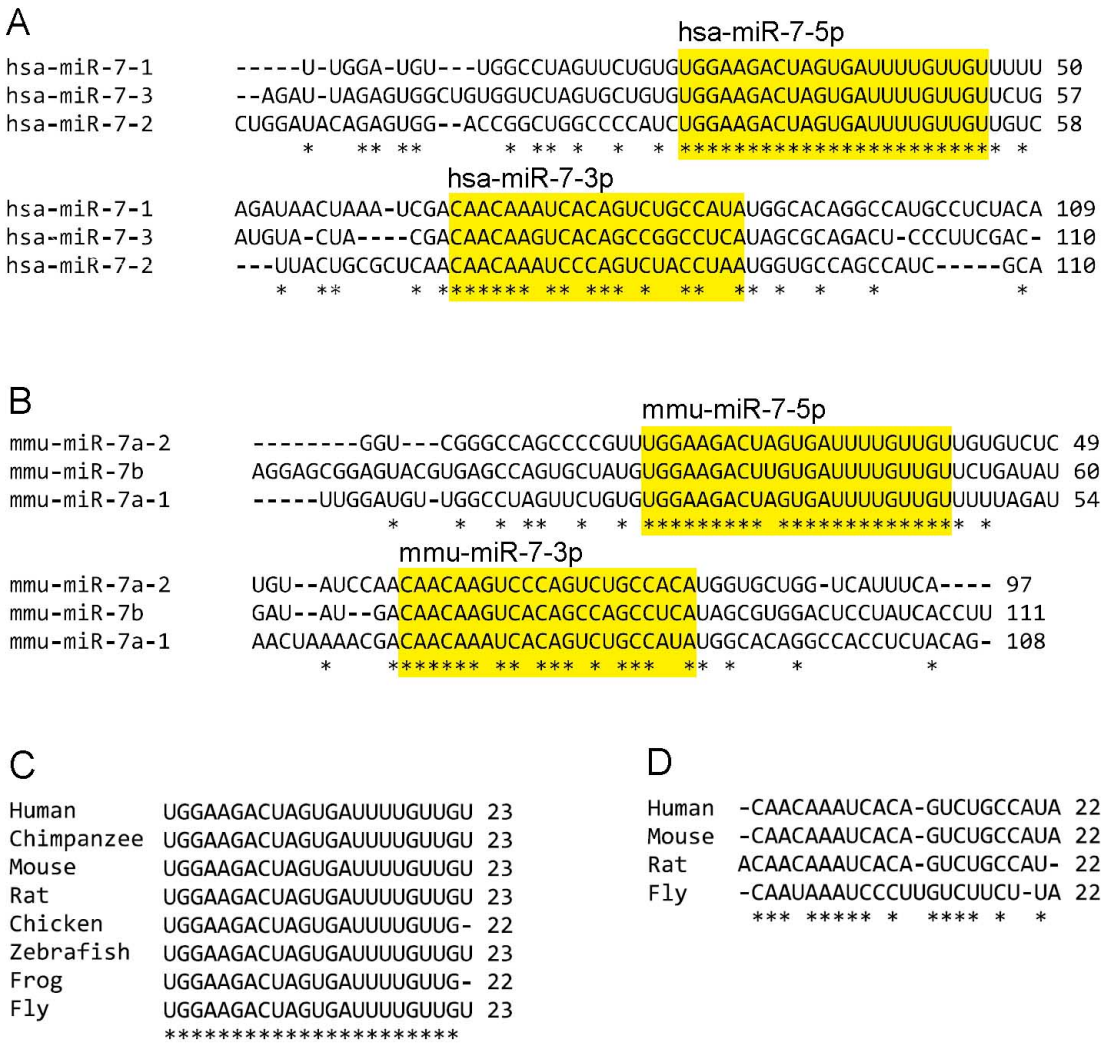


Figure 37. Conservation of miR-7-5p and miR-7-3p

(A) *Conservation of paralogous pre-miR-7 sequences in human.* The three miR-7 genes are named miR-7-1, -2, and -3. Yellow highlight is used to denote miR-7-5 and -3p mature miRNA sequences. (B) *Conservation of paralogous pre-miR-7 sequences in mouse*. In mouse, the miR-7 genes are named miR-7a-1, miR-7a-2, and miR-7b. (C) *Conservation of orthologous mature miR-7-5p* (D) *and miR-7-3p sequences.* MiR-7 is encoded by a single gene in fly and three different genes in vertebrates. Vertebrate sequences shown are for miR-7-1. The miR-7-3p sequence has not yet been identified experimentally in many species. All miRNA sequences were collected from miRBase release 21 [18] and alignments generated using ClustalW2.

Functional analysis of miR-7 MREs

Next, we sought to address the functionality of the identified miR-7 MREs *in vitro*. We generated plasmid vectors expressing a firefly luciferase reporter fused to the mouse *Pax6* 3'UTR. Mutations were introduced into the candidate MREs at 3'UTR positions 517, 626, 655 and 758 singly and in combination (Figure 36C, Figure 38A). We used the open source software tool ImiRP [353] to design mutants that did not introduce new predicted MREs for other miRNAs upon disruption of the desired MREs. WT *Pax6* 3'UTR and mutant 3'UTR luciferase plasmids were transiently transfected into αTC1-6 cells, which endogenously express miR-7. Disruption of the miR-7-5p MREs at positions 517 and 655 individually had a small impact on relieving repression of luciferase by miR-7. However, disruption of both together had a larger impact (Figure 36D).

Despite the low abundance of miR-7-3p relative to miR-7-5p [18], we wanted to address the functionality of the miR-7-3p MREs at *Pax6* 3'UTR positions 626 and 758 by luciferase assay. Disruption of either miR-7-3p MRE individually or in combination did not have a significant impact of luciferase activity relative the WT *Pax6* 3'UTR (Figure 38B). Interestingly, relative luciferase activity trended downward with disruption of the position 626 MRE, in opposition to the predicted de-repression. Additionally, disruption of all four MREs, for both miR-7-5p and -3p, in combination did not significantly elevate luciferase activity relative to WT. Together, these findings suggest that the miR-7-5p MREs at *Pax6* 3'UTR positions 517 and 655 function in an additive manner and that the miR-7-3p MREs at 3'UTR positions 626 and 758 do not mediate regulation of *Pax6*.

Given the lack of an effect associated with disruption of the miR-7-3p MREs *in vitro*, we chose to focus on the miR-7-5p MREs for *in vivo* analysis and did not investigate the miR-7-3p MREs further.

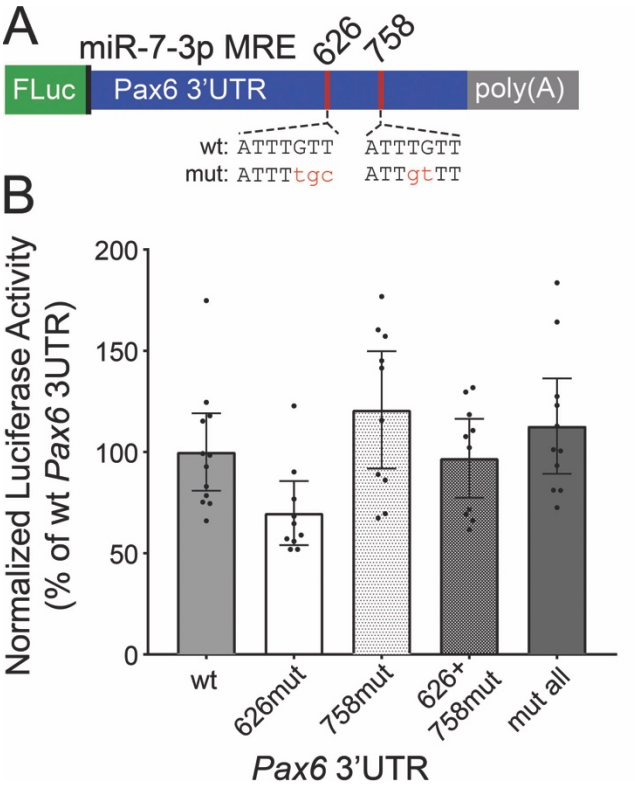


Figure 38. in vitro functional analysis of predicted *Pax6* 3'UTR miR-7-3p MREs

(A) Schematic of the firefly luciferase (FLuc) transcript containing the *Pax6* 3'UTR. Locations of MREs for miR-7-3p at 3'UTR positions 626 and 758 and mutations are shown. (B) Plasmid vectors were transiently transfected into αTC1-6 cells, and FLuc activity was normalized to an internal control (renilla luciferase) and expressed as a percentage of the wt 3'UTR control. "Mut all" refers to the *Pax6* 3'UTR containing mutations introduced into the positions 517 and 655 miR-7-5p MREs, and the positions 626 and 758 miR-7-3p MREs. +/- 95% CI. n=7-12. Statistics were performed using one-way ANOVA with Tukey's multiple comparisons test.

4.4.2 *in vivo* mutation of *Pax6* 3'UTR positions 517 and 655 miR-7-5p MREs

Generation of mice using CRISPR/Cas9 genome editing

To address the role of miR-7-5p regulation of *Pax6 in vivo*, we used CRISPR/Cas9 genome editing to generate mice harboring mutations in the miR-7-5p MREs at 3'UTR positions 517 and 655 (Figure 39A-D). We used the mutations that were introduced into the luciferase reporters (Figure 36C), which were designed to avoid introducing new MREs for other miRNAs. A total of three mouse lines were generated, two in which either position 517 or 655 miR-7-5p MRE were disrupted singly, and an additional line in which both MREs were disrupted in combination (Figure 39D).

Sequencing of the *Pax6* transcript was performed to ensure that only the desired mutations were introduced. The mouse *Pax6* cDNA sequence containing 13 exons was obtained from Ensembl (transcript Pax6-206, ENSMUST 00000111086.10). This transcript contains exons 0-12, plus exon 5a located between exons 5 and 6. *Pax6* has several 5'UTR variants resulting from the use of different promoters. *Pax6-206* is transcribed from P0 and contains exon 0 but not exons 1 or 2, which are transcribed from P1 [188]. Both transcript variants contain exon 3, and the coding region begins in exon 4 (Figure 39A).

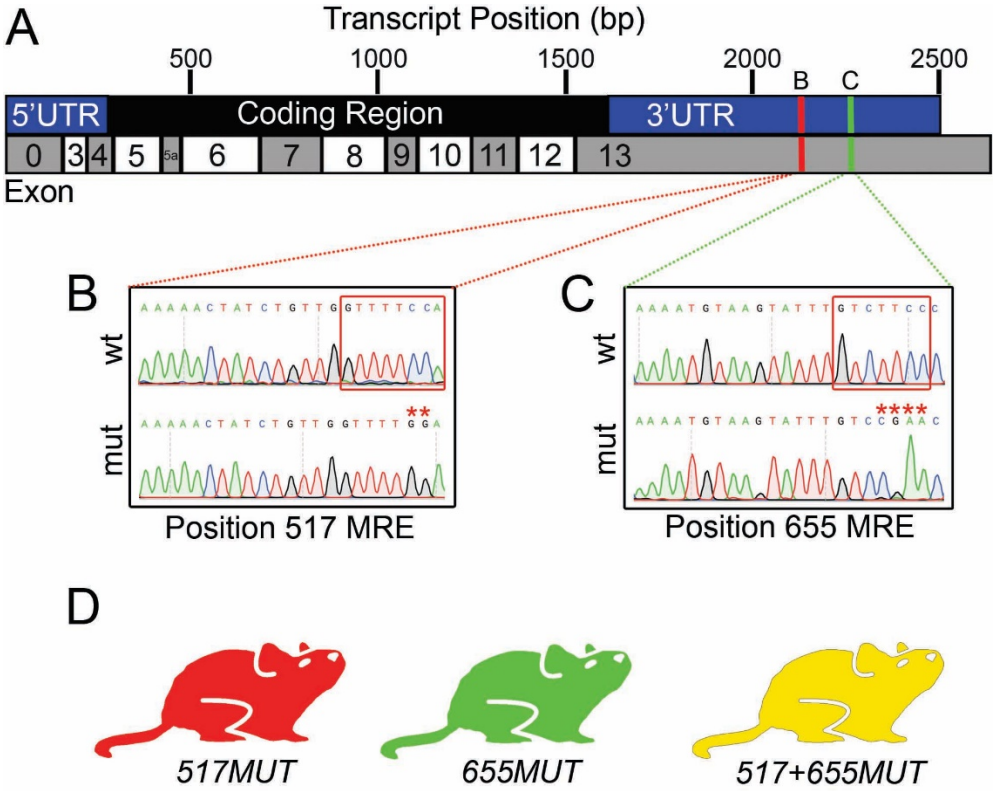


Figure 39. miR-7-5p MRE mutagenesis strategy

(A) Schematic representation of the mouse *Pax6* transcript (Ensembl, transcript Pax6-206). Locations of miR-7-5p MREs at positions 517 (B) and 655 (C) are shown. 21bp sequence of the wt *Pax6* 3'UTR representing the complete miR-7-5p MRE is shown, with nucleotides complementary to the miRNA "seed" indicated by red boxes and mutated nucleotides indicated by asterisks. ImiRP was used to design mutations such that new MREs for other miRNAs were not introduced, and mutants were generated using CRISPR/Cas9. Chromatograms show nucleotide sequences for wt (top) and homozygote mutants (bottom). (D) Schematic of mouse lines generated.

cDNA prepared from adult mouse retina was used to sequence exons 3-13 and genomic DNA (gDNA) was used to sequence the region of exon 13 containing the positions 517 and 655 miR-7-5p MREs. Mice harboring the desired miR-7-5p MRE mutations without off-target mutations in the *Pax6* transcript were backcrossed three times into the 129S1 background. Of note, sequencing data for exon 5a, a 42-bp

segment produced by alternative splicing [188], were not obtained. *Pax6* is more abundant than *Pax6(5a)* in embryonic mouse brain and eye [415], suggesting that *Pax6(5a)* may also be relatively less abundant than *Pax6* in the adult eye as well. Additionally, complete sequencing data were not obtained for the *Pax6* 5'UTR. The 5' end of exon 0 was incompletely sequenced and exons 1 and 2, 5'UTR variants resulting from alternative promoter use, were not sequenced. Overall, sequencing results reveal that only the intended miR-7-5p MRE mutations were introduced into the region of *Pax6* gene encoding canonical PAX6 and the 3'UTR by CRISPR/Cas9.

4.4.3 Impact of miR-7-5p MRE mutation on PAX6 levels

Following generation of the *Pax6* 3'UTR miR-7-5p MRE mutants, we wanted to address whether miR-7-5p MRE mutation alters levels of *Pax6* transcript and PAX6 protein. Since miRNA silencing is accomplished through both translational repression and mRNA destabilization [65, 66, 416], we predicted that levels of *Pax6* transcript and protein would be elevated upon loss of miR-7 regulation.

Gradient of Pax6 mRNA in the P1 V-SVZ

MicroRNA-7-5p may interact with the *Pax6* 3'UTR through the identified MREs and increase the rate of *Pax6* mRNA turnover and mutation the predicted *Pax6* 3'UTR miR-7-5p MREs is predicted to elevate relative levels of *Pax6*, particularly in double mutants. Quantification of *Pax6* transcript levels across the P1 V-SVZ axis was performed by dissecting the V-SVZ from 400 µm coronal sections and separating this region into D, DL and VL pieces (Figure 40A). RT-qPCR was used to quantify *Pax6*

transcript levels in the entire D and L V-SVZ, and in each region individually. *Pax6* transcript quantity is expressed relative to multiple reference genes, *Hprt* and *Tbp*. In WT animals, *Pax6* mRNA was expressed at 20% of dorsal levels in the dorsolateral V-SVZ and 8% of dorsal levels in the ventrolateral V-SVZ (Figure 40B). *Pax6* transcript quantity in the pooled dorsal and lateral V-SVZ regions was unaltered in *517+655MUT* animals relative to WT (Figure 40C). Plotting relative *Pax6* levels in the dorsolateral and ventrolateral regions of WT and miR-7-5p mutants reveals little change in *Pax6* transcript quantity (Figure 40D).

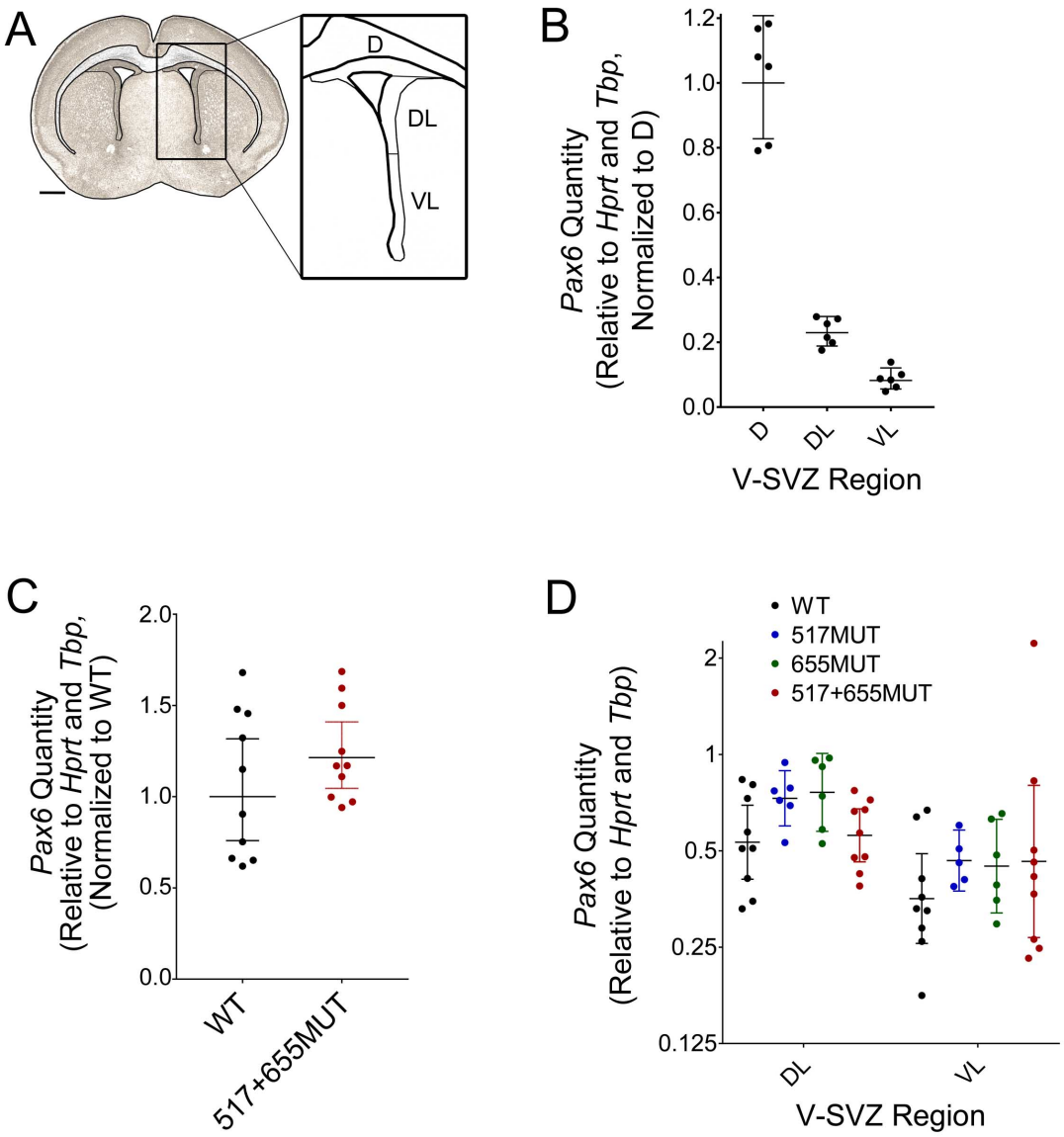


Figure 40. *Pax6* expression in the P1 V-SVZ with miR-7-5p MRE mutation

(A) Schematic of postnatal day 1 (P1) brain region harvested for *Pax6* mRNA quantification. V-SVZ was subdivided into dorsal (D), dorsolateral (DL) and ventrolateral (VL) as shown. Scale bar - 500 µm. (B-D) *Pax6* mRNA was quantified using SYBR green reverse transcriptase quantitative polymerase chain reaction (RT-qPCR). *Pax6* quantity is expressed relative to *Hprt* and *Tbp* internal control reference genes using Pfaffl's method with multiple reference genes [408]. (B) *Pax6* quantity in WT D, DL and VL V-SVZ normalized to D levels showing gradient in *Pax6* transcript. n=6 (C) Relative *Pax6* transcript quantity for WT and 517+655MUT genotypes in pooled D, DL and VL V-SVZ regions normalized to WT. n=10 (D) Relative *Pax6* quantity in DL and VL regions. n=9 for WT and 517+655MUT samples, and n=6 for 517MUT and 655MUT samples. Each data point represents 3 pups worth of pooled tissue. Data represents geometric mean +/- 95% C.I. D, dorsal; DL, dorsolateral; VL, ventrolateral

PAX6 protein gradient in the P1 V-SVZ

If miRNA-7-5p participates in regulating *Pax6* in the P1 V-SVZ, we also predict that PAX6 protein will be elevated with mutation of the predicted miR-7-5p MREs at 3'UTR positions 517 and 655. Specifically, we predict that mutation of either MRE on its own will have a small impact on PAX6 protein level, and that mutation of both simultaneously will produce a larger increase in PAX6 protein. PAX6 expression is lowest in the ventrolateral V-SVZ, where miR-7-5p is most abundantly expressed (Figure 28)[206]. Consequently, we predict that the level of PAX6 will be most impacted by disrupted miR-7-5p regulation ventrally.

The shape and extent of the PAX6 gradient in the V-SVZ was measured using immunofluorescence (IF). PAX6 IF was performed on coronal sections of P1 mouse brain, and fluorescence intensity was quantified in the dorsal VZ and along the length of the lateral VZ (Figure 41A). Mean pixel intensity for each VZ region was normalized to the dorsal region within each image. The PAX6 protein gradient for the dorsal-most half of the lateral VZ was similar between WT mice and the three mutants, with fluorescence intensity decreasing to approximately 25% of dorsal levels (Figure 41B). In the ventral-half of the lateral VZ, PAX6 fluorescence intensity declined to approximately 10% of dorsal levels in WT and single miR-7-5p MRE mutant animals but plateaued at 21% of dorsal levels in *517+655MUT* animals (Figure 41C). These results suggest that PAX6 protein is elevated approximately 2-fold in the ventral-lateral V-SVZ with disruption of both miR-7-5p MREs in the 3'UTR of *Pax6*, providing evidence that both miR-7-5p MREs are needed in combination for proper regulation of PAX6.

To address whether the observed elevation in PAX6 protein was an effect of the different mouse lines unrelated to the miR-7-5p MRE mutations, we compared the VZ PAX6 gradient between WT mice from the *517MUT*, *655MUT*, and *517+655MUT* lines (Figure 41D). The gradient of PAX6 IF is very similar between WT mice from the three miR-7-5p MRE mutant lines, providing validation that the observed increase in PAX6 IF in the ventral V-SVZ of *517+655MUT* homozygous mice represents a real change in PAX6 protein.

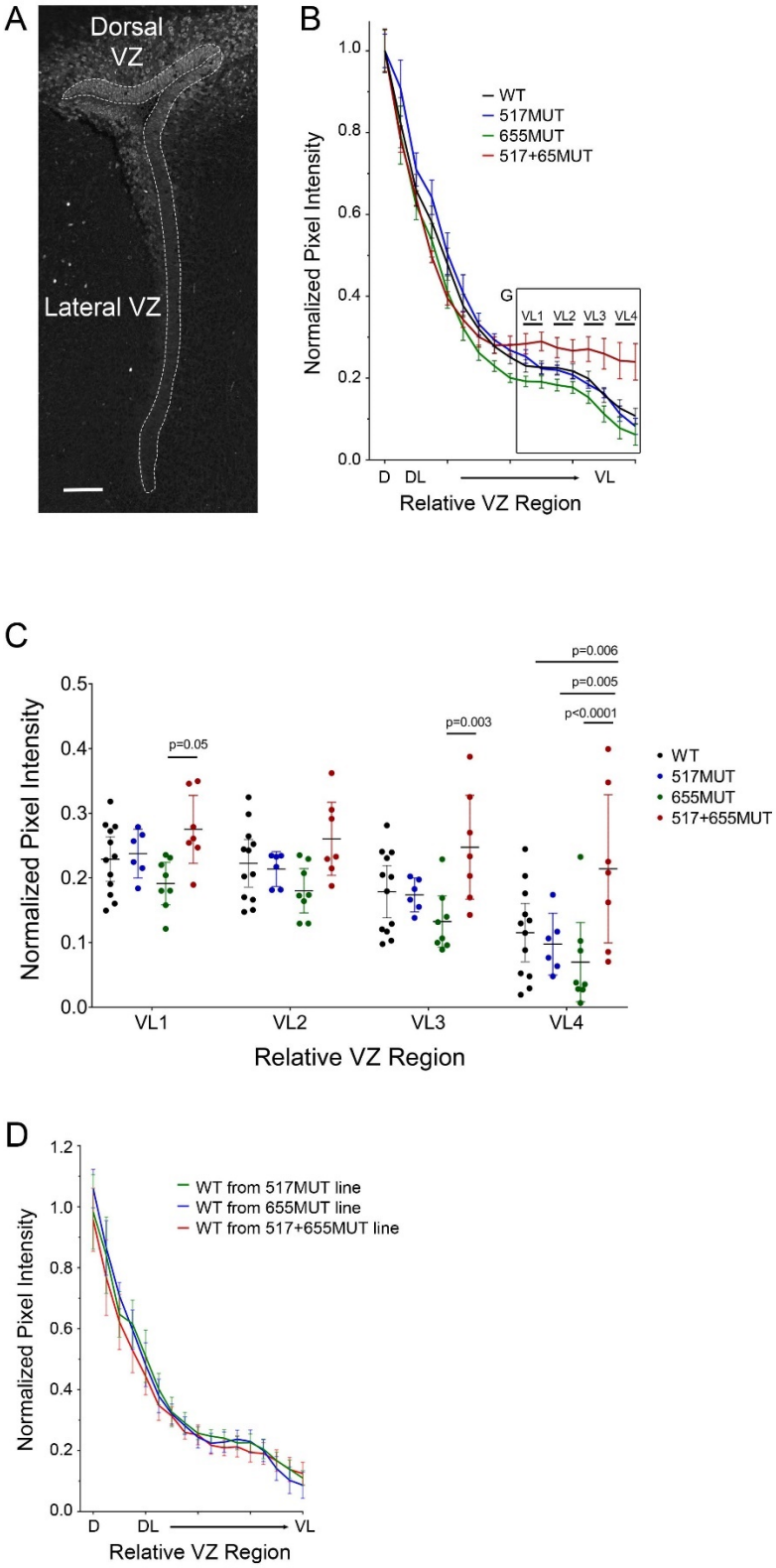


Figure 41. PAX6 immunofluorescence gradient in P1 V-SVZ with miR-7-5p MRE mutation

(A) Immunofluorescence (IF) for PAX6 was performed using P1 paraffin sections, and confocal microscopy was used to collect images. Dorsal and lateral VZ were defined by morphology, and mean pixel intensity of PAX6 IF was measured in the dorsal region and in 16 equal-area regions along the lateral length. Pixel intensity was averaged for 3-4 different LV images per animal and was normalized to the intensity of the dorsal-most region within each genotype. Scale bar – 100 μm. (B) Normalized pixel intensity for PAX6 IF across the length of the lateral VZ in WT (n=12), *517MUT* (n=6), *655MUT* (n=8) and *517+655MUT* (n=7) Mean ± SEM. (C) Normalized pixel intensity for PAX6 IF in the VL VZ. Two lateral regions are pooled for each VL1-4. Mean ± 95% C.I. (D) Normalized pixel intensity for PAX6 IF across the length of the lateral VZ compared between WT mice from the *517MUT*, *655MUT* and *517+655MUT* lines. n=4. Mean ± SEM. Data analyzed using 1-way ANOVA with Tukey's multiple comparisons. D, dorsal; DL, dorsolateral; VL, ventrolateral

Though *Pax6* is not implicated in the generation of OB interneurons derived from the medial VZ, we noticed dorsal-high to ventral-low expression of PAX6 in the medial VZ by IF. Quantification of PAX6 IF intensity in the medial VZ reveals remarkable similarity between the PAX6 gradients along the lateral and medial walls of the VZ. Additionally, mutation of the *Pax6* 3'UTR positions 517 and 655 miR-7-5p MREs in combination elevates PAX6 IF intensity in the mediolateral VZ similarly to the ventrolateral VZ. Though the specific role of PAX6 in the medial VZ has not been described, the gradient pattern of expression suggests that PAX6 may be important for specification of dorsomedial derived OB interneurons.

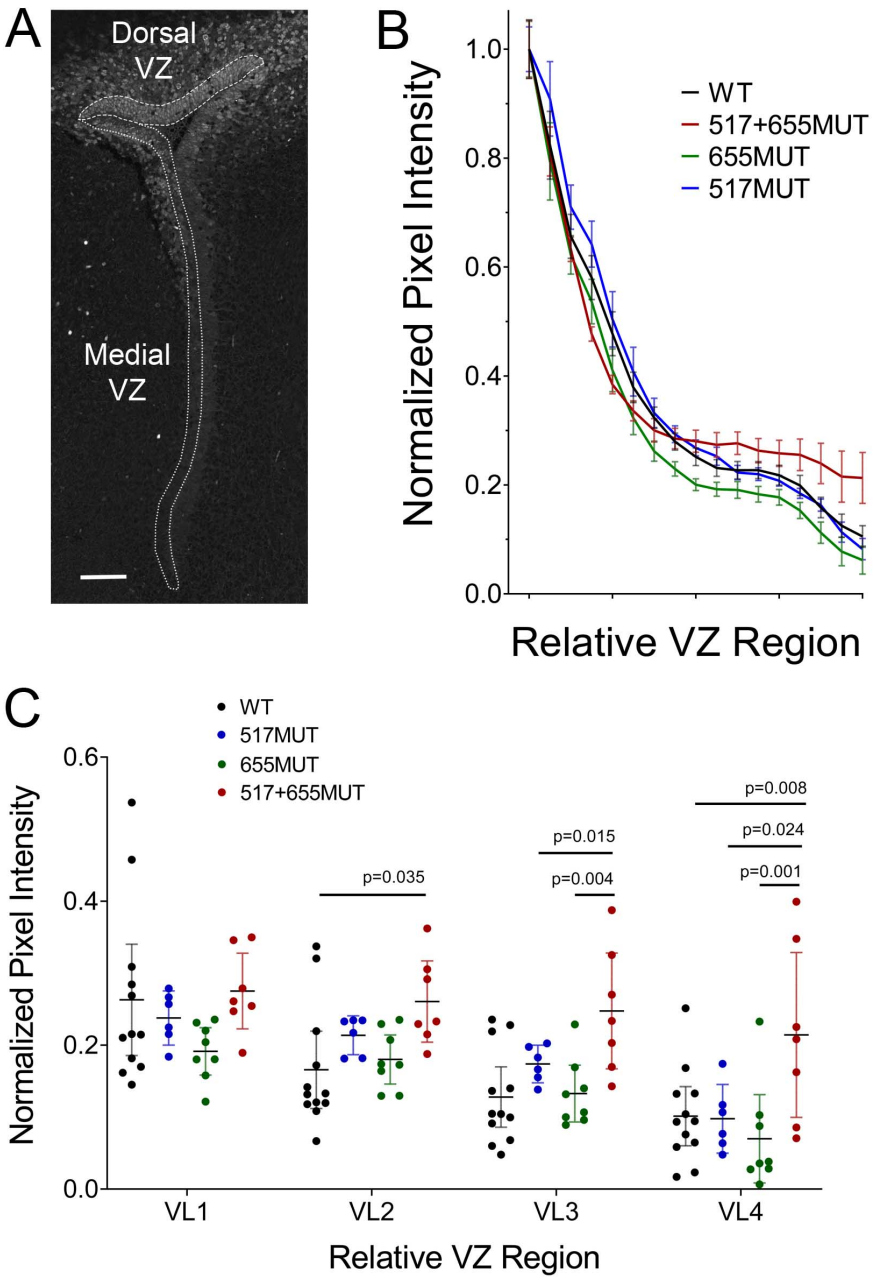


Figure 42. PAX6 protein in the P1 medial V-SVZ with miR-7-5p MRE mutation

(A) Immunofluorescence (IF) for PAX6 was performed using P1 paraffin sections, and confocal microscopy was used to collect images. Dorsal and medial VZ were defined by morphology, and mean pixel intensity of PAX6 IF was measured in the dorsal region and in 16 equal-area regions along the medial length. Pixel intensity was averaged for 3-4 different LV images per animal and was normalized to the intensity of the dorsal-most region within each genotype. Scale bar – 100 μm. (F) Normalized pixel intensity for PAX6 IF across the length of the medial VZ in WT (n=12), *517MUT* (n=6), *655MUT* (n=8) and *517+655MUT* (n=7) mice. Mean ± SEM. (G) Normalized pixel intensity for PAX6 IF in the ML VZ. Mean ± 95% C.I., 1-way ANOVA with Tukey's multiple comparisons.

There are several caveats associated with the IF approach to quantifying PAX6 protein in the P1 VZ. First, pixel intensity of PAX6 IF along the lateral wall of the VZ was normalized to dorsal levels. It is possible that PAX6 protein quantity is elevated in the dorsal VZ of miR-7-5p MRE mutants relative to WT, since miR-7-5p is present in this region. Consequently, the change in PAX6 protein associated with miR-7-5p MRE mutations in the *Pax6* 3'UTR may be larger than what was observed in Figure 41B and C.

Coronal P1 V-SVZ brain sections were prepared through the region of the brain featuring open, relatively straight lateral ventricles, approximately 1-1.5 mm into the brain starting at the rostral end (Appendix C: Figure 47A). However, sections were prepared through different positions along the rostral-caudal axis within this range (Appendix C: Figure 47B-D). During development of the cortex, PAX6 is expressed in a rostrolateral-high to caudomedial-low gradient in the VZ of the dorsal telencephalon [205]. It is therefore possible that PAX6 is also expressed in a rostral-caudal gradient in the P1 VZ in addition to its observed dorsal-ventral gradient. Consequently, there may be differences in PAX6 IF intensity associated with miR-7-5p MRE mutations at specific locations along the rostral-caudal axis. Sections taken from rostral-most positions generally have shorter LVs and a variety of LV lengths were observed for all genotypes (Appendix C: Figure 47A). Given how consistent the PAX6 IF gradient is between WT sections from the three mutant lines (Figure 41D) and between WT and single miR-7-5p MRE mutants (Figure 41B-C) variability introduced from rostral-caudal sectioning plane appears to be small.

Overall, we conclude that both miR-7-5p MREs at *Pax6* 3'UTR positions 517 and 655 appear to function cooperatively to regulate *Pax6*. Mutating either MRE singly has no effect of the intensity of PAX6 IF in the P1 V-SVZ relative to WT. However, mutation of both in combination produces an approximately two-fold increase in the intensity of PAX6 IF in the ventrolateral V-SVZ.

4.4.4 PGN phenotype in the main olfactory bulb with miR-7-5p MRE mutation

Given that our preliminary IF data suggests PAX6 protein is elevated in the V-SVZ of PAX6 3'UTR miR-7-5p MRE mutant mice, we next sought to address whether this increase in PAX6 protein is associated with a neuronal fate phenotype in the main olfactory bulb. To track the fate of the V-SVZ neural stem cells, we injected P1 pups with the thymidine analog 5-ethynyl-2'-deoxyuridine (EdU) to label cells in synthesis phase (S-phase) (Figure 26).

Given that PAX6 may be upregulated in the ventral-lateral V-SVZ of *517+655MUT* animals, we predicted that the number of TH-positive PGNs may be elevated as well. Previous evidence reveals that CalR and CalB-positive PGN generation is not impacted by either *Pax6* haploinsufficiency or overexpression of PAX6 in the lateral V-SVZ [206, 245], so we did not predict that these interneuron populations would be impacted by the *Pax6* 3'UTR miR-7-5p MRE mutations.

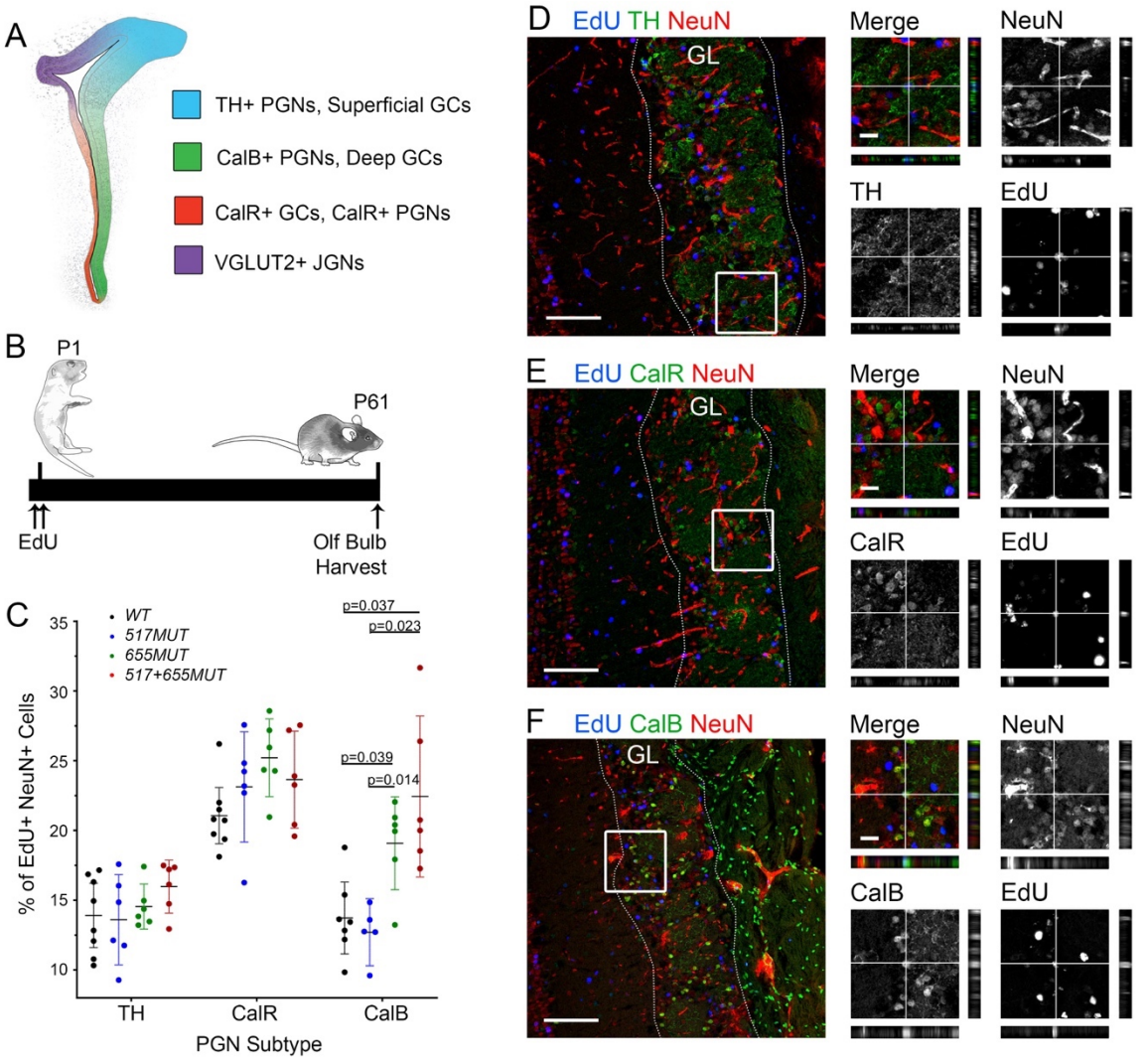


Figure 43. Periglomerular neuron phenotype in the olfactory bulb associated with *Pax6* 3'UTR miR-7-5p MRE mutation

(A) V-SVZ neural stem cells (NSCs) are spatially organized giving rise to distinct olfactory bulb (OB) interneuron subtypes [393]. Specifically, PAX6 is important for the specification of dopamine-producing tyrosine hydroxylase-positive (TH) OB interneurons [214]. (B) P1 pups were injected twice with 10 µg/g EdU to label V-SVZ stem cells in S-phase. OBs were harvested 60 days post-injection for sectioning, immunofluorescence (IF) and microscopy. (C) Quantification of EdU labelled TH, CalR and CalB-positive periglomerular neurons (PGNs) in the glomerular layer (GL) of the OB. Neurons were identified by NeuN IF. PGNs triple positive for EdU, NeuN and the subtype-specific marker were counted and expressed as a proportion of the total EdU+ NeuN+ double labelled cells. n=6-8, +/- 95% C.I., analyzed by mixed-effects two-way ANOVA with Sidak's multiple comparisons test. The proportion of counted PGNs differs significantly between genotypes ($p<0.0001$) and differs between PGN subtypes ($p<0.0001$). Specifically, the proportion of EdU and CalB positive PNGs may differ between WT and *655MUT* mice ($p=0.039$), WT and *517+655MUT* ($p=0.038$), *517MUT* and *655MUT* ($p=0.014$) and *517MUT* and *517+655MUT* ($p=0.023$) (D-F) Photomicrograph of P61 GL stained for TH, CalR or CalB (green), NeuN (red) and EdU (blue). Zoom shows colocalization of the three markers. Scale bar represents 100 µm and 30 µm for zoom panels. TH, tyrosine hydroxylase; CalB, calbindin; CalR, calretinin; PGC, periglomerular cell; GC, granule cell; VGLUT2, vesicular glutamate transporter; JGN, juxtaglomerular neuron.

V-SVZ NSC fate tracking and PGN phenotype

To track the fate of V-SVZ derived progenitors as they migrate and differentiate into mature main OB interneurons, we set up heterozygous crosses for each the *517MUT*, *655MUT* and *517+655MUT* mouse lines to generate WT and homozygous mutant littermates. We injected P1 mice twice with EdU, six hours apart, to label progenitors in S-phase (Figure 43B). OBs were harvested 60 days post-EdU injection to permit migration and differentiation of labeled cells. Notably, dopaminergic PGN differentiation peaks between 45 [392] and 90 [391] days post-labeling in adult mice and rats, respectively. We counted EdU positive PGNs labelled for different markers of distinct interneuron populations: TH and the calcium binding proteins CalR and CalB

[392, 395]. Though local stem cell populations exist in the rodent main OB, these local stem cells primarily contribute glia. The OB interneurons are primarily the contribution of V-SVZ-derived progenitors [417]. Thus, to ensure we were focused primarily on V-SVZ derived PGNs, we counted the glomerular cells positive for both EdU and the neuronal marker NeuN. It should be noted that TH+ PGNs represent two distinct interneuron populations based on coexpression with other markers and morphology [395, 418]. We did not differentiate between these subtypes.

We determined the proportion of the EdU+ PGNs that were also positive for the interneurons subtype markers TH, CalR and CalB. Unsurprisingly, the relative proportions of the three PGN subtypes analyzed differ from one another across all genotypes ($p<0.0001$), with 14.5±1.1% of EdU+ PGNs expressing the marker TH, 23.3±1.7% expressing CalR and 16.6±4.3 expressing CalB (Table 7). Similar relative proportions of PGNs expressing these markers were described previously [392]. Specifically, 13.9±2.8% of the EdU labeled PGNs were also immunopositive for TH in WT animals (Figure 43C-D). Contrary to predictions, disruption of the miR-7-5p MREs singly or in combination did not dramatically impact the proportion of TH positive PGNs generated. 16.0±1.8% of EdU+ PGNs were also positive for TH in *517+655MUT* animals. For WT animals, 21.1±2.4% of EdU labeled PGNs were positive for CalR (Figure 43C, E). This proportion may be slightly elevated to approximately 23-25% in single and double mutants. Finally, 13.7±2.8% of EdU+ PGNs colocalized with the marker CalB in WT animals. Surprisingly, this proportion may be elevated to 19.1±3.2% in *655MUT* animals and to 22.4±5.5% in *517+655MUT* animals. (Figure 43C, F). Generally, the proportions of

PGN subtypes analyzed differs between genotypes (two-way ANOVA, p<0.0001, Table 8). However, the presence of an interaction effect (two-way ANOVA, p=0.008, Table 8) suggests that the effect of genotype must be qualified based on which PGN subtype is being analyzed. Post hoc test (Sidak's) reveals a difference in the proportion of CalB positive EdU labeled PGNs between genotypes, with a specific differences between WT/*517MUT* and *655MUT* mice (p=0.039 and p=0.014, respectively) and between WT/*517MUT* and *517+655MUT* animals (p=0.0377 and p=0.023, respectively). In summary, miR-5p MRE mutations in the 3'UTR of *Pax6* may be associated with changes in the proportion of EdU labeled PGNs that also express markers for TH, CalR or CalB. Contrary to predictions, the proportion of TH positive PGNs does not differ between genotypes, though the proportion of CalB positive PGNs may be elevated with miR-7-5p MRE mutation.

Table 7. Summary of means for PGN cell counting experiment

	Genotype				
PGN marker	*WT*	*517MUT*	*655MUT*	*517+655MUT*	Mean
TH	13.9 ± 2.8	13.6 ± 3.1	14.6 ± 1.5	16.0 ± 1.8	14.5 ± 1.1
CalR	21.1 ± 2.8	23.1 ± 3.8	25.2 ± 2.6	23.6 ± 3.3	23.3 ± 1.7
CalB	13.7 ± 2.8	12.8 ± 2.0	19.0 ± 3.2	22.4 ± 5.5	17.0 ± 4.5
Mean	16.2 ± 4.2	16.5 ± 5.7	19.6 ± 5.3	20.7 ± 4.1	

Dark grey cells represent numbers of EdU+ PGNs expressing cell type specific markers (TH, CalR, CalB) as a proportion of the total number of EdU+ PGNs. Light grey cells are the two-way analysis of variance (ANOVA) means. Error represents standard deviations (S.D.).

Table 8. Two-way ANOVA summary for PGN cell counting experiment

Source of variation	Factor	Degrees of freedom	F-ratio	P-value
Factor A	PGN subtype	2	55.90	<0.0001
Factor B	Genotype	3	10.25	<0.0001
Interaction	PGN subtype x Genotype	6	3.22	0.0079

Two-way ANOVA was performed using GraphPad Prism 8.1.2. Mixed effects model was used because several values are missing from the CalB PGN subtype category.

Pax6 3'UTR miR-7-5p MRE mutations may alter proliferation of V-SVZ progenitors or newborn neuron cell death. To address this, we compared the number of EdU labeled neurons per mm^2 of the GL between WT and miR-7-5p MRE mutant mice two months post-EdU injection. For all genotypes, the GL contains approximately 300 EdU and NeuN double positive cells per mm^2 (Figure 44A). These results suggest that V-SVZ progenitor proliferation and differentiated PGN cell death are not impacted by mutation of the *Pax6* 3'UTR miR-7-5p MREs.

We also determined the number of EdU labeled neurons per mm^2 in the GL of the OB that were also positive for the markers TH, CalR or CalB. Given the variability in the numbers of EdU and NeuN double positive cells counted in the GL, we wanted to use this alternative analysis approach is to address whether apparent changes in the proportion of PGN subtypes between WT and mutant mice (Figure 43C) reflect changes in absolute cell numbers. Expressing the numbers of counted EdU and TH, CalR or CalB double labeled PGNs relative to GL area reveals little difference between genotypes (Figure 44B).

In summary, mutation of the *Pax6* 3'UTR miR-7-5p MREs is not associated with a measurable increase in the proportion of EdU labelled PGNs that are also positive for TH, or an increase in the absolute number of EdU and TH positive PGNs per mm^2.

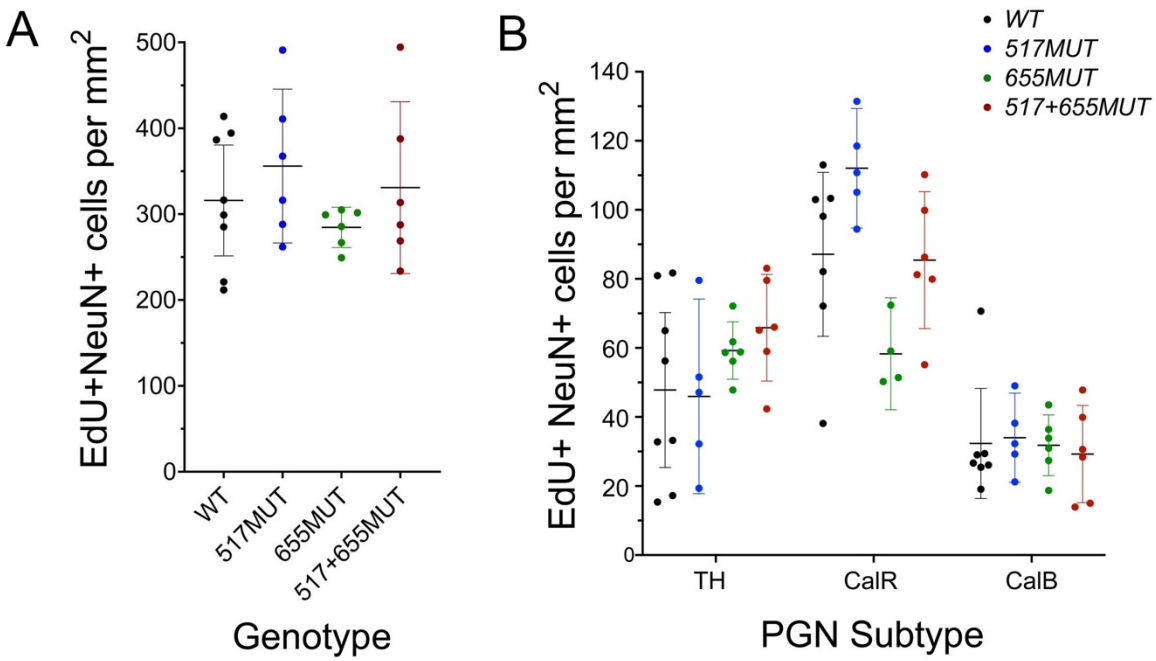


Figure 44. PGN cell numbers per mm2 in mice harbouring *Pax6* 3'UTR miR-7-5p MRE mutations

(A) Number of EdU positive neurons per mm^2 in the GL for WT and *Pax6* 3'UTR miR-7-5p MRE mutant mice. The GL of the OB was defined by morphology. (B) Number of EdU labelled TH, CalR and CalB-positive periglomerular neurons (PGNs) per mm^2 in the glomerular layer (GL) of the OB. Neurons were identified by NeuN IF. PGNs triple positive for EdU, NeuN and the subtype-specific marker were counted and expressed relative to GL area. n=6-8, +/- 95% C.I., analyzed by two-way ANOVA with Sidak's multiple comparisons test. The proportion of counted PGNs does not differ significantly between genotypes ($p=0.22$) but does differ between PGN subtypes ($p<0.0001$). Note incomplete data sets for the *655MUT* CalR condition and for the *517+655MUT* CalB condition. TH, tyrosine hydroxylase; CalB, calbindin; CalR, calretinin; PGN, periglomerular neuron. Data represents mean ± 95% C.I.

4.4.5 miRNA expression profile in the V-SVZ

Our results suggest that loss of miR-7-5p regulation of *Pax6* in the early postnatal V-SVZ may be insufficient to produce a dopaminergic neuronal fate phenotype in the olfactory bulb, in contrast to published results [206]. One explanation for this finding is that miR-7-5p may not be the only miRNA regulator of *Pax6* in the P1 V-SVZ. To address this, we performed RT-qPCR for miRNAs having predicted recognition elements in the mouse *Pax6* 3'UTR. First, we looked at expression of miR-7-3p in the P1 V-SVZ using miScript SYBR green assay. Next, we used TaqMan multiplex array cards to assay for many predicted miRNA regulators simultaneously.

In agreement with small RNA deep sequencing experiments [18], our qPCR data reveal that miR-7-5p is more abundant in P1 mouse V-SVZ relative to miR-7-3p (Figure 45B). MicroRNA-7-5p is approximately 10 to 100-fold more abundant than miR-7-3p in dorsal, dorsolateral and ventrolateral V-SVZ. Additionally, the ventral-high to dorsal-low gradient of miR-7-5p is more pronounced than the gradient of miR-7-3p. The relative level of miR-7-5p increases approximately 100-fold from the dorsal to ventrolateral V-SVZ, whereas the relative level of miR-7-3p only increases approximately 2-fold. These results further support the reasoning that miR-7-5p likely plays a more important role in regulating *Pax6* in comparison to miR-7-3p.

We assayed for 41 candidate miRNAs having conserved predicted MREs in the mouse *Pax6* 3'UTR. Of these, we identified 31 as being expressed in pooled dorsal and lateral P1 V-SVZ tissue (Figure 45C). Of these, three are expressed in a ventral-high to dorsal-low gradient: miR-7a-5p, miR-7b-5p and miR-675-3p (Figure 45D-G). These

results suggest that, in addition to miR-7a-5p, which has previously been implicated in negatively regulating *Pax6* in the ventral V-SVZ, miR-7b-5p and miR-675-3p may also be involved in regulation of *Pax6.* Additionally, in our miR-7-5p MRE mutants, which should disrupt binding of both miR-7a and miR-7b, regulation miR-675-3p may be able to compensate for the loss of miR-7-5p regulation. Notably, we identified three MREs for miR-675-3p in the 3'UTR of mouse *Pax6,* located at 3'UTR positions 175, 461 and 749 (Appendix B Table 10)[400]. It should be noted that miR-675-3p was detected at levels approximately 4-fold less than miR-7a-5p and 7.5-fold less than miR-7b (Figure 45C). In summary, many miRNAs predicted to regulate *Pax6* are expressed in the P1 V-SVZ and some of these may participate cooperatively with miR-7-5p to regulate *Pax6*.

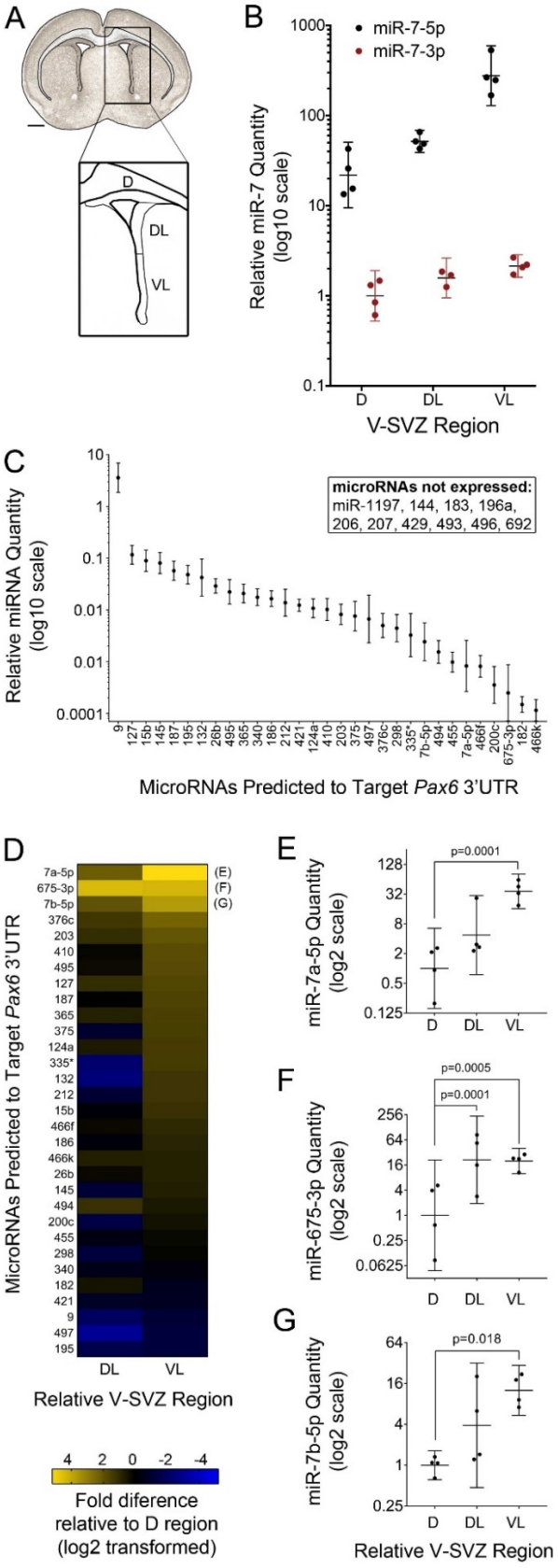


Figure 45. Expression profile of miRNAs predicted to target *Pax6* in the WT P1 V-SVZ

(A) Schematic of postnatal day 1 (P1) brain region harvested for miR-7 quantification. WT V-SVZ was subdivided into dorsal (D), dorsolateral (DL) and ventrolateral (VL) as shown. Scale bar - 500 µm. (B) miR-7-5p and -3p were quantified from WT P1 V-SVZ using miScript (Qiagen) reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) and were expressed relative to miR-26a-5p, miR-106a-5p and U6 as internal control reference genes using Pfaffl's method with multiple reference genes [408]. Relative miR-7 quantity was normalized to D miR-7-3p levels, revealing 10-100-fold greater quantities of miR-7-5p relative to miR-7-3p in all V-SVZ regions. (C) Relative abundance of miRNAs predicted to target the *Pax6* 3'UTR in WT P1 pooled D and L V-SVZ. Note miRNAs that were not detected. (D) Fold difference in expression of miRNAs predicted to target the *Pax6* 3'UTR in the DL and VL V-SVZ relative to D, with miR-7a-5p (E), miR-675-3p (F) and miR-7b (G) expressed in a D-low to V-high gradient. In log2 transformed data, 0 indicates no change in the level of expression relative to D. Yellow indicates elevated expression and blue indicates lower levels of expression relative to D. Data were analyzed by two-way ANOVA with Dunnett's multiple comparisons. miRNAs were quantified using TaqMan multiplex reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) array cards and were expressed relative to U6 snRNA, miR-16, miR-26a and miR-106a internal controls using Pfaffl's method with multiple reference genes [408]. n=4 for all. Each point represents 3 pups of pooled tissue. Data represents geometric mean +/- 95% C.I. D, dorsal; DL, dorsal-lateral; VL, ventral-lateral.

4.5 Discussion

4.5.1 Summary of findings

The functional role of miRNAs as regulators of *Pax6* has already begun to be characterized [206, 217, 272–279, 326, 327, 329, 331–333]. However, the direct biological relevance of *Pax6*-miRNA regulation *in vivo* is not fully understood. Here, we present evidence that miR-7-5p functions to refine the dorsal-high to ventral-low gradient of PAX6 in the early postnatal mouse V-SVZ by directly targeting two conserved MREs located at *Pax6* 3'UTR positions 517 and 655. Our results suggest that the consequence of miR-7-5p regulation of *Pax6* during PGN specification is likely subtle and that additional miRNAs may be important in shaping the PAX6 gradient in the V-SVZ.

4.5.2 Impact of miR-7-5p MRE mutation on levels of *Pax6* mRNA and PAX6 protein

MicroRNAs are known to impact levels of both targeted mRNA and protein, through a combination inhibition of translation initiation and mRNA destabilization [63, 65–67]. Given this, we predicted that disruption of the miR-7-5p MREs in the *Pax6* 3'UTR would be associated with an increase in *Pax6* transcript levels due to derepression. Specifically, we predicted that *Pax6* mRNA would be elevated in the ventrolateral V-SVZ, where the ratio of miR-7-5p to *Pax6* transcript is highest.

Though a clear dorsal-high to ventral-low gradient in *Pax6* transcript is visible in the P1 V-SVZ (Figure 25A), we did not observe a difference in *Pax6* mRNA quantity between genotypes by RT-qPCR (Figure 40C-D). There are several possible explanations for this observation. First, it is possible that miR-7-5p primarily functions to interfere with translation of PAX6 protein, without having a large impact on the *Pax6* transcript. Additionally, a single mRNA transcript could be used to synthesize multiple proteins. For example, in yeast, proteins tend to be 10-60,000-fold more abundant than their corresponding mRNAs [419]. Generally, qPCR can reliably detect relative differences in expression of two-fold or higher [420]. As such, small changes in mRNA that cannot be resolved by qPCR could translate to larger changes in corresponding protein.

Caveats associated with the experimental procedure and detection method may also explain the lack of an observed change in *Pax6* mRNA with mutation of the predicted miR-7-5p MREs. For relative qPCR analyses, the PCR reaction goes through 40 cycles and the DNA quantity is measured by interaction of a fluorescent molecule with

synthesized DNA. A cycle threshold (Ct) value is determined as the cycle number at which the fluorescence signal crosses a threshold of detection. Higher Ct values are associated with smaller initial template quantities [420]. Generally, Ct values between 15 and 30 can be measured reproducibly [420], and Ct values higher than 30 are associated with increased variability due to stochastic amplification and subsampling error [421]. *Pax6* mRNA is present at low levels in the ventral V-SVZ. Despite pooling tissue from three pups per sample and maximizing the quantity of RNA and cDNA used for RT and qPCR reactions, respectively, Ct values for the *Pax6* assay ranged between 30-32. The consequence of this is increased technical variability, which may mask genuine differences between genotypes [421]. Additionally, the dissection procedure is a possible source of random error. The V-SVZ is a thin strip of tissue, particularly in the ventral region, and PAX6 is not expressed in neighboring striatum (Figure 41E). Though dissections were performed with great care, the challenge associated with dissecting only the V-SVZ and not neighboring striatum could increase inter-individual variability. Laser-capture microdissection could be used as an approach to improve reproducibility of V-SVZ dissection [422]. These technical limitations may contribute to increased random variability, masking genuine differences in *Pax6* quantity between genotypes [421].

In addition to predicted changes in *Pax6* mRNA associated with mutation of the miR-7-5p MREs, we also predicted that PAX6 protein may be elevated in miR-7-5p MRE mutants. Specifically, we predicted that PAX6 quantity in the ventral V-SVZ may be elevated in miR-7-5p MRE mutants but that the dorsal V-SVZ would not be impacted.

Given that miR-7-5p and *Pax6* expression are anti-correlated along the D-V axis of the V-SVZ, PAX6 quantity is predicted to respond most sensitively to miR-7-5p regulation in the ventral region [64]. Additionally, we predicted that *Pax6* is regulated additively through the two miR-7-5p MREs. MicroRNAs are observed to function cooperatively when MREs are spaced 10-40 nt apart [35, 59, 138]. The spacing of the miR-7-5p MREs at 3'UTR positions 517 and 655 is consequently predicted to be too large to mediate cooperative regulation. In support of this, the *Pax6* 3'UTR luciferase reporter assay results suggest that the two miR-7-5p MREs are functioning additively (Figure 36D).

We used quantitative IF to measure PAX6 in the P1 VZ. Using quantitative microscopy, we were able to focus specifically on the VZ, which would be very challenging to dissect using conventional methods. We chose to focus on the VZ for several reasons. First, the P1 VZ consists entirely of radial glia-like neural progenitors, whereas the SVZ contains both transit amplifying cells and migrating neuroblasts [387, 388]. Additionally, VZ nuclei are very densely packed and have little cytoplasm (Appendix C: Figure 46B-C) whereas SVZ nuclei are less densely packed (Figure 26B). Given that PAX6 is a transcription factor and primarily detected in the nucleus (Appendix C: Figure 46C), we reasoned that focusing on the VZ would produce more reliable estimates of PAX6 quantity. Finally, the VZ is easily defined by nuclear morphology and its position adjacent to the lateral ventricle, whereas the SVZ-boundary is less clearly defined particularly in the ventral SVZ where PAX6 expression is lowest (Figure 41A).

Analysis of PAX6 protein quantity in the dorsal and lateral VZ suggest that miR-7-5p cooperatively regulates *Pax6* through the 3'UTR positions 517 and 655 MREs in the

ventrolateral but not dorsolateral VZ (Figure 41B-C). Contrary to our predictions, individual miR-7-5p MRE mutations are not associated with increased PAX6 IF intensity in the ventral VZ. Mutation of both MREs simultaneously is associated with an approximately two-fold increase in PAX6 IF intensity. Despite being more than 40 nt apart, the positions 517 and 655 miR-7-5p MREs appear to be functioning cooperatively to regulate PAX6 in the P1 VZ. One possible explanation for this finding is that miR-7-5p may be participating cooperatively with additional miRNAs having MREs spaced closely to either the 3'UTR positions 517 or 655 miR-7-5p MREs. Given the spacing of the miR-7-5p MREs, other MREs for coexpressed miRNAs located between *Pax6* 3'UTR positions 477-507, 536-557, 615-645 and 672-695 may be able to function cooperatively with miR-7-5p. Several miRNAs detected in the P1 V-SVZ (Figure 45C) have predicted MREs that reside within these ranges: miR-495, 376c, 200c, 182 and 466k. Notably, miR-376c has two predicted MREs located within cooperative distance of both miR-7-5p MREs, at 3'UTR positions 545 and 632 (Appendix B: Table 10). Consequently, cooperative repression of *Pax6* by miRNAs may persist with loss of either miR-7-5p MRE individually due to the presence of closely spaced MREs for other coexpressed miRNAs. Disruption of both miR-7-5p MREs simultaneously may be necessary to destabilize the association of other miRNAs with neighboring MREs and lead to derepression of PAX6.

Though PAX6 appears to be elevated in the ventral VZ of *517+655MUT* animals, PAX6 protein is still present in a steep D-V gradient in this region despite disrupted miR-7-5p regulation (Figure 41B). This observation suggests that additional layers of regulation are necessary for establishing the D-V gradient of PAX6. One possible

mechanism may involve regulation of *Pax6* by other miRNAs that function together with miR-7-5p. For example, we identified miR-675-3p as a miRNA expressed in a ventral-high to dorsal-low gradient along with miR-7a-5p and 7b-5p in the P1 V-SVZ (Figure 45D-G). Transcriptional regulation of *Pax6* is likely also involved in establishing the PAX6 gradient. As evidence for this, transgenic mice expressing GFP under the control of the human *PAX6* regulatory element show a gradient pattern of GFP expression in the P1 V-SVZ similar to PAX6, though less steep [206], suggesting that transcriptional control of *Pax6* is the primary method by which the D-V gradient is generated. The morphogen sonic hedgehog (Shh) may play a role in establishing the transcriptional pattern of *Pax6* in the P1 V-SVZ. Like the V-SVZ, PAX6 is also expressed in a dorsal-high to ventral-low gradient during spinal cord development, and Shh produced by the notochord adjacent to the ventral neural tube regulates *Pax6* expression [207]. Additionally, Shh signaling in the ventral V-SVZ of adult mice is important for specification of ventrally derived CalB OB interneurons [423]. Taken together, miR-7-5p may be important for refining the PAX6 gradient in the V-SVZ, though other regulatory mechanisms are also required for generating the PAX6 gradient.

PAX6 IF intensity serves as a proxy for PAX6 quantity. However, a two-fold increase in intensity may not represent a corresponding two-fold increase in protein. Multiple secondary antibodies can interact with a single primary antibody resulting in signal amplification. Though useful for detecting low quantity antigens, this means that differences in fluorescent signal between genotypes may not accurately represent changes in the protein of interest. Additionally, IF is associated with inter-sample

variability in signal and background intensity. Consequently, internal normalization is required when comparing many images to one another using a quantitative IF strategy. We chose to normalize fluorescence intensity in the lateral VZ to the dorsal-most region (Figure 34), where we predict PAX6 quantity will be unaffected by loss of miR-7-5p regulation. However, miR-7-5p is expressed dorsally (Figure 45 D, E, G), albeit at lower levels than in the ventral V-SVZ, suggesting that PAX6 could be elevated in the dorsal VZ of miR-7-5p MRE mutants as well. To address these limitations, alternative strategies could be used to quantify PAX6 protein in the V-SVZ between WT and miR-7-5p MRE mutants, such as quantitative mass-spectrometry and western blot.

4.5.3 Absence of a DAergic PGN phenotype in *Pax6* 3'UTR miR-7-5p MRE mutant mice

PAX6 is important for specification of the dopaminergic PGNs in the OB [206, 214, 215, 245] and miR-7 LOF in the lateral V-SVZ is associated with increased dopaminergic PGN production [206]. Given this, we predicted that derepression of PAX6 in *517+655MUT* mice would be associated with increased TH positive PGN generation. Contrary to these predictions, we observed little change in the proportion of EdU labeled PGNs positive for TH in miR-7-5p MRE mutants relative to WT (Figure 43C). There are several possible explanations for this observation.

EdU labels all V-SVZ progenitors in S-phase

First, the vast majority of dopaminergic (DAergic) PGNs are generated from dorsal VZ progenitors. Electroporation of dorsal progenitors with a GFP-expressing vector reveals that approximately 20% of these progenitors give rise to PGNs, half of

which are positive for TH. Conversely, approximately 10% of lateral progenitors give rise to PGNs, one third of which express TH [206]. Elevated PAX6 in the ventrolateral VZ associated with *Pax6* 3'UTR miR-7-5p MRE mutation may result in increased DAergic PGN production from the ventrolateral VZ wall. However, since the lateral V-SVZ overall contributes a small proportion of TH positive PGNs to the OB relative to the dorsal V-SVZ, elevated DAergic production from the lateral wall associated with increased PAX6 may be masked by the contribution from the dorsal V-SVZ. This may explain the small but statistically insignificant increase in the proportion of EdU positive PGNs also expressing TH in *517+655MUT* mice (Figure 43C).

Our choice to use EdU for V-SVZ fate tracking is associated with some limitations that may explain differences in our results compared to previous observations made by de Chevigny *et al*., (2012)[206]. EdU is incorporated into all V-SVZ and OB progenitors in S-phase, and consequently labels a heterogeneous population of cells. To avoid counting locally proliferating OB glia [417], we focused on PGNs expressing the neuronal marker NeuN. However, the V-SVZ is composed of radial glia-like progenitors, transit amplifying cells and migrating neuroblasts, all of which undergo cell divisions and may be labeled with EdU (Figure 26)[390]. An additional limitation of using EdU for V-SVZ fate tracking is that EdU incorporated into the DNA is diluted with each cell division. Consequently, cells that undergo several rounds of division after taking up EdU will be less strongly labeled than neuroblasts labelled in their final round of division prior to differentiating into OB interneurons.

Interestingly, we observed many EdU labeled cells in the GL that were positive for CalR or CalB but not NeuN (Appendix C: Figure 48A-B). In WT animals, approximately 62% of EdU and CalR double positive cells were also positive to NeuN, and 37% were NeuN negative. Conversely, approximately, 29% of EdU and CalB double positive cells were positive for NeuN, and 71% were negative for NeuN (Appendix C: Figure 48C). Though we do not know whether these NeuN- cells are neurons, there are several examples of neurons known to be NeuN negative, such as Purkinje cells of the cerebellum, mitral cells of the olfactory bulb and photoreceptor cells of the retina [424]. Immunolabeling for glial fibrillary acidic protein (GFAP) could be used to confirm whether these NeuN negative cells are glia. Whether these NeuN negative CalR or CalB positive cell populations differ in *Pax6* 3'UTR miR-7-5p MRE mutants is unknown.

Electroporation and viral transduction can also be used to label NSCs in specific spatial domains within the V-SVZ [206, 393]. However, since EdU labeling was used to track the fate of V-SVZ progenitors, all V-SVZ progenitors in S-phase were labeled and we could not focus on specific V-SVZ positional domains. Since most DAergic PGNs are derived from the dorsal wall and dorsal PAX6 levels are unlikely to be impacted by loss of miR-7-5p regulation, a small increase in the proportion of TH positive PGNs generated from the lateral wall may be masked by dorsally derived cells. To address this, we plan to electroporate a GFP expressing vector into the P1 lateral VZ wall to track fate the fate of progenitors derived specifically from this region in miR-7-5p MRE mutants.

MicroRNA-7 interaction with the miR-7-5p MRE mutant Pax6 3'UTR

Though PAX6 may be upregulated in the ventral VZ, this increase may be insufficient to produce a DAergic fate specification phenotype. As mentioned previously, the D-V gradient of PAX6 is still present in *Pax6* 3'UTR miR-7-5p MRE mutants. This may be the result of transcriptional regulation or regulation of *Pax6* by other ventrally expressed miRNAs. Another possibility is that miR-7-5p is continuing to regulate *Pax6* through unmutated MREs. In addition to the predicted miR-7-5p MREs at 3'UTR positions 517 and 655, miR-7-5p is also predicted to interact with an OS-6mer MRE at 3'UTR position 609 and a 6mer+G:U MRE at position 765. These weaker MREs could be providing redundant miR-7-5p mediated repression of *Pax6*. To address whether miR-7-5p is still capable of interacting with the *Pax6* 3'UTR in *517+655MUT* mice we can use an affinity purification strategy called miR-CATCH [405] to assay for endogenous mRNA-miRNA interactions. This approach requires binding biotinylated oligonucleotide against *Pax6* to magnetic streptavidin-coated beads. Paraformaldehyde (PFA) fixed cell lysate is then applied to the beads to affinity purify *Pax6* along with any interacting miRNAs. Following pulldown, RT-qPCR can be used to assay for *Pax6* and miRNAs of interest (Figure 32).

We tested three candidate oligonucleotides for their ability to affinity purify *Pax6* transcript from unfixed αTC1-6 cell lysate. Oligonucleotides 2 and 3 showed high affinity for *Pax6* relative to a non-specific negative control oligonucleotide (Appendix C: Figure 49A). Additionally, the non-specific target Hprt was not selectively purified using oligonucleotides against *Pax6* (Appendix C: Figure 49B). Oligo 2 was selected for affinity

purification experiments because it is not predicted to recognize any off-target transcripts by nucleotide basic local alignment search tool (BLAST) search, it yielded approximately 100-fold enrichment of *Pax6* relative to total RNA and did not enrich for the non-specific target *Hprt*.

Unfortunately, given the low level of *Pax6* expression in the V-SVZ, particularly in the ventral region, performing miR-CATCH from P1 V-SVZ is not feasible (data not shown). As an alternative, we plan to perform *Pax6* affinity purification from fixed E12.5 dorsal telencephalon. *Pax6* is expressed in radial glia-like progenitors of the VZ but not in differentiated pyramidal neurons during cortical development. VZ thickness increases between E10 and E12.5, after which point glutamatergic pyramidal neuron production peaks and VZ thickness decreases [204]. The quantity of *Pax6* expressing tissue can be maximized by collecting developing cortex at E12.5. Preliminary results suggest that the miR-CATCH affinity purification approach can be used to purify *Pax6* mRNA from E12.5 cortex (Appendix C: Figure 49C). Additionally, miR-7a-5p is detected in the VZ of the developing cortex by northern blot, RT-qPCR and in situ hybridization [425], suggesting that this tissue can be used to address whether miR-7-5p interaction with the *Pax6* 3'UTR is impacted by endogenous mutation of miR-7-5p MREs. We predict that miR-7-5p purification with the *Pax6* transcript will be reduced with disruption of either the position 517 or 655 miR-7-5p MREs singly relative to WT, and that mutation of both in combination will have a larger impact on interaction of miR-7-5p with the *Pax6* 3'UTR.

MicroRNA expression in miR-7-5p MRE mutants relative to WT

Given the lack of an observed dopaminergic PGN phenotype in our miR-7-5p MRE mutant mice, we also hypothesized that upregulation of PAX6 protein levels with miR-7-5p MRE mutation may be blunted by a corresponding increase in the expression of miRNA regulators. A proposed function of miRNAs within gene regulatory networks is to provide stability and robustness by participating in feedback and feedforward network motifs with their targets [107]. We speculated that PAX6 may participate in a transcriptional feedback network with miRNAs involved in directly regulating of *Pax6*. Consequently, increases in PAX6 protein may result in a corresponding increase in the expression of miRNAs which then feedback onto and repress *Pax6*.

To address this, we compared the level of several miRNAs predicted to interact with the *Pax6* 3'UTR between WT and miR-7-5p MRE mutant mice by RT-qPCR. We found that miR-7-5p expression in the P1 V-SVZ of the *Pax6* 3'UTR miR-7-5p MRE mutant mice was unaltered relative to WT (Appendix C: Figure 50A-B). Using TaqMan qPCR multiplex array cards, we identified miR-145-5p and miR-375-3p as being greater than two-fold differentially regulated in 517+655MUT P1 V-SVZ relative to WT (Appendix C: Figure 50C). Contrary to our predictions, miR-375-3p is downregulated in mutants relative to WT. However, when we attempted to confirm this finding using individual miRNA qPCR reactions and larger sample sizes, we found that the change in levels of miR-145-5p and miR-375-3p in the miR-7-5p MRE mutants observed by multiplex qPCR were likely false-positives (Appendix C: Figure 50D-E). We conclude that a miRNA-mediated compensatory feedback approach is unlikely to explain the observed

absence of a dopaminergic PGN phenotype in our miR-7-5p MRE mutant mice. It is, however, possible that other miRNAs that were not selected to assay for are playing an important role.

"Many-to-Many" regulation by miRNAs

The "many-to-many" principle of miRNA regulation may explain the lack of an obvious DAergic phenotype associated with specifically disrupting miR-7-5p regulation of *Pax6* (Figure 5). MicroRNA-7 LOF in the P1 lateral VZ is associated with increased TH positive PGN generation in the OB [206] presumably through regulation of *Pax6*. However, other transcription factors are known to be important for DAergic cell specification embryonically and postnatally, and others are known to be involved in forebrain patterning and are expressed in specific spatial domains along with *Pax6* in the early postnatal V-SVZ. These transcription factors may also be regulated by miR-7-5p to produce coherent cell fate decisions and phenotypes.

The embryonic lateral ganglionic eminence (LGE) is a source of OB interneurons embryonically [426] and gives rise to the lateral V-SVZ postnatally [388]. Embryonically, *Pax6* is expressed at high levels in the dorsal telencephalon and expressed at lower levels in the LGE [203]. Several transcription factors are also expressed in a similar domain (Table 9), Tbr1, Emx1 and Lhx2, suggesting that they are also involved in forebrain patterning [388, 427]. Other transcription factors are expressed in the LGE and are important for DAergic OB neurogenesis embryonically: Dlx1/2, Gsh1/2, Arx and Sp8 [426, 428–434]. Some of these transcription factors, such as Arx, Gsh1/2, Sp8, continue

to be expressed in the postnatal V-SVZ and DAergic PGNs of the olfactory bulb [388, 433, 434].

A number of transcription factors are important for differentiation of DAergic neurons postnatally in addition to PAX6 [214, 324] (Table 9), those being : Dlx2, Id2, Klf7, ER81 and Meis2 [315, 435–438]. Specifically, Meis2 interacts with PAX6 to promote DAergic specification of V-SVZ progenitors [315]. PAX6 also controls expression of CryαA, which is important for DAergic PGN survival [215].

Many of these transcription factors are predicted targets of miR-7-5p and miR-7-3p (Table 9). For example, the 3'UTRs of *CryαA*, *Gsh2*, *Sp8*, *Emx1* and *Klf7* contain predicted MREs for miR-7-5p, with *Emx1* containing three predicted miR-7-5p MREs. *Dlx2, Sp8, Tbr1, Lhx2, Id2, Klf7, Er81 and Meis2* are predicted targets of miR-7-3p. MicroRNA-7 may be involved in regulating many different targets important for OB and more specifically DAergic PGN specification. This may explain the more dramatic DAergic PGN phenotype observed by de Chevigny *et al.*, (2012)[206] following electroporation of a miR-7a sponge into the P1 lateral V-SVZ. MicroRNA-7 LOF is predicted to alter global miR-7 expression in electroporated cells, likely impacting many downstream targets of miR-7.

Table 9. Transcription factors involved in brain development and important for DAergic PGNs that may be targeted by miR-7

TF	miRNA	Type	Position	Role
Dlx1	NA			Expressed in the lateral adult V-SVZ. Important for the formation of GABAergic interneurons in the OB. *Dlx1/2*$^{-/-}$ mice lack GABA-expressing cells and TH+ cells (DAergic PGNs are also GABAeric).
Dlx2	7-3p	7mer-m8	932	Expressed in the embryonic LGE, lateral adult V-SVZ and PGNs. Important for the formation of GABAergic interneurons in the OB. *Dlx1/2*$^{-/-}$ mice lack GABA-expressing cells and TH+ cells (DAergic PGNs are also GABAeric). Required for formation of DAergic PGNs during embryonic development with *Dlx2*$^{-\backslash-}$ mice lacking TH+ PGNs. Interacts with PAX6 to promote DAergic PGN specification. Overexpression in the RMS promotes TH+ PGN formation at the expense of CalR+ PGNs.
Gsh1/ Gsx1	NA			Expressed in the lateral adult V-SVZ. Important for the formation of DAergic PGNs during the embryonic development. *Gsh1/2*$^{-/-}$ embryos lack TH+ PGNs.
Gsh2/ Gsx2	7a-5p/ 7b-5p	7mer-A1	48	Coexpressed with PAX6 in the dorsal LGE of the embryonic telencephalon. Expressed in the lateral adult V-SVZ. Important for the formation of DAergic PGNs during the embryonic development. *Gsh1/2*$^{-/-}$ embryos lack TH+ PGNs.
Arx	NA			Coexpressed with TH+ PGNs in the OB. Required for formation of DAergic PGNS. *Arx*$^{-/-}$ mice lack TH+ PGNs.
Sp8	7a-5p/ 7b-5p	7mer-A1	874	Expressed in the dorsal LGE and postnatal V-SVZ.
	7-3p	7mer-A1	2528	Coexpressed with differentiated PGNs including the TH+ PGNs. Reduced TH+ PGNs in embryonic OB with conditional *Sp8* deletion in the LGE.
	7-3p	7mer-m8	4276	

Tbr1	7-3p	7mer-m8	973	Expressed in similar domain as PAX6 in the developing telencephalon, and early postnatal/adult V-SVZ
	7-3p	7mer-A1	1295	
	7-3p	7mer-A1	1529	
Emx1	7a-5p/ 7b-5p	7mer-A1	1333	Expressed in similar domain as PAX6 in the developing telencephalon, and early postnatal/adult V-SVZ
		7mer-A1	1621	
		7mer-m8	1926	
Lhx2	7-3p	7mer-A1	439	Expressed in similar domain as PAX6 in the developing telencephalon
Id2	7-3p	7mer-A1	87	Important for DAergic PGN development in the OB. TH+ PGNs are dramatically reduced in *Id2*$^{-/-}$ mice.
Klf7	7-3p	8mer	1430	Important for DAergic PGN development in the OB. TH+ PGNs are dramatically reduced in *Klf7*$^{-/-}$ mice.
	7a-5p/ 7b-5p	7mer-A1	5480	
	7-3p	7mer-A1	6118	
	7-3p	7mer-m8	6526	
ER81/ Etv1	7-3p	7mer-m8	335	Expressed in DAergic PGNs. Required for Th expression in rodents.
	7-3p	7mer-m8	556	
	7-3p	7mer-A1	3578	
CryαA	7a-5p/ 7b-5p	7mer-m8	235	Downstream target of PAX6. Important for DAergic PGN survival.
Meis2	7-3p	7mer-A1	1480	PAX6 cofactor. Important for DAergic PGN specification. Expressed in most TH positive PGNs
	7-3p	7mer-A1	2439	

Predictions performed using TargetScan. Position refers to the 3'UTR position of the first nucleotide of the MRE. Position 1 is the first nucleotide after the stop codon. Type refers to the MRE type. List includes conserved and poorly conserved TargetScan predictions. TF, transcription factor.

Other miRNAs likely participate with miR-7 to regulate *Pax6*. To identify other miRNAs that may be important for regulating *Pax6* in the VZ, we can use miR-CATCH to purify WT *Pax6* from embryonic dorsal telencephalon. Multiplex RT-qPCR can then be used to assay for many miRNAs simultaneously. Using this approach, we can generate a

miRNA interaction landscape for the *Pax6* 3'UTR in the developing cortex. This would likely reveal a unique complement of miRNAs that may function in cooperative regulation of *Pax6* similar to the miRNAs identified as possible regulators of *Pax6* by miTRAP in aTC1-6 cells (Figure 24). For example, we could use this strategy to determine whether miR-675-3p interacts with the endogenous *Pax6* transcript.

MicroRNA-7 buffering gene regulatory networks

A proposed function of miRNAs is to provide robustness to gene regulatory networks (GRNs), buffering GRN output in response to stress and environmental perturbation [103, 107]. For example, miR-7-5p participates in GRNs involved in *Drosophila* sensory organ and photoreceptor development [130] and may help to buffer fluctuations in these developmental networks under conditions of temperature stress [115]. Additionally, many miRNA gene knockout phenotypes are revealed under conditions of stress [116–127]. Given this, disruption of functional MREs in the 3'UTR of an endogenous mRNA transcript may not be associated with overt phenotypes under stable laboratory conditions. This may explain the lack of an overt DAergic specification phenotype in our *Pax6* 3'UTR miR-7-5p MRE mutant mice. However, physiological stress or environmental perturbation may expose defective miR-7-5p regulation of *Pax6* during specification or differentiation of DAergic PGNs in the OB. Odorant exposure is important for regulation DAergic PGN production [439]. For example, TH+ PGN genesis is reduced following odor depravation by naris occlusion in mice [435, 440] and following olfactory nerve lesions in rats [396]. Additionally, olfactory enrichment

increases production and survival of PGNs in the mouse OB [441] and specifically increases TH positive PGN density in the GL [442]. Malnutrition during embryogenesis can be used to stress developmental GRNs. For example, maternal low protein diet reduces brain-derived neurotrophic factor (BDNF) production in neonatal rats [443] and reduces proliferation of NSCs during mouse cortical development [444]. Future experiments could examine whether production of DAergic PGNs is altered in mice harboring mutations in the *Pax6* 3'UTR miR-7-5p MREs relative to WT following exposure to deprived or enriched odorant environments, or in pups exposed to nutrient depravation during embryonic development.

4.5.4 General PGN phenotype in *Pax6* 3'UTR miR-7-5p MRE mutant mice

The V-SVZ fate tracking results suggest that the proportion of TH, CalR and CalB positive PGNs may differ between WT and miR-7-5p MRE mutant mice (Figure 43C, Table 8). *Pax6*$^{+/Sey}$ mice have reduced TH positive PGNs with no change in numbers of CalR or CalB positive PGNs [245]. Similarly, *Pax6* overexpression or electroporation of miR-7 sponge into the lateral VZ is associated with an increase in the proportion of TH positive PGNs with no large change in the proportions of CalR or CalB cells [206]. These findings suggest that increased PAX6 protein associated with loss of miR-7-5p regulation may increase the production of TH positive PGNs without impacting CalR or CalB PGN populations. Contrary to predictions, the proportions of CalR and CalB positive PGNs appear most impacted by mutation of the *Pax6* 3'UTR miR-7-5p MREs. There are several possible explanations for these findings.

We also observed that PAX6 is expressed in a dorsal-high to ventral-low gradient along the medial wall of the LV, mirroring the gradient observed in the lateral VZ (Figure 42A). This expression pattern suggests that PAX6 may be important for specification of medially derived OB interneurons. MicroRNA-7 may be expressed in an opposing gradient to PAX6 in the medial V-SVZ, though this remains to be tested. Interestingly, PAX6 may also be elevated in the ventromedial VZ in *517+655MUT* mice (Figure 42B-C). CalR positive PGNs are specified primarily from the medial VZ [388, 393]. Though heterozygous *Pax6* mutations do not impact CalR positive PGN generation [245], it is possible that increased PAX6 protein in medial V-SVZ progenitors may promote increased CalR PGN genesis.

In addition to its role in specifying DAergic PGNs in the dorsal V-SVZ, PAX6 expression in the rostral migratory stream (RMS) of adult mice is also important for PGN specification generally [324]. The RMS is composed of migrating neuroblasts, many of which express PAX6 and give rise to PGNs in the OB [324]. Additionally, PAX6 is maintained in PGNs but not granule cells (GCs) of the OB. Overexpression of *Pax6* in RMS precursors converts approximately 70% of GCs into PGNs and increases the proportion of TH positive neurons formed [324]. Though the expression pattern of miR-7-5p in the RMS has not been described, it is possible that miR-7 acts as a temporal switch, repressing *Pax6* in migrating neuroblasts to promote granule cell fate. If this is the case, mutation of the miR-7-5p MREs in the *Pax6* 3'UTR may interfere with adoption of GC fate by RMS progenitors and increase the proportion of PGNs formed. We could test this hypothesis by counting the numbers of EdU and NeuN double positive cells in

the GL versus granule cell layer (GCL). We predict that the proportion of EdU positive neurons in the GL will be elevated at the expense of EdU positive neurons in the GCL in *Pax6* 3'UTR miR-7-5p MRE mutants.

Finally, the possible PGN phenotype observed in *Pax6* 3'UTR miR-7-5p MRE mutant mice may be the consequence of small sample sizes and large variability in the numbers of counted EdU positive PGNs between individual mice (Figure 44B) and may not reflect actual changes in the production of different PGN subtypes. Both male and female mice were analyzed for our PGN fate tracking experiments, possibly accounting for some of the observed variability. Differences in V-SVZ neurogenesis [445], olfactory bulb neuronal composition [446] and odorant processing [447] have been observed between adult male and female mice and humans. We compared the number of EdU positive neurons in the GL between male and female mice, finding no large differences between the sexes. Similarly, we compared the PGN fate of the NSCs labelled with EdU between WT male and female mice and found no differences (Appendix C: Figure 51). Given the laborious nature of this cell fate tracking experiment, resolving possible changes in PGN specification in miR-7-5p MRE mutants by increasing the sample size may not be practical.

4.5.5 EdU and cell survival

Fate mapping of progenitor cells can be performed using the thymidine analogs, BrdU, CldU, or EdU, which are incorporated into newly synthesized DNA during S-phase. Ponti *et al.,* (2013)[390] published evidence suggesting that EdU may be cytotoxic to cells that attempt to divide a second time following EdU incorporation. Mice were

injected with either EdU or CldU and whole mount VZ-SVZ sections were analyzed at 12 hour increments over 3 days. No significant toxicity was observed within 24 hours, however 3 days after injection the number of EdU labeled nuclei was approximately 1/3 that of CldU. Given that proliferating cells in the SVZ may undergo multiple divisions before finally differentiating in the OB, the use of EdU as an S-phase marker may be problematic for use in our OB fate mapping experiments.

To address this concern, we performed injections of either BrdU or EdU into mice at P1 and harvested OBs at 2 months post-injection. EdU was detected using a fluorophore-conjugated azide and a "click" reaction [448], and BrdU was detected using immunofluorescence (IF). IF was also used to detect CalR and TH expressing interneurons in the GL, to address whether formation of these cell types is negatively impacted by EdU labeling. Absolute numbers of EdU or BrdU CalR or TH double-positive neurons per mm^2 in the GL is not impacted by the use of EdU relative to BrdU. However, absolute number of BrdU labeled neurons is reduced relative to EdU (Appendix C: Figure 52). This difference may reflect increased sensitivity of the click reaction for detecting EdU over IF detection of BrdU (Appendix C: Figure 52C-F). The proportion of thymidine analog positive neurons expressing TH is not altered when EdU versus BrdU is used, however the proportion of CalR positive neurons expressing BrdU is slightly increased relative to those expressing EdU (Appendix C: Figure 52B). Antigen retrieval was used for BrdU IF to permit antibody access to the DNA [410], but was not needed for EdU detection. CalR IF is improved by antigen retrieval, which may account for the increased proportion of CalR and BrdU double positive neurons that could be reliably counted

(Appendix C: Figure 53A-D). TH IF is not noticeably impacted by antigen retrieval (data not shown).

4.5.6 MicroRNA-7 MRE conservation and *in vitro* functionality

Though we chose not to investigate the functionality of predicted MREs for miR-7-3p *in vivo*, the 7mer-m8 miR-7-3p MRE at *Pax6* 3'UTR position 626 is well conserved between orthologous vertebrate *Pax6* 3'UTRs (Table 6, Figure 36A) suggesting that this region of the *Pax6* 3'UTR is playing an important functional role. The *Pax6* 3'UTR *in vitro* reporter assay results further suggest that this region of the 3'UTR may be conserved for reasons other than miR-7-3p binding (Figure 38B). One possible explanation for conservation of this region of the *Pax6* 3'UTR is binding of other miRNAs. For example, MREs for miR-495, 376c-3p, 182-5p and 96-5p are located within this region of the *Pax6* 3'UTR, and MREs for miR-495 and 376c-3p directly overlap the predicted miR-7-3p MRE. MicroRNA- 182-5p and 96-5p may participate in regulating *Pax6* during epidermal specification (Du et al., 2013), and miR-376c-3p and miR-182-5p may be able to interact with the *Pax6* 3'UTR [400](Figure 24). Given this, the presence of a conserved miR-7-3p seed match at *Pax6* 3'UTR position 626 may be a coincidence.

Curiously, mutation of the miR-7-3p MRE at 3'UTR position 626 is associated with a possible repressive effect on reporter expression *in vitro* (Figure 38B). *Pax6* 3'UTR genetic reporter experiments reveal that disruption of the position 626 miR-7-3p MRE singly may be associated with a small decrease in luciferase activity relative a reporter containing the WT 3'UTR. Similarly, mutation of the position 626 miR-7-3p MRE in combination with the positions 517 and 655 miR-7-5p MREs abolished the derepression

associated with mutation of the two miR-7-5p MREs alone (Figure 36D, Figure 38B). In addition to expressing miR-7-5p and -3p, the cultured mouse pancreatic α-cells (αTC1-6) used for the *in vitro* reporter assay also express miR-495 and 376c-3p [400](Figure 21). The predicted MREs for miR-495 and 376c-3p overlap the predicted miR-7-3p MRE at 3'UTR position 626 and are disrupted upon mutation of the miR-7-3p MRE. If the 3'UTR region containing the predicted miR-7-3p MRE is conserved because of overlap with these other MREs, the observed decrease in luciferase activity with the miR-7-3p MRE mutant 3'UTR opposes the predicted effect. One explanation for this observation is that the miR-7-3p MRE mutation may enhance regulation by other miRNAs. Though we used ImiRP to ensure no new MREs were created upon introduction of the miR-7-3p MRE mutation, it is possible that the introduced mutation enhances 3' pairing of miRNAs that interact with adjacent MREs (Figure 3). The miR-SNP identified in the human *PAX6* 3'UTR associated with extreme myopia provides an example of this phenomenon [331, 332]. Here, a SNP in the human *PAX6* 3'UTR may enhance binding of miR-328, not by altering the interaction between the miR-328 seed region and the 3'UTR, but by enhancing miR-328 3' pairing. The predicted MRE for miR-182-5p/96-5p is located just downstream of the 3'UTR position 626 miR-7-3p MRE, and the introduced mutation may stabilize miR-182-5p/96-5p binding by improving 3' pairing.

Other RNA binding proteins (RBPs), in addition to miR-RISC, are known to regulate mRNA stability and translation [449]. It is possible that the predicted miR-7-3p MRE overlaps a recognition element for a stabilizing RBP, such as HuR, a ubiquitously expressed protein related to *Drosophila* ELAV [450]. Alternatively, the miR-7-3p MRE

mutation may create a novel recognition element for an RBP that enhances mRNA decay, such as AUF1 [451]. Both of these RPBs are known to bind AU-rich elements (AREs) in 3'UTRs, and bioinformatics software tools exist to predict putative RBP binding sites in 3'UTR sequences [403]. The tool RBP database (RBPDB) predicts that the *Pax6* 3'UTR contains a binding site for ELAV-like protein 1 (Elavl1)/HuR, which is disrupted by the miR-7-3p MRE mutation [403]. Since HuR is known to stabilize mRNAs, disruption of this predicted binding site may explain the observed decrease in reporter activity *in vitro*.

4.5.7 Other predicted phenotypes associated with *Pax6* 3'UTR miR-7-5p MRE mutants

miR-7 regulation of PAX6 is predicted to impact development and adult maintenance of other tissues. Kredo-Russo *et al*., (2012)[217] present evidence that miR-7-5p regulation of *Pax6* in the developing endocrine pancreas is important for refining the proportion of different endocrine pancreas cell generated. Specifically, disrupted miR-7-5p regulation of *Pax6* is predicted to increase the proportion of insulin and glucagon positive endocrine pancreatic cells at the expense of ghrelin positive cells. In support of this hypothesis, conditional knock out of *Pax6* in adult mouse α-cells and β-cells causes *Pax6*-deficient cells to adopt a ghrelin positive fate.[251, 452]

MicroRNA-7 regulation of *Pax6* may also be important for normal function of the adult endocrine pancreas. The insulin gene is a direct transcriptional target of PAX6 [195] and Glut2, the glucose transporter responsible for glucose sensing in pancreatic β-cells, is downregulated in *Pax6* heterozygous mice [250]. Generally, reduced pancreatic insulin content associated with *Pax6* heterozygosity may be due to a combination of

decreased insulin transcription and reduced ability for β-cells to respond to elevated glucose [250]. *Pax6* heterozygous mice develop early onset diabetes in response to high fat diet, which may be in part due to impaired proinsulin processing resulting from reduced production of prohormone convertase 1/3 (PC1/3) [453]. miRNA-7 is expressed in the adult mouse and human endocrine pancreas [125, 367]. Mice lacking the miR-7a-2 gene display improved glucose tolerance when challenged with intraperitoneal glucose tolerance test (IPGTT) due to enhanced insulin secretion [125]. Though *Pax6* was not identified as a miR-7 target by Latreille *et al*., (2014)[125], it is possible that disruption of the miR-7-5p MREs in the *Pax6* 3'UTR may be associated with increased production and secretion of insulin in response to glucose stress. Additionally, mutation of the *Pax6* 3'UTR miR-7-5p MREs may protect against early onset diabetes in *Sey* mice fed high fat diet.

Pax6 3'UTR miR-7-5p MRE mutant mice may not show overt defects in endocrine pancreas cell type development under stable laboratory conditions. However, maternal low protein diet can also be used as a strategy to stress development of the endocrine pancreas [454, 455]. Maternal protein restriction reduces β-cell proliferation and β-cell mass [454] and causes increased expression of miR-375, another predicted miRNA regulator of *Pax6* [276–278], in rat embryonic islets [456]. Additionally, protein restriction during embryonic development causes insulin resistance in adult rats [455]. Maternal low protein diet provides a mechanism to stress endocrine pancreas development and could be used to reveal developmental or adult pancreatic phenotypes associated with disruption of miR-7-5p regulation of *Pax6.*

miR-7-5p regulation of *Pax6* may also be important for controlling the balance of neurogenesis versus gliogenesis in the cortex. *Pax6* expression in the V-SVZ and RMS promotes neurogenesis, and expression of Olig2 opposes the neurogenic role of PAX6 and promotes oligodendrogenesis [324, 457]. Overexpression of PAX6 in the adult subependymal zone (SEZ), the adult derivative to the early postnatal V-SVZ, causes progenitors to adopt a neuronal phenotype. Alternatively, transduction of SEZ progenitors with a dominant-negative form of PAX6 reduces the proportion of neuroblasts generated [324]. Overexpression of Olig2 in SEZ progenitors results in migration of the progenitors away from the SEZ-RMS and into the corpus callosum, where they differentiate into oligodendrocytes [324]. *Pax6* overexpression in the early post-natal V-SVZ has a similar effect of promoting neurogenesis at the expense of gliogenesis by directly binding and repressing transcription of *Olig2* [457]. These results suggest that both OB neurons and oligodendrocytes arise from the adult SEZ and that PAX6 and Olig2 have opposing roles in specifying neurons versus oligodendrocytes. MicroRNA-7 may play an important role in repressing *Pax6* in progenitors fated to become oligodendrocytes, helping to ensure the correct proportions of neurons and glia are generated during cortical development [274]. We predict that loss direct of miR-7-5p regulation of *Pax6* may upregulate neurogenesis at the expense of gliogenesis in the developing cortex.

Pax6 is important for development of the eye, and miR-7 may play an important role in regulating *Pax6* during this process. MicroRNA-7 is plays a role during development of the *Drosophila* eye where it is important for photoreceptor

differentiation [130]. Though TargetScan [49] predicts that *ey*, the *Drosophila* homologue of *Pax6*, is not a target of miR-7, it is possible that miR-7 has a conserved pattern of expression in vertebrates. We found that miR-7a-5p is expressed in whole adult mouse retina but not E12.5 retina, adult lens or adult cornea by RT-qPCR (Figure 21C), though it is present at low levels relative to other retina expressed miRNAs (Figure 22C). This suggests that miR-7a-5p may be expressed in a subset of neurons in the adult retina. Given that *Pax6* expression is absent from differentiated photoreceptors in the vertebrate retina [199], this raises that possibility that miR-7a-5p may be functioning to downregulate *Pax6* during differentiation of retinal cell types that lack *Pax6*. The expression pattern of miR-7a in the adult mouse retina could be determined using *in situ* hybridization or genetic reporters. For example, a mouse line expressing lacZ under the control of the miR-7a-2 transcriptional unit could be used to address whether miR-7a-2 and *Pax6* display anticorrelated patterns of expression in the mature retina [87]. Though the role of miR-7 during development of the vertebrate retina is not known, it is possible that miR-7a-5p plays an important role during differentiation of photoreceptors by negatively regulating *Pax6*.

4.5.8 Conclusions and a cautionary tale

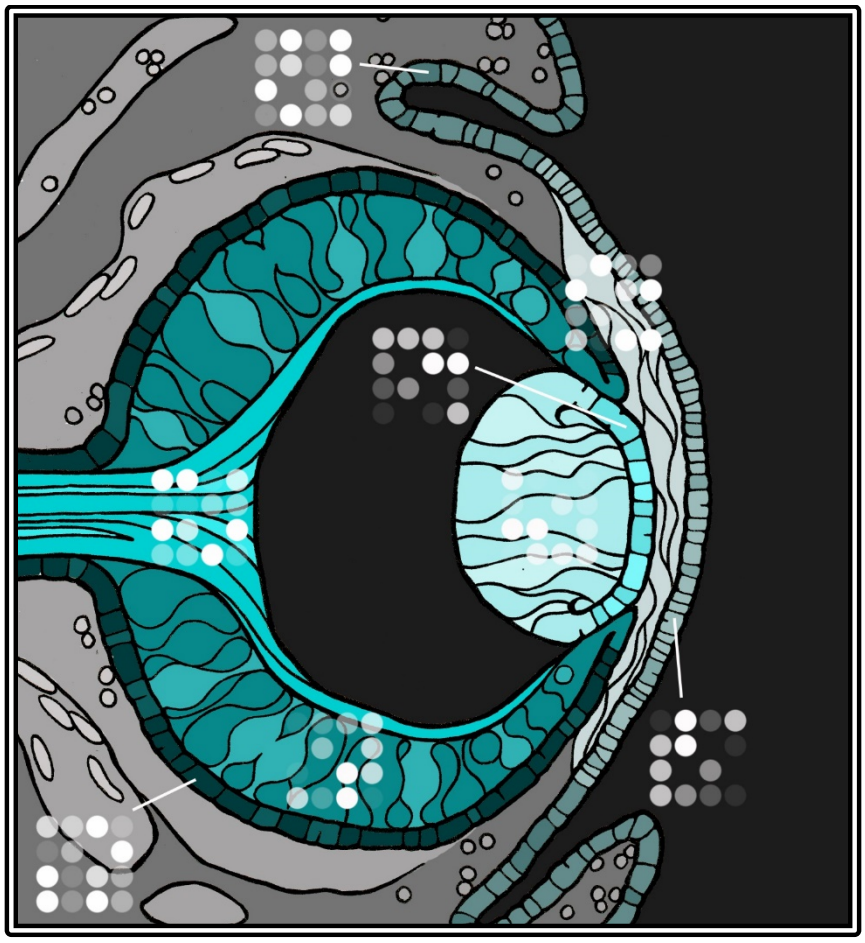
The direct functional consequences of miRNA regulation are typically studied using a combination of 3'UTR genetic reporter assays, and miRNA overexpression or LOF *in vitro* and *in vivo*. (Table 1). These approaches have utility for identifying likely mRNA targets of specific miRNAs and for identifying possible phenotypes associated with

miRNA regulation of a given target mRNA. However, these experimental approaches may suggest miRNA-target interactions that do not occur normally *in vivo*. Additionally, phenotypes observed following miRNA LOF or overexpression may exaggerate the functional consequence of one miRNA regulating one presumed target gene transcript [150, 156]. Given that miRNAs likely participate in complex GRNs with many different targets [40, 62, 103, 107, 131, 134, 135], miRNA LOF or overexpression may alter levels of specific gene products indirectly. Additionally, a single miRNA may target many gene products to produce a coherent functional outcome [90, 139–141]. Consequently, it is difficult to infer the function of a single miRNA-mRNA regulatory relationship from miRNA LOF or overexpression experiments.

Mutation of predicted MREs at the genomic level is necessary to characterize the specific phenotypic consequence of direct miRNA-target regulation. However, to our knowledge, direct genomic mutation of endogenous MREs in animals is rarely used to address this question [158, 159]. Many published findings suggest strong measurable phenotypes associated with regulation of a single gene transcript by a single miRNA, often through a single MRE. For example, loss of direct miR-7-5p regulation of *Pax6* is predicted to elevate PAX6 in regions of the brain and endocrine pancreas, increase DAergic PGN formation in the olfactory bulb, increase neurogenesis at the expense of gliogenesis in the developing cortex, and increase α-cell and β-cell differentiation at the expense of ghrelin producing ε-cells in the developing endocrine pancreas [206, 217, 274]. These hypothesized outcomes are testable by direct genetic mutation of candidate MREs. However, endogenous mutation of two predicted MREs for miR-7-5p in the 3’UTR

of *Pax6* produces subtle changes in the level of PAX6 protein but not *Pax6* mRNA in the early postnatal V-SVZ and is not associated with large changes in OB PGN cell fate specification. Though phenotypes associated with direct miR-7-5p MRE mutation in the *Pax6* 3'UTR remain to be fully characterized, our work illustrates that caution needs applied when drawing conclusions from miRNA LOF and overexpression experiments.

Chapter 5: Concluding remarks



"Visualizing miR-Codes" Digital illustration (2013)

5.1 Purpose

Cells can employ many mechanisms to regulate the spatial and temporal domains of gene expression and protein synthesis, as well as the dosage of synthesized proteins. We are interested in post-transcriptional regulation by microRNAs (miRNAs), which function by targeting miRNA recognition elements (MREs) located in mRNA 3' untranslated regions (3'UTRs) by complementary base pairing and negatively regulate levels of both target mRNA and protein product. Particularly, we wanted to investigate the physiological consequence of direct and specific regulation of an endogenous gene transcript by mutation of MREs *in vivo.* To accomplish this, we chose to focus on the transcription factor PAX6. PAX6 is dynamically regulated in space and time during development and maintaining the correct dosage of PAX6 is important for normal development and adult homeostasis of many tissues. Additionally, several miRNAs are implicated as regulators of *Pax6* in a variety of developmental and disease contexts. However, the physiological consequence of direct regulation of *Pax6* by miRNAs remains to be addressed. Here, we sought to disrupt predicted MREs in the 3'UTR of *Pax6* endogenously, with the goal of investigating the biological role of direct miRNA regulation *in vivo.*

5.2 ImiRP Summary

We started by developing a bioinformatics tool to facilitate design of MRE mutations. Given our plan to disrupt endogenous MREs for candidate miRNAs, we wanted to ensure any phenotypic effects observed following MRE mutation were specific to the miRNA of interest. One concern was that MRE mutation in the 3'UTR of

Pax6 would create new MREs for different miRNAs. If these miRNAs are coexpressed with *Pax6*, they could negatively regulate *Pax6* through these "illegitimate" MREs, complicating phenotypic analyses. Naturally occurring mutations that create functional MREs associated with phenotypes are documented in the literature. For example, a small nucleotide polymorphism (SNP) in the 3'UTR of the *myostatin* gene is associated with muscular hypertrophy observed in Texel sheep [336]. This SNP creates an illegitimate MRE for two miRNAs that are highly expressed in muscle, miR-1 and miR-206. Negative regulation of *myostatin* by miR-1/206 in Texel sheep may contribute to the meatiness of this breed [336]

Many bioinformatics tools are available to predict MREs that may function *in vivo.* However, none can generate mutations in MREs of interest and ensure that mutant sequences do not create new predicted MREs for other miRNAs. To address this, we built a software tool called the Illegitimate miRNA predictor (ImiRP). ImiRP can be used to specific MREs in a 3'UTR sequence of interest and generate mutant sequences that disrupt the MRE of interest while avoiding creation of new MREs.

5.3 *Pax6* 3'UTR miRNA regulatory landscape summary

Following this, we wanted to lay the foundation for future investigations into regulation of *Pax6* by miRNAs. Upon beginning this project, the site of *Pax6* transcript cleavage and polyadenylation was not known. We identified several conserved poly(A) signals in the 3'UTR of *Pax6* and used a combination of RNA sequencing and 3'RACE to determine the primary 3' cleavage site of the mouse *Pax6* message. With this knowledge, we predicted candidate MREs located in this 3'UTR sequence and

determined the expression patterns of the targeting miRNAs in several PAX6 expressing cells and tissues. Finally, using a *Pax6* 3'UTR reporter, we identified a cohort of miRNAs predicted to regulate *Pax6* that interact with the *Pax6* 3'UTR. This wholistic approach to investigating miRNA regulation of a single gene transcript enabled us to identify a short list of candidate miRNA regulators that could be studied in more detail. Additionally, we identified miRNAs that may be able to function cooperatively to regulate *Pax6.*

5.4 Endogenous MRE mutation summary

Finally, we wanted to engineer mice harboring mutations in predicted MREs in the 3'UTR of *Pax6* to address to role of miRNA regulation directly. We chose to focus on miR-7, which had previously been identified as a likely regulator of *Pax6* and was predicted to bind *Pax6* through multiple MREs. We confirmed that two predicted MREs for miR-7-5p, located at 3'UTR positions 517 and 655, likely mediate miR-7 regulation of *Pax6 in vitro.* We used genome engineering to disrupt these MREs in the context of the endogenous *Pax6* 3'UTR and found that they may function cooperatively to regulate levels of PAX6 protein in the early postnatal V-SVZ. Contrary to predictions, elevated PAX6 in the ventral V-SVZ of was not associated with altered specification of dopaminergic periglomerular neurons in the main olfactory bulb. Using *in vivo* mutation of predicted MREs in the 3'UTR of *Pax6* endogenously, we demonstrate that miR-7-5p may participate in refining the gradient of PAX6 in the V-SVZ, though other mechanisms are also required to regulate *Pax6.*

5.5 Future plans

Initially, we planned to examine miRNA regulation of endogenous *Pax6* more holistically, by either deleting the entire *Pax6* 3'UTR or by genomic mutation of many predicted MREs in the endogenous *Pax6* 3'UTR simultaneously. Ultimately, we opted not to use these strategies for several reasons. The 3'UTR has many functions in addition to miRNA regulation. A variety of RNA binding proteins are known to interact with 3'UTRs and regulate mRNA stability [449–451]. Additionally, poly(A) sequences are important for mRNA 3' cleavage and polyadenylation, which also impacts mRNA stability [5]. Given this, removing the entire *Pax6* 3'UTR and replacing it with a generic poly(A) signal, such as the SV40 poly(A) sequence, would likely produce effects that were independent of miRNA regulation.

We also designed a *Pax6* 3'UTR "master mutant" sequence in which many predicted MREs were disrupted simultaneously and attempted to generate mice carrying these mutations using a classical gene targeting approach. This strategy was problematic because mouse embryonic stem (ES) cell DNA repair mechanisms corrected many of the desired MRE substitution mutations following homologous recombination of the targeting vector. We ultimately abandoned generation of *Pax6* 3'UTR MRE master mutant mice because at the time we did not have information about expression patterns of the candidate miRNAs or information about the interaction of these miRNAs with the 3'UTR of *Pax6.* Moving forward, we could combine data on *Pax6* tissue-specific "miR-codes" with 3'UTR interaction landscape data from pulldown experiments to identify a short list likely miRNA regulators of *Pax6.* We define miR-codes as miRNAs

predicted to target a specific gene transcript, such as *Pax6,* that are expressed in tissue or cell type-specific patterns. Interaction landscapes define the subset of these miRNAs that can interact with the gene transcript of interest. In the future, we could use miR-CATCH to identify miRNAs predicted to regulate *Pax6* that also interact with the endogenous *Pax6* transcript. Then using ImiRP, we could design mutations that disrupt the predicted MREs without creating illegitimate MREs. Using CRISPR/Cas9 genome engineering, we could revisit generating *Pax6* 3'UTR "master mutant" mice, to study regulation of *Pax6* by miRNAs more holistically.

To reveal potential miRNA regulation through the *Pax6* 3'UTR *in vivo*, we designed a two-colour bi-directional reporter construct expressing GFP and mCherry reporters under the control of a bi-directional phosphoglycerate kinase (PGK) promoter. The *Pax6* 3'UTR is fused to the 3' end of the mCherry gene and TurboGFP contains a short synthetic poly(A) sequence. This approach would permit convenient internal normalization of mCherry fluorescence to that of GFP on a single cell basis. Mutations can be introduced into the *Pax6* 3'UTR and the mCherry:GFP fluorescence ratio compared between WT and mutant *Pax6* 3'UTR conditions.

This reporter was designed specifically for identifying miRNA interactions in the mouse retina, and several features of its design were optimized for this purpose. First, the PGK promoter was selected because it drives relatively low constitutive expression in mammalian cell types and expresses in the mouse retina [458, 459]. Since regulation of a target gene transcript by miRNAs responds most sensitively when the ratio of miRNA:target is low [64], we wanted to ensure that we did not use a constitutive

promoter that drives high levels of transcription. In addition to selecting a constitutive promoter that drives low levels of expression mammalian cells, we also inserted modified proline/glutamic acid/serine/ threonine (PEST) sequences to the C-terminal ends of both mCherry and GFP to promote rapid protein turnover [460]. We reasoned that stable fluorescent reporter proteins would accumulate in the cell and fluorescent signal may not reflect translational regulation of miRNAs. Consequently, we opted to use destabilized reporter proteins. Finally, we designed fluorescent reporter proteins to contain nuclear localization signals (NLS) [461] to enable easier quantification of fluorescent signal. Nuclear localized fluorescent reporters containing PEST sequences are subject to rapid turnover [462]. Quantitative fluorescence microscopy can be used to quantify GFP and mCherry fluorescence from transfected or electroporated cells. Alternatively, flow cytometry could be used to quantify fluorescence on a single cell basis for thousands of cells.

Pax6 is expressed at low levels in retinal progenitor cells (RPCs) of the developing retina and is absent from differentiated photoreceptors and bipolar cells [199]. We plan to electroporate these vectors into the P0 retina to identify MREs in the *Pax6* 3'UTR that may be mediating negative regulation of *Pax6*. Additionally, these bi-directional fluorescent reporters could be used to identify MREs in the *Pax6* 3'UTR that may be important for mediating negative regulation of *Pax6* during RPC differentiation into photoreceptors or bipolar cells. This approach provides an advantage over bulk luciferase assays, which involve lysing cells and quantifying reporter protein level using an enzymatic assay. The level of miRNA mediated repression of reporter protein

synthesis can vary dramatically between cells [64]. Consequently, analysing reporter expression on a single cell basis provides sensitivity that cannot be accomplished using bulk luciferase assays.

5.6 Significance

This work represents one of few studies documenting the consequence of genetic ablation of endogenous MREs *in vivo* [158, 159]. Despite this, many studies claim to demonstrate “direct” regulation of an individual target gene transcript by an individual miRNA often through a single MRE in various developmental or disease contexts. However, the conclusion of direct regulation using experimental approaches such as miRNA overexpression and loss of function, and 3’UTR genetic reporter assays. To illustrate this, a Pubmed search for ((("directly targets" OR "direct target" OR "directly regulates") AND miR) in the title or abstract yields 4349 hits (as of July 10, 2019). Including ((("genome engineering") OR "gene targeting") in the whole article yields only 24 hits (as of July 10, 2019). Closer inspection of these 24 articles reveals that none performed mutation of predicted MREs in the context of the endogenous gene. Generally, these results help to illustrate a gap in the miRNA literature. To confirm direct and specific regulation by a miRNA, mutation of predicted MREs at the genomic level is necessary.

Bibliography

1. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Res. 2011;21:381–95. doi:10.1038/cr.2011.22.

2. Lee TI, Young RA. Transcriptional regulation and its misregulation in disease. Cell. 2013;152:1237–51. doi:10.1016/j.cell.2013.02.014.

3. Ayoubi T, Van de Ven W. Regulation of gene expression by alternative promoters. FASEB. 1996;10:3.

4. Nilsen TW, Graveley BR. Expansion of the eukaryotic proteome by alternative splicing. Nature. 2010;463:457–63.

5. Proudfoot NJ. Ending the message: Poly(A) signals then and now. Genes Dev. 2011;25:1770–82.

6. Shyu A Bin, Wilkinson MF, Van Hoof A. Messenger RNA regulation: To translate or to degrade. EMBO J. 2008;27:471–81.

7. Zhang X, Virtanen A, Kleiman FE. To polyadenylate or to deadenylate: That is the question. Cell Cycle. 2010;9:4437–49.

8. Sonenberg N, Hinnebusch AG. Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. Cell. 2009;136:731–45. doi:10.1016/j.cell.2009.01.042.

9. Varshavsky A. The ribosome revealed. Trends Biocemical Sci. 2005;30:283–6.

10. Seet BT, Dikic I, Zhou M-M, Pawson T. Reading protein modifications with interaction domains. Nat Rev Mol Cell Biol. 2006;7:473–83. doi:10.1038/nrm1960.

11. Mattick JS, Makunin I V. Non-coding RNA. Hum Mol Genet. 2006;15 Spec No:R17-29. doi:10.1093/hmg/ddl046.

12. Hoagland M, Stepheson M, Scott J, Hecht L, Zamecnik P. A soluble ribonucleic acid intermediate in protein synthsis. J Biol Chem. 1958;231:241–57. http://www.ncbi.nlm.nih.gov/pubmed/16651366.

13. Bartel DP, Lee R, Feinbaum R. MicroRNAs : Genomics , Biogenesis , Mechanism , and Function Genomics : The miRNA Genes. 2004;116:281–97.

14. Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense comlementarity to lin-14. Cell. 1993;75:843–54.

15. Reinhart BJ, Slack FJ, Basson M, Pasquinelli a E, Bettinger JC, Rougvie a E, et al. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature. 2000;403:901–6.

16. Lim LP. Vertebrate MicroRNA Genes. Science (80-). 2003;299:1540–1540. doi:10.1126/science.1080372.

17. Berezikov E, Guryev V, Van De Belt J, Wienholds E, Plasterk RHA, Cuppen E. Phylogenetic shadowing and computational identification of human microRNA genes. Cell. 2005;120:21–4.

18. Kozomara A, Griffiths-Jones S. MiRBase: Annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. 2014;42:68–73.

19. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell. 2005;120:15–20.

20. Friedman RC, Farh KKH, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res. 2009;19:92–105.

21. Miranda KC, Huynh T, Tay Y, Ang YS, Tam WL, Thomson AM, et al. A Pattern-Based Method for the Identification of MicroRNA Binding Sites and Their Corresponding Heteroduplexes. Cell. 2006;126:1203–17.

22. Lee Y, Kim M, Han J, Yeom K-H, Lee S, Baek SH, et al. MicroRNA genes are transcribed by RNA polymerase II. EMBO J. 2004;23:4051–60. doi:10.1038/sj.emboj.7600385.

23. Lagos-Quintana M. Identification of Novel Genes Coding for Small Expressed RNAs. Science (80-). 2001;294:853–8. doi:10.1126/science.1064921.

24. Lau NC. An Abundant Class of Tiny RNAs with Probable Regulatory Roles in Caenorhabditis elegans. Science (80-). 2001;294:858–62. doi:10.1126/science.1065062.

25. Gregory RI, Yan KP, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, et al. The Microprocessor complex mediates the genesis of microRNAs. Nature. 2004;432:235–40.

26. Gregory RI, Chendrimada TP, Cooch N, Shiekhattar R. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. Cell. 2005;123:631–40.

27. Bernstein E, Caudy AA, Hammond SM, Hannon GJ. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature. 2001;409:363–6. doi:10.1038/35053110.

28. Neilsen CT, Goodall GJ, Bracken CP. IsomiRs - The overlooked repertoire in the dynamic microRNAome. Trends Genet. 2012;28:544–9. doi:10.1016/j.tig.2012.07.005.

29. Schwarz DS, Du T, Xu Z, Aronin N, Zamore PD. Asymmetry in the assembly of the RNAi enzyme complex. Cell. 2003;115:199–208.

30. Wang Y, Sheng G, Juranek S, Tuschl T, Patel DJ. Structure of the guide-strand-containing argonaute silencing complex. Nature. 2008;456:209–13.

31. Schirle NT, Sheu-Gruttadauria J, MacRae IJ. Structural basis for microRNA targeting. Science (80-). 2014;346:608–13. doi:10.1126/science.1258040.

32. Sheu-Gruttadauria J, MacRae IJ. Structural Foundations of RNA Silencing by Argonaute. J Mol Biol. 2017;429:2619–39. doi:10.1016/j.jmb.2017.07.018.

33. Salomon WEE, Jolly SMM, Moore MJJ, Zamore PDD, Serebrov V. Single-Molecule Imaging Reveals that Argonaute Reshapes the Binding Properties of Its Nucleic Acid Guides. Cell. 2015;162:517–20. doi:10.1016/j.cell.2016.06.048.

34. Gu S, Jin L, Zhang F, Sarnow P, Kay MA. Biological basis for restriction of microRNA targets to the 3′ untranslated region in mammalian mRNAs. Nat Struct Mol Biol. 2009;16:144–50. doi:10.1038/nsmb.1552.

35. Grimson A, Farh KKH, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP. MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing. Mol Cell. 2007;27:91–105.

36. Tay Y, Zhang J, Thomson AM, Lim B, Rigoutsos I. MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. Nature. 2008;455:1124–8. doi:10.1038/nature07299.

37. Forman JJ, Legesse-Miller A, Coller HA. A search for conserved sequences in coding regions reveals that the let-7 microRNA targets Dicer within its coding sequence. Proc Natl Acad Sci. 2008;105:14879–84. doi:10.1073/pnas.0803230105.

38. Duursma AM, Kedde M, Schrier M, Sage CLE, Agami R. Mir-148 targets human DNMT3b. Rna. 2008;14:872–7.

39. Elcheva I, Goswami S, Noubissi F, Spiegelman V. CRD-BP protects the coding region of betaTrCP1 mRNA from miR-183-mediated degradation. Mol Cell. 2009;35:240–6.

40. Clark PM, Loher P, Quann K, Brody J, Londin ER, Rigoutsos I. Argonaute CLIP-Seq reveals miRNA targetome diversity across tissue types. Sci Rep. 2014;4:1–11.

41. Ambros V, Lee R, Lavanway A, Williams P, Jewel D. MicroRNAs and other tiny endogenous RNAs in C. elegans. Curr Biol. 2003;13:807–18.

42. Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel DP. Prediction of Plant MicroRNA Targets. 2002;110:513–20.

43. Lewis BP, Shih I, Jones-Rhoades MW, Bartel DP, Burge CB. Prediction of Mammalian MicroRNA Targets. Cell. 2003;115:787–98. doi:10.1016/S0092-8674(03)01018-3.

44. Lai EC. Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. Nat Genet. 2002;30:363–4. doi:10.1038/ng865.

45. Stark A, Brennecke J, Russell RB, Cohen SM. Identification of Drosophila microRNA targets. PLoS Biol. 2003;1.

46. Lim LP, Lau NC, Weinstein EG, Abdelhakim A, Yekta S, Rhoades MW, et al. The microRNAs of Caenorhabditis elegans. Genes Dev. 2003;17:991–1008.

47. Doench JG, Sharp PA. Specificity of microRNA target selection in translational repression. Genes Dev. 2004;18:504–11.

48. Nielsen CB, Shomron N, Sandberg R, Hornstein E, Kitzman J, Burge CB. Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. Rna. 2007;13:1894–910.

49. Agarwal V, Bell GW, Nam J, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. Elife. 2015;4:1–38.

50. Baek D, Villén J, Shin C, Camargo FD, Steven P, Bartel DP. The impact of microRNAs on protein output. Nature. 2009;455:64–71.

51. Bartel DP. MicroRNAs: Target Recognition and Regulatory Functions. Cell. 2009;136:215–33.

52. Brennecke J, Stark A, Russell RB, Cohen SM. Principles of microRNA-target recognition. PLoS Biol. 2005;3:0404–18.

53. Didiano D, Hobert O. Molecular architecture of a miRNA-regulated 3' UTR. Rna. 2008;14:1297–317.

54. Chi SW, Hannon GJ, Darnell RB. An alternative mode of microRNA target recognition. Nat Struct Mol Biol. 2012;19:321–7. doi:10.1038/nsmb.2230.

55. Helwak A, Kudla G, Dudnakova T, Tollervey D. Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. Cell. 2013;153:654–65. doi:10.1016/j.cell.2013.03.043.

56. Chi SW, Zang JB, Mele A, Darnell RB. Argonaute HITS-CLIP decodes microRNA–mRNA interaction maps. Nature. 2009;460:479–86. doi:10.1038/nature08170.

57. Ellwanger DC, Büttner FA, Mewes HW, Stümpflen V. The sufficient minimal set of miRNA seed types. Bioinformatics. 2011;27:1346–50.

58. Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E. The role of site accessibility in microRNA target recognition. Nat Genet. 2007;39:1278–84. doi:10.1038/ng2135.

59. Hon LS, Zhang Z. The roles of binding site arrangement and combinatorial targeting in microRNA repression of gene expression. Genome Biol. 2007;8:R166. doi:10.1186/gb-2007-8-8-r166.

60. Yekta S, Shih I, Bartel DP. MicroRNA-Directed Cleavage of HOXB8 mRNA. Science (80-). 2004;304:594–6. doi:10.1126/science.1097434.

61. Rogers K, Chen X. Biogenesis, Turnover, and Mode of Action of Plant MicroRNAs. Plant Cell. 2013;25:2383–99.

62. Selbach M, Schwanhäusser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N. Widespread changes in protein synthesis induced by microRNAs. Nature. 2008;455:58–63.

63. Jonas S, Izaurralde E. Towards a molecular understanding of microRNA-mediated gene silencing. Nat Rev Genet. 2015;16:421–33. doi:10.1038/nrg3965.

64. Mukherji S, Ebert MS, Zheng GXY, Tsang JS, Sharp PA, van Oudenaarden A. MicroRNAs can generate thresholds in target gene expression. Nat Genet. 2011;43:854–9. doi:10.1038/ng.905.

65. Gebert LFR, Macrae IJ. Regulation of microRNA function in animals. Nat Rev Mol Cell Biol. 2019;20 January:21–37.

66. Guo H, Ingolia NT, Weissman JS, Bartel DP. Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature. 2010;466:835–40. doi:10.1038/nature09267.

67. Meijer H a, Kong YW, Lu WT, Wilczynska A, Spriggs R V, Robinson SW, et al. Translational Repression and MicroRNA-mediated Gene Regulation. Science (80-). 2013;340 April:82–5.

68. Vasudevan S, Tong Y, Steitz J. Switching from Repression to Activation : MicroRNAs Can Up-Regulate Translation. Science (80-). 2007;318 December:1931–4.

69. Wienholds E, Koudijs MJ, van Eeden FJM, Cuppen E, Plasterk RH a. The microRNA-producing enzyme Dicer1 is essential for zebrafish development. Nat Genet. 2003;35:217–8. doi:10.1038/ng1251.

70. Bernstein E, Kim SY, Carmell MA, Murchison EP, Alcorn H, Li MZ, et al. Dicer is essential for mouse development. Nat Genet. 2003;35:215–7. doi:10.1038/ng1253.

71. Lynn FC, Skewes-Cox P, Kosaka Y, McManus MT, Harfe BD, German MS. MicroRNA

expression is required for pancreatic islet cell genesis in the mouse. Diabetes. 2007;56:2938. doi:10.2337/db07-0175.Additional.

72. Melkman-Zehavi T, Oren R, Kredo-Russo S, Shapira T, Mandelbaum AD, Rivkin N, et al. miRNAs control insulin content in pancreatic β-cells via downregulation of transcriptional repressors. EMBO J. 2011;30:835–45. doi:10.1038/emboj.2010.361.

73. Mandelbaum AD, Melkman-Zehavi T, Oren R, Kredo-Russo S, Nir T, Dor Y, et al. Dysregulation of Dicer1 in beta cells impairs islet architecture and glucose metabolism. Exp Diabetes Res. 2012;2012.

74. De Pietri Tonelli D, Pulvers JN, Haffner C, Murchison EP, Hannon GJ, Huttner WB. miRNAs are essential for survival and differentiation of newborn neurons but not for expansion of neural progenitors during early neurogenesis in the mouse embryonic neocortex. Development. 2008;135:3911–21. doi:10.1242/dev.025080.

75. Iida A, Shinoe T, Baba Y, Mano H, Watanabe S. Dicer Plays Essential Roles for Retinal Development by Regulation of Survival and Differentiation. Investig Opthalmology Vis Sci. 2011;52:3008. doi:10.1167/iovs.10-6428.

76. Davis N, Mor E, Ashery-Padan R. Roles for Dicer1 in the patterning and differentiation of the optic cup neuroepithelium. Development. 2010;138:127–38.

77. Georgi SA, Reh TA. Dicer is required for the maintenance of notch signaling and gliogenic competence during mouse retinal development. Dev Neurobiol. 2011;71:1153–69.

78. Damiani D, Alexander JJ, O'Rourke JR, McManus M, Jadhav AP, Cepko CL, et al. Dicer Inactivation Leads to Progressive Functional and Structural Degeneration of the Mouse Retina. J Neurosci. 2008;28:4878–87. doi:10.1523/JNEUROSCI.0828-08.2008.

79. Cuellar TL, Davis TH, Nelson PT, Loeb GB, Harfe BD, Ullian E, et al. Dicer loss in striatal neurons produces behavioral and neuroanatomical phenotypes in the absence of neurodegeneration. Proc Natl Acad Sci. 2008;105:5614–9. doi:10.1073/pnas.0801689105.

80. Davis TH, Cuellar TL, Koch SM, Barker AJ, Harfe BD, McManus MT, et al. Conditional Loss of Dicer Disrupts Cellular and Tissue Morphogenesis in the Cortex and Hippocampus. J Neurosci. 2008;28:4322–30. doi:10.1523/JNEUROSCI.4815-07.2008.

81. Song MS, Rossi JJ. Molecular mechanisms of Dicer: Endonuclease and enzymatic activity. Biochem J. 2017;474:1603–18.

82. Miska EA. Microarray analysis of microRNA expression in the developing mammalian brain. Genome Biol. 2004;5:R68. doi:10.1186/gb-2004-5-9-r68.

83. Aboobaker AA, Tomancak P, Patel N, Rubin GM, Lai EC. Drosophila microRNAs exhibit diverse spatial expression patterns during embryonic development. Proc Natl Acad Sci. 2005;102:18017–22.

84. Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N, Aravin A, et al. A Mammalian microRNA Expression Atlas Based on Small RNA Library Sequencing. Cell. 2007;129:1401–14.

85. Martinez NJ, Ow MC, Reece-Hoyes JS, Barrasa MI, Ambros VR, Walhout AJM. Genome-scale spatiotemporal analysis of Caenorhabditis elegans microRNA promoter activity. Genome Res. 2008;18:2005–15.

86. Kato M, de Lencastre A, Pincus Z, Slack FJ. Dynamic expression of small non-coding RNAs, including novel microRNAs and piRNAs/21U-RNAs, during Caenorhabditis elegans development. Genome Biol. 2009;10:1–15.

87. Park CY, Jeker LT, Carver-Moore K, Oh A, Liu HJ, Cameron R, et al. A Resource for the Conditional Ablation of microRNAs in the Mouse. Cell Rep. 2012;1:385–91. doi:10.1016/j.celrep.2012.02.008.

88. Wienholds E, Kloosterman WP, Miska E, Alvarez-saavedra E, Berezikov E, Bruijn E De, et al. MicroRNA Expression in Zebrafish Embryonic Development. Science (80-). 2005;309:310–1.

89. Wienholds E, Plasterk RHA. MicroRNA function in animal development. FEBS Lett. 2005;579:5911–22.

90. Shenoy A, Blelloch RH. Regulation of microRNA function in somatic stem cell proliferation and differentiation. Nat Rev Mol Cell Biol. 2014;15:565–76. doi:10.1038/nrm3854.

91. Zhao T, Li G, Mi S, Li S, Hannon GJ, Wang XJ, et al. A complex system of small RNAs in the unicellular green alga. Genes Dev. 2007;:1190–203.

92. Bråte J, Neumann RS, Fromm B, Haraldsen AAB, Tarver JE, Suga H, et al. Unicellular Origin of the Animal MicroRNA Machinery. Curr Biol. 2018;28:3288-3295.e5.

93. Chen PY, Manninga H, Slanchev K, Chien M, Russo JJ, Ju J, et al. The developmental miRNA profiles of zebrafish as determined by small RNA cloning. Genes Dev. 2005;19:1288–93.

94. Lin Y, Zeng Y, Zhang F, Xue L, Huang Z, Li W, et al. Characterization of MicroRNA Expression Profiles and the Discovery of Novel MicroRNAs Involved in Cancer during Human Embryonic Development. PLoS One. 2013;8:1–11.

95. Zhan M, Miller CP, Papayannopoulou T, Stamatoyannopoulos G, Song CZ. MicroRNA

expression dynamics during murine and human erythroid differentiation. Exp Hematol. 2007;35:1015–25.

96. Hildebrand J, Rütze M, Walz N, Gallinat S, Wenck H, Deppert W, et al. A comprehensive analysis of microRNA expression during human keratinocyte differentiation in vitro and in vivo. J Invest Dermatol. 2011;131:20–9.

97. Arora A, Guduric-Fuchs J, Harwood L, Dellett M, Cogliati T, Simpson D a. Prediction of microRNAs affecting mRNA expression during retinal development. BMC Dev Biol. 2010;10:1.

98. Chen C-Z. MicroRNAs Modulate Hematopoietic Lineage Differentiation. Science (80-). 2004;303:83–6. doi:10.1126/science.1091903.

99. Kanellopoulou C, Muljo SA, Kung AL, Ganesan S, Drapkin R, Jenuwein T, et al. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev. 2005;19:489–501.

100. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, et al. MicroRNA expression profiles classify human cancers. Nature. 2005;435:834–8. doi:10.1038/nature03702.

101. Ji Z, Lee JY, Pan Z, Jiang B, Tian B. Progressive lengthening of 3’ untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. Proc Natl Acad Sci. 2009;106:7028–33. doi:10.1073/pnas.0900028106.

102. Wehrspaun CC, Ponting CP, Marques AC. Brain-expressed 3’UTR extensions strengthen miRNA cross-talk between ion channel/transporter encoding mRNAs. Front Genet. 2014;5 FEB:1–11.

103. Ebert MS, Sharp PA. Roles for microRNAs in conferring robustness to biological processes. Cell. 2012;149:515–24.

104. Stark A, Brennecke J. Article Animal MicroRNAs Confer Robustness to Gene Expression and Have a Significant Impact on 3 ’ UTR Evolution Results Target-Site Prediction with High Specificity and Coverage. 2005.

105. Pasquinelli a E, Reinhart BJ, Slack F, Martindale MQ, Kuroda MI, Maller B, et al. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature. 2000;408:86–9.

106. Leucht C, Stigloher C, Wizenmann A, Klafke R, Folchert A, Bally-Cuif L. MicroRNA-9 directs late organizer activity of the midbrain-hindbrain boundary. Nat Neurosci. 2008;11:641–8. doi:10.1038/nn.2115.

107. Tsang J, Zhu J, van Oudenaarden A. MicroRNA-Mediated Feedback and

Feedforward Loops Are Recurrent Network Motifs in Mammals. Mol Cell. 2007;26:753–67.

108. Miska EA, Alvarez-Saavedra E, Abbott AL, Lau NC, Hellman AB, McGonagle SM, et al. Most Caenorhabditis elegans microRNAs are individually not essential for development or viability. PLoS Genet. 2007;3:2395–403.

109. Pucella JN, Yen W-F, Kim M V., van der Veeken J, Socci ND, Naito Y, et al. miR-182 Is Largely Dispensable for Adaptive Immunity: Lack of Correlation between Expression and Function. J Immunol. 2015;194:2635–42. doi:10.4049/jimmunol.1402261.

110. Abbott AL, Alvarez-Saavedra E, Miska EA, Lau NC, Bartel DP, Horvitz HR, et al. The let-7 MicroRNA family members mir-48, mir-84, and mir-241 function together to regulate developmental timing in Caenorhabditis elegans. Dev Cell. 2005;9:403–14.

111. Ventura A, Young AG, Winslow MM, Lintault L, Meissner A, Erkeland SJ, et al. Targeted Deletion Reveals Essential and Overlapping Functions of the miR-17∼92 Family of miRNA Clusters. Cell. 2008;132:875–86.

112. Liu N, Bezprozvannaya S, Williams AH, Qi X, Richardson JA, Bassel-Duby R, et al. microRNA-133a regulates cardiomyocyte proliferation and suppresses smooth muscle gene expression in the heart. Genes Dev. 2008;22:3242–54.

113. Alvarez-Saavedra E, Horvitz HR. Many Families of C. elegans MicroRNAs Are Not Essential for Development or Viability. Curr Biol. 2010;20:367–73. doi:10.1016/j.cub.2009.12.051.

114. Chen YW, Song S, Weng R, Verma P, Kugler JM, Buescher M, et al. Systematic study of Drosophila MicroRNA functions using a collection of targeted knockout mutations. Dev Cell. 2014;31:784–800. doi:10.1016/j.devcel.2014.11.029.

115. Li X, Cassidy JJ, Reinke CA, Fischboeck S, Carthew RW. A MicroRNA Imparts Robustness against Environmental Fluctuation during Development. Cell. 2009;137:273–82. doi:10.1016/j.cell.2009.01.058.

116. Xin M, Small EM, Sutherland LB, Qi X, McAnally J, Plato CF, et al. MicroRNAs miR-143 and miR-145 modulate cytoskeletal dynamics and responsiveness of smooth muscle cells to injury. Genes Dev. 2009;23:2166–78.

117. Gaudet AD, Mandrekar-Colucci S, Hall JCE, Sweet DR, Schmitt PJ, Xu X, et al. miR-155 Deletion in Mice Overcomes Neuron-Intrinsic and Neuron-Extrinsic Barriers to Spinal Cord Repair. J Neurosci. 2016;36:8516–32. doi:10.1523/JNEUROSCI.0735-16.2016.

118. Hu Z, Cui Y, Qiao X, He X, Li F, Luo C, et al. Silencing miR-150 ameliorates experimental autoimmune encephalomyelitis. Front Neurosci. 2018;12 JUL:1–9.

119. Wu KC, Chen XJ, Jin GH, Wang XY, Yang DD, Li YP, et al. Deletion of miR-182 leads to retinal dysfunction in mice. Investig Ophthalmol Vis Sci. 2019;60:1265–74.

120. van Rooij E, Sutherland LB, Qi X, Richardson JA, Hill J, Olson EN. Control of Stress-Dependent Cardiac Growth and Gene Expression by a MicroRNA. Science (80-). 2007;316:575–9. doi:10.1126/science.1139089.

121. Khoshgoo N, Visser R, Falk L, Day CA, Ameis D, Iwasiow BM, et al. MicroRNA-200b regulates distal airway development by maintaining epithelial integrity. Sci Rep. 2017;7:1–12.

122. Remenyi J, van den Bosch MWM, Palygin O, Mistry RB, McKenzie C, Macdonald A, et al. miR-132/212 Knockout Mice Reveal Roles for These miRNAs in Regulating Cortical Synaptic Transmission and Plasticity. PLoS One. 2013;8:1–14.

123. Klein U, Lia M, Crespo M, Siegel R, Shen Q, Mo T, et al. The DLEU2/miR-15a/16-1 Cluster Controls B Cell Proliferation and Its Deletion Leads to Chronic Lymphocytic Leukemia. Cancer Cell. 2010;17:28–40. doi:10.1016/j.ccr.2009.11.019.

124. Poy MN, Hausser J, Trajkovski M, Braun M, Collins S, Rorsman P, et al. miR-375 maintains normal pancreatic - and -cell mass. Proc Natl Acad Sci. 2009;106:5813–8. doi:10.1073/pnas.0810550106.

125. Latreille M, Hausser J, Stutzer I, Zhang Q, Hastoy B, Gargani S, et al. MicroRNA-7a regulates pancreatic beta cell function. J Clin Invest. 2014;124:2722–35. doi:10.1172/JCI73066.

126. Gaudet AD, Fonken LK, Gushchina L V., Aubrecht TG, Maurya SK, Periasamy M, et al. MIR-155 deletion in female mice prevents diet-induced obesity. Sci Rep. 2016;6 March:1–13. doi:10.1038/srep22862.

127. Brenner JL, Jasiewicz KL, Fahley AF, Kemp BJ, Abbott AL. Loss of individual MicroRNAs causes mutant phenotypes in sensitized genetic backgrounds in C. Elegans. Curr Biol. 2010;20:1321–5. doi:10.1016/j.cub.2010.05.062.

128. Gibson G and GW, Wagner G. Canalization in Evolutionary Genetics. a Stablizing Theory? BIoessays. 2000;22:372–80. doi:10.1002/(SICI)1521-1878(200004)22:4<372::AID-BIES7>3.0.CO;2-J.

129. O'Donnell KA, Wentzel EA, Zeller KI, Dang C V., Mendell JT. c-Myc-regulated microRNAs modulate E2F1 expression. Nature. 2005;435:839–43. doi:10.1038/nature03677.

130. Li X, Carthew RW. A microRNA mediates EGF receptor signaling and promotes photoreceptor differentiation in the Drosophila eye. Cell. 2005;123:1267–77.

131. John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS. Human MicroRNA Targets. PLoS Biol. 2004;2:e363. doi:10.1371/journal.pbio.0020363.

132. Tu K, Yu H, Hua YJ, Li YY, Liu L, Xie L, et al. Combinatorial network of primary and secondary microRNA-driven regulatory mechanisms. Nucleic Acids Res. 2009;37:5969–80.

133. Zare H, Khodursky A, Sartorelli V. An evolutionarily biased distribution of miRNA sites toward regulatory genes with high promoter-driven intrinsic transcriptional noise. BMC Evol Biol. 2014;14:74. doi:10.1186/1471-2148-14-74.

134. Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS. MicroRNA targets in Drosophila. Genome Biol. 2003;5:R1. doi:10.1186/gb-2003-5-1-r1.

135. Krek A, Grün D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, et al. Combinatorial microRNA target predictions. Nat Genet. 2005;37:495–500. doi:10.1038/ng1536.

136. Du L, Schageman JJ, Subauste MC, Saber B, Hammond SM, Prudkin L, et al. miR-93, miR-98, and miR-197 Regulate Expression of Tumor Suppressor Gene FUS1. Mol Cancer Res. 2009;7:1234–43.

137. Wu S, Huang S, Ding J, Zhao Y, Liang L, Liu T, et al. Multiple microRNAs modulate p21Cip1/Waf1 expression by directly targeting its 3′ untranslated region. Oncogene. 2010;29:2302–8. doi:10.1038/onc.2010.34.

138. Sætrom P, Heale BSE, Snøve O, Aagaard L, Alluin J, Rossi JJ. Distance constraints between microRNA target sites dictate efficacy and cooperativity. Nucleic Acids Res. 2007;35:2333–42.

139. Linsley PS, Schelter J, Burchard J, Kibukawa M, Martin MM, Bartz SR, et al. Transcripts Targeted by the MicroRNA-16 Family Cooperatively Regulate Cell Cycle Progression □ †. Mol Cell Biol. 2007;27:2240–52.

140. Mestdagh P, Boström A, Impens F, Fredlund E, Peer V, Antonellis P De, et al. The miR-17-92 microRNA cluster regulates multiple components of the TGFb pathway in neuroblastoma. Mol Cell. 2011;40:762–73.

141. Uhlmann S, Mannsperger H, Zhang JD, Horvat E-A, Schmidt C, Ku M, et al. Global microRNA level regulation of EGFR-driven cell-cycle protein network in breast cancer. Mol Syst Biol. 2012;8:1–15.

142. Lai EC. Predicting and validating microRNA targets. Genome Biol. 2004;5:115. doi:10.1186/gb-2004-5-9-115.

143. Griffiths-Jones S, Saini HK, Van Dongen S, Enright AJ. miRBase: Tools for microRNA genomics. Nucleic Acids Res. 2008;36 SUPPL. 1:154–8.

144. Hammell M, Long D, Zhang L, Lee A, Carmack CS, Han M, et al. mirWIP: microRNA target prediction based on microRNA-containing ribonucleoprotein–enriched transcripts. Nat Methods. 2008;5:813–9. doi:10.1038/nmeth.1247.

145. Maragkakis M, Alexiou P, Papadopoulos GL, Reczko M, Dalamagas T, Giannopoulos G, et al. Accurate microRNA target prediction correlates with protein repression levels. BMC Bioinformatics. 2009;10:295. doi:10.1186/1471-2105-10-295.

146. Deveci M, Catalyürek U V, Toland AE. mrSNP: software to detect SNP effects on microRNA binding. BMC Bioinformatics. 2014;15:73. doi:10.1186/1471-2105-15-73.

147. Gaidatzis D, van Nimwegen E, Hausser J, Zavolan M. Inference of miRNA targets using evolutionary conservation and pathway analysis. BMC Bioinformatics. 2007;8:69. doi:10.1186/1471-2105-8-69.

148. Betel D, Wilson M, Gabow A, Marks DS, Sander C. The microRNA.org resource: Targets and expression. Nucleic Acids Res. 2008;36 SUPPL. 1:149–53.

149. Didiano D, Hobert O. Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. Nat Struct Mol Biol. 2006;13:849–51. doi:10.1038/nsmb1138.

150. Steinkraus BR, Toegel M, Fulga TA. Tiny giants of gene regulation: Experimental strategies for microRNA functional studies. Wiley Interdiscip Rev Dev Biol. 2016;5:311–62.

151. Hammell M. Computational methods to identify miRNA targets. Semin Cell Dev Biol. 2010;21:738–44. doi:10.1016/j.semcdb.2010.01.004.

152. Vo NK, Dalton RP, Liu N, Olson EN, Goodman RH. Affinity purification of microRNA-133a with the cardiac transcription factor, Hand2. Proc Natl Acad Sci. 2010;107:19231–6. doi:10.1073/pnas.1013162107.

153. Hassan T, Smith SGJ, Gaughan K, Oglesby IK, O'Neill S, McElvaney NG, et al. Isolation and identification of cell-specific microRNAs targeting a messenger RNA using a biotinylated anti-sense oligonucleotide capture affinity technique. Nucleic Acids Res. 2013;41:1–13.

154. Haraguchi T, Nakano H, Tagawa T, Ohki T, Ueno Y, Yoshida T, et al. A potent 2'-O-methylated RNA-based microRNA inhibitor with unique secondary structures. Nucleic Acids Res. 2012;40:1–13.

155. Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, et al. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature. 2005;433:769–73. doi:10.1038/nature03315.

156. Thomson DW, Bracken CP, Goodall GJ. Experimental strategies for microRNA target identification. Nucleic Acids Res. 2011;39:6845–53.

157. Staton AA, Giraldez AJ. Use of target protector morpholinos to analyze the physiological roles of specific miRNA-mRNA pairs in vivo. Nat Protoc. 2011;6:2035–49. doi:10.1038/nprot.2011.423.

158. Dorsett Y, McBride KM, Jankovic M, Gazumyan A, Thai TH, Robbiani DF, et al. MicroRNA-155 Suppresses Activation-Induced Cytidine Deaminase-Mediated Myc-Igh Translocation. Immunity. 2008;28:630–8.

159. Bassett AR, Azzam G, Wheatley L, Tibbit C, Rajakumar T, McGowan S, et al. Understanding functional miRNA–target interactions in vivo by site-specific genome engineering. Nat Commun. 2014;5:4640. doi:10.1038/ncomms5640.

160. Chalepakis G, Stoykova A, Wijnholds J, Tremblay P, Gruss P. Pax: Gene regulators in the developing nervous system. J Neurobiol. 1993;24:1367–84.

161. Bopp D, Burri M, Baumgartner S, Frigerio G, Noll MLB-PBSR 2320. Conservation of a large protein domain in the segmentation gene paired and in funcitonally related genes of Drosophila. Cell. 1986;47:1033–40.

162. Treisman J, Harris E, Desplan C. The paired box encodes a second DNA-binding domain in the Paired homeo domain protein. Genes Dev. 1991;5:594–604.

163. Frigerio G, Surri M, Baumgartner S. f the Segmentation Gene paired an rosophila PRD Gene Set as Part f a Gene Network. Cell. 1986;47:735–46.

164. Walther C, Guenet JL, Simon D, Deutsch U, Jostes B, Goulding MD, et al. Pax: A murine multigene family of paired box-containing genes. Genomics. 1991;11:424–34.

165. Stapleton P, Weith A, Urbanek P, Kozmik Z, Busslinger M. Chomosomal localization of seven PAX gene and cloning of a novel family member, PAX-9. Nat Genet. 1993;3:292–8.

166. Neubuser A, Koseki H, Balling R. Characterization and developmental expression of Pax9. Dev Biol. 1995;170:701–16.

167. Deutsch U, Dressler GR, Gruss P. Pax-1, a member of a paired box homologuos murine gene family is expressed in segmental structures during development. Cell. 1988;53:617–25.

168. Peters H, Neubüser A, Kratochwil K, Balling R. Pax9-deficient mice lack pharyngeal pouch derivatives and teeth and exhibit craniofacial and limb abnormalities. Genes Dev. 1998;12:2735–47.

169. Dressler GR, Deutsch U, Chowdhury K, Nornes HO, Gruss P. Pax 2, a new murine paired-box containing gene and its expression in the developing excretory system. Development. 1990;109:787–95.

170. Nornes HO, Dressler GR, Knapik EW, Deutsch U, Gruss P. Spatially and temporally restricted expression of pax 2 during murine neurogenesis. Development. 1990;109:797–809.

171. Goulding MD, Calepakis G, Deutsch U, Erselius JR, Gruss P. Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. Embo J. 1991;10:1135–47.

172. Sosa-Pineda B, Chowdhury K, Torres M, Oliver G, Gruss P. The Pax4 gene is essential for differentiation of insulin-producing beta cells in the mammalian pancreas. Nature. 1997;386:399–402.

173. Asano M, Gruss P. Pax-5 is expressed at the midbrain-hindbrain boundary during mouse development. Mech Dev. 1992;39:29–39.

174. Adams B, Dorfler P, Aguzzi A, Kozmik Z, Urbanek P, Maurer-Fogy I, et al. Pax-5 encodes the transcription factor BSAP and is expressed in B lymphocytes, the developing CNS, and adult testis. Genes Dev. 1992;6:1589–607.

175. Jostes B, Walther C, Gruss P. The murine paired box gene, Pax7, is expressed specifically during the development of the nervous and muscular system. Mech Dev. 1990;33:27–37.

176. Plachov D, Chowdhury K, Walther C, Simon D, Guenet JL, Gruss P. Pax8, a murine paired box gene expressed in the developing excretory system and thyroid gland. Development. 1990;110:643–51.

177. Walther C, Gruss P. Pax-6, a murine paired box gene, is expressed in the developing CNS. Development. 1991;113:1435–49.

178. Ton CCT, Hirvonen H, Miwa H, Weil MM, Monaghan P, Jordan T, et al. Positional cloning and characterization of a paired box- and homeobox-containing gene from the aniridia region. Cell. 1991;67:1059–74.

179. van Heyningen V, Williamson K a. PAX6 in sensory development. Hum Mol Genet. 2002;11:1161–7.

180. Tang HK, Chao LY, Saunders GF. Functional analysis of paired box missense mutations in the PAX6 gene. Hum Mol Genet. 1997;6:381–6.

181. Xu W, Rould MA, Jun S, Desplan C, Pabo CO. Crystal structure of a paired domain-DNA complex at 2.5 å resolution reveals structural basis for pax developmental mutations. Cell. 1995;80:639–50.

182. Xu HE, Rould MA, Xu W, Epstein JA, Maas RL, Pabo CO. Crystal structure of the human Pax6 paired domain-DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding. Genes Dev. 1999;13:1263–75.

183. Jun S, Desplan C. Cooperative interactions between paired domain and homeodomain. Development. 1996;122:2639–50. http://www.ncbi.nlm.nih.gov/pubmed/8787739.

184. Czerny T, Busslinger M. DNA-binding and transactivation properties of Pax-6: three amino acids in the paired domain are responsible for the different sequence recognition of Pax-6 and BSAP (Pax-5). Mol Cell Biol. 1995;15:2858–71.

185. Tabata H, Koinui A, Ogura A, Nishihara D, Yamamoto H. A novel nuclear localization signal spans the linker of the two DNA-binding subdomains in the conserved paired domain of Pax6. Genes Genet Syst. 2018;93:75–81.

186. Epstein JA, Glaser T, Cai J, Jepeal L, Walton DS, Maas RL. Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing. Genes Dev. 1994;8:2022–34.

187. Kozmik Z, Czerny T, Busslinger M. Alternatively spliced insertions in the paired domain restrict the DNA sequence specificity of Pax6 and Pax8. EMBO J. 1997;16:6793–803.

188. Shaham O, Menuchin Y, Farhy C, Ashery-Padan R. Pax6: A multi-level regulator of ocular development. Prog Retin Eye Res. 2012;31:351–76. doi:10.1016/j.preteyeres.2012.04.002.

189. Tang HK, Singh S, Saunders GF. Dissection of the transactivation function of the transcription factor encoded by the eye developmental gene PAX6. J Biol Chem. 1998;273:7210–21.

190. Mikkola I, Bruun J-A, Bjorkoy G, Holm T, Johansen T. Phosphorylation of the Transactivation Domain of Pax6 by Extracellular Signal-regulated Kinase and p38 Mitogen-activated Protein Kinase. J Biol Chem. 1999;274:15115–26. doi:10.1074/jbc.274.21.15115.

191. St-Onge L, Sosa-Pineda B, Chowdhury K, Mansouri A, Gruss P. Pax6 is required for differentiation of glucagon-producing alpha-cells in mouse pancreas. Nature. 1997;387:406–9.

192. Larsson LI, St-Onge L, Hougaard DM, Sosa-Pineda B, Gruss P. Pax 4 and 6 regulate gastrointestinal endocrine cell development. Mech Dev. 1998;79:153–9.

193. Jami A, Gadi J, Lee MJ, Kim EJ, Lee MJ, Jung HS, et al. Pax6 expressed in osteocytes inhibits canonical Wnt signaling. Mol Cells. 2013;35:305–12.

194. Kimura R, Yoshizaki K, Osumi N. Dynamic expression patterns of Pax6 during spermatogenesis in the mouse. J Anat. 2015;227:1–9.

195. Sander M, Neubüser A, Kalamaras J, Ee HC, Martin GR, German MS. Genetic analysis reveals that PAX6 is required for normal transcription of pancreatic hormone genes and islet development. Genes Dev. 1997;11:1662–73.

196. Tam PPL, Behringer RR. Mouse gastrulation: The formation of a mammalian body plan. Mech Dev. 1997;68:3–25.

197. Graw J. Eye development. In: Current Topics in Developmental Biology. 2010. p. 343–86.

198. Fuhrmann S. Eye morphogenesis and patterning of the optic vesicle. Curr Top Dev Biol. 2017;93:61–84. doi:10.1016/B978-0-12-385044-7.00003-5.Eye.

199. Macdonald R, Wilson SW. Distribution of Pax6 protein during eye development suggests discrete roles in proliferative and differentiated visual cells. Dev Genes Evol. 1997;206:363–9.

200. Davis N, Yoffe C, Raviv S, Antes R, Berger J, Holzmann S, et al. Pax6 dosage requirements in iris and ciliary body differentiation. Dev Biol. 2009;333:132–42. doi:10.1016/j.ydbio.2009.06.023.

201. Makarenkova HP, Ito M, Govindarajan V, Faber SC, Sun L, McMahon G, et al. FGF10 is an inducer and Pax6 a competence factor for lacrimal gland development. Development. 2000;127:2563–72.

202. Xu S, Sunderland ME, Coles BLK, Kam A, Holowacz T, Ashery-Padan R, et al. The proliferation and expansion of retinal stem cells require functional Pax6. Dev Biol. 2007;304:713–21.

203. Manuel M, Price DJ. Role of Pax6 in forebrain regionalization. Brain Res Bull. 2005;66:387–93.

204. Osumi N, Shinohara H, Numayama-Tsuruta K, Maekawa M. Concise Review: Pax6 Transcription Factor Contributes to both Embryonic and Adult Neurogenesis as a Multifunctional Regulator. Stem Cells. 2008;26:1663–72. doi:10.1634/stemcells.2007-0884.

205. Manuel MN, Mi D, Mason JO, Price DJ. Regulation of cerebral cortical neurogenesis by the Pax6 transcription factor. Front Cell Neurosci. 2015;9 March:1–21. doi:10.3389/fncel.2015.00070.

206. de Chevigny A, Coré N, Follert P, Gaudin M, Barbry P, Béclin C, et al. miR-7a regulation of Pax6 controls spatial origin of forebrain dopaminergic neurons. Nat

Neurosci. 2012;15:1120–6. doi:10.1038/nn.3142.

207. Ericson J, Rashbass P, Schedl A, Brenner-Morton S, Kawakami A, Van Heyningen V, et al. Pax6 controls progenitor cell identity and neuronal fate in response to graded Shh signaling. Cell. 1997;90:169–80.

208. Duan D, Fu Y, Paxinos G, Watson C. Spatiotemporal expression patterns of Pax6 in the brain of embryonic, newborn, and adult mice. Brain Struct Funct. 2013;218:353–72.

209. Panayiotou E, Panayi E, Lapathitis G, Francius C, Clotman F, Kessaris N, et al. Pax6 is expressed in subsets of V0 and V2 interneurons in the ventral spinal cord in mice. Gene Expr Patterns. 2013;13:328–34. doi:10.1016/j.gep.2013.06.004.

210. Treloar HB, Miller AM, Ray A, Greer CA. Development of the olfactory system. In: The Neurobiology of Olfaction. 2010. p. Chapter 5.

211. Guo Z, Packard A, Krolewski RC, Harris MT, Manglapus GL. Expression of Pax6 and Sox2 in adult olfactory epithelium. J Comp Neurol. 2010;518:4395–418.

212. López-Mascaraque L, de Castro F. The olfactory bulb as an independent developmental domain. Cell Death Differ. 2002;9:1279–86.

213. Stoykova A, Gruss P. Roles of Pax-Genes Expression Patterns in Developing and Adult Brain as Suggested by. J Neurosci. 1994;14:1395–412.

214. Kohwi M, Osumi N, Rubenstein JLR, Alvarez-Buylla A. Pax6 Is Required for Making Specific Subpopulations of Granule and Periglomerular Neurons in the Olfactory Bulb. J Neurosci. 2005;25:6997–7003. doi:10.1523/JNEUROSCI.1435-05.2005.

215. Ninkovic J, Pinto L, Petricca S, Lepier A, Sun J, Rieger MA, et al. The transcription factor Pax6 regulates survival of dopaminergic olfactory bulb neurons via crystallin αA. Neuron. 2010;68:682–94.

216. Bonal C, Herrera PL. Genes controlling pancreas ontogeny. Int J Dev Biol. 2008;52:823–35.

217. Kredo-Russo S, Mandelbaum AD, Ness A, Alon I, Lennox KA, Behlke MA, et al. Pancreas-enriched miRNA refines endocrine cell differentiation. Development. 2012;139:3021–31. doi:10.1242/dev.080127.

218. Puschel AW, Gruss P, Westerfield M. Sequence and expression pattern of pax-6 are highly conserved between zebrafish and mice. Development. 1992;651:643–51.

219. Callaerts P, Halder G, Gehring WJ. PAX-6 in development. Annu Rev Neurosci. 1997;20:483–532.

220. Quiring R, Walldorf U, Kloter U, Gehring W. Homology of the eyeless gene of Drosophila to the Small eye gene in mice and Aniridia in humans. Science (80-). 1994;265:785–9. doi:10.1126/science.7914031.

221. Gehring, Walter; Ikeo K. Pax 6 and eye evolution. Trends Genet. 1999;9525:371–7.

222. Kammandel B, Chowdhury K, Stoykova A, Aparicio S, Brenner S, Gruss P. Distinct cis-essential modules direct the time-space pattern of the Pax6 gene activity. Dev Biol. 1999;205:79–97. doi:10.1006/dbio.1998.9128.

223. Tyas DA, Simpson TI, Carr CB, Kleinjan DA, Van Heyningen V, Mason JO, et al. Functional conservation of Pax6 regulatory elements in humans and mice demonstrated with a novel transgenic reporter mouse. BMC Dev Biol. 2006;6:1–11.

224. Halder G, Callaerts P, Gehring W. Induction of ectopic eyes by targeted expression of the eyeless gene in Drosophila. Science (80-). 1995;267:1788–92. doi:10.1126/science.7892602.

225. Wawersik S, Maas RL. Vertebrate eye development as modeled in Drosophila. Hum Mol Genet. 2000;9:917–25. doi:10.1093/hmg/9.6.917.

226. Zuber ME, Gestri G, Viczian AS, Barsacchi G, Harris WA. Specification of the vertebrate eye by a network of eye field transcription factors. Development. 2003;130:5155–67. doi:10.1242/dev.00723.

227. Xu PX, Zhang X, Heaney S, Yoon a, Michelson a M, Maas RL. Regulation of Pax6 expression is conserved between mice and flies. Development. 1999;126:383–95.

228. Morgan R. Conservation of sequence and function in the Pax6 regulatory elements. Trends Genet. 2004;20:283–7.

229. Theiler K, Varnum DS. Anatomy and Embryology Development of Dickie ' s Small Eye , a Mutation. 1978;86:81–6.

230. Hill RE, Favor J, Hogan, Brigid LM, Ton CCT, Saunders GF, Hanson IM, et al. Mouse Small eye results from mutations in a paired-like homeobox-containing gene. Nat Genet. 1991;354:552–525. doi:10.1016/S1359-6454(02)00539-6.

231. Fujiwara M, Uchida T, Osumi-Yamashita N, Eto K. Uchida rat (rSey): a new mutant rat with craniofacial abnormalities resembling those of the mouse Sey mutant. Differentiation. 1994;57:31–8.

232. Schedl A, Ross A, Lee M, Engelkamp D, Rashbass P, Van Heyningen V, et al. Influence of PAX6 gene dosage on development: Overexpression causes severe eye abnormalities. Cell. 1996;86:71–82.

233. Shaw MW, Falls HF, Neel J V. Congenital Aniridia. Am J Hum Genet. 1960;12:389–415. http://www.ncbi.nlm.nih.gov/pubmed/16692741%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1316584.

234. Hingorani M, Hanson I, van Heyningen V. Aniridia. Eur J Hum Genet. 2012;44 June:1011–7.

235. Aalfs CM, Fantes JA, Wenniger-Prick LJJM, Sluijter S, Hennekam RCM, Van Heyningen V, et al. Tandem duplication of 11p12-p13 in a child with borderline development delay and eye abnormalities: Dose effect of the PAX6 gene product? Am J Med Genet. 1997;73:267–71.

236. Aradhya S, Smaoui N, Marble M, Lacassie Y. De novo duplication 11p13 involving the PAX6 gene in a patient with neonatal seizures, hypotonia, microcephaly, developmental disability and minor ocular manifestations. Am J Med Genet Part A. 2011;155:442–4.

237. Grindley JC, Davidson DR, Hill RE. The role of Pax-6 in eye and nasal development. Development. 1995;121:1433–42.

238. Philips GT, Stair CN, Lee HY, Wroblewski E, Berberoglu MA, Brown NL, et al. Precocious retinal neurons: Pax6 controls timing of differentiation and determination of cell type. Dev Biol. 2005;279:308–21.

239. van Raamsdonk CD, Tilghman SM. Dosage requirement and allelic expression of PAX6 during lens placode formation. Development. 2000;127:5439–48.

240. Ashery-Padan R, Marquardt T, Zhou X, Gruss P. Pax6 activity in the lens primordium is required for lens formation and for correct placement of a single retina in the eye. Genes Dev. 2000;14:2701–11.

241. Nomura T, Haba H, Osumi N. Role of a transcription factor Pax6 in the developing vertebrate olfactory system. Dev Growth Differ. 2007;49:683–90.

242. Schmahl W, Knoedlseder M, Favor J, Davidson D. Defects of neuronal migration and the pathogenesis of cortical malformations are associated with Small eye (Sey) in the mouse, a point mutation at the Pax-6-locus. Acta Neuropathol. 1993;86:126–35.

243. López-Mascaraque L, García C, Blanchart A, De Carlos JA. Olfactory epithelium influences the orientation of mitral cell dendrites during development. Dev Dyn. 2005;232:325–35.

244. Nomura T, Osumi N. Misrouting of mitral cell progenitors in the Pax6/small eyerat telencephalon. Development. 2004;131:787–96. doi:10.1242/dev.00984.

245. Dellovade TL, Pfaff DW, Schwanzel-Fukuda M. Olfactory bulb development is altered in small-eye (Sey) mice. J Comp Neurol. 1998;402:402–18.

246. Dellovade TL, Pfaff DW, Schwanzel-Fukuda M. The gonadotropin-releasing hormone system does not develop in small-eye (Sey) mouse phenotype. Dev Brain Res. 1998;107:233–40.

247. Stoykova a, Fritsch R, Walther C, Gruss P. Forebrain patterning defects in Small eye mutant mice. Development. 1996;122:3453–65. http://www.ncbi.nlm.nih.gov/pubmed/8951061.

248. Bishop KM, Goudreau G, O'Leary DDM. Regulation of Area Identity in the Mammalian Neocortex by Emx2 and Pax6. Science (80-). 2000;288:344–9. doi:10.1126/science.288.5464.344.

249. Hamasaki A, Yamada Y, Kurose T, Ban N, Nagashima K, Takahashi A, et al. Adult pancreatic islets require differential pax6 gene dosage. Biochem Biophys Res Commun. 2007;353:40–6.

250. Ashery-Padan R, Zhou X, Marquardt T, Herrera P, Toube L, Berry A, et al. Conditional inactivation of Pax6 in the pancreas causes early onset of diabetes. Dev Biol. 2004;269:479–88.

251. Hart AW, Mella S, Mendrychowski J, van Heyningen V, Kleinjan DA. The Developmental Regulator Pax6 Is Essential for Maintenance of Islet Cell Function in the Adult Mouse Pancreas. PLoS One. 2013;8:1–10.

252. Wen JH, Chen YY, Song SJ, Ding J, Gao Y, Hu QK, et al. Paired box 6 (PAX6) regulates glucose metabolism via proinsulin processing mediated by prohormone convertase 1/3 (PC1/3). Diabetologia. 2009;52:504–13.

253. Marquardt T, Ashery-Padan R, Andrejewski N, Scardigli R, Guillemot F, Gruss P. Pax6 is required for the multipotent state of retinal progenitor cells. Cell. 2001;105:43–55.

254. Oron-Karni V, Farhy C, Elgart M, Marquardt T, Remizova L, Yaron O, et al. Dual requirement for Pax6 in retinal progenitor cells. Development. 2008;135:4037–47. doi:10.1242/dev.028308.

255. Estivill-Torrus G, Pearson H, van Heyningen V, Price DJ, Rashbass P. Pax6 is required to regulate the cell cycle and the rate of progression from symmetrical to asymmetrical division in mammalian cortical progenitors. Development. 2002;129:455–66.

256. Quinn JC, Molinek M, Martynoga BS, Zaki PA, Faedo A, Bulfone A, et al. Pax6 controls cerebral cortical cell number by regulating exit from the cell cycle and specifies cortical cell identity by a cell autonomous mechanism. Dev Biol. 2007;302:50–65.

257. Bel-Vialar S, Medevielle F, Pituello F. The on/off of Pax6 controls the tempo of neuronal differentiation in the developing spinal cord. Dev Biol. 2007;305:659–73.

258. Hsieh Y-W, Yang X-J. Dynamic Pax6 expression during the neurogenic cell cycle influences proliferation and cell fate choices of retinal progenitors. Neural Dev. 2009;4:32. doi:10.1186/1749-8104-4-32.

259. Kokotas H, Petersen MB. Clinical and molecular aspects of aniridia. Clin Genet. 2010;77:409–20.

260. Glaser T, Walton DS, Maas RL. Genomic structure, evolutionary conservation and aniridia mutations in the human PAX6 gene. Nat Genet. 1992;2:232–9.

261. Chauhan BK, Yang Y, Cveklová K, Cvekl A. Functional Properties of Natural Human PAX6 and PAX6(5a) Mutants. Investig Ophthalmol Vis Sci. 2004;45:385–92.

262. Singh S, Chao LY, Mishra R, Davies J, Saunders GF. Missense mutation at the C-terminus of PAX6 negatively modulates homeodomain function. Hum Mol Genet. 2001;10:911–8.

263. Sisodiya SM, Free SL, Williamson KA, Mitchell TN, Willis C, Stevens JM, et al. PAX6 haploinsufficiency causes cerebral malformations and olfactory dysfunction in humans. Nat Genet. 2001;28:214–6.

264. Yasuda T, Kajimoto Y, Fujitani Y, Watada H, Yamamoto S, Watarai T, et al. PAX6 mutation as a genetic factor common to aniridia and glucose intolerance. Diabetes. 2002;51:224–30.

265. Netland PA, Scott ML, Boyle IV JW, Lauderdale JD. Ocular and systemic findings in a survey of aniridia subjects. J AAPOS. 2011;15:562–6. doi:10.1016/j.jaapos.2011.07.009.

266. Chow RL, Altmann CR, Lang R a, Hemmati-Brivanlou a. Pax6 induces ectopic eyes in a vertebrate. Development. 1999;126:4213–22.

267. Lagutin O, Zhu CC, Furuta Y, Rowitch DH, McMahon AP, Oliver G. Six3 promotes the formation of ectopic optic vesicle-like structures in mouse embryos. Dev Dyn. 2001;221:342–9.

268. Mathers PH, Grinberg a, Mahon K a, Jamrich M. The Rx homeobox gene is essential for vertebrate eye development. Nature. 1997;387 June:603–7.

269. Manuel M, Pratt T, Liu M, Jeffery G, Price DJ. Overexpression of Pax6 results in microphthalmia, retinal dysplasia and defective retinal ganglion cell axon guidance. BMC Dev Biol. 2008;8:59. doi:10.1186/1471-213X-8-59.

270. Manuel M, Georgala PA, Carr CB, Chanas S, Kleinjan DA, Martynoga B, et al.

Controlled overexpression of Pax6 in vivo negatively autoregulates the Pax6 locus, causing cell-autonomous defects of late cortical progenitor proliferation with little effect on cortical arealization. Development. 2006;134:545–55. doi:10.1242/dev.02764.

271. Yamaoka T, Yano M, Yamada T, Matsushita T, Moritani M, Ii S, et al. Diabetes and pancreatic tumours in transgenic mice expressing Pa x 6. Diabetologia. 2000;43:332–9. http://www.ncbi.nlm.nih.gov/pubmed/10768094.

272. Kaspi H, Chapnik E, Levy M, Beck G, Hornstein E, Soen Y. Brief Report: MiR-290-295 regulate embryonic stem cell differentiation propensities by repressing Pax6. Stem Cells. 2013;31:2266–72.

273. Du Z-W, Ma L-X, Phillips C, Zhang S-C. miR-200 and miR-96 families repress neural induction from human embryonic stem cells. Development. 2013;140:2611–8. doi:10.1242/dev.092809.

274. Zhao X, Wu J, Zheng M, Gao F, Ju G. Specification and maintenance of oligodendrocyte precursor cells from neural progenitor cells: involvement of microRNA-7a. Mol Biol Cell. 2012;23:2867–77. doi:10.1091/mbc.E12-04-0270.

275. Mo JL, Liu Q, Kou ZW, Wu KW, Yang P, Chen XH, et al. MicroRNA-365 modulates astrocyte conversion into neuron in adult rat brain after stroke by targeting Pax6. Glia. 2018;66:1346–62.

276. Dulcis D, Lippi G, Stark CJ, Do LH, Berg DK, Spitzer NC. Neurotransmitter Switching Regulated by miRNAs Controls Changes in Social Preference. Neuron. 2017;95:1319-1333.e5. doi:10.1016/j.neuron.2017.08.023.

277. Bhinge A, Namboori SC, Bithell A, Soldati C, Buckley NJ, Stanton LW. MiR-375 is Essential for Human Spinal Motor Neuron Development and May Be Involved in Motor Neuron Degeneration. Stem Cells. 2016;34:124–34.

278. Yongblah K, Alford SC, Ryan BC, Chow RL, Howard PL. Protecting Pax6 3′ UTR from MicroRNA-7 Partially Restores PAX6 in Islets from an Aniridia Mouse Model. Mol Ther - Nucleic Acids. 2018;13 December:144–53. doi:10.1016/j.omtn.2018.08.018.

279. Shalom-Feuerstein R, Serror L, De La Forest Divonne S, Petit I, Aberdam E, Camargo L, et al. Pluripotent stem cell model reveals essential roles for miR-450b-5p and miR-184 in embryonic corneal lineage specification. Stem Cells. 2012;30:898–909.

280. Fujita PA, Rhead B, Zweig AS, Hinrichs AS, Karolchik D, Cline MS, et al. The UCSC genome browser database: Update 2011. Nucleic Acids Res. 2011;39 SUPPL. 1:1–7.

281. Plaza S, Dozier C, Turque N, Saule S. Quail Pax-6 (Pax-QNR) mRNAs are expressed from two promoters used differentially during retina development and neuronal differentiation. Mol Cell Biol. 1995;15:3344–53.

282. Carriere C, Plaza S, Martin P, Quatannens B, Bailly M, Stehelin D, et al. Characterization of Quail Pax-6 (Pax-QNR) Proteins Expressed in the Neuroretina. Mol Cell Biol. 1993;13:7257–66.

283. Kim J, Lauderdale JD. Analysis of Pax6 expression using a BAC transgene reveals the presence of a paired-less isoform of Pax6 in the eye and olfactory bulb. Dev Biol. 2006;292:486–505.

284. Mikkola I, Bruun JA, Holm T, Johansen T. Superactivation of Pax6-mediated Transactivation from Paired Domain-binding Sites by DNA-independent Recruitment of Different Homeodomain Proteins. J Biol Chem. 2001;276:4109–18.

285. Anderson TR, Hedlund E, Carpenter EM. Differential Pax6 promoter activity and transcript expression during forebrain development. 2002;114:171–5.

286. Epstein J, Cai J, Glaser T, Jepeal L, Maas R. Identification of A Pax Paired Domain Recognition Sequence and Evidence for DNA-dependent Conformational Changes *. J Biol Chem. 1994;269:8355–61.

287. Yamaguchi Y, Sawada J, Yamada M, Handa H. Autoregulation of Pax6 transcriptional activation by two distinct DNA-binding subdomains of the paired domain. Genes to Cells. 1997;2:255–61.

288. Azuma N, Tadokoro K, Asaka A, Yamada M, Yamaguchi Y, Handa H, et al. The Pax6 isoform bearing an alternative spliced exon promotes the development of the neural retinal structure. 2005;14:2–10.

289. Pinson J, Mason JO, Simpson TI, Price DJ. Regulation of the Pax6: Pax6(5a) mRNA ratio in the developing mammalian brain. BMC Dev Biol. 2005;5:4–7.

290. Hanson IM, Fletcher JM, Jordan T, Brown A, Taylon D, Adams RJ, et al. Mutations at the PAX6 locus are found in heterogeneous anterior segment malformations including Peters' anomaly. Nat Genet. 1994;6:168–73. doi:10.1038/ng1294-340.

291. Azuma N. PAX6 missense mutation in isolated foveal hypoplasia. Nat Genet. 1996;13:141–2.

292. Williams SC, Altmann CR, Chow RL, Hemmati-Brivanlou A, Lang RA. A highly conserved lens transcriptional control element from the Pax-6 gene. Mech Dev. 1998;73:225–9.

293. Kleinjan DA, Seawright A, Childs AJ, Van Heyningen V. Conserved elements in Pax6 intron 7 involved in (auto)regulation and alternative transcription. Dev Biol. 2004;265:462–77.

294. Kleinjan DA, Seawright A, Mella S, Carr CB, Tyas DA, Simpson TI, et al. Long-range

downstream enhancers are essential for Pax6 expression. Dev Biol. 2006;299:563–81.

295. Griffin C, Kleinjan DA, Doe B, Van Heyningen V. New 3' elements control Pax6 expression in the developing pretectum, neural retina and olfactory region. Mech Dev. 2002;112:89–100.

296. Kleinjan DJ, Coutinho P. Cis-ruption mechanisms: Disruption of cis-regulatory control as a cause of human genetic disease. Briefings Funct Genomics Proteomics. 2009;8:317–32.

297. Fantes J, Redeker B, Breen M, Boyle S, Brown J, Fletcher J, et al. Aniridia-associated cytogenetic rearrangements suggest that a position effect may cause the mutant phenotype. Hum Mol Genet. 1995;4:415–22.

298. Kleinjan DA, Seawright A, Schedl A, Quinlan RA, Danes S, Heyningen V Van. hypersensitivity , sequence comparison and transgenic analysis redefine the functional domain of PAX6. Centrum. 2001;10:2049–60.

299. Crolla JA, van Heyningen V. Frequent Chromosome Aberrations Revealed by Molecular Cytogenetic Studies in Patients with Aniridia. Am J Hum Genet. 2002;71:1138–49.

300. Lauderdale JD, Wilensky JS, Oliver ER, Walton DS, Glaser T. 3' deletions cause aniridia by preventing PAX6 gene expression. Proc Natl Acad Sci. 2000;97:13755–9.

301. Lacomme M, Medevielle F, Bourbon HM, Thierion E, Kleinjan DJ, Roussat M, et al. A long range distal enhancer controls temporal fine-tuning of PAX6 expression in neuronal precursors. Dev Biol. 2018;436:94–107. doi:10.1016/j.ydbio.2018.02.015.

302. Kim EA, Yoon TN, Ryu MJ, Kim HT, Lee SE, Kim CH, et al. Phosphorylation and transactivation of Pax6 by homeodomain-interacting protein kinase 2. J Biol Chem. 2006;281:7489–97.

303. Wang Y, Yang F, Fu Y, Huang X, Wang W, Jiang X, et al. Spatial phosphoprotein profiling reveals a compartmentalized extracellular signal-regulated kinase switch governing neurite growth and retraction. J Biol Chem. 2011;286:18190–201.

304. Olsen J, Vermeulen M, Santamaria A, Kumar C, Miller M, Jensen L, et al. Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. Sci Signals. 2010;3:125.

305. Hornbeck P V., Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E. PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations. Nucleic Acids Res. 2015;43:D512–20.

306. Yan Q, Liu W Bin, Qin J, Liu J, Chen HG, Huang X, et al. Protein phosphatase-1

modulates the function of Pax-6, a transcription factor controlling brain and eye development. J Biol Chem. 2007;282:13954–65.

307. Akimov V, Barrio-Hernandez I, Hansen SVF, Hallenborg P, Pedersen AK, Bekker-Jensen DB, et al. Ubisite approach for comprehensive mapping of lysine and n-terminal ubiquitination sites. Nat Struct Mol Biol. 2018;25. doi:10.1038/s41594-018-0084-y.

308. Tuoc TC, Stoykova A. Trim11 modulates the function of neurogenic transcription factor Pax6 through ubiquitin-proteosome system. Genes Dev. 2008;22:1972–86.

309. Pfirrmann T, Jandt E, Ranft S, Lokapally A, Neuhaus H, Perron M, et al. Hedgehog-dependent E3-ligase Midline1 regulates ubiquitin-mediated proteasomal degradation of Pax6 during visual system development. Proc Natl Acad Sci. 2016;113:10103–8.

310. Yan Q, Gong L, Deng M, Zhang L, Sun S, Liu J, et al. Sumoylation activates the transcriptional activity of Pax-6, an important transcription factor for eye and brain development. Proc Natl Acad Sci. 2010;107:21034–9.

311. Uslupehlivan M, Şener E, Deveci R. In silico analysis of Pax6 protein glycosylation in vertebrates. Comput Biol Chem. 2018;77 September:116–22. doi:10.1016/j.compbiolchem.2018.09.016.

312. Lefebvre T, Planque N, Leleu D, Bailly M, Caillet-Boudin ML, Saule S, et al. O-glycosylation of the nuclear forms of Pax-6 products in quail neuroretina cells. J Cell Biochem. 2002;85:208–18.

313. Ninkovic J, Steiner-Mezzadri A, Jawerka M, Akinci U, Masserdotti G, Petricca S, et al. The BAF complex interacts with Pax6 in adult neural progenitors to establish a neurogenic cross-regulatory transcriptional network. Cell Stem Cell. 2013;13:403–18.

314. Kim CH, An MJ, Kim DH, Kim JW. Histone deacetylase 1 (HDAC1) regulates retinal development through a PAX6-dependent pathway. Biochem Biophys Res Commun. 2017;482:735–41. doi:10.1016/j.bbrc.2016.11.103.

315. Agoston Z, Heine P, Brill MS, Grebbin BM, Hau A. Meis2 is a Pax6 co-factor in neurogenesis and dopaminergic periglomerular fate specification in the adult olfactory bulb. Stem Cells. 2014;141:28–38.

316. Kamachi Y, Uchikawa M, Tanouchi A, Sekido R, Kondoh H. Pax6 and SOX2 form a co-DNA-binding partner complex that regulates initiation of lens development. Genes Dev. 2001;15:1272–86.

317. Andersen FG, Heller RS, Petersen HV, Jensen J, Madsen OD, Serup P. Pax6 and Cdx2/3 form a functional complex on the rat glucagon gene promoter G1-element. FEBS Lett. 1999;445:306–10.

318. Andersen FG, Jensen J, Heller RS, Petersen HV, Larsson LI, Madsen OD, et al. Pax6 and Pdx1 form a functional complex on the rat somatostatin gene upstream enhancer. FEBS Lett. 1999;445:315–20.

319. Bhatia S, Bengani H, Fish M, Brown A, Divizia MT, De Marco R, et al. Disruption of autoregulatory feedback by a mutation in a remote, ultraconserved PAX6 enhancer causes aniridia. Am J Hum Genet. 2013;93:1126–34. doi:10.1016/j.ajhg.2013.10.028.

320. Wawrocka A, Krawczynski MR. The genetics of aniridia — simple things become complicated. J Appl Genet. 2018;59:151–9.

321. Chen JH, Lin W, Sun G, Huang C, Huang Y, Chen H, et al. A novel PAX6 deletion in a Chinese family with congenital aniridia. Mol Vis. 2012;18:989–95.

322. Plaisancié J, Tarilonte M, Ramos P, Jeanton-Scaramouche C, Gaston V, Dollfus H, et al. Implication of non-coding PAX6 mutations in aniridia. Hum Genet. 2018;137:831–46. doi:10.1007/s00439-018-1940-x.

323. Zheng GXY, Ravi A, Calabrese JM, Medeiros LA, Kirak O, Dennis LM, et al. A latent pro-survival function for the Mir-290-295 cluster in mouse embryonic stem cells. PLoS Genet. 2011;7.

324. Hack MA, Saghatelyan A, de Chevigny A, Pfeifer A, Ashery-Padan R, Lledo P-M, et al. Neuronal fate determinants of adult olfactory bulb neurogenesis. Nat Neurosci. 2005;8:865–72. doi:10.1038/nn1479.

325. Li CG, Eccles MR. PAX genes in cancer; friends or foes? Front Genet. 2012;3 JAN:1–7.

326. Huang B, Luo Q, Han Y, Li X, Cao L, Wu L. microRNA-223 promotes the growth and invasion of glioblastoma cells by targeting tumor suppressor PAX6. Oncol Rep. 2013;30:2263–9. doi:10.3892/or.2013.2683.

327. Pavlakis E, Tonchev AB, Kaprelyan A, Enchev Y, Stoykova A. Interaction between transcription factors PAX6/PAX6-5a and specific members of miR-183-96-182 cluster, may contribute to glioma progression in glioblastoma cell lines. Oncol Rep. 2017;37:1579–92.

328. Bai S wei, Li B, Zhang H, Jonas JB, Zhao B wen, Shen L, et al. Pax6 regulates proliferation and apoptosis of human retinoblastoma cells. Investig Ophthalmol Vis Sci. 2011;52:4560–70.

329. Wang J, Wang X, Wu G, Hou D, Hu Q. MiR-365b-3p, down-regulated in retinoblastoma, regulates cell cycle progression and apoptosis of human retinoblastoma cells by targeting PAX6. FEBS Lett. 2013;587:1779–86.

330. Liu C, Zhang F, Li T, Lu M, Wang L, Yue W, et al. MirSNP, a database of polymorphisms altering miRNA target sites, identifies miRNA-related SNPs in GWAS SNPs and eQTLs. BMC Genomics. 2012;13. doi:10.1186/1471-2164-13-661.

331. Liang CL, Hsi E, Chen KC, Pan YR, Wang YS, Juo SHH. A functional polymorphism at 3'UTR of the PAX6 gene may confer risk for extreme myopia in the Chinese. Investig Ophthalmol Vis Sci. 2011;52:3500–5.

332. Chen KC, Hsi E, Hu CY, Chou WW, Liang CL, Juo SHH. MicroRNA-328 may influence myopia development by mediating the PAX6 gene. Invest Ophthalmol Vis Sci. 2012;53:2732–9.

333. Panjwani N, Wilson MD, Addis L, Crosbie J, Wirrell E, Auvin S, et al. A microRNA-328 binding site in *PAX6* is associated with centrotemporal spikes of rolandic epilepsy. Ann Clin Transl Neurol. 2016;3:512–22. doi:10.1002/acn3.320.

334. Wang X, Zhang J, Li F, Gu J, He T, Zhang X, et al. MicroRNA identification based on sequence and structure alignment. Bioinformatics. 2005;21:3610–4.

335. Bhattacharya A, Cui Y. MiR2GO: Comparative functional analysis for microRNAs. Bioinformatics. 2015;31:2403–5.

336. Clop A, Marcq F, Takeda H, Pirottin D, Tordoir X, Bibé B, et al. A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. Nat Genet. 2006;38:813–8. doi:10.1038/ng1810.

337. Wynendaele J, Böhnke A, Leucci E, Nielsen SJ, Lambertz I, Hammer S, et al. An illegitimate microRNA target site within the 3' UTR of MDM4 affects ovarian cancer progression and chemosensitivity. Cancer Res. 2010;70:9641–9.

338. Jensen KP, Covault J, Conner TS, Tennen H, Kranzler HR, Furneaux HM. A common polymorphism in serotonin receptor 1B mRNA moderates regulation by miR-96 and associates with aggressive human behaviors. Mol Psychiatry. 2009;14:381–9. doi:10.1038/mp.2008.15.

339. Tan Z, Randall G, Fan J, Camoretti-Mercado B, Brockman-Schneider R, Pan L, et al. Allele-Specific Targeting of microRNAs to HLA-G and Risk of Asthma (PII:S0002-9297(07)63059-6). Am J Hum Genet. 2008;82:251.

340. Lal A, Navarro F, Maher CA, Maliszewski LE, Yan N, O'Day E, et al. miR-24 inhibits cell proliferation by targeting E2F2, MYC and other cell-cycle genes via binding to "seedless" 3'UTR microRNA recognition elements. Mol Cell. 2009;35:610–25.

341. Brodersen P, Voinnet O. Revisiting the principles of microRNA target recognition and mode of action. Nat Rev Mol Cell Biol. 2009;10:141–8. doi:10.1038/nrm2619.

342. Jacobs GH, Chen A, Stevens SG, Stockwell PA, Black MA, Tate WP, et al. Transterm: A database to aid the analysis of regulatory sequences in mRNAs. Nucleic Acids Res. 2009;37 SUPPL. 1:72–6.

343. Cook KB, Kazan H, Zuberi K, Morris Q, Hughes TR. RBPDB: A database of RNA-binding specificities. Nucleic Acids Res. 2011;39 SUPPL. 1:301–8.

344. Papadopoulos GL, Reczko M, Simossis VA, Sethupathy P, Hatzigeorgiou AG. The database of experimentally supported targets: A functional update of TarBase. Nucleic Acids Res. 2009;37 SUPPL. 1:155–8.

345. Xiao F, Zuo Z, Cai G, Kang S, Gao X, Li T. miRecords: An integrated resource for microRNA-target interactions. Nucleic Acids Res. 2009;37 SUPPL. 1:105–10.

346. Hausser J, Berninger P, Rodak C, Jantscher Y, Wirth S, Zavolan M. MirZ: An integrated microRNA expression atlas and target prediction resource. Nucleic Acids Res. 2009;37 SUPPL. 2:266–72.

347. Karali M, Peluso I, Gennarino VA, Bilio M, Verde R, Lago G, et al. miRNeye: a microRNA expression atlas of the mouse eye. BMC Genomics. 2010;11:715. doi:10.1186/1471-2164-11-715.

348. Collinson JM, Quinn JC, Hill RE, West JD. The roles of Pax6 in the cornea, retina, and olfactory epithelium of the developing mouse embryo. Dev Biol. 2003;255:303–12.

349. Hogan BL, Horsburgh G, Cohen J, Hetherington CM, Fisher G, Lyon MF. Small eyes (Sey): a homozygous lethal mutation on chromosome 2 which affects the differentiation of both lens and nasal placodes in the mouse. J Embryol Exp Morphol. 1986;97:95–110. http://www.ncbi.nlm.nih.gov/pubmed/3794606.

350. Hill R, Favor J, Hogan B, Ton CCT, Saunders GF, Hanson IM, et al. Mouse Small eye results from mutations in a paired-like homeobox containing gene. Nature. 1991;354.

351. Jordan T, Hanson I, Zaletayev D, Hodgson S, Prosser J, Seawright A, et al. The human PAX6 gene is mutated in two patients with aniridia. Nat Genet. 1992;1:328–32. doi:10.1038/ng0892-328.

352. Du L, Schageman JJ, Subauste MC, Saber B, Hammond SM, Prudkin L, et al. miR-93, miR-98, and miR-197 Regulate Expression of Tumor Suppressor Gene FUS1. Mol Cancer Res. 2009;7:1234–43. doi:10.1158/1541-7786.MCR-08-0507.

353. Ryan BC, Werner TS, Howard PL, Chow RL. ImiRP: a computational approach to microRNA target site mutation. BMC Bioinformatics. 2016;17:190. doi:10.1186/s12859-016-1057-y.

354. Frohman MA. Rapid amplification of complementary DNA ends for generation of

full-length complementary DNAs: Thermal RACE. Methods Enzymol. 1993;24:340–56.

355. Jurica MS, Licklider LJ, Gygi SR, Grigorieff N, Moore MJ. Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. RNA. 2002;8:426–39. doi:10.1017/S1355838202021088.

356. Golding I, Cox EC. RNA dynamics in live Escherichia coli cells. Proc Natl Acad Sci. 2004;101:11310–5. doi:10.1073/pnas.0404443101.

357. Zhou Z, Licklider LJ, Gygi SP, Reed R. Comprehensive proteomic analysis of the human spliceosome. Nature. 2002;419:182–5. doi:10.1038/nature01031.

358. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics. 2012;13:134. doi:10.1186/1471-2105-13-134.

359. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 2001;29:45e – 45. doi:10.1093/nar/29.9.e45.

360. Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biol. 2007;8:R19. doi:10.1186/gb-2007-8-2-r19.

361. Mangone M, Manoharan AP, Thierry-mieg D, Thierry- J, Han T, Mackowiak SD, et al. The landscape of C. elegans 3 prime UTRs. Science (80-). 2010;329:432–5. doi:10.1126/science.1191244.The.

362. Xia Z, Clark P, Huynh T, Loher P, Zhao Y, Chen H-W, et al. Molecular dynamics simulations of Ago silencing complexes reveal a large repertoire of admissible “seed-less” targets. Sci Rep. 2012;2:1–10.

363. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The Human Genome Browser at UCSC. Genome Res. 2002;12:996–1006. doi:10.1101/gr.229102.

364. Hamaguchi K, Leiter E. Comparison of cytokine effects on mouse pancreatic alpha-cell and beta-cell lines. Viability, secretory function, and MHC antigen expression. Diabetes. 1990;39:415–25.

365. Poitout V, Stout L, Armstrong M, Walseth T, Sorenson R, Robertson P. Morphological and Functional Characterization of βTC-6 Cells—an Insulin-Secreting Cell Line Derived From Transgenic Mice. Diabetes. 1995;44:306–13.

366. Mizusawa N, Hasegawa T, Ohigashi I, Tanaka-Kosugi C, Harada N, Itakura M, et al. Differentiation phenotypes of pancreatic islet β- and α-cells are closely related with homeotic genes and a group of differentially expressed genes. Gene. 2004;331:53–63.

367. Klein D, Misawa R, Bravo-Egana V, Vargas N, Rosero S, Piroso J, et al. MicroRNA Expression in Alpha and Beta Cells of Human Pancreatic Islets. PLoS One. 2013;8.

368. Poy MN, Eliasson L, Krutzfeldt J, Kuwajima S, Ma X, MacDonald PE, et al. A pancreatic islet-specific microRNA regulates insulin secretion. Nature. 2004;432:226–30. doi:10.1038/nature03076.

369. van de Bunt M, Gaulton KJ, Parts L, Moran I, Johnson PR, Lindgren CM, et al. The miRNA Profile of Human Pancreatic Islets and Beta-Cells and Relationship to Type 2 Diabetes Pathogenesis. PLoS One. 2013;8:1–7.

370. Bravo-Egana V, Rosero S, Molano RD, Pileggi A, Ricordi C, Domínguez-Bendala J, et al. Quantitative differential expression analysis reveals miR-7 as major islet microRNA. Biochem Biophys Res Commun. 2008;366:922–6.

371. Hackler L, Wan J, Swaroop A, Qian J, Zack DJ. MicroRNA profile of the developing mouse retina. Investig Ophthalmol Vis Sci. 2010;51:1823–31.

372. Karali M, Peluso I, Marigo V, Banfi S. Identification and characterization of microrna expressed in the mouse eye. Investig Ophthalmol Vis Sci. 2007;48:509–15.

373. Xu S, Witmer PD, Lumayag S, Kovacs B, Valle D. MicroRNA (miRNA) transcriptome of mouse retina and identification of a sensory organ-specific miRNA cluster. J Biol Chem. 2007;282:25053–66.

374. Moreau MP, Bruse SE, Jornsten R, Liu Y, Brzustowicz LM. Chronological Changes in MicroRNA Expression in the Developing Human Brain. PLoS One. 2013;8.

375. Braun J, Misiak D, Busch B, Krohn K, Hüttelmaier S. Rapid identification of regulatory microRNAs by miTRAP (miRNA trapping by RNA in vitro affinity purification). Nucleic Acids Res. 2014;42.

376. Peabody DS. The RNA binding site of bacteriophage MS2 coat protein. EMBO J. 1993;12:595–600. http://www.ncbi.nlm.nih.gov/pubmed/8440248%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC413242.

377. Yoon JH, Srikantan S, Gorospe M. MS2-TRAP (MS2-tagged RNA affinity purification): Tagging RNA to identify associated miRNAs. Methods. 2012;58:81–7. doi:10.1016/j.ymeth.2012.07.004.

378. Park CY, Choi Y, Mcmanus MT. Analysis of microRNA knockouts in mice. Hum Mol Genet. 2010;19:R169-175.

379. van Rooij E, Quiat D, Johnson BA, Sutherland LB, Qi X, Richardson JA, et al. A Family of microRNAs Encoded by Myosin Genes Governs Myosin Expression and Muscle

Performance. Dev Cell. 2009;17:662–73. doi:10.1016/j.devcel.2009.10.013.

380. Kloosterman WP, Lagendijk AK, Ketting RF, Moulton JD, Plasterk RHA. Targeted inhibition of miRNA maturation with morpholinos reveals a role for miR-375 in pancreatic islet development. PLoS Biol. 2007;5:1738–49.

381. Joglekar M V., Parekh VS, Mehta S, Bhonde RR, Hardikar AA. MicroRNA profiling of developing and regenerating pancreas reveal post-transcriptional regulation of neurogenin3. Dev Biol. 2007;311:603–12.

382. Tang X, Muniappan L, Tang G, Ozcan S. Identification of glucose-regulated miRNAs from pancreatic {beta} cells reveals a role for miR-30d in insulin transcription. RNA. 2009;15:287–93. doi:10.1261/rna.1211209.

383. El Ouaamari A, Baroukh N, Martens G a, Lebrun P, Pipeleers D, van Obberghen E. miR-375 targets 3'-phosphoinositide-dependent protein kinase-1 and regulated glucose-induced biological responses in pancreatic beta-cells. Diabetes. 2008;57:2708–17.

384. Ramachandran D, Roy U, Garg S, Ghosh S, Pathak S, Kolthur-Seetharam U. Sirt1 and mir-9 expression is regulated during glucose-stimulated insulin secretion in pancreatic ??-islets. FEBS J. 2011;278:1167–74.

385. Lovis P, Gattesco S, Regazzi R. Regulation of the expression of components of the exocytotic machinery of insulin-secreting cells by microRNAs. Biol Chem. 2008;389:305–12.

386. de Chevigny A, Core N, Follert P, Wild S, Bosio A, Yoshikawa K, et al. Dynamic expression of the pro-dopaminergic transcription factors Pax6 and Dlx2 during postnatal olfactory bulb neurogenesis. Front Cell Neurosci. 2012;6 February:1–8. doi:10.3389/fncel.2012.00006.

387. Tramontin AD, García-Verdugo JM, Lim DA, Alvarez-Buylla A. Postnatal development of radial glia and the ventricular zone (VZ): A continuum of the neural stem cell compartment. Cereb Cortex. 2003;13:580–7.

388. Fiorelli R, Azim K, Fischer B, Raineteau O. Adding a spatial dimension to postnatal ventricular-subventricular zone neurogenesis. Development. 2015;142:2109–20. doi:10.1242/dev.119966.

389. Merkle FT, Tramontin AD, García-Verdugo JM, Alvarez-Buylla A. Radial glia give rise to adult neural stem cells in the subventricular zone. Proc Natl Acad Sci U S A. 2004;101:17528–32. doi:10.1073/pnas.0407893101.

390. Ponti G, Obernier K, Guinto C, Jose L, Bonfanti L, Alvarez-Buylla A. Cell cycle and lineage progression of neural progenitors in the ventricular-subventricular zones of adult mice. Proc Natl Acad Sci. 2013;110:E1045–54. doi:10.1073/pnas.1219563110.

391. Winner B, Cooper-Kuhn CM, Aigner R, Winkler J, Kuhn HG. Long-term survival and cell death of newly generated neurons in the adult rat olfactory bulb. Eur J Neurosci. 2002;16:1681–9.

392. Kohwi M, Petryniak MA, Long JE, Ekker M, Obata K, Yanagawa Y, et al. A Subpopulation of Olfactory Bulb GABAergic Interneurons Is Derived from Emx1- and Dlx5/6-Expressing Progenitors. J Neurosci. 2007;27:6878–91. doi:10.1523/JNEUROSCI.0254-07.2007.

393. Merkle FT, Mirzadeh Z, Alvarez-Buylla A. Mosaic organization of neural stem cells in the adult brain. Science (80-). 2007;317:381–4. doi:10.1126/science.1144914.

394. Merkle FT, Fuentealba LC, Sanders TA, Magno L, Kessaris N, Alvarez-Buylla A. Adult neural stem cells in distinct microdomains generate previously unknown interneuron types. Nat Neurosci. 2014;17:207–14. doi:10.1038/nn.3610.

395. Parrish-Aungst S, Shipley M, Erdelyi F, Szabo G, Puche A. Quantitative Analysis of Neuronal Diversity in the Mouse Olfactory Bulb. J Comp Neurol. 2007;501:825–36. doi:10.1002/cne.

396. McLean JH, Shipley MT. Postmigrational Olfactory Bulb Dopaminergic Expression Neurons of Tyrosine Hydroxylase in Olfactory Bulb Dopaminergic Neurons. Methods. 1988;8 October:3658–69.

397. Needhamsen M, White RB, Giles KM, Dunlop SA, Thomas MG. Regulation of human PAX6 expression by miR-7. Evol Bioinforma. 2014;10 October 2013:107–13.

398. LAGOS-QUINTANA M. New microRNAs from mouse and human. Rna. 2003;9:175–9. doi:10.1261/rna.2146903.

399. Tessmar-Raible K, Raible F, Christodoulou F, Guy K, Rembold M, Hausen H, et al. Conserved Sensory-Neurosecretory Cell Types in Annelid and Fish Forebrain: Insights into Hypothalamus Evolution. Cell. 2007;129:1389–400.

400. Ryan BC, Lowe K, Hanson L, Gil T, Braun L, Howard PL, et al. Mapping the Pax6 3' untranslated region microRNA regulatory landscape. BMC Genomics. 2018;19:820.

401. Matsuda T, Cepko CL. Electroporation and RNA interference in the rodent retina in vivo and in vitro. Proc Natl Acad Sci. 2004;101:16–22. doi:10.1073/pnas.2235688100.

402. Chang TH, Huang HY, Hsu JBK, Weng SL, Horng JT, Huang H Da. An enhanced computational platform for investigating the roles of regulatory RNA and for identifying functional RNA motifs. BMC Bioinformatics. 2013;14 Suppl 2:S4. doi:10.1186/1471-2105-14-S2-S4.

403. Cook KB, Kazan H, Zuberi K, Morris Q, Hughes TR. RBPDB: A database of RNA-

binding specificities. Nucleic Acids Res. 2011;39 SUPPL. 1:301–8.

404. Grillo G, Turi A, Licciulli F, Mignone F, Liuni S, Banfi S, et al. UTRdb and UTRsite (RELEASE 2010): A collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs. Nucleic Acids Res. 2009;38 SUPPL.1:75–80.

405. Vencken S, Tidi H, McElvaney N, Smith S, Greene C. miR-CATCH: MicroRNA Capture Affinity Technology. RNA Interf Challenges Ther Oppor. 2014;1218:365–73.

406. Lorenz R, Bernhart SH, Höner C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA Package 2.0. Algorithms Mol Biol. 2011;6:1–14.

407. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 2003;31:3406–15.

408. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. 2002;3:RESEARCH0034. doi:10.1186/gb-2002-3-7-research0034.

409. Liu XS, Chopp M, Zhang RL, Tao T, Wang XL, Kassis H, et al. MicroRNA profiling in subventricular zone after stroke: MiR-124a regulates proliferation of neural progenitor cells through notch signaling pathway. PLoS One. 2011;6:1–11.

410. Tang X, Falls DL, Li X, Lane T, Luskin MB. Antigen-Retrieval Procedure for Bromodeoxyuridine Immunolabeling with Concurrent Labeling of Nuclear DNA and Antigens Damaged by HCl Pretreatment. J Neurosci. 2007;27:5837–44. doi:10.1523/JNEUROSCI.5048-06.2007.

411. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: An open-source platform for biological-image analysis. Nat Methods. 2012;9:676–82.

412. Preibisch S, Saalfeld S, Tomancak P. Globally optimal stitching of tiled 3D microscopic image acquisitions. Bioinformatics. 2009;25:1463–5.

413. Waters JC. Accuracy and precision in quantitative fluorescence microscopy. J Cell Biol. 2009;185:1135–48.

414. Friedman Y, Balaga O, Linial M. MicroRNA Cancer Regulation. 2013;774. doi:10.1007/978-94-007-5590-1.

415. Pinson J, Mason JO, Simpson TI, Price DJ. Regulation of the Pax6: Pax6(5a) mRNA ratio in the developing mammalian brain. BMC Dev Biol. 2005;5:4–7.

416. Ameres SL, Zamore PD. Diversifying microRNA sequence and function. Nat Rev Mol Cell Biol. 2013;14:475–88. doi:10.1038/nrm3611.

417. Fukushima N, Yokouchi K, Kawagishi K, Moriizumi T. Differential neurogenesis and gliogenesis by local and migrating neural stem cells in the olfactory bulb. Neurosci Res. 2002;44:467–73.

418. Kosaka T, Kosaka K. “ Interneurons” in the olfactory bulb revisited. Neurosci Res. 2011;69:93–9.

419. Marguerat S, Schmidt A, Codlin S, Chen W, Aebersold R, Bähler J. Quantitative analysis of fission yeast transcriptomes and proteomes in proliferating and quiescent cells. Cell. 2012;151:671–83.

420. Karlen Y, McNair A, Perseguers S, Mazza C, Mermod N. Statistical significance of quantitative PCR. BMC Bioinformatics. 2007;8:1–16.

421. Taylor SC, Nadeau K, Abbasi M, Lachance C, Nguyen M, Fenrich J. The Ultimate qPCR Experiment : Producing Publication Quality , Reproducible Data the First Time. Trends Biotechnol. 2018;xx:1–14. doi:10.1016/j.tibtech.2018.12.002.

422. Espina V, Milia J, Wu G, Cowherd S, Liotta LA. Laser capture microdissection. In: Methods in Molecular Biology. 2011. p. 213–29.

423. Ihrie RA, Shah JK, Harwell CC, Levine JH, Guinto CD, Lezameta M, et al. Persistent Sonic Hedgehog Signaling in Adult Brain Determines Neural Stem Cell Positional Identity. Neuron. 2011;71:250–62. doi:10.1016/j.neuron.2011.05.018.

424. Mullen RJ, Buck CR, Smith AM. NeuN, a neuronal specific nuclear protein in vertebrates. Development. 1992;116:201–11. http://www.ncbi.nlm.nih.gov/pubmed/1483388.

425. Pollock A, Bian S, Zhang C, Chen Z, Sun T. Growth of the Developing Cerebral Cortex Is Controlled by MicroRNA-7 through the p53 Pathway. Cell Rep. 2014;7:1184–96. doi:10.1016/J.CELREP.2014.04.003.

426. Stenman J, Toresson H, Campbell K. Identification of two distinct progenitor populations in the lateral ganglionic eminence: implications for striatal and olfactory bulb neurogenesis. J Neurosci. 2003;23:167–74. http://www.ncbi.nlm.nih.gov/pubmed/12514213.

427. Yun K, Potter S, Rubenstein JL, May D, Golonzhka O, Hoch R, et al. Gsh2 and Pax6 play complementary roles in dorsoventral patterning of the mammalian telencephalon. Development. 2001;128:193–205. doi:10.1016/J.CELL.2012.12.041.

428. Qiu M, Bulfone A, Martinez S, Meneses J, Shimamura K, Pedersen RA, et al. Null mutation of Dlx-2 results in abnormal morphogenesis of proximal first and second branchial akch derivatives and abnormal differentiation in the forebrain. Genes Dev. 1995;9:2523–38.

429. Anderson SA, Qiu M, Bulfone A, Eisenstat DD, Meneses J, Pedersen R, et al. Mutations of the homeobox genes Dlx-1 and Dlx-2 disrupt the striatal subventricular zone and differentiation of late born striatal neurons. Neuron. 1997;19:27–37.

430. Bulfone A, Wang F, Hevner R, Anderson S, Cutforth T, Chen S, et al. An olfactory bulb sensory map develops in the absence of normal projection neurons or GABAergic interneurons. Neuron. 1998;21:1273–82.

431. Corbin JG, Gaiano N, Machold RP, Langston A, Fishell G. The Gsh2 homeodomain gene controls multiple aspects of telencephalic development. Development. 2000;127:5007–20. http://www.ncbi.nlm.nih.gov/pubmed/11060228.

432. Yun K, Garel S, Fischman S, Rubenstein JLR. Patterning of the lateral ganglionic eminence by the Gsh1 and Gsh2 homeobox genes regulates striatal and olfactory bulb histogenesis and the growth of axons through the basal ganglia. J Comp Neurol. 2003;461:151–65.

433. Yoshihara S, Omichi K, Yanazawa M, Kitamura K, Yoshihara Y. Arx homeobox gene is essential for development of mouse olfactory system. Development. 2005;132:751–62.

434. Waclaw RR, Allen ZJ, Bell SM, Erdélyi F, Szabó G, Potter SS, et al. The zinc finger transcription factor Sp8 regulates the generation and diversity of olfactory bulb interneurons. Neuron. 2006;49:503–16.

435. Saino-Saito S, Cave JW, Akiba Y, Sasaki H, Goto K, Kobayashi K, et al. ER81 and CamKIV identify anatomically and phenotypically defined subsets of mouse olfactory bulb interneurons. J Comp Neurol. 2007;502:485–96.

436. Brill MS, Snapyan M, Wohlfrom H, Ninkovic J, Jawerka M, Mastick GS, et al. A Dlx2- and Pax6-Dependent Transcriptional Code for Periglomerular Neuron Specification in the Adult Olfactory Bulb. J Neurosci. 2008;28:6439–52. doi:10.1523/JNEUROSCI.0700-08.2008.

437. Havrda MC, Harris BT, Mantani A, Ward NM, Paolella BR, Cuzon VC, et al. Id2 Is Required for Specification of Dopaminergic Neurons during Adult Olfactory Neurogenesis. J Neurosci. 2008;24:14074–87.

438. Caiazzo M, Colucci-D'Amato L, Volpicelli F, Speranza L, Petrone C, Pastore L, et al. Krüppel-like factor 7 is required for olfactory bulb dopaminergic neuron development. Exp Cell Res. 2011;317:464–73. doi:10.1016/j.yexcr.2010.11.006.

439. Bonzano S, Bovetti S, Gendusa C, Peretto P, De Marchis S. Adult born olfactory bulb dopaminergic interneurons: Molecular determinants and experience-dependent plasticity. Front Neurosci. 2016;10 MAY:1–8.

440. Baker H, Liu N, Chun HS, Saino S, Berlin R, Volpe B, et al. Phenotypic differentiation

during migration of dopaminergic progenitor cells to the olfactory bulb. J Neurosci. 2001;21:8505–13. doi:21/21/8505 [pii].

441. Bovetti S, Veyrac A, Peretto P, Fasolo A, de Marchis S. Olfactory enrichment influences adult neurogenesis modulating GAD67 and plasticity-related molecules expression in newborn cells of the olfactory bulb. PLoS One. 2009;4.

442. Bonzano S, Bovetti S, Fasolo A, Peretto P, De Marchis S. Odour enrichment increases adult-born dopaminergic neurons in the mouse olfactory bulb. Eur J Neurosci. 2014;40:3450–7.

443. Marwarha G, Claycombe-Larson K, Schommer J, Ghribi O. Maternal low-protein diet decreases brain-derived neurotrophic factor expression in the brains of the neonatal rat offspring. J Nutr Biochem. 2017;45:54–66. doi:10.1016/j.jnutbio.2017.03.005.

444. Gould JM, Smith PJ, Airey CJ, Mort EJ, Airey LE, Warricker FDM, et al. Mouse maternal protein restriction during preimplantation alone permanently alters brain neuron proportion and adult short-term memory. Proc Natl Acad Sci. 2018;115:E7398–407.

445. Ponti G, Farinetti A, Marraudino M, Panzica G, Gotti S. Sex Steroids and Adult Neurogenesis in the Ventricular-Subventricular Zone. Front Endocrinol (Lausanne). 2018;9 April:1–8.

446. Oliveira-Pinto A V, Santos RM, Coutinho RA, Oliveira LM, Santos GB, Alho ATL, et al. Sexual Dimorphism in the Human Olfactory Bulb : Females Have More Neurons and Glial Cells than Males. PLoS One. 2014;9:e111733.

447. Kass MD, Czarnecki LA, Moberly AH, Mcgann JP. Differences in peripheral sensory input to the olfactory bulb between male and female mice. Nat Publ Gr. 2017;7:1–15. doi:10.1038/srep45851.

448. Chehrehasa F, Meedeniya ACB, Dwyer P, Abrahamsen G, Mackay-Sim A. EdU, a new thymidine analogue for labelling proliferating cells in the nervous system. J Neurosci Methods. 2009;177:122–30.

449. Dassi E. Handshakes and Fights: The Regulatory Interplay of RNA-Binding Proteins. Front Mol Biosci. 2017;4 September:1–8.

450. Loflin P, Chen CYA, Shyu A Bin. Unraveling a cytoplasmic role for hnRNP d in the in vivo mRNA destabilization directed by the AU-rich element. Genes Dev. 1999;13:1884–97.

451. Brennan CM, Steitz JA. HuR and mRNA stability. Cell Mol Life Sci. 2001;58:266–77.

452. Ahmad Z, Rafeeq M, Collombat P, Mansouri A. Pax6 inactivation in the adult pancreas reveals ghrelin as endocrine cell maturation marker. PLoS One. 2015;10:1–29.

453. Chen Y, Feng R, Wang H, Wei R, Yang J, Wang L, et al. High-fat diet induces early-onset diabetes in heterozygous Pax6 mutant mice. Diabetes Metab Res Rev. 2014;32:13–23. doi:10.1002/dmrr.

454. Dumortier O, Blondeau B, Duvillié B, Reusens B, Bréant B, Remacle C. Different mechanisms operating during different critical time-windows reduce rat fetal beta cell mass due to a maternal low-protein or low-energy diet. Diabetologia. 2007;50:2495–503.

455. Petry CJ, Dorling MW, Pawlak DB, Ozanne SE, Hales CN. Diabetes in old male offspring of rat dams fed a reduced protein diet. Int J Exp Diabetes Res. 2001;2:139–43. doi:10.1155/EDR.2001.139.

456. Dumortier O, Hinault C, Gautier N, Patouraux S, Casamento V, Obberghen V. Maternal protein restriction leads to pancreatic failure in offspring: Role of misexpressed microRNA-375. Diabetes. 2014;63:3416–27.

457. Jang ES, Goldman JE. Pax6 expression is sufficient to induce a neurogenic fate in glial progenitors of the neonatal subventricular zone. PLoS One. 2011;6.

458. Qin JY, Zhang L, Clift KL, Hulur I, Xiang AP, Ren BZ, et al. Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. PLoS One. 2010;5:3–6.

459. Kostic C, Chiodini F, Salmon P, Wiznerowicz M, Deglon N, Hornfeld D, et al. Activity analysis of housekeeping promoters using self-inactivating lentiviral vector delivery into the mouse retina. Gene Ther. 2003;10:818–21.

460. Li X, Zhao X, Fang Y, Duong T, Fan C, Huang C, et al. Generation of Destabilized Green Fluorescent Protein as a Transcription Reporter Generation of Destabilized Green Fluorescent Protein as a Transcription Reporter. J Biol Chem. 1998;273:34970–5.

461. Dingwall C, Laskey RA. Nuclear targeting sequences - a consensus? TIBS. 1991;16 December:478–81.

462. Cheng HYM, Alvarez-Saavedra M, Dziema H, Choi YS, Li A, Obrietan K. Segregation of expression of mPeriod gene homologs in neurons and glia: Possible divergent roles of mPeriod1 and mPeriod2 in the brain. Hum Mol Genet. 2009;18:3110–24.

Appendix

A. Equations

Equation 1. ROI subdivision macro

```
//divides current roi into equal pieces, either horizontally or vertically
macro "divide roi [F2]"{
	startMacro();

	// Ensures we have a selected ROI in the ROI manager
	function startMacro(){
		var selectedROI = roiManager("index");
		if(selectedROI < 0){
			Dialog.create("Unable to slice ROI.");
			Dialog.addString("Unable to Slice:", "You must have an ROI selected in
the ROI Manager.");
			Dialog.show();
		}else{
			doSlice(selectedROI);
		}
	}

	function doSlice(roiIdx){
		// Figure out the area of the initial ROI
		roiManager("Select", roiIdx);
		getSelectionBounds(x, y, width, height);
		getRawStatistics(totalArea);

		// Ask how many parts and if we want horizontal or vertical slicing
		Dialog.create("ROI Division Parameters");
		Dialog.addChoice("Slice Type:", newArray("horizontal", "vertical"));
		Dialog.addNumber("Number of Slices:", 2);
		Dialog.show();

		// Get the user input results
		sliceType = Dialog.getChoice();
		numSlices = abs(Dialog.getNumber());

		// calculate the target area size that we want
		var targetArea = totalArea / numSlices;

		// Setup our 2 exclusion regions
		//  this one moves
		var movingRegion = newArray(x,y,width,height);
		//  This one excludes areas we already sliced
		var previouslyIncludedRegion = newArray(0,0,0,0);
```

```
		// Figure out where we end
		var sliceEnd = 0;
		if(sliceType == "horizontal"){
			sliceEnd = width; // horizontal, so we move along the x-axis
		}else{
			sliceEnd = height; // vertical, so we move along the y-axis
		}

		// Advance region 1
		for(i=0; i < numSlices; i++){
			// Store our "best fit" region
			var bestFitRegion = newArray(4);

			// Threshold for when our slice is "close enough" to being
totalArea/numSlices in area
			minOff = 1e9;
			var timesNoErrorImproved = 0;
			for(j=0; j < sliceEnd; j++){
				// Expand moving region
				if(sliceType == "horizontal"){
					movingRegion[2] = j;
				}else{
					movingRegion[3] = j;
				}

				// Make the slice and find its area
				sliceROI(roiIdx, movingRegion, previouslyIncludedRegion);
				getRawStatistics(sliceArea);

				// Calculate how far off the area is to being
totalArea/numSlices
				off = abs(targetArea - sliceArea);
				if(off < minOff){
					timesNoErrorImproved = 0;
					// We found a tighter fit, advance best region
					minOff = off;
					bestFitRegion[0] = movingRegion[0];
					bestFitRegion[1] = movingRegion[1];
					bestFitRegion[2] = movingRegion[2];
					bestFitRegion[3] = movingRegion[3];
				}else{
					if(timesNoErrorImproved++ > 10){
						j = sliceEnd; // terminate search
					}
				}
			}

			// Store a slice using our last, best fit region
			addROISlice(roiIdx, bestFitRegion, previouslyIncludedRegion);

			// Update previously included region to include the latest area that we
captured
			if(sliceType == "horizontal"){
```

```
                                        // calculate end coordinate of last best fit region
                                        previouslyIncludedRegion[0] = bestFitRegion[0] +
bestFitRegion[2];
                                        previouslyIncludedRegion[1] = y;
                                        // set height to everything the last best fit region didn't
cover
                                        previouslyIncludedRegion[2] = width - bestFitRegion[2];
                                        previouslyIncludedRegion[3] = height;
                                }else{
                                        // calculate end coordinate of last best fit region
                                        previouslyIncludedRegion[0] = x;
                                        previouslyIncludedRegion[1] = bestFitRegion[1] +
bestFitRegion[3];
                                        previouslyIncludedRegion[2] = width;
                                        // set height to everything the last best fit region didn't
cover
                                        previouslyIncludedRegion[3] = height - bestFitRegion[3];
                                }
                        }
                }

                // Makes a "slice" by subtracting 2 rectangles from the selected ROI
                function sliceROI(roiIdx, rect1, rect2){
                        roiManager("select", roiIdx);
                        if(rect1[2] > 0 && rect1[3] > 0){
                                setKeyDown("alt");
                                makeRectangle(rect1[0],rect1[1],rect1[2],rect1[3]);
                        }
                        if(rect2[2] > 0 && rect2[3] > 0){
                                setKeyDown("alt");
                                makeRectangle(rect2[0],rect2[1],rect2[2],rect2[3]);
                        }
                }

                function addROISlice(roiIdx, rect1, rect2){
                        sliceROI(roiIdx, rect1, rect2);
                        run("Copy"); // TODO: See if this is necessary
                        run("Add to Manager");
                        run("Measure");
                        wait(50); // TODO: See if this is necessary
                }
}
```

Equation 2. Pseudocode of the ImiRP mutation generation algorithm

```
do:
        copy input_sequence to mutant
        for each site in mutationites:
                for each position in site:
                        set nuc to select_random(allowable_nucleotides)
                        overwrite mutant[position] to nuc
                end
        end
        if mutant acceptable:
                output mutant to callback
        end
while(further mutations are needed)
```

Equation 3. Pseudocode for ImiRP's microRNA target site prediction algorithm

```
for each five_prime in mirbase:
        do:
                set sequence_chunk to sequence[window_start, window_end]
                set match_result to check_conditions(five_prime, sequence_chunk)
                increment window_start
                increment window_end
        while(window_end < sequence_end)
end

function check_conditions(five_prime, sequence_chunk):

end
```

Equation 4. Number of mutant sequences generated using different mutagenesis strategies

"N" represents the number of nucleotide positions to mutate per predicted MRE. ImiRP introduces mutations into 6 nucleotide regions complementary to the miRNA "seed region" (positions 2-7 of the miRNA, starting at the miRNA 5' end). The number of nucleobases to use for mutagenesis is represented by "n", with n≤4. "x" and "y" designate the number of specified MREs per independent region and the number of independent regions, respectively. (A) The number of possible mutant sequences produced when at least two nucleotide substitutions are introduced per MRE using any permutation of the four nucleobases. In this example, three non-overlapping MREs are present in a single independent region. The number of possible mutants is greater than 60 billion. (B) When exactly two adjacent nucleotide substitutions are introduced per MRE using any of the four nucleobases to generate mutations, the number of possible mutants for a single independent region containing three non-overlapping MREs is reduced to approximately 90,000. (C) If two independent regions containing three MREs each are mutated simultaneously by introducing two adjacent nucleotide changes using any of the four nucleobases, over 8 billion mutant sequences are possible. (D) However, mutating predicted MREs that are more than seven nucleotides apart independently of one another significantly reduces the number of redundant mutations generated. When two independent regions each containing three non-overlapping MREs are mutated in isolation, the number of mutant sequences that need to be processed is reduced to 180,000.

A $$Total\ Mutant\ Sequences = \left[\sum_{i=2}^{N} \binom{N}{i} (n-1)^i\right]^x$$

$$= \left[\sum_{i=2}^{6} \binom{6}{i} (4-1)^i\right]^3$$

$$= \left[\binom{6}{2}3^2 + \binom{6}{3}3^3 + \binom{6}{4}3^4 + \binom{6}{5}3^5 + \binom{6}{6}3^6\right]^3$$

$$\approx 6.8 \times 10^{10}$$

B $$Total\ Mutant\ Sequences = \left[\binom{N-1}{1} (n-1)^2\right]^x$$

$$= \left[(6-1)(4-1)^2\right]^3$$

$$\approx 9.1 \times 10^4$$

C $$Total\ Mutant\ Sequences = \left[\left[\binom{N-1}{1}(n-1)^2\right]^x\right]^y$$

$$= \left[\left[(6-1)(4-1)^2\right]^3\right]^2$$

$$\approx 8.3 \times 10^9$$

D $$Total\ Mutant\ Sequences = \left[\left[\binom{N-1}{1}(n-1)^2\right]^x\right] \times y$$

$$= \left[\left[(6-1)(4-1)^2\right]^3\right] \times 2$$

$$\approx 1.5 \times 10^5$$

B. Tables

Table 10. MicroRNAs predicted to target *Pax6*

miRNA	Position	Type	MRE	Conservation (%)	Predicted By
759-5p	44	7mer-m8	CACTCTG	poor	P
182-5p	49	6mer	TGTCAA	poor	P, ImiRP
96-5p	49	6mer	TGTCAA	poor	P, ImiRP
376c-3p	58	6mer	CTATGT	poor	P, ImiRP
683-3p	68	7mer	ACAGCAG	poor	P
365-3p	77	7mer-m8+G:U	GGGTATT	poor	P, ImiRP
691-3p	82	8mer	TCAGGAAA	poor	P
692-3p	92	8mer	AAAGAGAA	poor	P
302c-3p	111	7mer-m8	AAGCACT	78	Ts, P, ImiRP
291a-3p	112	7mer-m8	AGCACTT	70	P, ImiRP
294-3p	112	7mer-m8	AGCACTT	70	P, ImiRP
295-3p	112	7mer-m8	AGCACTT	70	P, ImiRP
302a-3p	112	7mer-m8	AGCACTT	70	Ts, P, ImiRP
302b-3p	112	7mer-m8	AGCACTT	70	Ts, P
302d-3p	112	7mer-m8	AGCACTT	70	Ts, P, ImiRP
106a-5p	119	6mer	CACTTT	poor	P, ImiRP
340-5p	121	8mer+G:U	CTTTGTAA	mouse	P, ImiRP
145-5p	137	7mer-m8	AACTGGA	mouse	P, ImiRP
9-5p	163	8mer+G:U	CCAAGGA	mouse	P, ImiRP
493-3p	168	7mer-m8+G:U	GACCTTT	90	P
540-5p	168	7mer+G:U	GACCTTT	90	P, ImiRP
675-3p	175	7mer-m8+G:U	GCGTACA	mouse and rat	P, ImiRP
511-5p	182	7mer-m8+G:U	GAAGGCA	78	P, ImiRP
187-3p	184	7mer-m8+G:U	AGGCACG	62	P, ImiRP
1190-3p	194	7mer+G:U	TCAGTTG	92	P
375-3p	201	7mer-A1	GAACAAA	98	Mir, P, Ts
1194-3p	232	7mer-m8+G:U	ATTCATT	92	P
500-3p	242	7mer-m8+G:U	GGTGTAT	95	P, ImiRP
501-3p	242	7mer-m8+G:U	GGTGTAT	95	P
466f-3p	243	6mer	GTGTAT	95	P
369-3p	245	7mer-m8	GTATTAT	100	Di, P, ImiRP
340-5p	251	7mer-A1+G:U	TTTGTAA	98	ImiRP
365-3p	260	7mer-m8	GGGCATT	98	Mc, Mir, P, Ts, ImiRP
369-3p	271	7mer-A1+G:U	TGTTATA	83	ImiRP
335-3p	278	8mer+G:U	ATGAAGAA	mouse and rat	P, ImiRP

375-3p	288	6mer	GAACAA	95	P
132-3p	298	7mer-m8+G:U	GGCTGTT	mouse and rat	P, ImiRP
212-3p	298	7mer-m8+G:U	GGCTGTT	mouse and rat	P, ImiRP
421-3p	300	8mer+G:U	CTGTTGGA	mouse and rat	P, ImiRP
127-3p	312	7mer-m8+G:U	GGATCTG	mouse and rat	P, ImiRP
376c-3p	315	7mer-m8+G:U	TCTGTGT	68	P, ImiRP
15a-5p	319	7mer-m8+G:U	TGTTGCT	mouse	P, ImiRP
15b-5p	319	7mer-m8+G:U	TGTTGCT	mouse	P, ImiRP
16-5p	319	7mer-m8+G:U	TGTTGCT	mouse	P, ImiRP
195-5p	319	7mer-m8+G:U	TGTTGCT	mouse	P, ImiRP
497-5p	319	7mer-m8+G:U	TGTTGCT	mouse	P, ImiRP
693-5p	328	7mer-m8+G:U	TGTGGTT	mouse and rat	P
710-5p	359	7mer-m8+G:U	AGATTTG	poor	P
183-5p	364	6mer	TGCCAT	28	P, ImiRP
499-5p	379	7mer-m8+G:U	GTTTTA	92	P, ImiRP
496-3p	399	7mer-m8+G:U	AATACTT	90	P, ImiRP
455-3p	428	8mer	TGGACTGA	88	P, ImiRP
381-3p	445	7mer-m8+G:U	TTGTAT	88	P, ImiRP
1a-3p	452	6mer	CATTCC	90	P, ImiRP
206-3p	452	6mer	CATTCC	90	P, ImiRP
298-5p	456	7mer-m8+G:U	CCTTTGC	83	P
300-3p	458	8mer+G:U	TTTGCATA	95	P
675-3p	461	8mer+G:U	GCATATAA	95	P, ImiRP
466k	483	7mer-m8	CACACAC	mouse	P
466d-5p	484	7mer-m8	CACACAC	mouse	P
878-5p	489	8mer	ACTAGATA	mouse and rat	P, ImiRP
448-3p	496	8mer+G:U	ATATGTAA	mouse and rat	P, ImiRP
450b-5p	499	7mer-A1+G:U	TGTAAAA	mouse and rat	ImiRP
129-5p	500	7mer-m8+G:U	GTAAAAA	mouse and rat	P, ImiRP
196a-5p	506	7mer-m8+G:U	ACTATCT	mouse	P, ImiRP
196b-5p	506	7mer-m8+G:U	ACTATCT	mouse	P, ImiRP
421-3p	511	7mer-m8+G:U	CTGTTGG	80	P, ImiRP
742-3p	515	7mer+G:U	TGGTTTT	95	P, ImiRP
7a-5p	517	8mer+G:U	GTTTTCCA	88	P, ImiRP
7b-5p	517	8mer+G:U	GTTTTCCA	88	P
450b-3p	519	8mer+G:U	TTTCCAAA	100	P, ImiRP
9-5p	522	6mer	CCAAAG	100	Mir, P, ImiRP
376c-3p	545	7mer-m8+G:U	TTTATGT	100	P, ImiRP
669i	547	7mer+G:U	TATGTGC	90	P
448-3p	548	8mer+G:U	ATGTGCAA	85	P, ImiRP
450b-5p	551	7mer-A1	TGCAAAA	85	P, ImiRP

129-5p	552	8mer	GCAAAAAA	85	Mir, P, Ts, ImiRP
207-3p	579	7mer-m8	GGAGAAG	mouse and rat	P
511-5p	596	7mer-m8+G:U	AAAGGTA	92	P, ImiRP
329-3p	605	6mer	GTGTGT	85	Mir, P, ImiRP
362-3p	605	6mer	GTGTGT	85	Mir, P, ImiRP
1197	607	7mer-m8+G:U	GTGTCTT	98	P
124-3p	607	7mer-m8+G:U	GTGTCTT	98	P, ImiRP
495	627	6mer	TTTGTT	90	Mir, P
376c-3p	632	7mer-m8+G:U	TTTATGT	92	P, ImiRP
182-5p	636	7mer-A1+G:U	TGTCAAA	92	ImiRP
96-5p	636	7mer-A1+G:U	TGTCAAA	92	ImiRP
200b-3p	648	6mer	AGTATT	95	Mir, P, ImiRP
200c-3p	648	6mer	AGTATT	95	Mir, P, ImiRP
429-3p	648	6mer	AGTATT	95	Mir, P, ImiRP
7a-5p	655	7mer-m8	GTCTTCC	100	Mc, Mir, P, Ts, ImiRP
7b-5p	655	7mer-m8	GTCTTCC	100	Mc, Mir, P, Ts
876-5p	665	7mer-m8	GAAATCC	100	Mir, P
590-3p	690	8mer+G:U	TAAAGTTA	95	P
203-3p	698	6mer	ATTTCA	95	Mir, P, ImiRP
340-5p	704	6mer	TTTATA	100	Mir, P, ImiRP
410-3p	705	6mer	TTATAT	100	Di, Mir, P, ImiRP
26a-5p	708	7mer-m8+G:U	TATTTGA	100	P, ImiRP
26b-5p	708	7mer-m8+G:U	TATTTGA	100	P, ImiRP
496-3p	719	7mer-m8	AATACTC	74	P, ImiRP
494	731	8mer+G:U	ATGTTTTA	88	P
340-5p	734	7mer-m8+G:U	TTTTATA	100	P, ImiRP
1187	739	7mer-m8+G:U	TACACAT	95	P
694-3p	744	7mer-m8	ATTTTCA	79	P
675-3p	749	7mer-A1+G:U	CATGCAA	82	ImiRP
186-5p	764	6mer	TTCTTT	85	P, ImiRP
129-5p	778	7mer-m8+G:U	GCAAAAG	66	P, ImiRP
878-5p	795	7mer-A1+G:U	TTAGATA	95	P, ImiRP
101-3p	811	6mer	TACTGT	95	P, ImiRP
144-3p	811	6mer	TACTGT	95	P, ImiRP
132-3p	812	6mer	ACTGTT	95	P, ImiRP
212-3p	812	6mer	ACTGTT	95	P, ImiRP
679-5p	818	7mer-m8	CACAGTC	72	P, ImiRP
500-3p	835	6mer	GTGCAT	100	P, ImiRP
501-3p	835	6mer	GTGCAT	100	P
590-3p	842	7mer-m8+G:U	TAGAATT	82	P
1a-3p	849	6mer	CATTCC	84	P, ImiRP

206-3p	849	6mer	CATTCC	84	P, ImiRP
450b-3p	851	7mer-m8+G:U	TTCCTAA	84	P, ImiRP

Position refers to the 3'UTR position of the first nucleotide of the MRE. Position 1 is the first nucleotide after the stop codon. Type refers to the MRE type. MRE is the 3'UTR sequence predicted to bind the miRNA seed. Conservation refers to the % conservation of the MRE sequence in orthologous mammalian *Pax6* 3'UTR sequences calculated using the UCSC genome browser. Several miRNA prediction software tools were used: PITA (P), TargetScan (Ts), Miranda (Mir), DianaMicroT (Di), MicroCosm (Mc) and ImiRP.

C. Figures

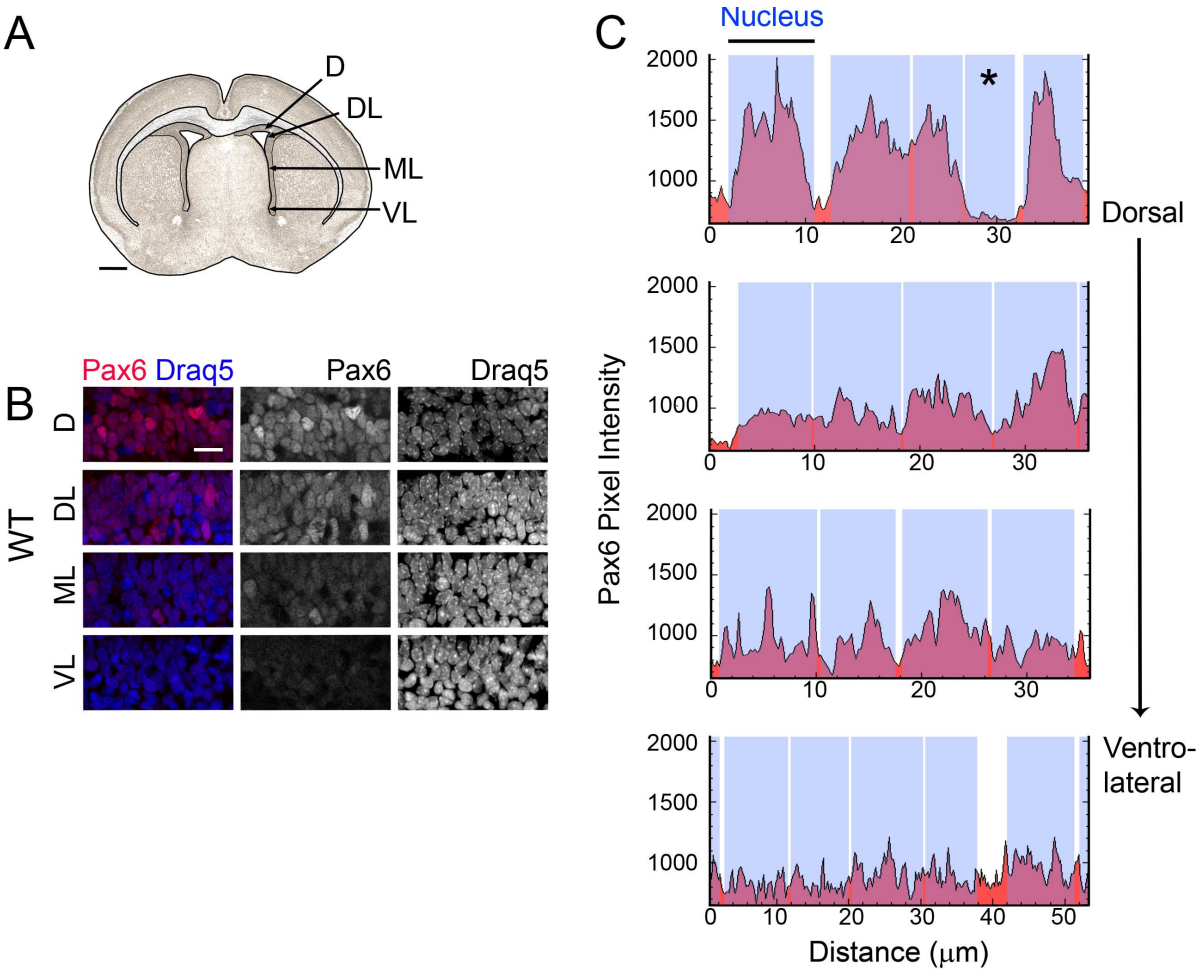


Figure 46. High magnification images of PAX6 immunofluorescence in the P1 VZ

(A) High magnification images were collected from the D, DL, mid-lateral (ML) and VL VZ of WT P1 brain sections following IF for PAX6. Scale bar – 500 µm. Images were collected using the 60X oil objective. (B) Representative WT images. Draq5 labels all nuclei, scale bar – 10 µm. (C) Nuclear localization of PAX6 protein in the D to VL P1 VZ. Pixel intensity of PAX6 IF from high magnification images is shown in red with the location of nuclei from Draq5 staining shown in blue. Note occasional nuclei lacking PAX6 (*).

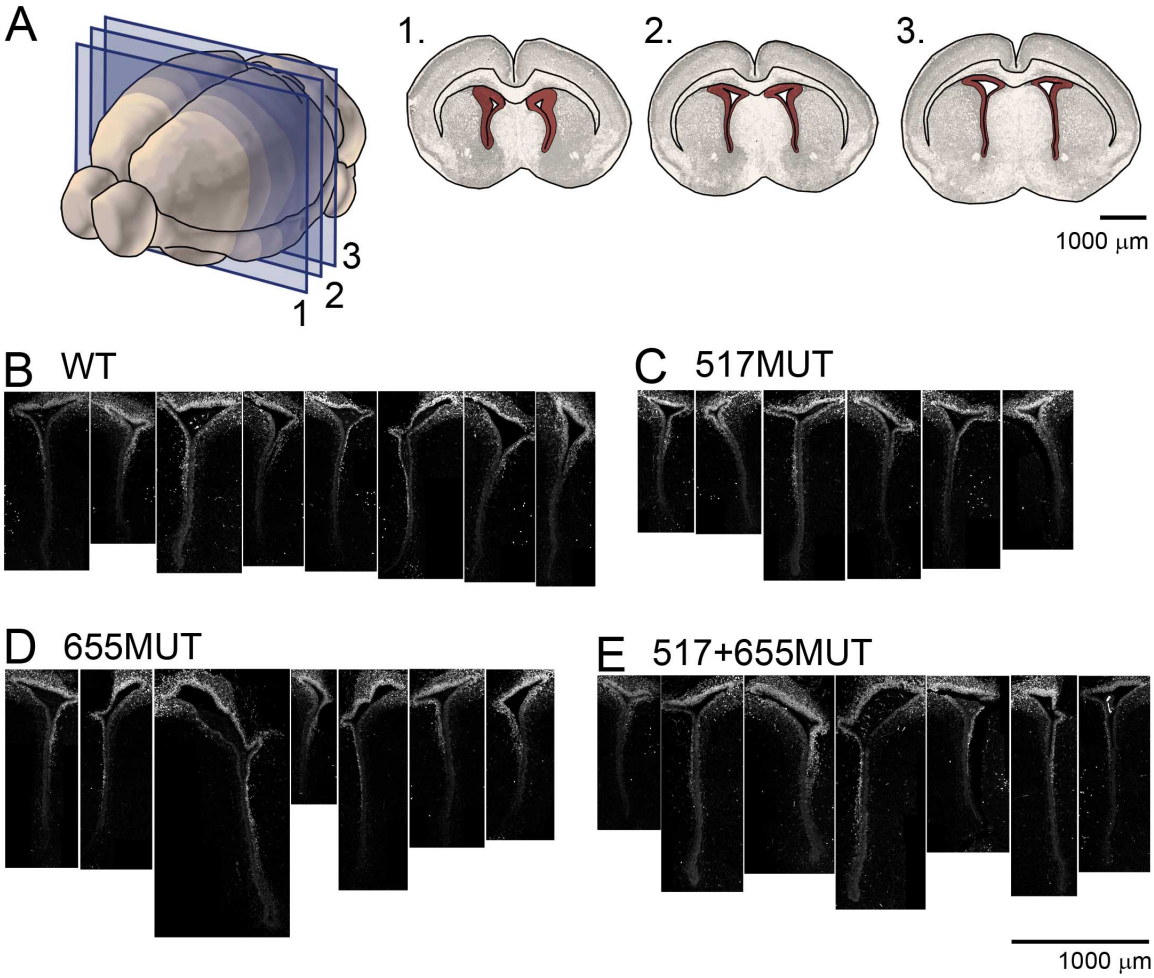


Figure 47. Rostral-caudal V-SVZ sectioning plane and PAX6 immunofluorescence intensity

(A) Schematic of the P1 brain showing with coronal sections prepared at different positions along the rostral-caudal axis reveal differences in LV length. Red indicates V-SVZ. 1. Represents approximately 1 mm into the brain, starting from the rostral end (not including OBs). 3. Represents approximately 1.5 mm into the brain. Scale bar – 1000 μm. Coronal section schematics were generated using images from the Allen developing mouse brain reference atlas. Representative PAX6 IF images from WT (B), 517MUT (C), 655MUT (D) and 517+655MUT (E) P1 VZ. Images were flat-field corrected and the background fluorescence was subtracted using Fiji ImageJ software. Image brightness was adjusted post-analysis in photoshop to improve visibility, and all images were adjusted identically. Note differences in ventricle size, because of different relative rostral-caudal positions of the prepared sections. Scale bar – 1000 μm.

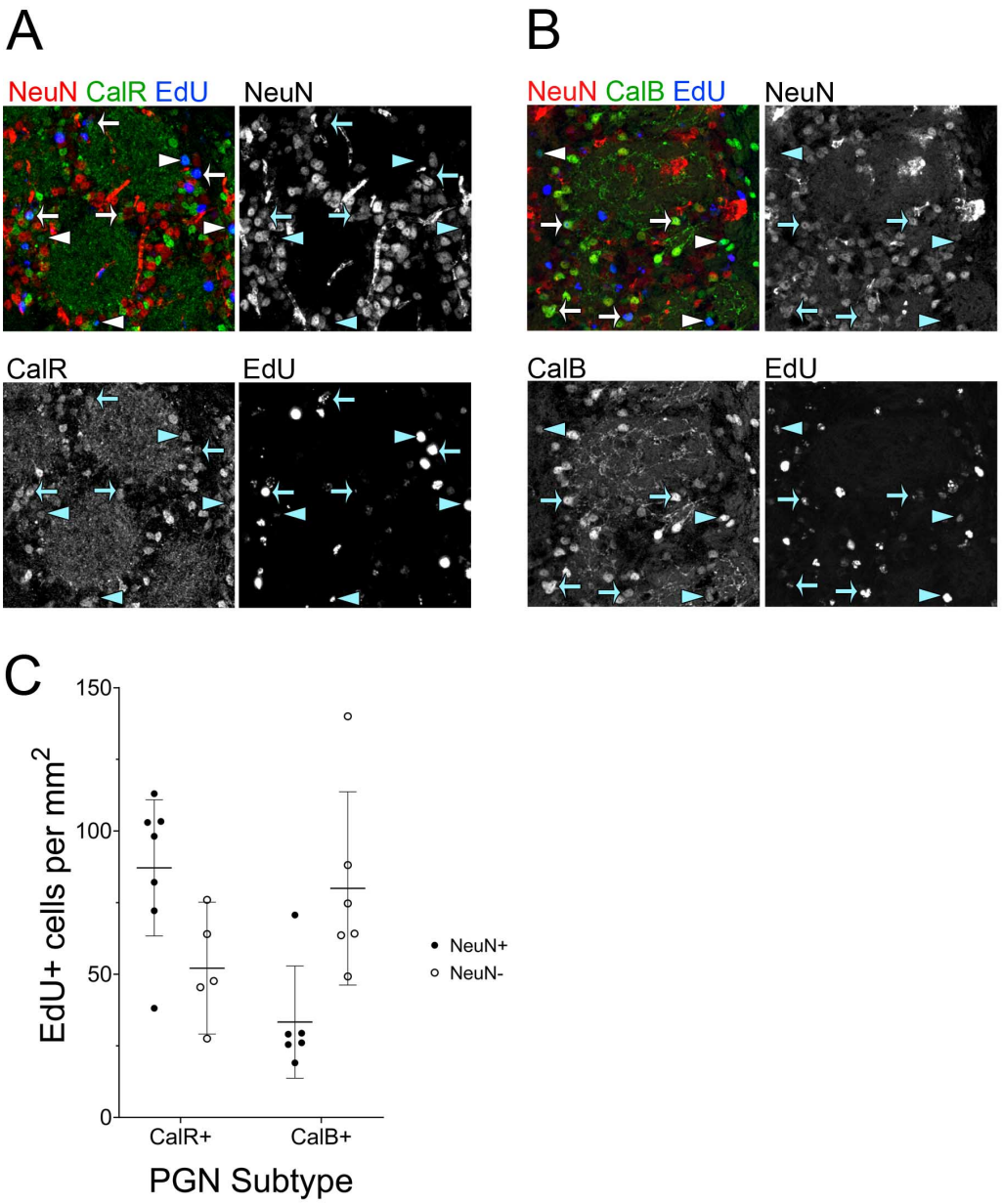


Figure 48. A population of CalR and CalB-positive cells in the GL lacking NeuN expression

(A-B) P1 pups were injected twice with 10 µg/g EdU to label V-SVZ stem cells in S-phase. Olfactory bulbs were harvested 60 days post-injection for sectioning, immunofluorescence (IF) and microscopy. Photomicrograph of P61 GL stained for NeuN (red), CalR or CalB (green) and EdU (blue). Arrows indicate EdU labeled cells immunopositive for both CalR/CalB and NeuN. Arrowheads indicate CalR/CalB positive cells that lack NeuN. (C) Quantification of CalR and CalB-positive cells in the GL of the olfactory bulb. Cells are expressed relative to GL area. n=5-7, +/- 95% C.I.

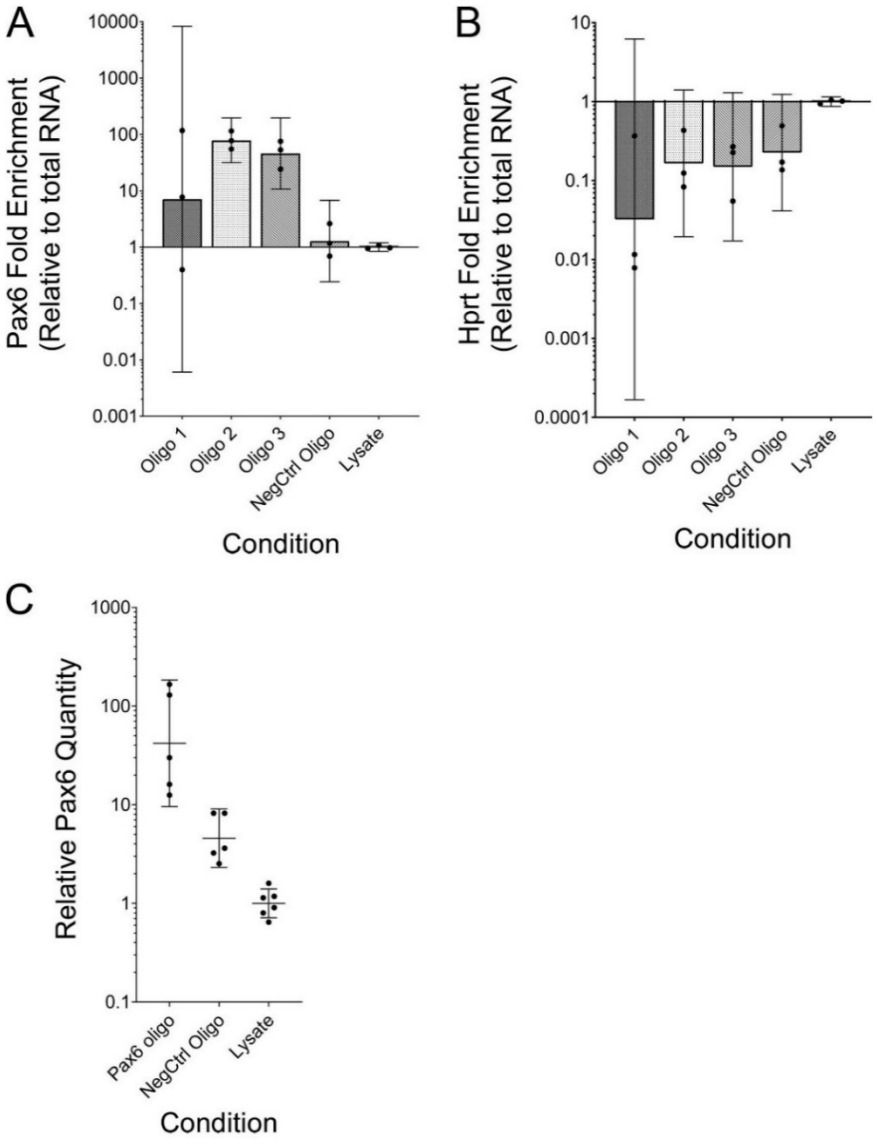


Figure 49. Testing of oligonucleotides for affinity purification of Pax6 mRNA for miR-CATCH

Affinity purification of *Pax6* from unfixed αTC1-6 cells. Three biotinylated oligonucleotides complementary to *Pax6* (Oligos 1, 2 and 3), and one non-specific oligo (NegCtrl oligo) were used for affinity purification. Quantitative PCR (qPCR) was used to assay for *Pax6* (A) and a non-specific, ubiquitously expressed target, *Hprt* (B). Total RNA from αTC1-6 cell lysate was used as a normalizing control. Fold enrichment of the targets *Pax6* and *Hprt* with affinity purification was expressed relative to the amount of target in total RNA, using Pfaffl's method. (C) Affinity purification of *Pax6* from PFA-fixed E12.5 dorsal telencephalon. The *Pax6* oligo represents Oligo 2 from A and B. *Pax6* was expressed relative to the non-specific target *Tbp* using Pfaffl's method [359]. Note lower enrichment of *Pax6* in the PFA-fixed condition. Data represents the geometric mean +/- 95% C.I.

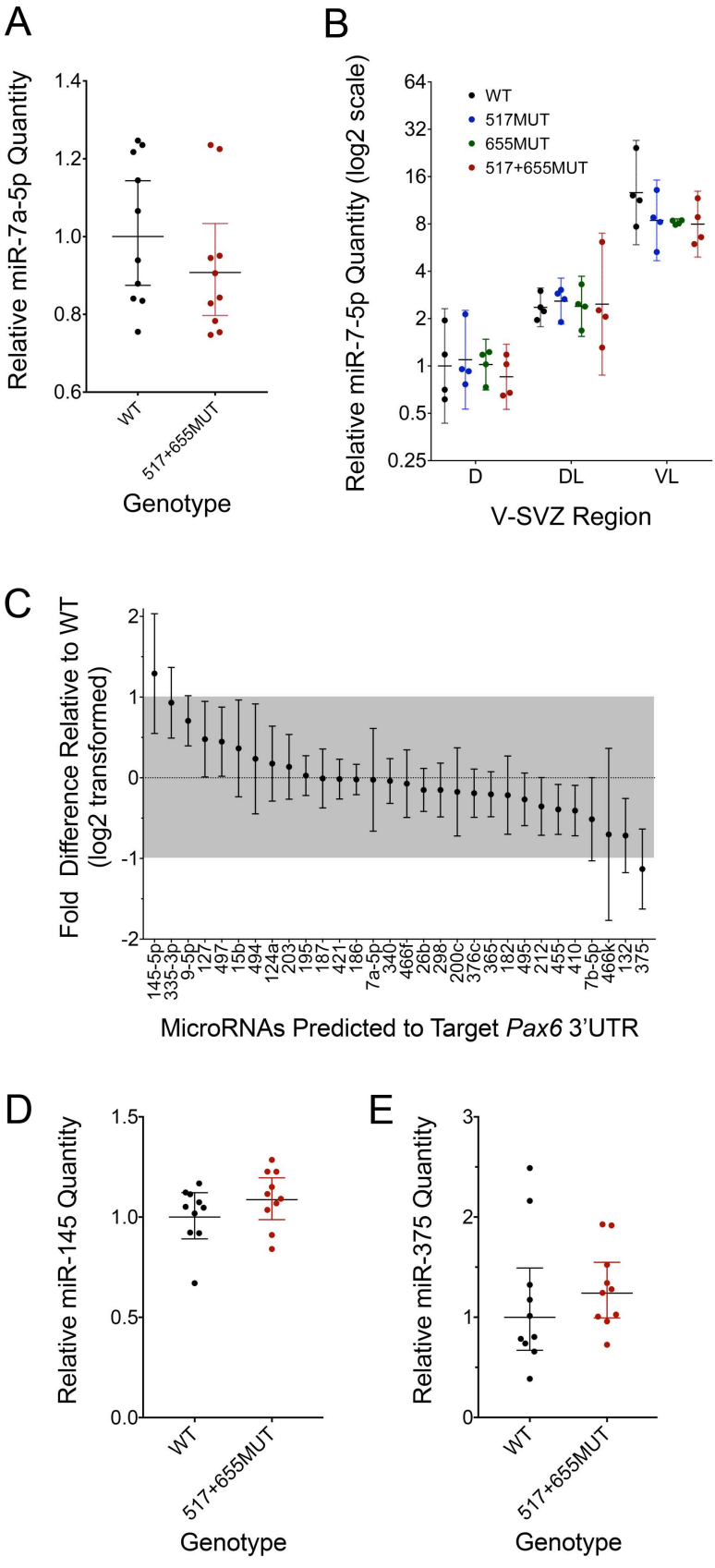

A
Relative miR-7a-5p Quantity
1.4
1.2
1.0
0.8
0.6
WT
517+655MUT
Genotype
B
Relative miR-7-5p Quantity (log2 scale)
64
32
16
8
4
2
1
0.5
0.25
WT
517MUT
655MUT
517+655MUT
D
DL
VL
V-SVZ Region
C
Fold Difference Relative to WT
(log2 transformed)
2
1
0
-1
-2
145-5p
335-3p
9-5p
127
497
15b
494
124a
203
195
187
421
186
7a-5p
340
466f
26b
298
200c
376c
365
182
495
212
455
410
7b-5p
466k
132
375
MicroRNAs Predicted to Target Pax6 3'UTR
D
Relative miR-145 Quantity
1.5
1.0
0.5
0.0
WT
517+655MUT
Genotype
E
Relative miR-375 Quantity
3
2
1
0
WT
517+655MUT
Genotype

Figure 50. Expression of miRNAs predicted to target Pax6 in 517+655MUT mice relative to WT

(A) Relative abundance of miR-7-5p in P1 pooled D and L V-SVZ and (B) in D, DL and VL V-SVZ from WT and miR-7-5p MRE mutant mice. miR-7a-5p was quantified using SYBR Green RT-qPCR and expressed relative to U6 snRNA, miR-26a and miR-106a internal controls. Relative miR-7-5p quantities were normalized to WT. n=10 for (A) and n=4 for (B). (C) Abundance of miRNAs predicted to target the *Pax6* 3'UTR in *517+655MUT* P1 pooled D and L V-SVZ relative to WT. In log2 transformed data, 0 (dotted line) indicates no change in the level of expression relative to WT. miRNAs that were less than two-fold up or down regulated (grey box) were not considered further. miRNAs were quantified using TaqMan multiplex reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) array cards and were expressed relative to U6 snRNA, miR-16, miR-26a and miR-106a internal controls using Pfaffl's method with multiple reference genes [408]. n=4. Relative abundance of miR-145-5p (D) and miR-375-3p (E) in P1 pooled D and L V-SVZ from WT and miR-7-5p MRE mutant mice. miRNAs were quantified using SYBR Green RT-qPCR and expressed relative to U6 snRNA, miR-26a and miR-106a internal controls. Relative miRNA quantities were normalized to WT. n=10. Each point represents 3 pups of pooled tissue. Data represents geometric mean +/- 95% C.I.

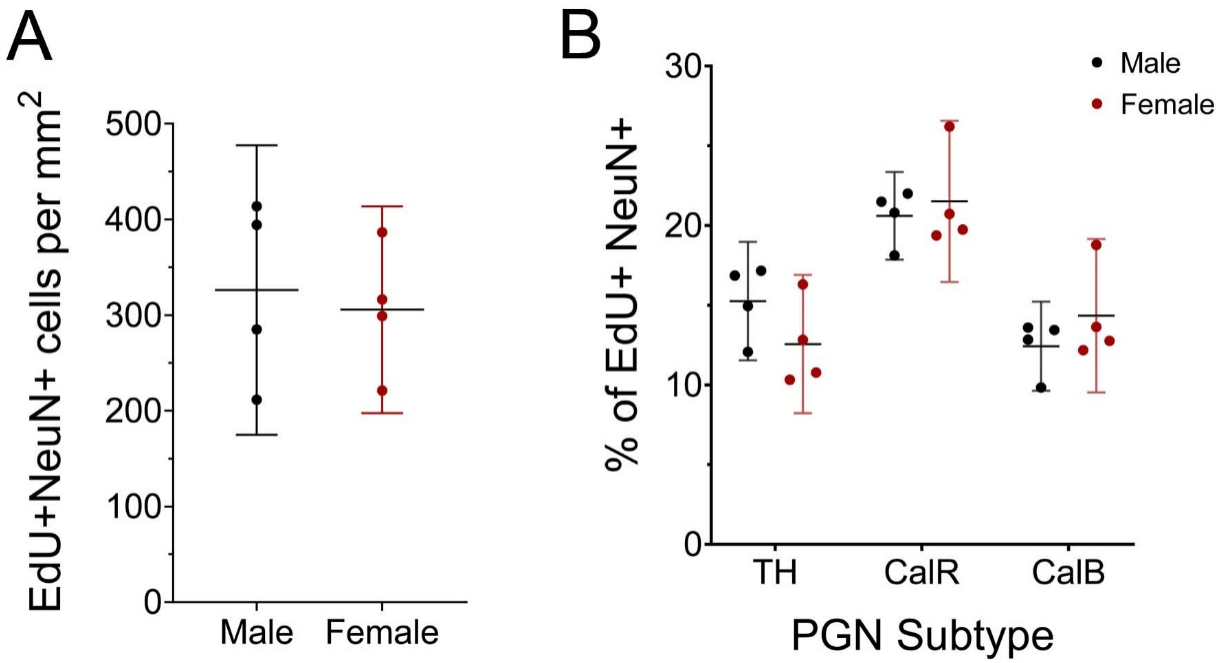
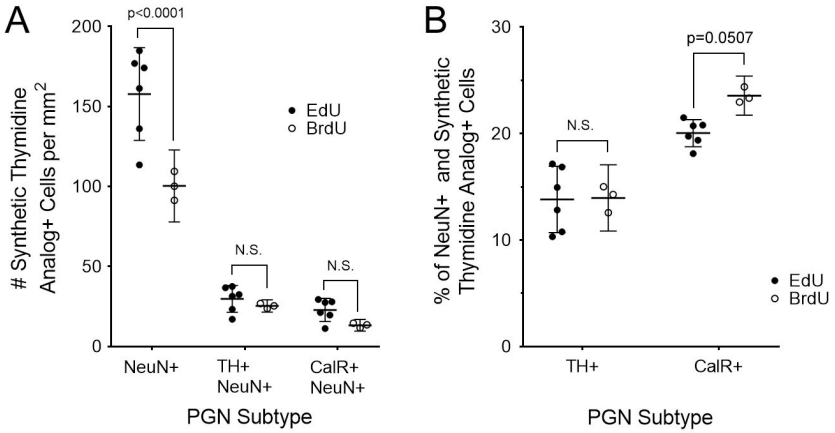
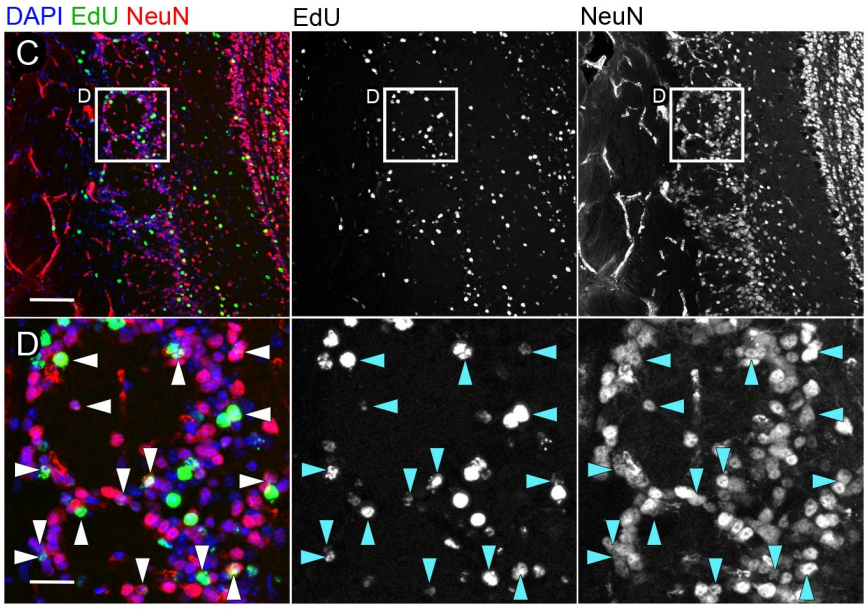


Figure 51. Sex differences in the PGN fate of V-SVZ NSCs

Male and female P1 pups were injected twice with 10 µg/g EdU to label V-SVZ stem cells in S-phase. Olfactory bulbs were harvested 60 days post-injection for sectioning, immunofluorescence (IF) and microscopy. (A) Number of EdU labeled neurons in the GL in WT male and female mice. (B) Proportion of TH, CalR and CalB-positive PGNs that are positive for EdU expressed relative to the total number of EdU positive PGNs compared between WT male and female mice. Neurons were identified by NeuN IF. n=4, ± 95% C.I.


A
p<0.0001
Synthetic Thymidine Analog+ Cells per mm2
200
150
100
50
0
EdU
BrdU
N.S.
N.S.
NeuN+
TH+ NeuN+
CalR+ NeuN+
PGN Subtype
B
% of NeuN+ and Synthetic Thymidine Analog+ Cells
30
20
10
0
p=0.0507
N.S.
EdU
BrdU
TH+
CalR+
PGN Subtype


DAPI EdU NeuN
EdU
NeuN
C
D
D
D
D

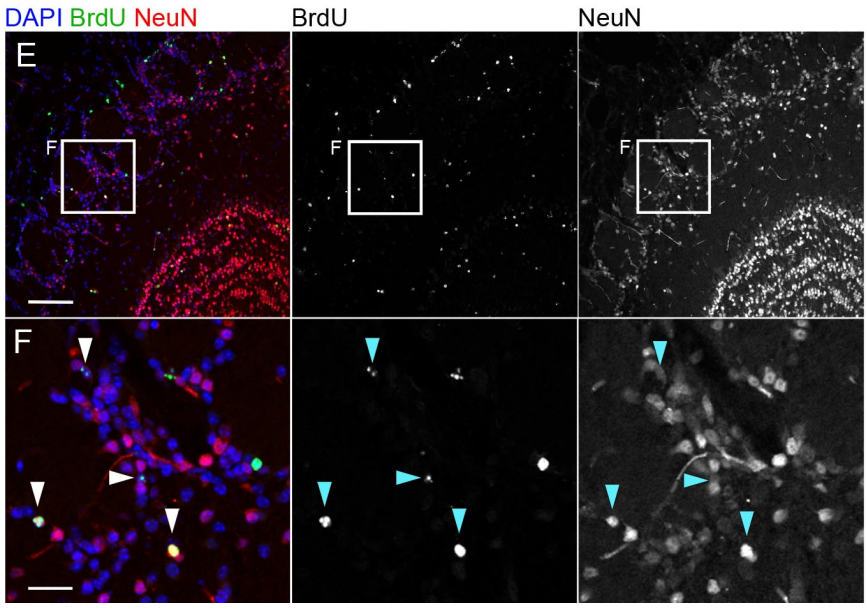

DAPI BrdU NeuN
BrdU
NeuN
E
F
F
F
F

Figure 52. V-SVZ NSC fate tracking using EdU versus BrdU

Mice were injected twice at postnatal day 1 (P1) with either the synthetic thymidine analog EdU or BrdU to label ventricular-subventricular zone (VZ-SVZ) neuronal progenitors in S-phase. OBs were harvested two months post-injection and labeled for EdU or BrdU, NeuN, and tyrosine hydroxylase (TH) or calretinin (CalR). (A) Absolute numbers of EdU or BrdU positive periglomerular neurons (PGNs) per mm^2. (B) Relative numbers of EdU or BrdU positive PGNs co-expressing either TH or CalR. Data represents mean +/- 95% C.I. and was analyzed using two-way ANOVA with Sidak's multiple comparisons test. n=6 for EdU injected samples, and n=3 for BrdU injected samples. A click reaction using an azide-conjugated Alexa 647 fluorophore was used to detect EdU (A, B) and an antibody against BrdU was used for BrdU labeling (C, D). NeuN labels neurons and DAPI labels all nuclei. B and D show a zoom of the glomerular layer. Arrowheads indicate EdU or BrdU positive neurons. Scale bar in A and C represents 100 µm, and scale bar in zoom panels (B and D) represents 25 µm.

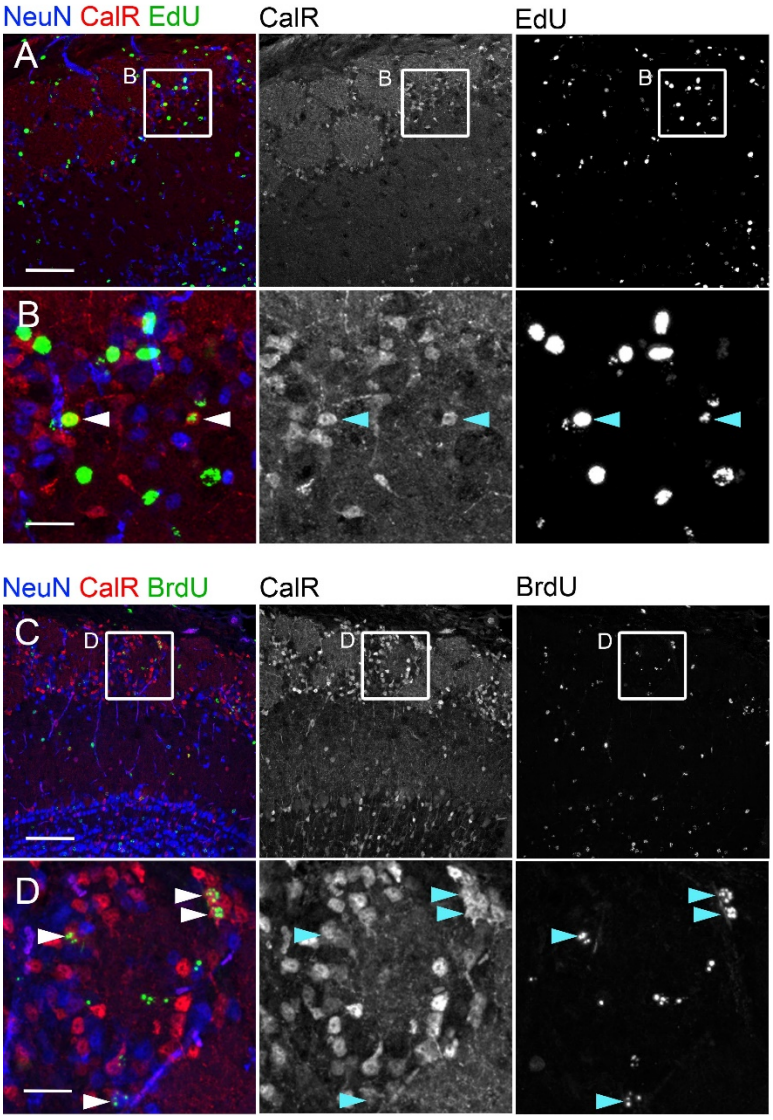


Figure 53. Impact of citrate antigen retrieval on calretinin immunofluorescence in the mouse olfactory bulb

(A, B) CalR IF without sodium citrate antigen retrieval and (C, D) with antigen retrieval. P1 pups were injected twice with either EdU or BrdU and OBs were harvested 2 months post-injection. A click reaction using an azide-conjugated Alexa 647 fluorophore was used to detect EdU (A, B) and an antibody against BrdU was used for BrdU labeling (C, D). To permit antibody access to BrdU in the DNA, BrdU-labeled sections underwent antigen retrieval using sodium citrate. NeuN labels neurons. B and D show a zoom of the GL. Arrowheads indicate CalR and EdU or BrdU double-positive cells. Scale bar in A and C represents 100 µm, and scale bar in zoom panels (B and D) represents 25 µm.