Targets and Functions of the Microrna-200 Family in the Developing Skin and Hair Follicle

Jaimee Elizabeth Hoefert
University of Colorado at Boulder, jaimee.hoefert@gmail.com

Follow this and additional works at: https://scholar.colorado.edu/mcdb_gradetds

Part of the Cell Biology Commons, Developmental Biology Commons, and the Molecular Biology Commons

Recommended Citation
Hoefert, Jaimee Elizabeth, "Targets and Functions of the Microrna-200 Family in the Developing Skin and Hair Follicle" (2018). Molecular, Cellular, and Developmental Biology Graduate Theses & Dissertations. 78.
https://scholar.colorado.edu/mcdb_gradetds/78

This Dissertation is brought to you for free and open access by Molecular, Cellular, and Developmental Biology at CU Scholar. It has been accepted for inclusion in Molecular, Cellular, and Developmental Biology Graduate Theses & Dissertations by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
TARGETS AND FUNCTIONS OF THE MICRORNA-200 FAMILY
IN THE DEVELOPING SKIN AND HAIR FOLLICLE

by

JAIMEE ELIZABETH HOEFERT

Bachelor of Arts, Saint Olaf College, 2011

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirement for the degree of
Doctor of Philosophy

Department of Molecular, Cellular and Developmental Biology

2017
This thesis entitled
TARGETS AND FUNCTIONS OF THE MICRORNA-200 FAMILY
IN THE DEVELOPING SKIN AND HAIR FOLLICLE
written by Jaimee Elizabeth Hoefert
has been approved for the Department of
Molecular, Cellular and Developmental Biology

Leslie Leinwand, PhD
Committee chair

Rui Yi, PhD
Thesis advisor

Date: ________________

The final copy of this thesis has been examined by the signatories, and we
find that both the content and the form meet acceptable presentation standards
of scholarly work in the above mentioned discipline.

IACUC Protocol Yi 2545
Hoefert, Jaimee E. (PhD, Molecular, Cellular and Developmental Biology)

Targets and functions of the microRNA-200 family in the developing skin and hair follicle
Thesis directed by Associate Professor Rui Yi, PhD

The microRNA-200 (miR-200) family is well known for preventing epithelial-to-mesenchymal transition in cancer. However, the targets and functions of this family in normal epithelial tissues remain unclear. This five-member microRNA (miRNA) family also presents a unique platform for studying miRNA-mediated regulation, as they share two nearly-identical seed sequences. The results presented within this dissertation establish a role for these miRNAs in governing hair follicle morphogenesis and fine-tuning cell specification by regulating cell adhesion, polarity, and signaling pathways. By directly ligating miRNAs to their targeted mRNA regions, numerous miR-200 family targets are identified, many of which are involved in the regulation of focal adhesions, actin cytoskeleton, cell cycle and Hippo/Yap signaling. In addition, the experiments presented within this dissertation show that members of the miR-200 family interact with some unique genes, but that they also frequently share targets. Mouse models for miR-200 over-expression and loss-of-function show that the miR-200 family regulates cell adhesion, polarity and reduces cell division, leading to precise cell fate specification and hair morphogenesis. These findings demonstrate that combinatorial targeting of many genes is critical for miR-200 family functions, and provide new insights into this family’s functions in the developing skin and hair follicle.
Dedication

This dissertation is dedicated to Mr. Stewart, who taught me that you are never too old to find wonder in the world.

And to Dr. Eric Cole, who simply knew I could.
Acknowledgements

I would like to start by thanking my advisor, Rui Yi, for providing such a challenging and engaging place to mature as a scientist. Thank you for your constant belief in my abilities, for your unwavering determination in moving my project forward, and for your commitment to my scientific growth.

Thank you to members of the Yi lab, past and present, who have contributed not only to my scientific training, but who have also made the lab a wonderful place to work. This experience would not have been the same without you.

Thank you to my friends, from graduate school and beyond, who were my emotional support during this long process. Having you in my life during graduate school helped me to get through, and I will be eternally grateful for your friendship.

Thank you to my family, who still have no idea what I do every day, but who love and support me nonetheless. Especially to my parents, who always listened and offered advice, even if there was nothing to be done.

Brian, I would have left graduate school long ago if not for you. Thank you for making me stay, and for being my truly constant support.

There were so many other people who contributed to my experience and growth in graduate school whose names do not appear here. To all, thank you for being there for me and with me.
TABLE OF CONTENTS

CHAPTER ONE ............................................................................................................................... 1
INTRODUCTION TO MICRORNAS, THE MIR-200 FAMILY, AND SKIN BIOLOGY .......... 2
a. microRNAs ............................................................................................................................ 2
   i. The discovery of microRNAs ......................................................................................... 2
   ii. microRNA genomic locations and biogenesis ......................................................... 4
   iii. Processing of clustered miRNAs ............................................................................. 7
   iv. microRNA-mediated repression ............................................................................. 8
   v. Canonical microRNA targeting & verified exceptions ............................................. 9
   vi. microRNA families & family targeting paradigms ...................................................... 10
   vii. microRNA regulatory networks ............................................................................. 12
   viii. microRNAs in disease and development ............................................................... 15
b. The microRNA-200 family ................................................................................................. 18
   i. miR-200 family evolution & conservation ............................................................... 18
   ii. miR-200 family function in cancer ......................................................................... 20
   iii. miR-200 function in diabetes .................................................................................. 22
   iv. miR-200 family function in development ............................................................... 23
c. Skin biology .......................................................................................................................... 25
   i. Skin as a model system .............................................................................................. 25
   ii. Appendages of the skin ............................................................................................. 26
   iii. Skin and hair follicle morphogenesis .................................................................... 27
   iv. Molecular mechanisms governing hair follicle morphogenesis ............................ 30
   v. microRNAs in the skin and hair follicle .................................................................... 34
d. The study of microRNA function ....................................................................................... 39
   i. Methods for microRNA target identification ........................................................... 39
   ii. Pitfalls of traditional approaches to microRNA studies ............................................ 42
   iii. Rationale and approach to project ........................................................................... 43

CHAPTER 2 ..................................................................................................................................... 45
DETAILED METHODS .................................................................................................................. 45
a. RNA in situ hybridization ................................................................................................. 45
b. Tough Decoy (TuD) RNAs ................................................................................................. 47
c. Mice ...................................................................................................................................... 47
   i. Straight transgenic (K14-Tg) ..................................................................................... 47
ii. Inducible transgenic (Tg) ................................................................. 47
iii. Double knock-out (dKO) ................................................................. 48
d. Primary keratinocyte harvesting and cell culture .................................. 48
e. Covalent Ligation of Endogenous Argonaute-bound RNAs with Cross-linked Immunoprecipitation (CLEAR-CLIP) .......................................................... 49
i. CLEAR-CLIP .................................................................................... 49
ii. Samples for CLEAR-CLIP ............................................................... 53
iii.Assigning chimeric reads and genome annotation ................................... 54
iv. Area selection for control-only analysis ............................................... 55
v. Area selection for combined analysis ................................................ 55
vi. Unbiased motif finding using HOMER ............................................. 56
f. Fluorescence Activated Cell Sorting and RNA-seq .................................. 56
i. FACS ............................................................................................... 56
ii. RNA-seq ......................................................................................... 57
iii. RNA-seq analysis ............................................................................. 57
iv. GO-term analysis ............................................................................. 58
g. Colony formation, scratch assays & cell adhesion assays .......................... 58
i. Colony formation assays .................................................................. 58
ii. Scratch assays ............................................................................... 59
iii. Cell adhesion assays ..................................................................... 59
h. Focal adhesion and adherens junction assays ....................................... 60
i. Cell plating, staining and imaging ....................................................... 60
ii. Quantification .................................................................................. 61
i. Cloning and 3’ UTR luciferase assays ................................................ 62
j. qRT-PCR .......................................................................................... 63
k. Immunofluorescence, H&E and Alkaline Phosphatase .......................... 63
i. Immunofluorescence ....................................................................... 63
ii. H&E ............................................................................................... 64
iii. Alkaline phosphatase ..................................................................... 64
iii. Imaging and related analyses ......................................................... 65

CHAPTER THREE ................................................................................. 66

MIR-200 FAMILY ENDOGENOUS EXPRESSION AND DEVELOPMENT OF MOUSE MODELS FOR MIR-200 FAMILY OVER-EXPRESSION AND LOSS-OF-FUNCTION .......... 66
a. miR-200 endogenous expression pattern .......................................... 66
i. miR-200s are highly expressed in the skin.......................................................... 66
ii. miR-200s become enriched in developing hair follicles during skin morphogenesis ......................................................... 68

b. Methods for miR-200 family level manipulation in vitro ......................................... 70
   i. Over-expression ........................................................................................................... 70
   ii. Loss-of-function......................................................................................................... 71

c. Mouse models for miR-200 family over-expression ................................................ 73
   i. Straight transgenic ...................................................................................................... 73
   ii. Inducible transgenic .................................................................................................. 74

d. Mouse model for miR-200 family loss-of-function .................................................. 75
   i. Global and skin-specific knock-out models............................................................... 75

e. Chapter Three Summary ............................................................................................ 77

CHAPTER FOUR ............................................................................................................. 78

GENOME-WIDE IDENTIFICATION OF MI R-200 FAMILY TARGETS ............................ 78

a. CLEAR-CLIP on murine keratinocytes ...................................................................... 79
   i. CLEAR-CLIP shows specific interactions between miRNAs and mRNAs ................. 79
   ii. Cross-seed binding between miR-200 family members is rare ................................ 81
   iii. miR-200s often interact with the same targets ..................................................... 83
   iv. miR-200 target identification using CLEAR-CLIP alone ......................................... 86

b. RNA-seq to probe miR-200 family functions in skin ............................................. 87
   i. FACS and RNA-seq .................................................................................................. 87
   ii. miR-200 over-expression is not sufficient to induce hair germ fate ....................... 89
   iii. CLEAR-CLIP targets are down-regulated in RNA-seq ........................................... 89
   iv. miR-200s largely target the same processes in hair germ and interfollicular epidermis 91
   v. Identification of high-confidence miR-200 family targets with CLEAR-CLIP and RNA-seq 93
   vi. miR-200s cooperate in target repression but there is specificity in their targeting 94

c. Chapter Four Summary ............................................................................................. 97

CHAPTER FIVE ............................................................................................................... 99

PHENOTYPES ASSOCIATED WITH MI R-200 FAMILY OVER-EXPRESSION .......... 99

a. miR-200 over-expression leads to neonatal lethality and decreased hair follicle number .... 100
   ii. Neonatal lethality .................................................................................................... 100
   ii. Hair follicle staging ................................................................................................. 101

b. miR-200 over-expression reduces proliferation and migration in vitro ..................... 103
   i. Reduced proliferation ............................................................................................... 103
ii. Reduced migration ........................................................................................................ 104

c. miR-200 over-expression affects cell adhesion and differentiation .............................. 105
   i. Disrupted focal adhesion formation ...................................................................... 105
   ii. Impaired adherens junction formation .................................................................. 106
   iii. Disrupted cell adhesion in vivo ........................................................................... 109
   iv. miR-200s regulate components of the Notch signaling pathway ......................... 110

d. miR-200s regulate components of the Hippo signaling pathway ................................. 111
   i. Nuclear retention of Yap ...................................................................................... 111
   ii. Reduced Yap phosphorylation and disrupted cell-cell contacts ......................... 113
   iii. miR-200s target processes and genes downstream of Yap .................................. 115

e. miR-200 over-expression severely disrupts hair follicle development ....................... 116
   i. Inverted hair germs .............................................................................................. 116
   ii. Imprecise cell fate specification ......................................................................... 118
   iii. Hair follicle arrest .............................................................................................. 120
   iv. Disrupted hair follicle orientation ...................................................................... 121

f. Chapter 5 Summary .................................................................................................. 123

CHAPTER SIX ............................................................................................................... 127

PHENOTYPES ASSOCIATED WITH COMPLETE ABLATION OF THE MIRO-200 FAMILY 127

a. Global or skin-specific loss of the entire miR-200 family leads to neonatal lethality ........ 128
   i. Fertility of single knock-outs .............................................................................. 128
   ii. Neonatal lethality .............................................................................................. 129
   iii. Hair follicle staging ......................................................................................... 130

b. Loss of the miR-200 family does not cause EMT in keratinocytes ............................... 131
   i. miR-200 dKO keratinocytes display typical cell morphology .............................. 131
   ii. Normal E-cadherin expression in vivo .................................................................. 133
   iii. No up-regulation of mesenchymal genes with loss of miR-200s ....................... 133

c. Loss of the miR-200 family enhances proliferation and migration .............................. 134
   i. Enhanced proliferation ....................................................................................... 134
   ii. Enhanced migration ......................................................................................... 135

d. Loss of the miR-200 family reduces cell adhesion .................................................... 136
   i. Disrupted focal adhesion formation .................................................................... 136
   ii. Disrupted adherens junction formation .............................................................. 138
   iii. No effect on cell adhesion in vivo ..................................................................... 139
   iv. No change in localization of Yap ...................................................................... 140
Appendixes

Bibliography

CHAPTER SEVEN

DISCUSSION AND LOOKING TO THE FUTURE

Discussion

Results

Materials & Methods

Introduction

Intr.

200s perform distinct functions in the skin than in cancer cells

200 family role in wound healing in vivo

f. Chapter Six Summary

Appendix 1: miR-200 expression in tissues beyond the skin

Appendix 2: RNA-seq on whole epidermis at P4.5

Appendix 3: miR-200 family role in wound healing in vivo

x
LIST OF TABLES

Table A1: Oligos used for cloning and library preparation .................................................. 187
Table A2: Genotypes, cell numbers, and mapping statistics for RNA-seq .......................... 193
Table A3: FIJI Macro for focal adhesion quantification ....................................................... 194
Table A4: Full list of miR-200 family targets identified with CLEAR-CLIP and Po.5 RNA-seq ........................................................................................................................................................................ 195
Table A5: List of miR-200 family targets identified with CLEAR-CLIP and P4.5 RNA-seq ........................................................................................................................................................................ 201
LIST OF FIGURES

Figure 1.1 Canonical microRNA biogenesis pathway .................................................. 5
Figure 1.2 microRNAs reinforce cell fate and reduce transcriptional noise .................. 14
Figure 1.3 miR-200 family sequences ........................................................................ 19
Figure 1.4 Hair follicle development and associated signaling pathways ..................... 28
Figure 1.5 microRNAs with known functions in the skin and hair follicle ................... 35
Figure 3.1 miR-200 family expression levels ............................................................... 67
Figure 3.2 miR-200 family expression pattern in developing skin ............................... 69
Figure 3.3 Confirmation of miR-200b cluster over-expression and knock-down .......... 73
Figure 3.4 K14-Tg validation ...................................................................................... 74
Figure 3.5 Inducible Tg validation .............................................................................. 75
Figure 3.6 miR-200 family knock-out target sites and validation ................................. 76
Figure 4.1 CLEAR-CLIP schematic and read locations ............................................. 80
Figure 4.2 CLEAR-CLIP is more specific than HITS-CLIP ........................................ 81
Figure 4.3 Cross-seed binding between miR-200 family members is rare ................... 82
Figure 4.5 Many miR-200 targets interact with more than one family member ........... 85
Figure 4.6 GO-term analysis on miR-200 family targets identified with CLEAR-CLIP .... 86
Figure 4.7 RNA-seq on hair germ and interfollicular epidermis ................................. 88
Figure 4.8 Genes with miR-200 family CLEAR-CLIP reads are down-regulated with miR-
200 over-expression ................................................................................................. 90
Figure 4.9 Identification of high-confidence miR-200 family targets with CLEAR-CLIP
and RNA-seq .............................................................................................................. 92
Figure 4.10 Cooperative targeting by the miR-200 family .......................................... 94
Figure 4.11 Heterologous luciferase assay ................................................................. 96
Figure 5.1 miR-200 over-expression is neonatal lethal ............................................... 101
Figure 5.2 Epidermal stratification and hair follicle staging with miR-200 over-expression.
........................................................................................................................................ 102
Figure 5.3 miR-200 over-expression reduces proliferation ........................................ 103
Figure 5.4 miR-200 over-expression reduces cell migration ...................................... 105
Figure 5.5 miR-200 over-expression disrupts focal adhesion formation. ................................106
Figure 5.6 miR-200 over-expression impairs cell junction formation in vitro. .....................109
Figure 5.7 miR-200 over-expression disrupts adherens junction formation in vivo.............110
Figure 5.8 miR-200 over-expression reduces Notch activity in vitro...................................111
Figure 5.9 miR-200 over-expression induces retention of nuclear Yap............................112
Figure 5.10 miR-200 over-expression reduces Yap phosphorylation and disrupts cell
junctions. ....................................................................................................................114
Figure 5.11 miR-200 over-expression dampens the effects of nuclear Yap..........................116
Figure 5.12 miR-200 over-expression causes inverted hair germs....................................118
Figure 5.13 Inverted hair germs display imprecise cell fate specification..........................119
Figure 5.14 Inverted hair germs arrest during development..............................................120
Figure 5.15 miR-200 over-expression disrupts hair follicle orientation...............................123
Figure 6.1 Loss of miR-200s causes neonatal lethality.........................................................129
Figure 6.2 Epidermal differentiation and hair follicle staging in dKO animals......................130
Figure 6.3 Loss of all five miR-200s does not cause EMT in keratinocytes.........................132
Figure 6.4 Loss of miR-200s enhances proliferation in vitro but not in vivo.......................135
Figure 6.5 Loss of miR-200s enhances cell migration and reduces cell adhesion.................136
Figure 6.6 Loss of miR-200s disrupts focal adhesion formation........................................137
Figure 6.7 Loss of miR-200s disrupts cell junction formation............................................139
Figure 6.8 miR-200 ablation does not affect Yap localization.............................................141
Figure 6.9. Complete loss of miR-200s disrupts the actin cytoskeleton in hair germs......142
Figure 6.10 Loss of miR-200s disrupts cell polarity..............................................................144
Figure 6.11 Loss of miR-200s increases variability in Left specification..............................146
Figure 6.12 miR-200 ablation increases variability in Sox9 specification...............................147
Figure 7.1 Model of miR-200 family function during hair follicle development..................159
Figure 7.2 Coordinated targeting by miR-200s regulates hair follicle morphogenesis.........160
Figure A1 miR-200 expression in various neonatal tissues..................................................179
Figure A2 Signatures of miR-200 targets identified at P4.5...............................................182
Figure A3 miR-200 over-expression slows wound healing four days post wound............185
Abbreviations

**RNA biology**
- 3’ UTR—3’ untranslated region
- Ago—Argonaute protein
- dsRNA—double-stranded RNA
- LNA—locked nucleic acid
- miRISC—microRNA-induced silencing complex
- miRNA—microRNA
- mRNA—messenger RNA
- siRNA—short interfering RNA
- stRNA—short temporal RNA

**Mouse models**
- CRISPR—Clustered Regularly Interspersed Short Palindromic Repeats
- dKO—double knock-out (genotype miR-200bcl/-/-/miR-200cclfl/fl/K14cre)
- gdKO—global double knock-out (genotype miR-200bcl/-/-/miR-200cclfl/fl/Ellacre)
- K14Tg—straight transgenic (genotype K14-200bcl)
- rtTA—reverse tetracycline transactivator
- Tg—inducible transgenic (genotype pTRE2-200bcl/K14rtTA)

**Target identification**
- CLEAR-CLIP—Covalent Ligation of Endogenous Argonaute-bound RNAs with Cross-Linked Immuno-Precipitations
- FACS—Fluorescence-Activated Cell Sorting
- HITS-CLIP—High Throughput Sequencing of Cross-Linked Immuno-Precipitations

**Cancer & development**
- EMT—epithelial-to-mesenchymal transition
- ESCs—embryonic stem cells
- iPSCs—induced pluripotent stem cells
- MET—mesenchymal-to-epithelial transition

**Skin biology**
- DP—dermal papilla
- HFSC—hair follicle stem cell
- HG—hair germ
- IFE—interfollicular epidermis
- K5—keratin-5
- K14—keratin-14
- SHH—sonic hedgehog
CHAPTER ONE

The first microRNA (miRNA), lin-4, was discovered in *Caenorhabditis elegans* by Victor Ambros’ and Gary Ruvkun’s groups in 1993 (Lee et al. 1993; Wightman et al. 1993). Over the following decades, the field of miRNA biology has rapidly expanded, leading to the discovery of thousands of miRNAs throughout evolution. These small, non-coding RNAs are only 19-22 nucleotides in length, and yet exert a broad and diverse effect on gene expression. It is estimated that up to 60% of the human genome is regulated by miRNAs, highlighting the widespread influence of these small RNAs. miRNAs have been shown to regulate cell fate, reinforce developmental programs, and in some cases have even been speculated to function as signaling molecules between cells. A wide array of discoveries has begun to elucidate miRNAs’ functions and mechanism of action; however, our understanding of miRNA biology is still remarkably incomplete. This dissertation aims to further the field of miRNA biology by thorough study of a family of miRNAs, the miR-200 family, during murine hair follicle development. The first chapter reviews miRNAs’ functions and mechanisms of action, known functions of the miR-200 family, and skin biology.¹ The second chapter outlines the materials and methods used for experiments, and the third chapter details the generation of mouse models with miR-200 family over-expression or loss-of-function to explore the role of this miRNA family in the skin. Chapter four focuses on genome-wide identification of miR-200 family targets using novel

---

¹ Portions of this chapter have been published as a peer-reviewed review under the title “Not miR-ly micromanagers: the functions and regulatory networks of microRNAs in mammalian skin” in the journal WIREs RNA (Riemondy, Hoefert & Yi, 2014).
² This idea has been challenged by at least one paper suggesting that EMT is not required
methods that are both high-throughput and high-confidence. Chapters five and six include an in-depth analysis of miR-200 family functions in cellular processes as well as hair follicle development. The final chapter of the dissertation summarizes the findings, discusses their significance, and looks to the future of the field of miRNAs in skin biology.

INTRODUCTION TO MICRORNAS, THE MIR-200 FAMILY, AND SKIN BIOLOGY

This chapter begins with a historical look at miRNAs and their discovery. From there, I outline current understandings of miRNA biogenesis, mechanisms of action, and functions in disease and development. This is followed by an introduction to the miR-200 family. This family of five miRNAs provides a unique opportunity to study not only miRNA function in the skin, but also the interactions between miRNAs with two similar-but-distinct seed sequences. The next portion of the chapter introduces skin as a model system for studying miRNAs and development, and outlines known functions of miRNAs in the skin. The chapter concludes with a look at traditional methods for identifying miRNA targets, the pitfalls associated with those approaches, and closes with the rationale for the project described in this dissertation.

a. microRNAs

i. The discovery of microRNAs

In 1993, Victor Ambros and colleagues discovered that an essential C. elegans gene, lin-4, did not encode for protein (Lee et al. 1993). Despite its important function in the nematode, the lin-4 locus does not contain a canonical open reading frame—it is only transcribed to RNA. The Ambros group, along with Gary Ruvkun’s lab, went on to show
that the lin-4 RNA repressed the expression of a critical protein, LIN-14, through an RNA-RNA interaction with the mRNA’s 3’ UTR (Lee et al. 1993; Wightman et al. 1993). While other, larger RNAs, such as XIST and H19 had been shown to have important developmental functions, lin-4 was the smallest RNA yet recognized to have a regulatory role. Ambros and Ruvkun speculated that lin-4 may represent a new class of regulatory RNAs, and indeed, seven years later, a similar repressive function was identified for let-7 (Reinhart et al. 2000).

Initially referred to as small temporal RNAs (stRNAs), both lin-4 and let-7 are differentially expressed throughout the stages of *C. elegans* development, and are essential regulators of developmental time-points (Lee et al. 1993; Reinhart et al. 2000). At this early stage of their discovery, however, it was not known whether their repressive functions were related, and it was suspected that the observed regulatory functions for such endogenous small RNAs may be unique to *C. elegans*. By the time let-7 was identified as a repressive small RNA, double-stranded RNA (dsRNA) had been shown as a trigger for RNA interference (RNAi) in plants, nematodes and various other systems, representing a new pathway by which small RNAs regulate gene expression (Ruiz et al. 1998; Pal-Bhadra et al. 1997; Romano & Macino 1992; Fire et al. 1991; Vaucheret et al. 1998). However, all instances of RNAi had been shown via exogenous introduction of dsRNA, whereas the newly discovered stRNAs were endogenously expressed, and it was not yet clear how closely related their mechanism of repression may be to that of the short interfering RNAs (siRNAs) of RNAi.
The Ruvkun lab quickly showed that let-7 is conserved throughout many invertebrate and vertebrate species (Pasquinelli et al. 2000), sparking the search for other stRNAs. In a landmark issue of Science, the Ambros, Tuschl and Bartel groups each identified many more examples of stRNAs in various species, including *C. elegans*, *Drosophila melanogaster*, *Mus musculus* and *Homo sapiens* (Lagos-Quintana et al. 2001; Lau et al. 2001; Lee & Ambros 2001). With these discoveries it became clear that stRNAs were not unique to *C. elegans*, and in fact embodied an incredibly important and evolutionarily widespread class of small RNAs. In addition, the discovery that the RNase-III enzyme Dicer is involved in the processing of both siRNAs and miRNAs linked the RNAi pathway to this newfound group of small, endogenous RNAs (Bernstein et al. 2001; Ketting et al. 2001). This new, exciting class of RNAs was re-termed as microRNAs, creating a new field of study that would have many implications for RNA biology, development and disease pathology, and even therapeutic and diagnostic applications.

**ii. microRNA genomic locations and biogenesis**

To date, nearly 2,600 miRNAs have been annotated within the human genome (miRBase). The expansive search for novel miRNAs revealed that the majority of these small RNAs are derived from distinct genomic loci, separate from other known genes, and around a quarter are found within the introns of protein-coding genes (Lau et al. 2001; Bartel 2004). miRNAs that are embedded within introns are typically in the same orientation as the host gene, suggesting that they are concurrently processed with the mRNA, rather than transcribed independently (Bartel 2004).
From their locations within the genome, most miRNAs are first transcribed by RNA polymerase II into primary miRNAs, or pri-miRNAs (Fig. 1.1; Lee et al. 2004). Compared with the mature miRNA of just 19-22 nucleotides, the pri-miRNA for a single miRNA is a long transcript, often >1kb in length. Each pri-miRNA is composed of single-stranded RNA at the 5’ and 3’ ends, flanking a stem-loop structure which houses the mature miRNA sequence (Ha & Kim 2014). The pri-miRNA is further processed within the nucleus by the microprocessor, comprised of the RNase-III enzyme Drosha and two molecules of its binding partner, DGCR8 (Lee et al. 2003; Gregory et al. 2004; Han et al. 2004; Yeom et al. 2006). DGCR8 recognizes and binds the miRNA hairpin structure, the base of which is cleaved by the RNase-III domain of Drosha. This releases a double-stranded precursor miRNA (pre-miRNA) of 60-65 nt in length with a 2 nucleotide 3’ single-stranded overhang at the base of the stem. The pre-miRNA is then exported to the
cytoplasm via Exportin-5, which recognizes the overhang left by microprocessor cleavage (Yi et al. 2003; Bohnsack et al. 2004; Lund et al. 2004).

Once in the cytoplasm, the same overhang is recognized by another RNase-III enzyme, DICER, which cleaves off the hairpin of the pre-miRNA. This produces a ~22nt double-stranded RNA duplex with 2 nucleotide overhangs on each end (Ketting et al. 2001; Hutvágner et al. 2001; Bernstein et al. 2001; Hofacker et al. 1989). In addition to a dsRNA-binding domain and two RNase-III domains which bind and cleave the pre-miRNA, DICER also contains an RNA-helicase domain thought to assist in unwinding the miRNA duplex (Ketting et al. 2001).

Following unwinding, a single strand of the miRNA duplex is loaded onto one of four mammalian Argonautes (AGO), the miRNA-interacting protein of the miRNA-induced silencing complex (miRISC) (Hammond et al. 2001). Strand selection for loading into miRISC is biased by the thermodynamic stability of the duplex; the strand which has the lowest stability at it’s 5’ terminus will be preferentially loaded into AGO (Khvorova et al. 2003; Schwarz et al. 2003). Whereas the AGO proteins found in flies display strong bias toward loading certain miRNAs, those in mice and humans are unbiased in which miRNAs they load (Wang et al. 2012; Dueck et al. 2012). The loaded strand is known as the guide strand; the opposite strand, referred to as the passenger or miR* strand, is typically degraded. Rarely, a miRNA duplex has nearly identical stability on both ends; in that case, either strand can be loaded with equal frequency (Schwarz et al. 2003).
iii. Processing of clustered miRNAs

Early studies to identify miRNAs showed that some of these small RNAs are clustered together in the genome, and appear to be co-expressed (Lagos-Quintana et al. 2001; Lau et al. 2001). These clustered miRNAs are processed in much the same way as single miRNAs; however, the polycistronic pri-miRNA for a clustered group of miRNAs is up to several kb in length, and contains an individual stem-loop for each miRNA within the cluster. It has recently been shown that some polycistronic miRNAs actually require the presence of the other stem-loops in order to be processed, suggesting that the scaffold-like structure of these long transcripts plays a role in processing (Truscott et al. 2016). An example of the importance of pri-miRNA structure for clustered miRNAs can be seen in the miR-17–92 cluster. Here, six miRNAs are transcribed as one long pri-miRNA. The tertiary structure formed by this pri-miRNA affects processing efficiency: the miRNAs that are embedded within the structure are processed less efficiently, leading to differential expression of miRNAs derived from the same transcript (Chaulk et al. 2011).

More recent work with the miR-17–92 cluster has shown that differential expression of the six miRNAs can further be achieved through the processing of a biogenesis intermediate, known as a progenitor or “pro-” miRNA (Du et al. 2015). In this example, two cis-regulatory elements within pri-miR-17–92 block the microprocessor at certain points in development. Rather, the pro-miRNA, which includes five of the six clustered miRNAs, is processed by an endonuclease (CPSF3) and a spliceosome-associated protein (ISY1), independently of the microprocessor. In this way, the expression of the miRNAs within the cluster is differentially expressed throughout development.
Identifying the patterns of polycistronic miRNA processing has been difficult because these transient structures are not well annotated. Recently, however, an effort has been made to better define pri-miRNAs genome-wide, which has revealed additional mechanisms by which polycistronic miRNAs may be differentially expressed (Chang et al. 2015). Despite some notable exceptions, however, it is widely accepted that the majority of miRNAs that derive from the same transcript are expressed within the same cells and at similar levels.

**iv. microRNA-mediated repression**

Regardless of whether it was derived from a single pri-miRNA or a polycistronic transcript, a miRNA that has been loaded into miRISC functions to guide the complex to target mRNAs to repress their expression. The miRISC is formed as a complex between the guide strand miRNA, an AGO protein, and GW182 (TNRC6A, -B or -C in mice and humans). In the most extreme examples of miRNA repression, perfect complementarity between the miRNA and its target leads to direct mRNA cleavage by AGO, a mechanism often seen in plants (Ameres & Zamore 2013). Animal systems, however, more often rely upon indirect mechanisms such as mRNA degradation or translational repression.

Within miRISC, GW182 is responsible for recruiting proteins involved in mRNA degradation and deadenylation, including PABP and the CCR4-NOT complex (Chekulaeva et al. 2011; Huntzinger et al. 2013). Though it is clear that these GW182 interactions are involved in miRNA-mediated mRNA decay, their involvement in miRNA-mediated translational repression has only recently been suggested, and the mechanisms by which this might occur are less well-defined (Fabian & Sonenberg 2012). It is well
established, however, that mammalian miRNAs repress expression of their target genes through a combination of mRNA degradation and translational repression. Recent studies examining the prominence of mRNA degradation in comparison with translational repression have found that the majority of miRNA suppression (60-90%) can be explained through mRNA degradation (Eichhorn et al. 2014).

v. Canonical microRNA targeting & verified exceptions

When lin-4 was first discovered, Ambros and colleagues noted that it had sequence complementarity to regions within the 3’ UTR of its mRNA target, lin-14 (Wightman et al. 1991; Lee et al. 1993; Wightman et al. 1993). When let-7 was discovered years later, it was also found to have complementarity to sites within the 3’ UTR of its mRNA target, lin-41 (Reinhart et al. 2000). Thus, it was suspected that miRNAs target through sequence complementarity within the 3’ UTR of mRNAs, and subsequent studies confirmed that relationship.

Whereas many miRNA targets in plants proved easy to predict because the target sites within their 3’ UTRs contain sequence complementarity to the entire miRNA, mammalian miRNA targeting is more complex (Rhoades et al. 2002; Lewis et al. 2003). Extensive efforts to define mammalian miRNA targeting paradigms revealed that the 5’ end of the miRNA is essential for targeting (Lewis et al. 2003; Lewis et al. 2005; Brennecke et al. 2005). This region, dubbed the seed sequence, is defined as nucleotides 2-8 from the 5’ end of the miRNA, and has continued to guide target predictions for miRNAs (Lewis et al. 2003). In addition to being essential for targeting, the seed region of the miRNA is the most strongly evolutionarily conserved region, further underscoring its importance to
miRNA function. Thus, canonical miRNA targeting has long been thought to occur primarily through seed binding.

However, as efforts to understand miRNA targeting have advanced, several other modes of miRNA:target interaction have been identified. Early efforts to predict miRNA targets identified another class of miRNA targeting relationship in which 3’ binding is important. Termed 3’ compensatory interactions, these types of targeting events involve weak 5’ binding but strong association between the target mRNA and the miRNA’s 3’ end (Brennecke et al. 2005). This early discovery showed that miRNA targeting events are more diverse than originally thought. Indeed, evidence has been found for a variety of functional miRNA targeting events that do not rely upon canonical seed sequence—from bulged seed regions, to central binding sites, and even targeting events lacking seed sites altogether (Lal et al. 2009; Shin et al. 2010; Chi et al. 2012). As miRNA target prediction methods become more sophisticated, more non-canonical targeting interactions are being identified and confirmed. New methods used to predict and define miRNA targets will be discussed in more detail in section d.

vi. microRNA families & family targeting paradigms

Whereas many mammalian miRNAs are unique in their sequence and genomic location, a large percentage of these small RNAs can be grouped into a miRNA family. In fact, some definitions place up to 75% of mammalian miRNAs within a family (miRBase). A miRNA family is generally defined as a group of miRNAs that are clustered together in the genome and/or share sequence identity. Sequence conservation among family members arises from gene duplication events, and can extend the length of the full miRNA, or may
include only the seed sequence (Hertel et al. 2006). Studies have shown that miRNAs are added to metazoan genomes over time, and are rarely lost following duplication events. Diversity within these duplicated miRNAs is accrued through rare mutations within the mature sequence, leading to miRNA family members with nearly-identical sequences (Wheeler et al. 2009).

Because the seed sequence has been shown to play such a crucial role in canonical targeting, miRNA family members that share the same or similar seed sequences have been thought to repress many of the same mRNA targets. It has been speculated that the presence of multiple copies of the same or similar miRNA leads to more robust target repression, potentially explaining the acquisition of miRNA families over evolutionary time (Ebert & Sharp 2012).

miRNA family targeting, however, appears to be much more complex than can be explained through simple seed sequence determination. Despite family members sharing seeds, and therefore presumably being able to interact with the same genes, some targets have been associated with just one member of a miRNA family. Some of these cases may be explained by incomplete analysis of whether other family members are able to regulate the same target; however, the prevalence of these instances in the literature suggests that within miRNA families, determinants other than the seed region play an important role in target selection. For example, miR-183 and miR-96, which have identical seed regions but a number of mismatches in their 3’ ends, have been shown to differentially regulate many targets (Dambal et al. 2015). In contrast, another member of this family, miR-182, has a single nucleotide mismatch within its seed region compared with that of miR-183 and -96.
Despite the seed mismatch, it has been shown to regulate many of the same targets as the other two members of its family. Thus, seed sequence is not the single determining factor in family targeting interactions, and mismatches in the seed may still allow miRNAs to interact with the same targets. Given the diversity of miRNA targeting events that have been found to occur independently of the seed, the ability of miRNA family members both to share targets and to distinguish between them is not surprising. However, the precise mechanisms by which miRNA family members are able to differentiate their targets are not known. In addition, how often miRNA family members target unique genes versus share targets has not been well described. Further complicating the dissection of miRNA family targeting paradigms is the fact that targeting might differ between miRNA families; some families with the same seed may share targets whereas others do not. Thus, miRNA family targeting paradigms are complex and warrant further investigation.

vii. microRNA regulatory networks

Due to the relative frequency with which a given 7-mer motif appears in the genome, each individual miRNA can potentially interact with hundreds of targets. However, the influence of miRNA-mediated repression on any particular gene is small—a single miRNA:target interaction results in only a mild reduction in protein (Baek et al. 2008). As such, miRNAs must be considered as components of regulatory networks, in which they simultaneously repress tens to hundreds of targets. Work examining the loss of all miRNAs versus loss of individual miRNAs supports this idea. Studies in mice to examine global loss of miRNAs by impairing miRNA biogenesis revealed that such ablation is
lethal early in embryonic development (Bernstein et al. 2003; Morita et al. 2007). Indeed, it has been well-established that the miRNA processing machinery must be intact for the development and survival of many organisms. Interestingly, however, studies in *C. elegans* and mice have revealed that the majority of individual miRNAs are dispensable for normal development (Miska et al. 2007; Park et al. 2012). Even more striking is the fact that, in *C. elegans*, ablation of only three out of 15 miRNA families tested resulted in any discernable developmental defect (Alvarez-Saavedra & Horvitz 2010). These results suggest a level of functional redundancy between miRNAs that extends beyond sequence similarity. Supporting this view, many mRNAs appear to be regulated by multiple miRNAs (TargetScan, unpublished observation from CLEAR-CLIP). As expected, this cooperative regulation sometimes occurs among family members; often, however, miRNAs interacting with the same target appear to be unrelated, at least in terms of sequence. Thus, there are likely to be redundant functions not only between members of miRNA families, but also between unrelated miRNAs.

The stark contrast between phenotypes observed with global versus individual miRNA deletion also suggests functional redundancy between miRNAs and other regulatory networks, which may have evolved to guarantee developmental resilience. To test this hypothesis in *C. elegans*, J. Brenner and colleagues employed the use of a sensitized genetic background, in which the organisms were missing one but not both of the worm Argonaute-like (ALG) proteins (Brenner et al. 2010). This partial AGO knockout results in worms which are viable but have lower levels of mature miRNAs compared with wild-type. When crossed with worms in which individual miRNAs were genetically
deleted, mutant phenotypes were revealed for many miRNAs that had previously shown no overt defects. In fact, individual loss of 25 of the 31 miRNAs studied in the sensitized background resulted in a developmental defect, a striking difference from the lack of phenotypes observed with loss of the miRNAs alone.

miRNA-mediated regulatory networks have been proposed to serve numerous roles, including buffering phenotypic and genetic variation, imparting robustness to developmental systems, or acting as critical rheostats for other components of regulatory networks (Peláez & Carthew 2012; Ebert & Sharp 2012). Mechanistically, miRNAs have been shown to function in setting thresholds for processes, suppressing transcriptional noise, or reducing network activity by targeting multiple members of the same pathway. Thus, many miRNAs are thought to reinforce cell fate (Fig. 1.2).

**Figure 1.2 microRNAs reinforce cell fate and reduce transcriptional noise.**

A) miRNAs can act to reinforce a cellular state by promoting negative regulation of factors. Here, miR-203 antagonizes ΔNp63, one of its many targets, during epidermal differentiation.

B) By negatively regulating the expression of many transcripts, miRNAs can buffer biological noise inherent in gene expression networks. Through this suppression, miRNAs can contribute to the formation of thresholds which must be crossed by target genes in order to promote a relevant cellular outcome.
Importantly, miRNA networks are organized distinctly from other regulatory molecules such as transcription factors. Transcription factors tend to group into key network hubs and thus act as master regulators in establishing cell-state specific gene expression programs (Neph et al. 2012). In contrast, miRNAs act downstream of master regulatory factors to fine-tune gene expression levels and reinforce cell-state specific gene expression (Fig. 1.2). It should be noted that while there are ubiquitously expressed miRNAs, many of these small RNAs exhibit tissue-specific expression, further supporting the idea that they serve to reinforce existing cellular states. Indeed, while miRNAs themselves do not function as master regulators, miRNA targets are enriched in genes that act as network hubs such as transcription factors, regulatory metabolic enzymes, or signaling cascade modifiers, indicating that miRNAs play important roles in each of these networks (Cui et al. 2007).

Lending credence to the importance of miRNA regulatory networks, many miRNA:target interactions are evolutionarily ancient, with key components maintained over millions of years of evolution. For example, the let-7 family has been shown to regulate RAS in C. elegans—a regulatory relationship conserved in mammals, including humans (Johnson et al. 2005). Additionally, miRNA target sites themselves within 3’ UTRs are often highly conserved, suggesting that the miRNA:target relationships are also preserved (TargetScan).

**viii. microRNAs in disease and development**

miRNAs have been reported to play crucial roles in many diseases and developmental processes. These small RNAs have been shown to be dysregulated in a wide variety of
pathologies, including Parkinson's disease, Alzheimer’s disease, diabetes, and various infectious diseases (Chandra et al. 2017). Perhaps the most studied of roles for miRNAs, however, is in cancer. Individual miRNAs have been shown to act as either tumor-suppressors or oncogenes—or sometimes both, depending on the cellular context. An excellent example of such interplay can be found in miR-155, which has been shown to function as a tumor-suppressor in pancreatic cancers, but as an oncogene in B-cell lymphomas (Greene, Herschkowitz, et al. 2010). Other miRNAs are differentially expressed throughout the various stages of cancer, further reflecting context-specific functions. For example, miR-205 was shown to inhibit growth and proliferation when over-expressed in breast cancer cells, but to promote the same processes when over-expressed in normal mammary tissue (Iorio et al. 2009; Greene, Gunaratne, et al. 2010). In line with these findings, miR-205 levels in breast cancer correlate with prognosis. Higher levels of miR-205 within a mammary tumor generally correspond with better outcomes, and low miR-205 levels are often found in breast tumors that have become metastatic (Iorio et al. 2005; Sempere et al. 2007). Several other miRNAs have been shown to correlate with survival and prognosis; however, whether miRNA dysregulation is a driving force for tumorigenesis, or is merely a secondary effect of cancer development is not clear.

Intriguingly, cancers originating from different tissues have unique miRNA signatures. Therefore, miRNA sequencing can be used to determine a tumor’s tissue of origin (Lu et al. 2005; Volinia et al. 2006). This finding led to the development of an exciting use for miRNAs as clinical biomarkers. miRNAs circulating in the blood can be
screened as a non-invasive diagnostic tool to detect and stage various diseases, including cancer and malaria (Montagnana et al. 2017; Santangelo et al. 2017; Chamnanchanunt et al. 2017). In addition to their potential to act as biomarkers, miRNAs have also shown promise as potential therapeutic targets for cancer and other diseases. The often tissue-specific expression of miRNAs makes them attractive targets for therapies. In fact, the company miRagen has an ongoing clinical trial testing the efficacy of inhibiting miR-155 for treating cutaneous T-cell lymphoma. Thus, miRNAs play important roles in many diseases, and also provide novel targets for the development of treatments.

The first miRNA, lin-4, was discovered because it plays an essential role in C. elegans development. However, this proved to be a rare function for a miRNA—as mentioned previously, individual ablation of many miRNAs does not result in overt developmental phenotypes. Often, loss of a miRNA results in more subtle defects within a specific aspect of development. For example, a recent study on the role of miRNAs in zebrafish vascular development revealed that loss of either miR-139 or the entire miR-24 family did not result in severe defects (Kasper et al. 2017). Rather, while miR-139 or miR-24 knockout fish ultimately developed the same vasculature as wild-type, loss of the mentioned miRNAs resulted in increased phenotypic variation during vascularization. This study confirms an expected role for miRNAs in buffering phenotypic variation, and highlights the fact that miRNA functions may be best examined during developmental processes or under stress conditions.

An excellent example of a miRNA whose functions were only revealed under stress and disease conditions is miR-206. This skeletal muscle-specific miRNA was found to be
among the most up-regulated of miRNAs in a mouse model of amyotrophic lateral sclerosis (ALS, Williams et al. 2009). Loss of miR-206 alone showed no overt phenotype during development. However, re-innervation of injured muscle was delayed in mice lacking miR-206. Additionally, ablating miR-206 in the mouse model of ALS accelerated symptoms of the disease, suggesting that the observed miR-206 up-regulation in ALS may be protective against disease symptoms. As mentioned, the conventional approach to studying miRNA loss-of-function is to examine the development of an organ or a given tissue; had disease-state- or injury-related functions not been examined for miR-206, important roles for this miRNA may have been overlooked. Thus, in some instances where no easily discernible phenotype is seen with loss of a miRNA, stress or disease may be an important factor needed to contextualize that miRNA’s function.

miRNA roles in disease and development are numerous and diverse. In addition, they may provide novel and unique targets for therapeutic applications. While a discussion of all known miRNA functions and therapies based upon these small RNAs is outside the scope of this dissertation, further discussion of miRNA functions in mammalian skin can be found in section c.

b. The microRNA-200 family

i. miR-200 family evolution & conservation

The miR-200 family of miRNAs is comprised of five members: miR-200b, -200a, -200c, -141 and -429 (Fig. 1.3). This family is found within two separate genomic clusters—the miR-200b cluster, (miR-200b, -200a, and -429, on chromosome 4 in mice), and the miR-200c cluster, (miR-200c and -141, on chromosome 6). Each of these clusters is under the
control of its own promoter (unpublished H3K4me3 ChIP), and is transcribed into a polycistronic primary transcript. Interestingly, among the five members of the miR-200 family, there are two different seed sequences that differ by just one nucleotide at the fourth position (Fig. 1.3). The positioning of this mutation is important, because this area has previously been shown to be critical for miRNA:target interactions (Lewis et al. 2003). Therefore, despite the sequence similarity between miR-200 family members, the two sub-families may interact with distinct target sets. However, whether or not the two sub-families share targets has not been well addressed.

![miR-200 family sequences.](image)

The miR-200 family is comprised of five family members with high sequence identity. The miR-200b seed has a U at the 4th position, whereas the miR-200a seed has a C. Here, the miR-200b cluster is shown in blue, and the miR-200c cluster in red. Both clusters contain at least one example of both seeds. The Drosophila melanogaster homolog of this family, miR-8, also shares the miR-200b seed.

The miR-200 family is highly conserved throughout evolution. This family is found in invertebrates (miR-8 in Drosophila melanogaster), deuterostomes, and all vertebrate classes, from fish to humans (Wheeler et al. 2009). Expansion of the miR-200 family over time from one member in invertebrates to five members in vertebrates suggests that these miRNAs play important roles in many organisms. Within the expanded family,
there is high conservation of the seed region (other than the point mutation between the two sub-families) as well as nucleotides 12-15, suggesting that the 3’ end may also be important for miR-200 function (Fig. 1.3). However, the specific targeting paradigms of this family, as well as the importance of the various regions within their sequences, have not yet been thoroughly examined.

**ii. miR-200 family function in cancer**

The discovery that miRNAs are dysregulated in cancer generated great interest in understanding the contributions of individual miRNAs to this disease. In particular, researchers wanted to understand miRNA functions during critical transitions in cancer progression, as such shifts may be driven by epigenetic changes beyond the primary mutations related to tumorigenesis. Epithelial-to-mesenchymal transition (EMT) has been proposed to be a critical step in the metastasis of tumors originating from epithelial tissue² (Singh et al. 2017), and is a simple process to induce in culture by treating cells with TGF-β. Thus, many early efforts to understand miRNAs in cancer focused on this process.

In 2008, several groups published studies examining the role of miRNAs in regulating EMT. These studies noted that miR-200 family expression seemed to correlate with epithelial identity—when cultured epithelial cells were forced to undergo EMT, miR-200 levels dramatically decreased (Park et al. 2008; Gregory et al. 2008; Korpal et al.

---

² This idea has been challenged by at least one paper suggesting that EMT is not required for metastasis of lung tumors (Fischer et al. 2015). However, epithelial tumors commonly undergo EMT, and is it thought to contribute to chemotherapy resistance of some cancers, making it an important process even if it is not involved with metastasis.
They discovered that miR-200s robustly repress expression of the transcription factors ZEB1 and ZEB2 (SIP1)—each of these genes has many highly conserved sites for both miR-200 seeds, suggesting strong regulation. These transcription factors themselves are strong repressors of E-cadherin; thus, miR-200 repression of ZEB1/2 promotes E-cadherin expression and epithelial identity. In fact, two groups showed that loss of miR-200s alone was sufficient to induce EMT—an effect that was seen even with downregulation of just the miR-200b seed (Gregory et al. 2008; Park et al. 2008). Additionally, forced expression of either cluster or any of the individual family members prevented TGF-β-driven EMT, although miR-200b had the strongest effect (Korpai et al. 2008). miR-200 regulation of EMT through ZEB1/2 repression was observed in cells from various species, suggesting that miR-200-mediated promotion of epithelial identity is highly conserved. Adding an additional layer of complexity to the relationship between ZEB proteins and the miR-200 family, expression of both the miR-200b cluster and the miR-200c cluster is repressed by these transcription factors in mesenchymal cells (Burk et al. 2008; Bracken et al. 2008). In this way, a double-negative feedback loop is formed between the miR-200 family and ZEB1/2, potentially as a way to ensure that the process of EMT is reversible (Bracken et al. 2008).

Since these landmark studies on the role of the miR-200 family in EMT, many other functions for this miRNA family in cancer have been identified. Like many miRNAs, the miR-200 family has been shown to play distinct roles during different stages of cancer progression. In fact, the miR-200 family has been shown to play a role in every step of metastasis, from formation and vascularization of the primary tumor to colonization of
distant organs (Humphries & Yang 2015). In addition to their well-established role in preventing EMT, miR-200s have been shown to repress tumor growth (Saydam et al. 2009), inhibit angiogenesis (Choi et al. 2011), prevent cell invasion (Bracken et al. 2014), promote apoptosis (Uhlmann et al. 2010) and even to promote metastatic colonization (Korpal et al. 2011). Thus, miR-200 expression early in cancer may suppress tumor development, whereas expression of this family in cells that have metastasized may actually advance tumor progression. Indeed, high levels of miR-200s in primary tumors correlates with positive outcomes, whereas high levels of circulating miR-200s, i.e. in tumor cells that may be traveling to distal sites to establish metastases, correlates with poor prognosis (Lee et al. 2017). The correlations between miR-200 expression and tumor outcomes make this miRNA family an ideal prognostic biomarker. miR-200 functions throughout tumorigenesis and metastasis also make them an appealing target for therapies, although care must be taken given the opposing roles they have been shown to play in preventing and promoting tumor progression (Chen & Zhang 2017).

The number of studies examining miR-200 function in cancer is vast, and deeply exploring every identified target and function for this miRNA family in tumorigenesis is outside the scope of this dissertation. However, there is at least one example of a well-executed study of the role of miR-200s in cancer beyond EMT that warrants highlighting, which will be discussed in detail in section d of this chapter.

**iii. miR-200 function in diabetes**

While many miR-200 studies have focused on the role of this miRNA family in cancer, a growing number of publications are considering their role in other diseases, including
diabetes. miR-200s’ role in diabetes was first identified in the fibrosis related to diabetic nephropathy. This role is not surprising, given that tubular EMT is thought to play a role in its development (Lan 2012). As would be expected, miR-200s prevent fibrosis-related EMT through repression of ZEB1/2, just as they regulate cancer-related EMT (Xiong et al. 2012).

Recently, miR-200s have been shown to play a more direct role in progression of type-2 diabetes through targets other than ZEB1/2. Belgardt and colleagues showed that miR-200 family members are highly expressed in the pancreatic islets of diabetic mice, and that miR-200 over-expression in beta cells induced apoptosis by repressing several anti-apoptotic targets (Belgardt et al. 2015). Additionally, loss of miR-200s is actually protective against beta cell apoptosis, and improves type-2 diabetes. Thus, miR-200s may play a critical role in promoting the beta cell apoptosis characteristic of type-2 diabetes, and could provide a promising therapeutic target for its treatment. In addition to providing evidence for a novel function for miR-200s, this study was also the first in which all five miR-200s were completely ablated in a mouse model, representing an important tool that will facilitate further study of this family.

iv. miR-200 family function in development

Given the extensive focus on miR-200s in cancer and other diseases, relatively few studies have examined the role of this family in mammalian development. Many studies investigating the developmental functions of miR-200s have actually focused on the family’s homolog in Drosophila, miR-8, which shares the b-type seed. Drosophila miR-8 has been shown to regulate innate immunity (Choi & Hyun 2012), pigmentation (Kennell
et al. 2012), and body size (Hyun et al. 2009; Jin et al. 2012), as well as to modify Wnt signaling (Kennell et al. 2008). A more recent study showed that miR-8 over-expression disrupts epithelial organization and induces apoptosis in the *Drosophila* wing-disk by modulating the actin cytoskeleton (Bolin et al. 2016). With the exception of Wnt signaling, none of these regulatory relationships has yet been shown in mammals or other vertebrates, perhaps suggesting that not all miR-8 functions are conserved in miR-200s. On the other hand, the *Drosophila* homolog of Zeb1 (Zfh1) has a binding site for miR-8, indicating that some key regulatory relationships are conserved (TargetScan). Therefore, further study into miR-200s during development is required to determine whether identified functions for miR-8 are relevant in mammals.

Of the studies that have been performed examining miR-200 developmental functions in vertebrates, many have focused on neurogenesis. Choi & Zakhary and colleagues combined an olfactory epithelium-specific mouse knockout of DICER1 with miR-200 antisense morpholinos in zebrafish to establish a role for this miRNA family in regulating the terminal differentiation of olfactory neurons (Choi et al. 2008). Several subsequent studies using rats, mice, and cultured cells supported a role for miR-200s in promoting neural differentiation and preventing proliferation (Peng et al. 2012; Pandey et al. 2015; Beclin et al. 2016).

Interestingly, miR-200s also play important roles in both induced pluripotent stem cells (iPSCs) and embryonic stem cells (ESCs). miR-200s promote mesenchymal-to-epithelial transition (MET) during the reprogramming of fibroblasts into iPSCs (Samavarchi-Tehrani et al. 2010; G. Wang et al. 2013). They were also shown to regulate
EMT and germ layer specification during differentiation of embryonic stem cells (ESCs), perhaps indicating that these miRNAs are important during embryogenesis (Gill et al. 2011).

miR-200s have been shown to have a myriad of other functions in mammals, including in regulating embryo implantation and fertility (Renthal et al. 2010; Hasuwa et al. 2013; Jimenez et al. 2016), preventing obesity (Tao et al. 2016), and promoting epithelial differentiation and branching morphogenesis in mammary tissue (Hilmarsdottir et al. 2014). Despite the well-established role for miR-200s in maintaining epithelial identity, and their high expression in the skin (Yi et al. 2006), a very small amount of work has examined endogenous functions for these miRNAs in epithelial tissues. One of the few studies done on miR-200s in epithelial tissues established a role for miR-200c in slowing wound healing in aged skin (Aunin et al. 2017), but the endogenous functions of these miRNAs during skin development are unknown.

**c. Skin biology**

**i. Skin as a model system**

As the largest organ in the body, the skin fulfills a number of essential roles. Principally, it serves as the critical first barrier between an organism and its environment. The skin also provides fundamental protection from pathogens, assists in maintaining body temperature, and plays an essential role in social interactions. Mammalian skin is mainly comprised of cells from two distinct origins: epidermal cells, which originate in the ectoderm, and dermal cells, derived from the mesoderm. In fully mature skin, the squamous stratified epidermis makes up the surface, wherein cells become progressively
more cornified to form a protective layer. Epidermis-associated appendages are embedded with the collagenous dermis, which provides structure and subcutaneous fat deposits (Blanpain & Fuchs 2006). There are a wide variety of cell-types which make up the skin, ranging from the keratinocytes of the epidermis to the fibroblasts and immune cells within the dermis.

The high accessibility of the skin, combined with the diverse number of cell types it contains, makes it an ideal model for the study of many processes. As will be explained further in the following sections, the skin is a highly dynamic tissue. Developmental processes of the skin and its appendages are well defined. In addition, it contains at least one population of stem cells that are relatively abundant and continuously activated. The skin is also highly susceptible to wounding and to treatments which induce tumorigenesis. Thus, mammalian skin is an ideal system in which to study development, stem cell dynamics, injury and disease.

ii. Appendages of the skin

The skin contains a variety of appendages which serve diverse functions. Sweat glands cool the skin’s surface and reduce overall body temperature by releasing sweat for evaporation. Whereas human sweat glands are found throughout the surface of the skin, mice develop these structures exclusively on the pads of their feet (Lu & Fuchs 2014). Another appendage of the skin is the sebaceous gland, which secretes an oily substance known as sebum to lubricate mammalian skin and hair.

The most obvious and well-studied appendage of mammalian skin is the hair follicle. This “mini-organ” senses touch, maintains organism temperature, and plays a
large part in mammalian social interactions (Schneider et al. 2009). The mouse hair follicle has been studied for over a century (Oyama 1904; Dry 1926). As such, the basic processes of hair follicle development and cycling are well defined.

iii. Skin and hair follicle morphogenesis

The epidermis begins as a single basal layer of ectodermal progenitors overlaying the dermis. Around embryonic day 9 (E9), these progenitor cells begin to be specified as epidermal cells, turning on expression of the intermediate filaments keratin-5 (K5) and keratin-14 (K14), which form an obligate heterodimer (Byrne et al. 1994). As embryonic development continues, epidermal cell division occasionally occurs asymmetrically. In this case, a progenitor located in the basal layer produces one cell which remains in the same layer, and one cell which detaches from the basement membrane to enter the suprabasal layer (Lechler & Fuchs 2005). The epidermal cells in the suprabasal layers then begin the process of terminal differentiation, turning on the expression of the keratins K1 and K10, and stratifying to form the layers of the skin (Byrne et al. 1994). The cells farthest away from the basal layer eventually become fully cornified, forming a protective barrier known as the stratum corneum. Following its initial development, the epidermis maintains homeostasis by continuously replenishing suprabasal layers with basal progenitors (Clayton et al. 2007).
Hair follicle development is guided by signaling cross-talk between the epidermis (epithelial lineage) and the dermis (mesenchymal lineage). The timing shown here applies to the first wave of hair follicle development, which results in the formation of guard hairs.

Hair follicle development begins a few days later than initial epidermal development, and involves complex cross-talk between the epithelial and mesenchymal lineages of the skin (Fig. 1.4). Around E14.5, signals from the dermis cause the layer of epidermal cells above to thicken and form evenly spaced pre-follicle structures called placodes (Millar 2002). These placodes form through directed migration, rather than proliferation, of epidermal cells (Ahtiainen et al. 2014). After initial placode formation, epithelial cells signal to the underlying dermal cells to form a dermal condensate, which in turn signals back to epithelial placode cells, causing them to proliferate and migrate.
down into the dermis. This process is dependent upon the migration of otherwise stationary epidermal cells, and results in the formation of hair germs around E15.5. By E18.5, the epidermal cells have extended further down into the dermis, forming bulbous pegs. During this period of downward growth and movement, hair follicle stem cells (HFSCs) are specified, and a subset then further differentiate to form the complex hair follicle structure (Nowak et al. 2008). The dermal condensate eventually becomes enveloped by epidermal tissue and becomes the dermal papilla. By postnatal day 6 (P6), all hair follicles are mature and have fully grown down to the bottom of the dermis.

The mature hair follicle contains a number of different cell lineages, displayed in Figure 1.5 (below). After initial morphogenesis, HFSCs are maintained in a specialized stem cell region known as the bulge (Cotsarelis et al. 1990). Throughout an adult animal’s lifetime, starting around P17, the hair follicle cycles between states of growth (anagen, when HFSCs are activated), destruction (catagen, in which the hair follicle regresses), and rest (telogen, when HFSCs are quiescent) (Schneider et al. 2009). Whereas human hair follicles cycle independently, the hair follicle cycle is synchronous in mice, at least for the first several instances (Schneider et al. 2009).

To fully develop an adult hair coat, mice undergo three separate waves of pelage hair follicle morphogenesis (Schlake 2007). The timing described above occurs for guard hairs, which represent less than five percent of the adult hair coat. Primary guard hair follicles are the largest and have the longest hair shaft of the various types in murine backskin. The second wave occurs at E16.5, when awl and auchene placodes are specified evenly between existing guard follicles. These hair follicles, which are smaller and shorter
than guard hairs, comprise up to twenty percent of the mature hair coat. The final wave of hair morphogenesis begins at E18.5, when zig-zag hairs are specified. These hairs are the same size as both awl and auchene, but as their name suggests, they have a zig-zag appearance. Zig-zag hairs comprise the majority of the adult hair coat in mice.

The existence of three separate waves of hair follicle morphogenesis in mice, the last of which occurs primarily postnatally, makes the murine hair follicle an ideal model for the study of epithelial-mesenchymal interactions during development. At P0.5, for example, backskin hair follicles can be found at a variety of stages of development, facilitating the study of such stages. In addition, the formation of a dermal condensate and subsequent growth to form a bulb or peg before proceeding to morphogenesis are common amongst all ectodermal appendages, including teeth, feathers, and mammary glands,. Thus, discoveries made in relation to hair follicle morphogenesis may carry over to other appendages as well (Biggs & Mikkola 2014).

iv. Molecular mechanisms governing hair follicle morphogenesis

Early studies to understand the molecular basis of ectodermal appendage development relied upon tissue recombination. These experiments, which combined dissected epidermis and dermis from different areas of the body, different animals, and even different species, showed that the dermal tissue is responsible for initiating ectodermal appendage development (Sengel 1986). Subsequently, more sophisticated micro-dissection studies showed that induction is actually due to a group of specialized dermal cells, termed the dermal papilla (Oliver 1970). Thus, it has been known for some time that morphogenesis of hair and other ectodermal appendages relies upon some sort of signal
passed between mesenchymal and epithelial lineages, although the mechanisms
governing hair follicle morphogenesis have still not been completely delineated.
Dissection of epithelial-mesenchymal interactions during hair follicle development in
mice has taken place over the course of decades, relying upon a wide variety of
techniques ranging from antibody staining and genetic fluorescence reporters to
spontaneous mutant models and tissue- or lineage-specific knockouts (Sennett & Rendl
2012, Fig. 1.5).

The earliest event in hair follicle formation is widespread activation of Wnt
throughout the epidermis (Chen et al. 2012). This occurs before hair follicle placodes are
specified, and is required for the “first signal” from the dermis to be released. Despite
extensive efforts to define this signal, which instructs placode formation, its identity
remains unknown. It has been well-established, however, that Wnt signaling is required
for the subsequent steps of placode formation. High Lef-1 expression, a transcription
factor binding partner of β-catenin, is characteristic of placodes and subsequent
developmental stages of hair follicles (Zhou et al. 1995). Ablation of β-catenin or
activation of its inhibitor Dkk1 in the epidermis blocks both placode development and
dermal condensate formation (Huelsken et al. 2001; Andl et al. 2002; Zhang et al. 2009).
Conversely, activation of Wnt signaling is sufficient to induce de novo hair follicle
morphogenesis. Stabilized β-catenin, deletion of its inhibitor APC, or forced β-catenin
expression in the dermis all promoted the formation of hair follicles (Gat et al. 1998;
Zhang et al. 2008; Chen et al. 2012). The β-catenin inhibitor Dkk4 provides one intriguing
explanation for how hair follicle spacing may be regulated. Endogenously, Dkk4 is
expressed only in first-wave hair placodes, and its over-expression throughout the epidermis only prevented the second wave of development (Cui et al. 2010). Thus, endogenous Dkk4 may serve to prevent aberrant Wnt activation in first wave placodes during second wave development. Interestingly, Wnt signaling appears to be required for the development of other ectodermal appendages as well: animals which lack Lef-1 are also unable to form teeth or mammary glands beyond the primary stages (van Genderen et al. 1994). Ectodysplasin (Eda) signaling is also required for hair placode formation, and is particularly important for the specification of guard hairs. Mutations in Eda or its receptor Edar led to sparse hair coat development, and a lack of guard hairs altogether (Mikkola 2011).

As mentioned previously, hair follicles develop from evenly-spaced placodes whose positioning may be regulated through inhibition from certain signaling pathways. Indeed, BMP signaling has been shown to impede hair induction and to regulate hair follicle spacing. Loss of BMP receptor 1A accelerated placode formation, and BMP inhibition with Noggin caused the formation of excess placodes (Andl et al. 2004; Plikus et al. 2004). Thus, Wnt and Eda signaling are required for hair placode induction, and BMP signaling plays an inhibitory role in the same process.

After placode formation, epidermal cells respond to further dermal cues by migrating and proliferating into the dermis. This process is dependent upon sonic hedgehog (SHH) signaling. Loss of SHH leads to arrested hair follicles early in development, after initial placode formation (St-Jacques et al. 1998; Chiang et al. 1999). In addition, loss of SHH does not disrupt expression of Wnt or BMP components, suggesting
that SHH is downstream or independent of those pathways during hair morphogenesis (St-Jacques et al. 1998; Chiang et al. 1999). Elegant experiments which disrupted primary cilia formation in either the epidermal or dermal populations demonstrated that the SHH required for hair follicle formation is secreted from the dermal condensate (Lehman et al. 2009; Croyle et al. 2011). Tgf-β signaling may also play a role in promoting hair growth, as loss of Tgf-β2 caused disruptions similar to those seen with loss of SHH (Heine et al. 1987; Pelton et al. 1991; Paus et al. 1997). Therefore, SHH and Tgf-β are dispensable for hair follicle specification, but are required for later-stage development.

The specification of hair follicle stem cells takes place during hair follicle morphogenesis. The earliest known marker of HFSCs is the transcription factor Lhx2, which becomes up-regulated in the placode and remains high within leading edge cells throughout morphogenesis (Rhee et al. 2006). Later in development, Lhx2 is expressed within bulge HFSCs, although it is not exclusive to that population. Another transcription factor, Sox9, is a marker of adult HFSCs and is required for their specification (Nowak et al. 2008). During morphogenesis, Sox9 is specified within the hair germ, one layer behind the leading edge (Nowak et al. 2008). Lhx2 appears to be upstream of Sox9, as forced Lhx2 expression throughout the epidermis and developing hair follicle aberrantly turned on Sox9 expression in those populations (Rhee et al. 2006). In addition, loss of Sox9 did not affect Lhx2 specification, although it did reduce Lhx2 expression over time, suggesting that Sox9 is required for Lhx2 maintenance (Nowak et al. 2008). Recently, HFSC specification was shown to be further controlled by antagonism between the Wnt and SHH signaling pathways (Ouspenskaia et al. 2016a). Through asymmetric cell divisions in
the basal layer soon after placode initiation, Wnt and SHH activation are compartmentalized within the developing hair follicle. Cells at the leading edge maintain high Wnt activation, and secrete SHH. Their progeny, which have low Wnt expression and receive SHH signals, become Sox9+ stem cell progenitors and expand through symmetric division to form the hair follicle (Ouspenskaia et al. 2016a). Previous studies had suggested that HFSC specification occurs once the bulge niche has formed; however, this new set of experiments establishes that HFSC specification occurs much earlier, and relies upon interplay between two signaling pathways.

v. microRNAs in the skin and hair follicle

The skin is one of only a few mammalian organs in which most of the individual components of the canonical miRNA biogenesis pathway have been ablated. Mice with skin-specific, K14cre-driven deletion of Dicer suffer pleiotropic skin defects and die neonatally (Yi et al. 2006). Hair placodes are specified, but they quickly arrest and form cysts within a few days, failing to produce mature hair follicles. This phenotype is mirrored in skin-specific knockouts of components at every step of the canonical miRNA biogenesis pathway, including Dgcr8, Ago and Xpo5 (Yi et al. 2009; Wang et al. 2012; unpublished data from J. Lee, Yi lab). Thus, the collective expression of miRNAs is required for proper hair follicle development. Given the critical importance of the miRNA pathway for regulation of hair follicle morphogenesis and maintenance, elucidating the functions of individual miRNAs in these processes has been a key focus of investigation (Fig. 1.5).
Figure 1.5 microRNAs with known functions in the skin and hair follicle.

Murine skin is comprised of a variety of cell lineages within the interfollicular epidermis and hair follicle. Within the differentiated cells of the epidermis, miR-203 and miR-24 regulate proliferation and actin-cytoskeletal dynamics, respectively. The balance of quiescence and differentiation in HFSCs is maintained in part by miR-205 and miR-125b. Within the inner root sheath, miR-24 ensures that self-renewal is repressed and differentiation is preferred. Finally, within the sebaceous gland, miR-125b contributes to the control of proliferation.

miR-203 was one of the first miRNAs to be comprehensively examined in the skin, and is the most highly expressed miRNA in normal murine skin (Riemondy et al. 2015). This miRNA is specifically expressed within the differentiated, post-mitotic keratinocyte populations of the interfollicular epidermis, as well as the differentiated layers of the hair follicle (Yi et al. 2008). Manipulation of miR-203 expression by K14-driven over-expression or by in vivo ablation via antagonir delivery demonstrates that this miRNA acts to
suppress the proliferation of epidermal progenitor cells during differentiation (Yi et al. 2008). Mechanistically, miR-203 represses cell proliferation by regulating diverse mRNA targets, as was demonstrated using genetic over-expression and loss-of-function models (Jackson et al. 2013; Riemondy et al. 2015). Recently, the Yi lab also identified a role for this miRNA in preventing Hras-mediated tumorigenesis (Riemondy et al. 2015). Thus, miR-203 plays an important role in normal skin development, and is also involved with tumor repression.

Another epithelial-specific miRNA, miR-205, has been shown to play a role in HFSCs. This miRNA is strongly and specifically expressed in stratified epithelia, and is among the most highly expressed miRNAs in the skin. In developing skin, miR-205 is detectable as early as E10 in ΔNp63+ skin progenitors (D. Wang et al. 2013). After hair follicles have matured, miR-205 is highly enriched in HFSCs and interfollicular progenitors. While the vast majority of individual miRNAs are dispensable for normal development, the Yi lab was the first to show that miR-205 is unique in that its ablation results in lethality within two weeks of birth (D. Wang et al. 2013; Farmer et al. 2013). Loss of miR-205 results in thinner epidermis, shorter and mis-angled hair follicles, and defects in proliferation. In developing hair follicles, miR-205 represses several negative regulators of the PI(3)K/Akt pathway, sustaining a high level of pAKT and thereby promoting the proliferation and expansion of progenitors and HFSCs (D. Wang et al. 2013). HFSCs lacking miR-205 expression prematurely exit the cell cycle to become quiescent. Therefore, miR-205 plays a critical role in maintaining HFSCs.
Although miR-203 and miR-205 are specifically expressed in epithelia, and have been shown to have important functions during skin morphogenesis, there are examples of more ubiquitous miRNAs that also have roles in the skin. A well-characterized example of this is miR-125b, the mammalian homolog of *C. elegans* lin-4. This miRNA is highly expressed in many tissues in both mice and humans (Landgraf et al. 2007). In the skin, miR-125b is modestly expressed within basal epidermal cells, but has significantly reduced expression in hair follicle populations (L. Zhang et al. 2011). Sustained levels of miR-125b through an inducible transgenic model led to grossly enlarged sebaceous glands at the expense of normal hair follicle growth (L. Zhang et al. 2011). miR-125b was found to target Blimp1 in sebaceous gland progenitors, leading to additional rounds of proliferation when its levels are sustained. When miR-125b levels are maintained in stem cell progeny, the cells preferentially retain their stemness, blocking differentiation of HFSCs into different lineages and preventing hair growth. However, this effect was completely reversible when miR-125b was restored to normal levels, even after several months of induction. Thus, miR-125b dynamically regulates stemness, rather than preventing differentiation altogether.

Another miRNA with interesting roles in skin and hair follicle development is miR-24. This miRNA is part of the miR-23~24~27 cluster, a group of clustered miRNAs located within two separate genomic loci. Whereas miR-23 and miR-27 each have two slightly different family members (miR-23a/b and miR-27a/b), the sequence of miR-24 produced from each locus is identical. Thus, within the cluster, they are known as miR-24-1 or miR-24-2; however, since their mature sequences are indistinguishable, they are often studied
as a single miRNA. Although all of the miR-23~24~27 clustered miRNAs are expressed in the skin (Yi et al. 2009), miR-24 is the only one which has been extensively studied. This miRNA has an expression pattern similar to miR-203, as it is enriched within the differentiated cells of the suprabasal layer of the epidermis and of the hair follicle (Amelio et al. 2012). miR-24 was shown to function during keratinocyte differentiation by regulating actin-cytoskeletal remodeling (Amelio et al. 2012). In addition, miR-24 over-expression led to stunted, cyst-like hair follicles with reduced proliferative capacity (Amelio et al. 2013). miR-24 is hypothesized to regulate hair follicle development by directly repressing the transcription factor Tcf3, which in combination with Tcf4 maintains epidermal and hair follicle progenitor self-renewal (Nguyen et al. 2009).

Recently, miR-214 was also shown to play a role in regulating hair follicle development. This miRNA is expressed within hair follicles throughout their development, and displays a dynamic pattern during hair follicle cycling (Ahmed et al. 2014). K14-driven over-expression of miR-214 resulted in thinner epidermis, reduced proliferation of keratinocytes, and a decrease in hair follicle numbers (Ahmed et al. 2014). miR-214 was found to directly repress the expression of β-catenin, and thus to repress Wnt signaling in the developing epidermis and hair follicle, leading to the phenotypes described above.

Overall, individual miRNAs have been shown to regulate diverse aspects of skin and hair follicle development, from differentiation to stem cell expansion. Despite this, our understanding of miRNA-mediated regulation in the skin remains incomplete—there
are many highly-expressed miRNAs whose expression patterns and functions have not yet been examined.

**d. The study of microRNA function**

**i. Methods for microRNA target identification**

To fully understand miRNA function, the targets of these small RNAs must be defined. Thus, miRNA target prediction has been a main focus within the field of miRNA biology. Because it was clear early on that the seed region is important for miRNA:target interactions, the first target prediction methods for miRNAs relied upon seed sequence. Online databases such as TargetScan greatly facilitate miRNA target searches by allowing users to search for a miRNA or target of interest, and are still widely used (Friedman et al. 2009). These databases are perhaps the most useful way to assess target site conservation. However, target prediction through seed sequence alone has major shortcomings. Such methods provide no context, requiring the individual validation of each potential target’s relevance in the system of interest. Further complicating this approach is the fact that the presence of a seed site does not necessarily result in repression (Wolter et al. 2014). Alternately, there is widespread evidence of seed-independent miRNA targeting (see section a). Thus, miRNA target prediction through seed sequence alone is likely to result both in many false positives as well as many targets missed.

Combining seed sequence searches with expression profiling, such as RNA sequencing (RNA-seq) or microarrays, can provide context for target prediction. The combined use of small RNA-seq with mRNA-seq can provide useful information as to which miRNAs and which targets are abundant in a given system. These techniques can
also be adapted to profile rare cell populations, or even single cells (Tang et al. 2010; Faridani et al. 2016). However, although most miRNA-mediated repression results in a reduction of mRNA levels, 10-40% does not (Eichhorn et al. 2014); thus, some miRNA targets are still likely to be missed. In order to capture more targets, some studies have utilized ribosome profiling (Eichhorn et al. 2014; Riemondy et al. 2015). This technique measures transcriptome-wide ribosome occupancy, allowing miRNA targets repressed only through translational repression to be defined as well.

Perhaps the most useful way to identify context-specific targets for miRNAs is with high throughput sequencing of cross-linked immunoprecipitations (HITS-CLIP; Licatalosi et al. 2008; Chi et al. 2009). HITS-CLIP identifies genome-wide binding sites for RNA-binding proteins by first cross-linking protein:RNA complexes, after which the protein of interest is immunoprecipitated. Next, the associated RNAs are subjected to high-throughput sequencing. The ability to pull down miRISC to identify its interacting RNAs was first shown in C. elegans (Ding et al. 2005), and has subsequently been adapted for use with high-throughput sequencing. Thus, AGO HITS-CLIP provides specific sequence information and allows for the identification of which mRNAs are interacting with miRISC in a given system. Identifying which miRNA was guiding RISC to that target, however, still relies upon searching for seed sites or other potential miRNA binding sites within the miRISC-associated mRNA fragments. Therefore, despite providing excellent context for which targets are relevant in a given system, HITS-CLIP has many of the same drawbacks as seed-sequence searching. This is particularly problematic for a miRNA family like the miR-200s, which has multiple members sharing identical seed sequences—
even with HITS-CLIP as a guide, using seed sequence prediction alone to distinguish between family members is impossible.

Recently, however, the Goodall lab provided an excellent example of how HITS-CLIP can be used to distinguish between miR-200 family members without relying solely on seed prediction (Bracken et al. 2014). Using a breast cancer cell line naturally lacking miR-200 expression, they transfected either miR-200b or miR-200a, and subsequently performed HITS-CLIP. To identify targets exclusive to one or the other family member, they looked for reads unique to the addition of miR-200b or miR-200a. Using this approach, the Goodall lab identified a role for the miR-200 family in regulating actin cytoskeletal dynamics and preventing the formation of metastasis-related invadopodia. More importantly, they were able to begin dissecting the complicated network targeted by this miRNA family with more specificity than was possible with previous techniques. They identified thousands of potential miR-200 targeting sites, highlighting the ability of these miRNAs to interact with a diverse set of targets. They also showed that while some targets appear to be unique to miR-200a or -200b, many are shared. This study is the first to begin to systematically determine targeting differences between miR-200 family members; however, these experiments only examined the effects of two of the five miR-200 family members, and still relied upon seed sequence to some extent for target identification.

During the generation of HITS-CLIP libraries, a tiny percentage of reads results from the formation of a chimera between a miRNA and its mRNA target. Recently, several labs generated methods to promote ligation between these two RNAs, in order to increase
the number of chimeric reads produced (Helwak et al. 2013; Moore et al. 2015a). This technique, called covalent ligation of endogenous Argonaute-bound RNAs with cross-linked immunoprecipitation (CLEAR-CLIP), provides a high-throughput method for examining direct miRNA:target interactions. While the chimeric reads generated using this method represent a small portion of overall reads, CLEAR-CLIP still provides a variety of advantages over other methods. For example, CLEAR-CLIP provides evidence of specific miRNA:mRNA interactions without relying on seed sequence, allowing for the determination of seed-independent or non-canonical targeting events. In addition, this novel technique is able to distinguish between miRNAs with the same seed, as long as they diverge elsewhere in their sequence. For example, using this technique in *C. elegans*, the Pasquinelli lab was recently able to show that 3’ end differences dictate differential targeting among let-7 family members (Broughton et al. 2016). Thus, CLEAR-CLIP provides the best current method to faithfully identify high-confidence targets for miRNAs which share seeds, and also provides the best tool for identifying seed-independent and other non-canonical miRNA:target interactions.

**ii. Pitfalls of traditional approaches to microRNA studies**

Despite the number of studies on miRNA targets and functions performed over the last twenty-five years, our understanding of these small RNAs is lacking. For example, a major shortcoming of many miRNA studies is the tendency to attribute a miRNA’s entire role to its interaction with a small number of targets. It is clear that miRNAs interact with hundreds of transcripts; therefore, reducing their regulatory ability to repression of just one or two targets results in an incomplete understanding of their function.
Additionally, many methods for target identification have relied upon seed sequence prediction. Thus, non-canonical and seed-independent targeting by miRNAs has often been overlooked. These methods have also been unable to distinguish between miRNAs which share seeds, making the study of miRNA families difficult.

Finally, many studies of miRNAs, such as those on miR-24 and miR-214 in the skin, were done exclusively with over-expression. Studies such as these can establish miRNA functions, but they do not determine how much control a miRNA has over a process or a group of targets. Therefore, such over-expression studies are less informative than they would be in combination with loss-of-function experiments, particularly those performed \textit{in vivo}.

\textbf{iii. Rationale and approach to project}

The miR-200 family is among the most highly studied miRNA families—nearly 600 papers have been published examining the functions and targets of this family (Pubmed). Despite their popularity, however, miR-200 family targeting paradigms and the role of this family in mammalian development remain remarkably poorly understood. Given the strong regulatory role that miR-200s have been shown to have in promoting epithelial identity, it is surprising that more work has not been done examining the functions of this family in endogenous epithelial tissues. Most of the hundreds of studies have focused on the role of miR-200s during tumorigenesis and metastasis. A major shortcoming of many of these studies and the few that have been conducted on miR-200s in development is their reliance upon just one or two miR-200 family targets to explain the function of these miRNAs.
Despite the wide range of functions the miR-200s have been shown to have in many organisms from invertebrates to mammals, our understanding of how these miRNAs truly function and through which targets remains incomplete. Many studies examining miR-200 roles in development focused on the effects of mis-regulation of these miRNAs, and established only a few, if any, new targets. Thus, a complete understanding of miR-200 regulatory networks is lacking. In addition, given the strong regulatory role that this family plays in promoting epithelial identity, it is particularly surprising that more work has not been done examining endogenous miR-200 targets and functions in epithelial tissues.

This dissertation fills a large gap in the current understanding of miR-200 family biology. To understand the functions of these miRNAs in the skin and hair follicle, I first determined their expression levels and endogenous expression pattern in those populations. In order to be able to best study the functions of these miRNAs, I wanted to be able to manipulate their expression levels in vivo. To that end, I generated two mouse models for skin-specific over-expression of miR-200s in the skin, as well as a double knockout mouse in which expression of all five miR-200 family members is ablated. Using a combination of CLEAR-CLIP and RNA-seq, I examined previously unexplored miR-200 family targeting paradigms and identified the first comprehensive list of miR-200 family targets in the skin. Guided by those targets, I revealed a complex regulatory network for this miRNA family in coordinately regulating cell adhesion, actin cytoskeleton and signaling pathways, that ultimately allows the miR-200 family to influence hair follicle development.
CHAPTER 2

DETAILED METHODS

This chapter describes the materials and methods used to complete the experiments performed as part of this dissertation. Methods are described in sufficient detail for qualified readers to repeat the experiments. Relevant references are included where appropriate. Tables of oligos used for cloning and library preparation can be found in the appendix.

a. RNA in situ hybridization

Fresh sections were cut from tissue embedded in O.C.T. (Tissue-Tek, VWR, catalog #25608-930) to a thickness ranging from 12-14 um. Sections were fixed in 4% PFA in DEPC-treated 1X PBS for 10 minutes, and subsequently washed 3x5min in DEPC-treated 1X PBS. Sections were then washed 2x10min in freshly prepared imidazole buffer (0.33 M NaCl, 0.1% 1-methylimidazole, 0.002% conc. HCl in DEPC-treated water) before being cross-linked in fresh EDC solution for 1 hour at room temperature (1.95mg EDC in 3mL imidazol buffer). Sections were washed 3x5min in DEPC-treated 1X PBS, treated with 20mg/mL Proteinase K for 5 min at room temp, and then post-fixed for 10 minutes in 4% PFA in DEPC-treated 1X PBS. Sections were washed 3x5min in DEPC-treated 1X PBS, and then incubated for 10 min in freshly prepared A1 buffer (0.0132% triethanolamine, 0.00175% conc. HCl in DEPC-treated water) followed by 10 min in freshly-prepared A2 buffer (0.0132% triethanolamine, 0.00175% conc. HCl, 0.0025% acetic anhydride in DEPC-treated water). Sections were washed 3x3min in DEPC-treated 1X PBS before pre-
hybridization for 1 hour at room temperature in *in situ* hybridization buffer (50% RNase-free formamide, 5X SSC, 5X Denhardt's solution, 0.5% salmon sperm DNA, 0.025% yeast tRNA in DEPC-treated water). The LNA probe against miR-200b (Exiqon) was diluted to 1pmol/50ul of *in situ* hybridization buffer and heat-denatured for 2 min at 85° C. The denatured probe was then hybridized to the tissue for 2 hours at 51° C in a hybridization chamber. The sections were then washed 2x10 min in 50% formamide/5X SSC and 2x10 min in 0.2X SSC at 53° C. For NBT-BCIP amplification, sections were washed at room temperature for 5 min in B1 solution (100mM Tris pH 7.5, 150mM NaCl), blocked for 30 min with 5% NSS in B1 at room temperature, and then incubated with anti-DIG-AP (1:1000, Roche, catalog #11093274910) overnight at 4° C. The following day, the sections were washed 3x5 min in B1 solution, 1x5 min in B3 solution (100mM Tris pH 7.5, 50mM NaCl, 25mM MgCl2), and then developed with B4 solution (24 ug fresh levamisole, 7.5 ul NBT, 5.6 ul BCIP in 1.47mL B3) for 2-4 hrs. Coverslips were mounted with 100% glycerol and sealed with nail polish. For TSA amplification, following washes at 53° C, sections were treated with 3% hydrogen peroxide for 10 minutes at room temperature. The sections were then washed in B1 solution for 5 minutes, blocked with 5% NSS in B1 for 30 min at room temperature, and incubated with anti-DIG-POD (1:1000, Roche, catalog #11207733910) for 4° C overnight. The following day, sections were washed 2x5 min in TNT buffer (100mM Tris pH 7.5, 180mM NaCl, 0.3% Triton X-100), 2x5 min in B1 solution, and then fluorescence signal was developed using the TSA Plus Fluorescein system (Perkin-Elmer) for 10 minutes. Where shown, sections were incubated after TSA amplification with a β4-integrin primary antibody (1:100, BD Biosciences, catalog #553745, clone 346-
11A) followed by an Alexa-Fluor-594 conjugated secondary (1:2,000, Invitrogen/Molecular probes, catalog #A11012 and #A11007).

**b. Tough Decoy (TuD) RNAs**

Tough decoy (TuD) RNAs were assembled from oligos derived from the following citation: (Haraguchi et al. 2009). A primary insert assembled from annealed oligos was cloned into the pLL3.7 lentiviral vector. TuDs with control sequence or sponges containing miR-203 or miR-200a/miR-200b were subsequently cloned into the primary sequence. TuDs were tested by transfecting into high passage wild-type keratinocytes with miR-200 family reporters for luciferase assays.

**c. Mice**

**i. Straight transgenic (K14-Tg)**

The straight transgenic (K14-Tg) mouse line was generated through standard transgenic injection of the linearized K14-200bcl DNA into an FVB background. One male founder was produced. He was bred with wild-type females to produce the F1 generation used for experiments.

**ii. Inducible transgenic (Tg)**

The inducible transgenic (Tg) mouse line was generated through standard transgenic injection of the linearized pTRE2-200bcl DNA into an FVB background. Founders were bred with mice harboring a keratin14-rtTA allele (reverse tetracycline trans-activator) to produce mice with skin-specific doxycycline-inducible over-expression of the miR-200b cluster (pTRE2-200bcl/K14rtTA). Doxycycline chow used in experiments was 625mg/kg.
(Teklad Rodent Diet TD-7012). Primer sequences used for miR-200bcl cloning can be found in Table A1.

iii. Double knock-out (dKO)

The miR-200b cluster knock-out allele was generated using two sgRNAs with the CRISPR-Cas9 system. In vitro-transcribed Cas9 mRNA and the two sgRNAs were injected into the cytoplasm of a fertilized mouse oocyte (FVB) and cultured overnight. Those that developed to the two-cell stage were subsequently transplanted into pseudo-pregnant females. Founders were bred with miR-200c/141fl/fl animals (Jackson laboratories, stock number 013706, Mirc13tm1Mtm/Mmjax) and then to E2a-cre animals (Jackson Laboratories Stock No. 003724) or to keratin14-cre animals (gift from E. Fuchs). This produced animals with global miR-200 family ablation (E2a-cre) or loss of miR-200s specifically in the skin (K14-cre). Sequences for miR-200bcl genotyping, and sgRNAs can be found in Table A1. Mice were housed and bred in a pathogen-free facility according to the guidelines of the IACUC at the University of Colorado Boulder.

d. Primary keratinocyte harvesting and cell culture

Primary keratinocytes were isolated from animals between the ages of Po and P4. Mice were sacrificed, back and belly skin was removed, and excess fat was scraped from the tissue. The skin was then placed dermal-side down in 1x Dispase for 1 hour at 37°C. The epidermis was subsequently removed from the dermis with forceps and placed into 0.05% Trypsin-EDTA for 10 minutes at 37°C. Trypsin was quenched with culture media, cells
were strained through a 40 um filter and subsequently plated onto dermal feeder cells (Yi et al. 2008). Keratinocytes were cultured in E-low calcium media.

e. Covalent Ligation of Endogenous Argonaute-bound RNAs with Cross-linked Immunoprecipitation (CLEAR-CLIP)

i. CLEAR-CLIP

CLEAR-CLIP and related analyses were performed by Glen Bjerke, PhD. The CLEAR-CLIP protocol was adapted from the following citations: (Moore et al. 2014; Moore et al. 2015a; Zarnegar et al. 2016). Mouse keratinocytes of the designated genotype were maintained in E-low calcium medium. pTRE2-200bcl/K14rtTA cells were treated with 3 ug/ml final concentration doxycycline for 24 hours before performing CLEAR-CLIP. One 15cm dish of confluent cells was used per sample. Cells were washed once with cold PBS. 10mls of cold PBS was added and cells were irradiated with 300mJ/cm2 UVC (254nm wavelength). Cells were then scraped from the plates in cold PBS and pelleted by centrifugation at 1,000g for 2 minutes. Pellets were frozen at -80°C until needed. Cells were then lysed on ice with occasional vortexing in 1ml of lysis buffer (50mM Tris-HCl pH 7.4, 100mM NaCl, 1mM MgCl2, 0.1 mM CaCl2, 1% NP-40, 0.5% Sodium Deoxycholate, 0.1% SDS) containing 1X protease inhibitors (Roche #88665) and RNaseOUT (Invitrogen) at 4ul/ml final concentration. Next, TurboDNase (10U), RNase A (0.66ug) and RNase T1 (0.66U) were added and samples were incubated at 37°C for 5 minutes with occasional mixing. Samples were immediately placed on ice and then centrifuged at 16,160g at 4°C for 20 minutes to clear lysate. 25ul of Protein-G Dynabeads were used per IP. Dynabeads were pre-washed
with lysis buffer and pre-incubated with 3μl of Wako Anti-Mouse-Ago2 (2D4) antibody. The Dynabead/antibody mixture was added to the lysate and rocked for 2 hours at 4°C. All steps after the IP were done on bead until samples were loaded into the polyacrylamide gel. Beads were captured on a magnetic stand and the supernatant removed, then washed 3 times with cold High Salt Clip Wash Buffer (50mM Tris-HCl pH 7.4, 1M NaCl, 1mM EDTA, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) for 3 minutes with rocking. Samples were then washed 2 times with PNK wash buffer (20mM Tris-HCl pH 7.4, 10mM MgCl2, 0.2% Tween-20). Samples were then phosphorylated at 37°C for 20 minutes in 50ul of PNK mixture: 41.8ul H2O, 5 ul 10X PNK buffer (NEB), 1 ul RNaseOUT, 1.67 ul ATP (30 mM), 0.5 ul T4 PNK - 3’ phosphatase minus (NEB M0236L). Samples were then washed 3 times on a magnetic rack with PNK wash buffer. miRNA-mRNA ligation was then carried out overnight at room temperature in 100ul of mixture: 49.25 ul H2O, 30 ul 50% PEG-8000, 10 ul 10X T4 RNA ligation buffer (NEB), 2.5 ul RNaseOUT, 1 ul ATP (100mM), 1 ul BSA (10 mg/ml), 6.25 ul T4 RNA ligase 1 (10U/ul – NEB M0204). Next morning, an additional 2.5 ul T4 RNA ligase 1 (10U/ul) and 1 ul ATP (100mM) were added and ligation was continued for another 5 hours. Samples were then washed 2 times with lysis buffer, once with PNK/EDTA/EGTA (50 mM Tris pH 7.4, 10 mM EDTA, 10 mM EGTA, 0.5% Igepal) and twice more with PNK wash buffer. Next, samples were treated with phosphatase at 37°C for 20 minutes with 50ul of mix: 41ul H2O, 5ul 10X FastAP buffer, 3ul FastAP enzyme (Thermo Fisher #EF0651) and 1ul RNaseOUT. Samples were then washed two times with PNK wash buffer. Next, 3’ adapter ligation was performed on beads overnight in 40 ul of mixture: 17ul H2O, 4ul 10X T4 RNA ligase buffer (NEB), 1ul of
3’ linker (5’-Adenylated & 3’ blocked - custom ordered from IDT), 16ul 50% PEG-8000, 1 ul RNaseOUT and 1ul T4 RNA ligase 2 truncated K227Q (NEB M0351). Samples were then washed twice with PNK buffer. Next, samples were then radiolabeled on bead with 50ul of the following mix: 5 ul 10X PNK buffer (NEB), 1 ul RNaseOUT, 1.5 ul y-p32-ATP (15µCi), 1 ul PNK enzyme (NEB Mo201) and 41.5 ul H2O. Radiolabeling was carried out for 5 minutes at 37°C, after which another 2 ul of cold ATP (10mM) was added and the samples were then incubated for another 5 minutes at 37°C. Samples were washed three times with PNK buffer, and re-suspended in 25 ul of 1.2X LDS NuPAGE Loading buffer (Thermo Fisher #NP0007) with 60 mM DTT added. Samples were heated to 70°C for 10 minutes with occasional agitation and supernatant was separated from beads on a magnetic stand.

Samples were loaded on an 8% Bis-Tris gel and run at 200V for 2 hours on ice. Protein-RNA complexes were then transferred to nitrocellulose at 90V for 90min. The membrane was washed with PBS and exposed to a phosphor-screen for 1 hour at -20°C. Fragments corresponding to the Argonaute complex with the miRNA & mRNA (~110 kDa to 160 kDa) were then excised and RNA was isolated by the following method: 15 ul of Proteinase K at 20 mg/ml was added to 285 ul of Proteinase K/SDS buffer (100 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM EDTA, 0.2% SDS). This solution was heated to 37°C for 20 minutes to inactivate any RNases, added to nitrocellulose membrane fragments and incubated at 50°C for 1 hour. Samples were briefly centrifuged and then 375 ul of saturated phenol/choloform/isoamyl alcohol (25:24:1) was added and incubated for 10 minutes at 37°C. Samples were then centrifuged at 16,000g at room temperature for 3 minutes, the aqueous layer was removed to a new tube and precipitated overnight at -20°C with 2 ul
Glycoblue (ThermoFisher #AM9516) and 900 ul of 100% ethanol. RNA was pelleted by centrifugation at 16,160g at 4°C for 20 minutes and the supernatant removed. Pellet was washed with 70% ethanol and left to air dry at room temperature for 5 minutes. Next, 5’ adapters were ligated by adding 8ul of the following mix to the pellet and thoroughly resuspending: 2 ul 50% PEG-8000, 1ul 10X NEB RNA ligation buffer, 1 ul 10mM ATP, 4 ul H2O. Samples were then heated briefly to 95°C, placed on ice and additional components were added: 0.5 ul RNaseOUT, 1 ul T4 RNA ligase 1 (10U/ul – NEB Mo204), 0.5 ul 100 uM 5’ RNA linker (Blocked at 5’ end & contains NNNN at 3’ end for barcoding). Ligation was carried out at 37°C for 4 hours with rocking. RT-PCR was then carried out in ligation buffer by adding 8.5 ul of the following mix to the sample: 4 ul 5X first-strand buffer (Invitrogen), 1.5 ul 100mM DTT, 2ul 1 uM RT primer, 1 ul 10mM dNTPs. Samples were heated to 65°C for 5 minutes, transferred to a PCR tube and then enzymes were added: 1 ul Superscript III (Thermo Fisher #18080093) and 0.5 ul RNaseOUT. RT reaction was then performed in a thermocycler: 50°C for 1 hour, 85°C for 5 minutes, and then hold at 4°C. Libraries were then amplified from the cDNA by PCR taking aliquots after cycles 12, 17 and 22. PCR mix: 18.8 ul H2O, 8 ul 5X HF buffer (NEB), 1 ul 25 uM library 1st round forward primer, 1 ul 25 uM RT primer, 0.8 ul 10 mM dNTPs, 0.4 ul Phusion polymerase (NEB Mo530) and 10 ul cDNA. Cycling parameters: Initial denaturation at 98°C for 30 seconds and then amplification cycles at 98°C for 15 seconds, 56°C for 30 seconds and then 72°C for 20 seconds. PCR products were run on a 9% acrylamide gel and then stained with SYBR gold (ThermoFisher #211494) at 1:10,000 for 10 minutes. The area corresponding to approximately 73 to 150 base pairs was excised from the lowest cycles.
condition that showed a product. Gel pieces were then frozen at -80°C for one hour and centrifuged through a hole in the tube made with a 20G needle to break up the gel. 400 ul HSCB buffer (25 mM Tris-HCl pH 7.5, 400 mM NaCl, 0.1% SDS) was added and samples were then rocked overnight at 4°C. The next day the gel slurry was transferred to a 0.22 um filter tube and spun at 16,000g for 20 minutes at room temperature. Samples were then precipitated overnight at -20°C in the presence of 1 ml 100% ethanol and 2ul Glycoblue. The next day samples were centrifuged at 16,160g at 4°C for 20 minutes. The pellet was washed with 70% ethanol and then air dried for 5 minutes. The pellet was then resuspended in 20 ul of H2O. Next, high throughput sequencing barcodes were added by PCR using the following mix: 10 ul previous PCR product, 3.84 ul H2O, 4 ul 5X HF buffer, 0.66 ul 10mM dNTPs, 0.5 ul 25 uM Illumina Index Primer, 0.5 ul 25 uM Illumina RP1 primer and 0.5 ul Phusion. Cycling conditions: Initial Denaturation at 98°C for 30 seconds, 2 cycles of: 98°C for 15 seconds, 50°C for 20 seconds, 72°C for 45 seconds and then 4 cycles of: 98°C for 15 seconds and 72°C for 50 seconds and then a final extension at 72°C for 3 minutes. PCR products were then run on a 9% acrylamide gel and stained with SYBR gold as above. Products corresponding to sizes approximately 144-200 base pairs were excised and isolated from the gel as above. Libraries were mixed in equal amounts and sequenced on an Illumina HiSeq 4000 by the Microarray and Genomics Core at the University of Colorado Anschutz Medical Campus.

**ii. Samples for CLEAR-CLIP**

CLEAR-CLIP was performed in two large experiments. The first experiment (Set 1) consisted of 6 replicates of Control cells (K14-Cre only keratinocytes) and 3 replicates of
miR-200 dKO keratinocytes. The second experiment (Set 2) consisted of 3 replicates of Control cells (pTRE2-200bcl/K14rtTA keratinocytes that were not induced with Doxycycline), 3 replicates of pTRE2-200bcl/K14tTA keratinocytes induced with 3 ug/ml Doxycycline (final concentration) for 24 hours and 3 replicates of miR-200 dKO keratinocytes.

iii. Assigning chimeric reads and genome annotation

Fastq files were obtained from the Genomics core and quality filtered using the FastX toolkit fastq_filter.pl requiring the first 25 base pairs to have a mean quality of 20. The 5’ and 3’ adapter sequences were then removed using Cutadapt. Next the 2 base pairs at the 3’ end of the read that correspond to the NN of the 3’ adapter were trimmed and only reads at least 43 base pairs long were kept (Minimum of 19 base pairs for the miRNA, 20 for mapping the mRNA and 4 base pairs for the barcode at the 5’ end). Next the barcode was trimmed from the read, but kept as a separate entry in the file using stripbarcode.pl that is part of the CIMS package from the Darnell lab. Reads were then mapped using BLAST to a Fasta file containing the mouse miRNA database (BLAST settings: -word_size 11, -outfmt 6 –strand plus). BLAST results were filtered requiring at least 19 matching base pairs matching and E < 0.05. Only reads containing a miRNA that mapped to the 5’ end of the read were selected and the resulting 3’ end sequence was then mapped to the mouse genome (mm10) using Novoalign. Reads between 20-24 base pairs were mapped using Novoalign and requiring an exact match to the genome. Reads 25+ base pairs were mapped using Novoalign and the -t 85 setting which allows two deletions, two substitutions, or one deletion and one substitution. Novoalign files were parsed using the
novoalign2bed.pl program from the Darnell lab CIMS package and then mapped reads corresponding to 20-24 and 25+ base pairs were combined into one file for each sample. Reads were annotated to regions of the genome using Bedtools intersect against annotated regions of the genome using annotation files downloaded from USCS table browser for the mouse mm10 genome. Similarly, reads were annotated to genes using Bedtools intersect against gene annotation files for the mouse mm10 genome.

**iv. Area selection for control-only analysis**

Reads mapped to any miR-200 family member from all control samples were combined and the genome coverage was computed using Bedtools GenomeCoverageBed. Areas where 3 or more reads overlapped were selected and merged into one area. Because merging can reduce the area significantly, making seed matching difficult, areas less then 50 base pairs were expanded symmetrically up to 50 base pairs. Areas of interest were then intersected with reads from individual miR-200 family members to determine which family members were interacting with the area. For each miR-200 family member areas that mapped to 3’UTRs of were then used for motif analysis.

**v. Area selection for combined analysis**

Individual samples were combined by sample type into sets for each experiment (Set 1 samples were Controls (6 replicates) and 200dKO (3 replicates) and Set 2 samples included Control (3 replicates) pTRE2-200bcl/K14rtTA (3 replicates) and 200dKO (3 replicates)). Reads mapped to any miR-200 family member were combined into one dataset per sample. For analysis of controls and pTRE2-200bcl/K14rtTA, areas of interest
were required to be found in 2 of the library sets between Set 1 Controls, Set 2 Controls and Set 2 \textit{pTRE2-200bcl/K14rtTA}. Overlapping areas were selected and then areas less than 40 base pairs were extended to 40 base pairs in the direction of the 3’ end of the mRNA (This increases the percent of reads with seed matches). Areas of interest were then intersected with a database of mouse (mm10) gene 3’UTRs obtained from the UCSC table browser.

\textbf{vi. Unbiased motif finding using HOMER}

HOMER was downloaded from (http://homer.ucsd.edu/homer/motif/). Motif finding was performed for each miR-200 family member individually only on reads found in controls. Areas corresponded to where that family member had a read overlapping areas where 3+ miR-200 family members overlapped (See area selection). The background for motif finding was all mouse mm10 3’UTRs. Settings for motif finding were: -size given, -rna, -chopify, -len 4,5,6,7,8,9,10, -noweight and –nlen 0. The top ranked motif is shown for each miR-200 family member.

\textbf{f. Fluorescence Activated Cell Sorting and RNA-seq}

\textbf{i. FACS}

Females harboring the \textit{pTRE2-200bcl} allele and a \textit{K14RFP} allele were bred with males harboring \textit{K14rtTA} and fed doxycycline chow throughout pregnancy. Pups were collected at P0.5 and whole back-skin epidermis keratinocytes were isolated as described above (see “Primary keratinocyte harvesting and cell culture”). After being filtered through a 40 um filter, cells were spun down for 10 minutes at 300g and resuspended in 1x PBS with 3% chelexed FBS on ice for one hour with the following antibodies: \textit{p-cadherin} (R&D
Systems, 1:100, FAB761P), α6-integrin (1:75, AbDSerotec, catalog #MCA699A647, clone NKI-GoH3). Cells were again spun down for 10 minutes at 300g and re-suspended in 1x PBS with 3% chelexed FBS containing Hoescht 33342 (1:10,000, Invitrogen). Sorting for hair germ and interfollicular epidermis populations was performed on a MoFlow XDP cell sorter (Beckman Coulter). The hair germ population was gated as K14RFP+/p-cad\textsuperscript{hi}/α6\textsuperscript{hi} and the interfollicular epidermis was gated as K14RFP+/p-cad\textsuperscript{low}/α6\textsuperscript{hi}. Cells were sorted into 1x PBS with 3% chelexed FBS, and subsequently spun down for 10 minutes at 300g and resuspended in 200ul of Trizol reagent. A total of four controls (wild-type, pTRE2-200bcl or K14rtTA) and four pTRE2-200bcl/K14rtTA samples were sorted. Control genotypes and numbers of sorted cells can be found in Table A2.

ii. RNA-seq

Libraries for RNA-seq were assembled using the SMARTer\textsuperscript{®} Stranded Total RNA Sample Prep Kit- Low Input Mammalian (Clontech, catalog #634861). For hair germ populations, the entire sample was used because the number of sorted cells was low. For interfollicular epidermis, 50 ng of starting RNA was used. Sequencing of completed libraries was performed on an Illumina HiSeq-4000.

iii. RNA-seq analysis

RNA-seq raw reads were first trimmed to remove 3’ adaptor sequence (GATCGGAAGAGCACACGTCTGAACTCCAGTC, CutAdapt, default settings). Reads were then aligned to the mouse genome with Tophat with a supplied .gtf transcript annotation file (Illumina iGenomes, mm10) (settings = --bowtie1 --library-type fr-firststrand).
Mapping statistics for each library can be found in Table A2. Alignments uniquely overlapping coding sequences were counted with HTSeq Count (settings = -s reverse -t CDS). Differential expression was calculated with DESeq using default settings.

**iv. GO-term analysis**

GO-term analysis was performed using the “Investigate Gene Sets” option in the Molecular Signatures Database (MSigDb) from the Broad Institute. Additional GO-term analysis was performed using DAVID Bioinformatics Resources (NIAID, NIH, versions 6.7 and 6.8.) Background genesets were either the provided Mus musculus genome (MSigDb and DAVID version 6.8) or, where possible, the genes detected in the hair germ RNA-seq (DAVID version 6.7). Background used is indicated in relevant figures.

**g. Colony formation, scratch assays & cell adhesion assays**

**i. Colony formation assays**

Cultured keratinocytes derived from pTRE2-200bcl/K14rtTA, 200dKO or K14cre back-skin were plated at a low density in a 6-well plate (2,000 cell/well for pTRE2-200bcl/K14rtTA cells at P9; 5,000 cells/well for 200dKO and K14cre at P5). For colony formation assays done with TuD RNAs, wildtype keratinocytes were transfected with pLL3.7-TuD-200a/b or pLL3.7-TuD-ctrl constructs, and FACS was subsequently used to isolate successfully transfected cells. In the case of pTRE2-200bcl/K14rtTA, cells were treated with PBS or 3ug/mL doxycycline every 24 hours, starting 24 hours after plating. Colonies were allowed to grow for 6-8 days; media was changed every two days. Colonies were then fixed with 100% methanol and stained with 0.25% crystal violet for 1 hour. Following rinsing with deionized water, colonies were imaged with a scanner. Each well was cropped to an
independent image. For signal intensity measurements, images were converted to 8-bit grayscale in ImageJ (Image>Mode>8-bit), inverted (Edit>Invert), and then the signal intensity was measured (Analyze>Measure). Colony number and size were measured manually.

**ii. Scratch assays**

Low passage keratinocytes derived from pTRE2-20obcl/K14rtTA, 200dKO, or K14cre back-skin were plated in 6-well plates and cultured until confluence. pTRE2-20obcl/K14rtTA cells were treated with PBS or 3ug/mL doxycycline every 24 hours for 48 hours prior to scratching. Once confluent, a scratch was made using a 200uL pipette tip. Scratches were imaged every 2 hours until the scratch was closed. The number of cells migrated into the scratch was manually counted in images taken at 8 hours.

**iii. Cell adhesion assays**

Low-passage keratinocytes derived from pTRE2-20obcl/K14rtTA, 200dKO, or K14cre back-skin were plated on 24-well plates (uncoated, or coated with 1ug/mL, 5ug/mL or 10ug/mL fibronectin; 50,000 cells/well). pTRE2-20obcl/K14rtTA cells were treated with PBS or 3ug/mL doxycycline for 24 hours prior to plating. After one hour, wells were washed three times with 1x PBS to removed non-adhered cells, and then fixed with 4% PFA over-night at 4°C. The next day cells were stained with 0.25% crystal violet in 70% ETOH for 1 hour and subsequently rinsed with deionized water. Wells were allowed to air dry for 30 minutes, and then 300 ul of 100% methanol was added to each well for 20 minutes to
solubilize the crystal violet. 200ul of the crystal violet mix from each well was transferred to a 96-well plate and absorbance was measured at 595 nm on a plate reader.

**h. Focal adhesion and adherens junction assays**

**i. Cell plating, staining and imaging**

Low-passage keratinocytes derived from pTRE2-200bcl/K14rtTA, 200dKO, or K14cre back-skin were sparsely plated on coverslips coated with 1ug/mL fibronectin. Tg keratinocytes were treated with PBS or 3ug/mL doxycycline for 24 hours after plating. Cells were fixed with 4% PFA, washed twice in 1x PBS and permeabilized for 10 minutes in 1x PBS with 0.01% Triton x-100. Blocking was performed with 5% normal goat serum in 1x PBS for 1 hour at room temperature. Cover-slips were then incubated with primary antibody for 1 hour at room temperature. Primary antibodies used were vinculin (mouse, 1:200, Sigma catalog #V9131) or e-cadherin (rat, 1:200, gift from E. Fuchs). Cover-slips were washed three times for five minutes in 1x PBS, and then incubated with secondary antibody (Alexa-Fluor-594-, Alexa-Fluor-488- or Alexa-Fluor-647-conjugated secondary antibodies (1:2000, Invitrogen/Molecular Probes)) and Phalloidin (1:50, Alexa-488 conjugated, Invitrogen catalog #A12379) for 1 hour at room temp. Nuclei were stained with Hoescht 33342 (1:5000, Invitrogen). Cover-slips were mounted onto slides with Prolong Gold Antifade (Invitrogen catalog #P36930) and left at room temperature overnight to cure. Cells were then imaged on a Nikon A1 laser-scanning confocal microscope in the CU-Boulder Light Microscopy Core Facility. 6-9 optical slices were taken at an interval of 0.125 microns, and each z-stack was made into a maximum-intensity projection for quantification.
ii. Quantification

A custom FIJI macro was used for focal adhesion and adherens junction quantification. Maximum intensity projections of phalloidin were auto contrasted (“Enhance contrast”, “saturated=0.35”) and auto-thresholded (setAutoThreshold(“Mean dark”). Images were converted to masks (“Convert to mask”), and the following binary commands were used to identify cellular regions: (“Dilate”), (“Close-“) and (“Fill holes”). The Analyze Particles function was then used to exclude regions below a size threshold (“Analyze Particles…”, “size=1200-Infinity circularity=0.00-0.99”). These masked regions were converted to selections (“Create selection”), added to the roiManager (roiManager(“ADD”)), and overlaid onto the maximum intensity projections for Hoescht and vinculin or E-cadherin staining using (roiManager(“AND”)). The number of nuclei per image was counted in the following manner: local contrast of maximum intensity projections of Hoescht was enhanced using CLAHE (“Enhance Local Contrast (CLAHE)”, “blocksize-75 histogram=256 maximum=6 mask=*None* fast_(less_accurate)”). Overall contrast was auto-contrasted (“Enhance Contrast”, “saturated=0.35”) and images were auto-thresholded (setAutoThreshold(“Default dark”). Next, the Analyze Particles function was used to count the number of nuclei above a size threshold (“Analyze Particles…”, “size=25-Infinity circularity=0.00-1.00 display summarize”). Nuclei accounts were manually inspected following analysis and corrected when errors occurred. Finally, the number of focal adhesions or adherens junctions were quantified using the maximum intensity projection for vinculin or E-cadherin in the following manner. Background was subtracted (“Subtract Background…”, “rolling=50 sliding”), and local contrast was
enhanced with CLAHE ("Enhance Local Contrast (CLAHE)", "blocksize=19 histogram=256 maximum=6 mask=*None* fast_(less_accurate)"). Exponentiation was used ("Exp") followed by auto contrasting ("Enhance Contrast", "saturated=0.35"). LoG 3D filtering was performed ("LoG 3D", "sigmax=5 sigmay=5"), images were converted to 8-bit ("8-bit"), auto-thresholded (setAutoThreshold("Default dark")) and the Analyze Particles function was used to count focal adhesions or adherens junctions ("Analyze Particles...", "size=1.2-2000 circularity=0.00-0.99 show=Outlines display summarize exclude"). The script for the FIJI macro, which can be run on a multi-channel image with Hoescht in channel 1, phalloidin in channel 2 and vinculin or e-cadherin in channel 3, can be found in Appendix 1.

i. Cloning and 3’ UTR luciferase assays

3’ UTR fragments were generated by PCR amplification from cDNA. Primers used for cloning can be found in Table A1. These were then cloned into the pGL3-control vector (Promega, Madison, WI). 2 ng renilla luciferase control, 20 ng of pGL3 reporter, and 380 ng of MIGR, MIGR-200bcl or K14-200ccl were co-transfected into miR-200 dKO keratinocytes in a 24-well plate using the Mirus LT-1 transfection reagent (Mirus Bio LLC, Madison, WI). For luciferase assays with both clusters, 190 ng of each MIGR-200bcl and K14-200ccl were transfected. Cell lysates were collected after 48 hours, and Renilla and firefly activity were measured using the Dual-Glo Luciferase Assay system (Promega).
j. qRT-PCR

cDNA was generated from 500-1000 ng starting RNA using the miScript® II RT kit (miRNA detection, Qiagen, catalog #218160) or the SuperScript® III First-Strand Synthesis Supermix for qRT-PCR (mRNAs, Thermo Fisher Scientific, catalog #11752-050). cDNA was diluted 1:10 for subsequent reactions. qPCR was performed using iQ™ SYBR® Green Supermix (Bio-Rad, catalog #170-8880) and quantified using the ΔΔC(t) method. miRNAs are normalized to sno25 and mRNAs are normalized to HPRT. Primers used can be found in Table A1.

k. Immunofluorescence, H&E and Alkaline Phosphatase

i. Immunofluorescence

OCT-embedded tissues were sectioned to 10-14 um and fixed with 4% PFA for 10 minutes at room temperature. Sections were permeabilized for 10 minutes at room temperature with 0.1% Triton-X 100 in 1X PBS. When staining with mouse monoclonal antibodies, we used the M.O.M. Basic Kit (Vector labs, catalog #BMK-2202). Otherwise blocking was performed in 5% normal serum of the same species the secondary antibody was raised in. Sections were incubated with primary antibody overnight at 4° C. The following antibodies and concentrations were used: cytokeratin 5 (chicken, 1:2000, Covance, catalog #SIG-3475), β4-integrin (rat, 1:500, BD Biosciences, catalog #553745, clone 346-11A), lef-1 (rabbit, 1:500, Cell Signaling, catalog #2230), p-cadherin (rat, 1:500, R&D Systems, catalog #MAB761), e-cadherin (rat, 1:200, gift from Elaine Fuchs), caspase-3 (rabbit, 1:500, R&D Systems, catalog #AF835), α-catenin (rabbit, 1:200, Cell Signaling, catalog #3236S), sox9 (rabbit, 1:500, Millipore, catalog #AB5535), yap1 (rabbit, 1:500, Cell Signaling, catalog
#14074, clone D8H1X), pericentrin (rabbit, 1:500, Covance, catalog #PRB-432C), β-catenin (mouse, 1:2000, BD Biosciences, catalog #610153), lhx2 (rabbit, 1:5000, gift from Elaine Fuchs). Following incubation with primary antibodies, sections were washed three times in 1x PBS, and then incubated for 1 hour at room temperature with Alexa-Fluor-594-, Alexa-Fluor-488- or Alexa-Fluor-647-conjugated secondary antibodies (1:2000, Invitrogen/Molecular Probes). EdU incorporation was detected using the Click-iT® EdU Alexa Fluor® 488 Imaging Kit (Thermo Fisher, catalog #C10337). Nuclei were stained with Hoescht 33342 (1:5000, Invitrogen).

**ii. H&E**

OCT-embedded tissues were sectioned to 10-12um and fixed with 4% PFA for 10 minutes at room temperature. Sections were permeabilized for 10 minutes at room temperature with 0.1% Triton-X 100 in 1X PBS. Next, in a fume hood, sections were incubated for 3 minutes in hematoxylin, followed by 1 minute each in water, clarifier and 95% ethanol. Sections were placed in eosin for 10 seconds, followed by 1 minute each in 100% ethanol (x2) and xylene. Sections were dried in the fume hood and mounted with Permount (Fisher, catalog #SP15).

**iii. Alkaline phosphatase**

Alkaline phosphatase staining was performed as above until permeabilization. After permeabilizing, sections were incubated with NBT-BCIP (7.5 ul NBT, 5.6 ul BCIP in B3 solution (100mM Tris pH 7.5, 50mM NaCl, 25mM MgCl2)) for 15-30 minutes at room temperature. Then, sections were placed in eosin for 10 seconds, followed by 1 minute
each in 100% ethanol (x2) and xylene. Sections were dried and mounted with Permount (Fisher, catalog #SP15).

iii. Imaging and related analyses

Widefield imaging was performed on a Leica DM5500B microscope with an attached Hamamatsu C10600-10B camera and Metamorph software (version 7.7; MDS Analytical Technologies). Confocal imaging was performed using a Nikon A1 laser scanning confocal in the CU-Boulder Light Microscopy Core Facility. Quantification of cell angle or hair follicle orientation was performed using the “angle” tool in ImageJ.
CHAPTER THREE

MIR-200 FAMILY ENDOGENOUS EXPRESSION AND DEVELOPMENT OF MOUSE MODELS

FOR MIR-200 FAMILY OVER-EXPRESSION AND LOSS-OF-FUNCTION

A critical step toward understanding miRNA functions is defining the cells and tissues in which they are endogenously expressed. This chapter begins with a presentation of miR-200s’ expression levels in the skin, and continues with the previously undefined endogenous expression pattern of these miRNAs in the skin and hair follicle. To better understand the role of these miRNAs in those tissues, I used various methods to manipulate miR-200 family levels in both tissue culture and animal models. The methods for miR-200 family manipulation in vitro and the mouse models described here were used to conduct the experiments described in this thesis.

a. miR-200 endogenous expression pattern

i. miR-200s are highly expressed in the skin

The miR-200 family is comprised of five separate miRNAs sharing two nearly-identical seed sequences, and located in two separate genomic clusters (Fig. 3.1a). Previous studies showed that miR-200 family members are highly expressed in murine skin, but their expression relative to other miRNAs and to each other was not well determined (Yi et al. 2006). Additionally, methods used to sequence miRNAs had proven to be biased toward certain sequences, making the determination of relative miRNA levels unreliable. Thus, accurate expression levels for the miR-200 family in the skin had not previously been defined. A new small RNA sequencing method optimized in the Yi lab confirmed that the
miR-200 family is indeed highly expressed in neonatal skin, and allowed me to further dissect the expression of this family of five miRNAs (Fig. 3.1b) (Zhang et al. 2013).  

![miR-200 family sequences](image)

Figure 3.1 miR-200 family expression levels.

A) miR-200 family sequences. miR-200b cluster is in red, miR-200c cluster in blue. Seed mismatch is highlighted in red. B) miRNA-seq on whole epidermis at P4.5; n=3; error bars represent standard deviation. C) Percentage of miR-200 family reads from the miR-200b and miR-200a seeds. D) Percentage of miR-200 family reads from the miR-200b or miR-200c cluster. E) Relative expression levels of all miR-200 family members. Error bars represent standard deviation.

Taken together, this family comprises nearly 10% of total miRNA reads. As mentioned, the unbiased nature of this sequencing method also allowed for the accurate determination of the levels of miR-200s relative to one another. I noted that the miR-200a seed was slightly more highly represented than the miR-200b seed (miR-200a seed 58%, miR-200b seed 41%), but that the miR-200b cluster was slightly more highly expressed

---

3 Small RNA-seq method optimized by Dr. Jerome Lee and Dr. Zhaojie Zhang; small RNA-seq libraries for neonatal skin were generated by Dr. Jerome Lee.
(miR-200b cluster 55%, miR-200c cluster 44%, Fig. 3.1c-d). However, neither of these differences was statistically significant. Interestingly, there did appear to be some differential expression of some of the miR-200s derived from a single cluster. There is a statistically significant difference in expression between miR-200b and miR-429, both derived from the miR-200b cluster (Fig. 3.1e). There are examples within the literature of differential expression of miRNAs derived from the same polycistronic transcript (see Chapter 1). However, the mechanisms by which miR-429 expression is regulated to be distinct from miR-200a and -200b are unknown, and answering that question is outside the scope of this dissertation. Despite this minor difference in expression levels, all members of the miR-200 family are highly expressed in the skin—all are individually among the top 25 most highly expressed miRNAs—and as such represent a potentially important group of regulators in murine skin.

**ii. miR-200s become enriched in developing hair follicles during skin morphogenesis**

To gain a better understanding of where the miR-200s are expressed within the skin, I performed *in situ* hybridization. Given that our miRNA-seq data indicated that the entire miR-200 family is expressed in the skin, and at relatively similar levels, I opted to use a probe against only miR-200b as representative of the entire family. Using fluorescence *in situ* as well as NBT-BCIP, I evaluated miR-200 family expression during prenatal and neonatal development.⁴

---

⁴ *In situ* data from E15.5 and E17.5 was generated by Evan O'Loughlin.
**Figure 3.2** miR-200 family expression pattern in developing skin.

A) NBT-BCIP or fluorescence n situ hybridization with a probe against miR-200b at E12.5, B) E14.5, C) E15.5, D) E17.5, E) P0.5, F) P1.5 and G) P4.5. Scale bars are 50 um.

miR-200s are detectable as early as E12.5, when the epidermis is still a single layer of cells (Fig. 3.2a). From this early stage to E15.5, miR-200s are expressed throughout the developing epidermis (Fig. 3.2a-c). By E17.5, when the second wave of hair follicle morphogenesis is occurring, miR-200 expression is enhanced within developing hair follicles (Fig. 3.2d). This expression is maintained postnatally, at which point miR-200s become further restricted to just the leading edge of developing hair follicle (Fig. 3.2e-f).
By postnatal day 4, miR-200s display highest expression within the matrix region, particularly within those cells that surround the dermal papilla (Fig. 3.2g).

Whereas the Yi lab had previously examined the expression patterns of many miRNAs in murine skin, none thus far evaluated had displayed such a restriction to developing hair follicles. As mentioned in Chapter 1, complete ablation of Dicer and other components of the miRNA biogenesis pathway leads to disruptions in hair follicle development (Yi et al. 2006; Yi et al. 2009; Wang et al. 2012). This phenotype has not been attributed to an individual or group of miRNAs, and the unique expression pattern of the miR-200s indicated that they may be involved in hair follicle morphogenesis. To further explore this family of miRNAs and their functions in murine skin, I utilized several methods to manipulate miR-200 family levels in vitro and in vivo.

b. Methods for miR-200 family level manipulation in vitro

i. Over-expression

Over-expression of miRNAs in cultured cells can be accomplished in several ways. Often, a synthetic miRNA mimic is transfected into cells to simulate over-expression. This approach has been widely used to study miRNAs in various cellular contexts, and has been important for determining the functions of individual miRNAs. However, we were interested in creating an expression vector that could be used for longer periods of time than would be allowed with a miRNA mimic. In addition, since miR-200s are expressed together endogenously, we were interested in examining the effects of simultaneous over-expression of multiple miR-200 family members. Thus, I first attempted to clone individual miR-200 family members miR-200a and miR-200b to be driven by a Keratin-14
(K14) promoter. Interestingly, when these miRNAs were cloned as individual stem loop regions, neither was expressed. When we amplified the entire miR-200b cluster and placed it under control of the K14 promoter, however, we saw robust over-expression of all three miR-200s within the locus (Fig. 3.3, below). This unexpected result indicates that the secondary structure formed by the primary transcript for the miR-200b cluster may be important for processing the associated miRNAs. Because the miR-200b cluster contains over half of the miR-200 family, and also contains at least one example of both seeds present within the family, we proceeded to use miR-200b cluster over-expression for the majority of our functional studies.

ii. Loss-of-function

Performing loss-of-function experiments for miRNAs has historically been more complicated than over-expression. miRNAs that are within their own genomic locus can be globally or conditionally knocked out using traditional methods or CRISPR. However, studying the loss of the miR-200 family is further complicated by the fact that its five members are spread across two genomic loci, requiring two separate knock-outs to generate cells lacking miR-200s. In addition, when I first began performing the experiments for this dissertation, CRISPR had not yet been developed as a gene editing method, and so we initially turned to other approaches.

In vitro, miRNA sponges can be used to knock down miRNA expression. Typically a single- or double-stranded RNA or LNA, sponges contain perfect binding sites for the

---

5 miR-200b cluster cloning and qPCR to confirm over-expression was performed by Dongmei Wang, PhD
miRNA of interest. As such, sponges bind miRNAs with a high affinity, titrating them away from interacting with their endogenous targets. Thus, the first approach I took in order to examine the effects of loss of miR-200s in cultured keratinocytes was the use of a sponge known as a Tough Decoy RNA, or TuD (Haraguchi et al. 2009). I designed a TuD with perfect binding sites for both miR-200a and miR-200b. Because published studies with TuD RNAs indicated successful knock-down of several let-7 family members with a single sponge (Xie et al. 2012), I reasoned that the presence of both seeds may be able function as a sponge for the entire miR-200 family. Indeed, I observed repression of all five miR-200 family members with this approach (Fig. 3.3). The fact that the entire miR-200 family was able to be repressed using binding sites specific to just miR-200a and miR-200b indicates that there is some sharing of targets between this miRNA family, although how often this occurs endogenously cannot be answered with this synthetic approach. Ultimately, the experiments done examining miR-200 family loss-of-function in vitro were performed using the miR-200a/200b TuD, as well as cells derived from a loss-of-function mouse model described below.
**Figure 3.3** Confirmation of miR-200b cluster over-expression and knock-down.

Relative luciferase activity for reporters with two perfect binding sites for each of the miR-200 family members in keratinocytes transfected with TuD-200a/b, TuD-203 or K14-200bcl. Dashed line represents relative luciferase activity for cells transfected with TuD-control. Error bars represent standard deviation. *P>0.05, **P>0.01.

c. Mouse models for miR-200 family over-expression

   i. Straight transgenic

To constitutively over-express miR-200 family members in murine skin, we generated a K14 driven over-expression model of the miR-200b cluster (genotype K14-200bcl, referred to as K14Tg). Germline transmission was achieved with one founder animal, leading to the generation of progeny with constitutive over-expression of the miR-200b cluster beginning during embryonic development. qRT-PCR and in situ hybridization confirmed strong over-expression of the miR-200b cluster in the skin of these animals (Fig. 3.4a-b). Unexpectedly, K14-200bcl animals were smaller than their littermate controls, and constitutive over-expression of miR-200s led to neonatal lethality (see Chapter 5). Due to
neonatal lethality of the F1 generation, a continuous line of K14-20obcl animals could not be established. I was able to collect a limited number of samples from these animals, but in order to perform an in depth study of miR-200 family over-expression the skin, I needed to generate a more sustainable mouse model.

![Figure 3.4 K14-Tg validation.](image)

**Figure 3.4 K14-Tg validation.**

A) qRT-PCR from keratinocytes derived from control and K14-Tg animals. B) *In situ* hybridization with a probe against miR-200b in wildtype and K14-Tg backskin at P9. Error bars represent SEM; scale bar is 200 um.

**ii. Inducible transgenic**

Because constitutive over-expression of the miR-200 family causes neonatal lethality, I chose to generate a mouse model with inducible over-expression. To that end, we cloned the miR-200b cluster into the pTRE2 expression construct, which allows for doxycycline-inducible expression when in the presence of a reverse tetracycline trans-activator, or
rtTA. Animals with a pTRE2-200bcl allele were bred with animals harboring a K14rtTA allele, generating progeny with doxycycline-inducible over-expression of the miR-200b cluster (genotype pTRE2-200bcl/K14rtTA, referred to as Tg). Feeding females doxycycline chow continuously throughout pregnancy generated progeny with robust miR-200 over-expression, as confirmed by both qRT-PCR and in situ hybridization (Fig. 3.5a-b).

**Figure 3.5. Inducible Tg validation.**

A) qRT-PCR on whole backskin of control and Tg epidermis at P0.5; n=3. B) In situ hybridization with a probe against miR-200b in wildtype and Tg backskin at P0.5. Error bars represent standard deviation. Scale bar is 100 µm.

d. *Mouse model for miR-200 family loss-of-function*

i. *Global and skin-specific knock-out models*

As mentioned above, studying loss of the miR-200 family is complicated by the presence of two genomic clusters. In order to create a loss-of-function model for all five miR-200s, we first used CRISPR to delete the entire miR-200b/a/429 locus. To accomplish this, we co-injected Cas9 mRNA and two sgRNAs with predicted cleavage sites flanking the miR-

---

6 sgRNAs for CRISPR were generated and validated by by Dongmei Wang, PhD
200b cluster (Fig. 3.6a, for details see Chapter 2). Three independently-generated founder lines were obtained, and each had a similar deletion pattern as expected by Cas9-mediated double-stranded DNA cleavage followed by ligation of cleaved ends (Fig. 3.6b). To minimize potential off-target effects, each founder line was bred for at least three generations before phenotypic analysis was conducted.

**Figure 3.6 miR-200 family knock-out target sites and validation.**

A) Target site, sgRNAs and DSB/repair site for deletion of the miR-200b cluster. B) Genotyping results for three miR-200b cluster founders. C) In situ hybridization with a probe
against miR-200b in control and dKO backskin at P0.5. D) qRT-PCR on whole epidermis from control and dKO animals at P0.5; n=3. Scale bars are 50 um. Error bars represent standard deviation.

To generate animals lacking expression of all miR-200s, we bred mice heterozygous for the miR-200b/a/429 cluster (miR-200b/a/429+/−) with mice harboring an existing floxed allele for the miR-200c cluster (Park et al. 2012) and also harboring E2a-cre or K14-cre. This resulted in progeny globally lacking all five miR-200 family members (genotype miR-200b/a/429−/−/200c/141fl/fl/E2acre, referred to as gdKO) or specifically in the skin (genotype miR-200b/a/429−/−/200c/141fl/fl/K14cre, referred to as dKO). I validated the loss of miR-200s with both qRT-PCR and in situ hybridization (Fig. 3.6c-d).

e. Chapter Three Summary

The data outlined in this chapter establishes the miR-200 family as a potentially important group of miRNAs in the skin. Small RNA-seq shows that they are among the most highly expressed miRNAs in neonatal epidermis. In addition, in situ hybridization indicates that miR-200s’ expression is dynamic throughout development, and uniquely enhanced within developing hair follicles. Thus, further examination of their roles in the skin and hair follicle is warranted. To that end, I developed several methods for miR-200 family over-expression and loss-of-function in both tissue culture and animal models. These methods allow for specific examination of miR-200s’ targets and functions within the skin and hair follicle, and are used to conduct the experiments found in the remainder of this thesis.
CHAPTER FOUR

GENOME-WIDE IDENTIFICATION OF MIR-200 FAMILY TARGETS

This chapter details the use of high-throughput, high-confidence methods for identifying miR-200 family targets. Historically, miRNA target prediction has relied upon seed sequence, muddling target prediction for miRNAs that share seeds. In addition, many studies have shown that non-canonical binding events, such as 3’ end binding, seed bulges, and centralized binding, are both functional and important for miRNA-mediated target repression. Further complicating reliable target identification is evidence that each individual miRNA is likely to repress the expression of several hundred genes; thus, the precise functions of these small RNAs cannot be understood without examining the potential contributions of many targets. A newly developed method, CLEAR-CLIP, allows for the precise determination of miRNA:mRNA interactions without the need to rely upon seed sequence predictions. Therefore, to most comprehensively study miR-200 family targeting events within the skin, we utilized CLEAR-CLIP performed on keratinocytes7, combined with RNA-seq performed on skin populations from control and Tg animals. The combination of these two techniques allowed for a precise look at miR-200 family targeting in the skin, as we were able to distinguish between targeting by different family members, as well as to determine which targets are relevant in vivo.

7 Glen Bjerke, PhD, performed the CLEAR-CLIP experiments and associated analysis, including cumulative distribution plots on RNA-seq data.
a. CLEAR-CLIP on murine keratinocytes

i. CLEAR-CLIP shows specific interactions between miRNAs and mRNAs

To gain the most comprehensive view of miR-200 family target regulation in murine skin, as well as to distinguish between targeting by different family members, we optimized Covalent Ligation of Endogenous Argonaute-bound RNAs with Cross-Linked Immuno-Precipitation (CLEAR-CLIP, (Moore et al. 2015b)) for use on mouse keratinocytes. This recently developed technique involves first cross-linking the miRNA-induced silencing complex (miRISC) to interacting mRNAs. miRISC is then pulled down using an antibody against AGO, and the complex is incubated with T4 RNA ligase to promote ligation between the miRNA embedded within miRISC and its interacting mRNA target (Fig. 4.1a). Resulting reads can still be comprised of either the miRNA or the mRNA alone (i.e. HITS-CLIP); however, a higher percentage of the reads than would be produced by HITS-CLIP alone are chimeras formed between the miRNA and its mRNA target. Sequencing of these miRNA:mRNA chimeras reveals both the mRNA target and the interacting miRNA, bypassing the requirement to search for seed sequences in order to predict targets.

AGO2 is the most highly expressed AGO protein in the skin, accounting for ~60% of all Ago protein expression (Wang et al. 2012). In addition, AGO2 shares a similar miRNA loading profile with AGO1 and AGO3, the two more minor AGO proteins expressed in the skin (Wang et al. 2012). Therefore, we used an antibody against AGO2 to perform CLEAR-CLIP on wildtype and miR-200 dKO keratinocytes. We recovered 1,521,535 miRNA:mRNA chimeras (~1.9% of total reads). As expected, the largest category of chimeric reads for miR-200s fell within 3’ UTRs, and we focused subsequent analyses
on those regions (Fig. 4.1b).

A)

![Diagram A](image)

B)

![Diagram B](image)

**Figure 4.1 CLEAR-CLIP schematic and read locations.**

A) Covalent ligation of endogenous argonaute-bound RNAs with cross-linked immunoprecipitation (CLEAR-CLIP) is performed by Ago2 IP followed by RNA ligation; this results in chimeric reads containing both the miRNA and the associated mRNA fragment. **B)** Percentage of miR-200 family CLEAR-CLIP reads within each genomic region

Comparison of HITS-CLIP and CLEAR-CLIP reads derived from the same library showed that CLEAR-CLIP provides much more robust sequence recovery (Fig. 4.2). Importantly, CLEAR-CLIP allowed us to distinguish between different miRNAs. For example, in WT CLEAR-CLIP reads within the 3’ UTR of the Quaking mRNA, we observed miRNA:mRNA chimeras for all five miR-200 family members in addition to miR-31, an unrelated miRNA also highly expressed in the skin. In CLEAR-CLIP performed on miR-200 dKO keratinocytes, however, we detected loss of signal only for miR-200s, but no loss of miR-31 signal (Fig. 4.2).
**Figure 4.2** CLEAR-CLIP is more specific than HITS-CLIP.

IGV tracks showing reads from HITS-CLIP, CLEAR-CLIP and miR-200 family-specific CLEAR-CLIP. CLEAR-CLIP allows for specific identification of miRNA:mRNA interactions.

**ii. Cross-seed binding between miR-200 family members is rare**

As mentioned previously, the miR-200 family shares two nearly-identical seed sequences: miR-200b, -c, and 429 have the b-type seed, and miR-200a and -141 have the a-type seed (Fig. 4.3a). Cross-seed binding has been observed when individual miR-200s were replaced in breast cancer cell lines lacking expression of the family (Bracken et al. 2014); however, the question of whether cross-seed binding normally occurs between family members in endogenous interactions remains unresolved. To determine the specificity of seed-sequence binding, we performed HOMER motif analysis on the mRNA portion of all chimeric reads recovered from each of the miR-200 family members. The most highly represented motif for each family member was its own seed sequence (Fig. 4.3b). Within those motifs, the most invariable sequences for each miRNA were derived from nucleotides 2-7, completely overlapping with the seed region. The nucleotide in the 8th position also showed a strong preference for matching the seed. In addition, miR-200s with the a-type seed had further enrichment for A or U in the 9th position. Overall, the
invariability within the seed region shown by HOMER analysis indicates that cross-seed binding between a- and b-type miR-200s is rare.

![miR-200 family sequences](image)

**Figure 4.3 Cross-seed binding between miR-200 family members is rare.**

A) miR-200 family member sequences shown by genomic cluster with the seed region highlighted. B) Top HOMER motif derived from the mRNA fragment portion of CLEAR-CLIP reads from each miR-200 family member is the family member’s own seed. C) Percentage of CLEAR-CLIP reads from each family member that contains a perfect seed match, a cross seed match, or no seed match.

Indeed, when we directly searched miR-200-interacting mRNAs for seed sequences, we found that the vast majority of reads contained the cognate seed—just a small percentage of chimeric reads appear to be generated from a miR-200 family member interacting with a target containing only a binding site for the opposite seed (Fig. 4.3c).m

Interestingly, for each miR-200 family member, ~30% of the identified interactions did
not contain a seed sequence within or near the captured mRNA region (Fig. 4.3c), suggesting that miR-200 targeting also occurs independently of canonical seed binding.

iii. miR-200s often interact with the same targets

![miRNA sequence vs Qk 3' UTR sequence diagram]

**Figure 4.4** Chimeric reads generated with CLEAR-CLIP.

CLEAR-CLIP chimeric reads from the 3’ UTR of Quaking (Qk) show that miR-200s with the same seed often share binding sites, but not in every case.

Another open question has been whether the miR-200 family truly cooperates in target repression. It has been previously suggested that miRNAs that share the same seed sequence are likely to repress many of the same mRNA targets. However, the fact that
most target identification methods have relied upon the seed itself has limited our ability to distinguish targets between miRNAs that do share seeds. Using CLEAR-CLIP, we were able to not only distinguish binding by different miR-200s, but also to determine how often miR-200s interact with the same target.

In many cases, we observed concurrent interaction of multiple miR-200 family members on the same transcript, and even at the same binding site (Fig. 4.4). As expected, miR-200s sharing the same seed were more likely to bind the same site, consistent with the rare occurrence of cross-seed binding we observed. From CLEAR-CLIP reads that passed the initial filter (see Chapter 2), we identified 3,601 mRNAs interacting with miR-200s. Of those, 52% (1,892) were found to be interacting with just one miR-200 family member, indicating that some mRNA regulation is specific to just one family member (Fig. 4.5a). However, 47% (1,709) of the mRNAs identified were found to be interacting with at least two miR-200s. Of those, 49% (847) were found to be interacting with two miR-200s, 27% (477) were interacting with three, 15% (262) were interacting with four, and 7% (123) were found to be interacting with all five miR-200s (Fig. 4.5b).

This result allowed us to further dissect interactions between the two seeds present within the miR-200 family. For example, when we examined chimeric regions interacting with two miR-200s, that is when two miR-200s were found to be interacting at the same binding site, both an a- and a b-type seed was present 20% of the time (Fig. 4.5c). The incidence of interaction with both seeds increased to 40% in regions interacting with three miR-200s (Fig. 4.5d). And of course, the mRNA regions with chimeric reads from
four or five miR-200 family members are inherently interacting with both seeds. These results highlight various aspects of miR-200 family targeting. First, although cross-seed binding appears to be rare when searching for individual miR-200s interacting with the opposite seed site, there is still a relatively high number of shared target sites between miR-200 family members that do not share seeds. Therefore, miR-200s with different seeds may be more likely to share targets through non-canonical, rather than seed-mediated, binding. In addition, concurrent miR-200 regulation occurs with relative frequency in endogenous contexts. And, though the number of genes interacting with all five miR-200s is relatively small, they still represent a significant portion of the interactions we identified, further highlighting the importance of cooperative regulation for the miR-200 family.

**Figure 4.5** Many miR-200 targets interact with more than one family member.
A) Proportion of miR-200 family chimeric reads interacting with one or two+ family members. B) Proportion of miR-200 family chimeric reads interacting with one to five family members. C) Regions interacting with two miR-200 family members that share the same seed (within seed) or have different seeds (across seed). D) as in C) but regions interacting with three miR-200 family members.

**iv. miR-200 target identification using CLEAR-CLIP alone**

For highest confidence miR-200 family target identification, we required that mRNA regions from chimeric reads fall within the 3’ UTR and be found within at least two libraries (for detailed description of how CLEAR-CLIP regions were selected, see Chapter 2). Using this method, we identified 1149 mRNAs as high-confidence miR-200 family targets. Gene ontology analysis for KEGG pathways and biological processes revealed that miR-200 targets are diverse, and are involved in processes from tissue development to pathways in cancer (Fig. 4.6a-b). While some of these genes and biological processes had previously been identified as miR-200 targets, none had yet been identified in the skin.

![GO Biological Process enriched in CLEAR-CLIP](image)

**Figure 4.6 GO-term analysis on miR-200 family targets identified with CLEAR-CLIP.**

A) KEGG pathways enriched in miR-200 family targets identified with CLEAR-CLIP. B) GO biological process enriched in miR-200 family targets identified with CLEAR-CLIP.
b. RNA-seq to probe miR-200 family functions in skin

i. FACS and RNA-seq

Given the enrichment for miR-200s within developing hair follicles, I wanted to determine the specific targets for this miRNA family in both early stage hair follicles as well as interfollicular epidermis. To accomplish this, I used a previously published method for isolating hair germ and interfollicular epidermis populations with Fluorescence-Activated Cell Sorting (FACS) (Rhee et al. 2006). This method takes advantage of the precise localization of p-cadherin, a cell adhesion protein highly expressed within developing hair follicles but not interfollicular epidermis. Using \( K14RFP \) to mark cells derived from the epidermal lineage, \( \alpha6 \)-integrin to mark cells located at the basement membrane, and p-cadherin to label hair germs, I was able to sort for hair germs (\( K14RFP^+ /p\text{-cad}^{hi}/\alpha6^{hi} \)) and interfollicular epidermis (\( K14RFP^+/p\text{-cad}^{low}/\alpha6^{hi} \)) (Fig. 4.7a).

Because miR-200s become restricted to developing hair follicles after the first wave of development (i.e. after P15.5), I chose to isolate hair germ and interfollicular populations at P0.5. I also wanted to be able to capture the differential expression that would result from miR-200 family over-expression; to that end, I bred \( pTRE2-200bcl/K14rtTA \) animals to \( K14-RFP \) animals, and fed pregnant females doxycycline chow. I then performed FACS, followed by RNA-seq on each population.
Figure 4.7 RNA-seq on hair germ and interfollicular epidermis.

A) Sorting strategy used to isolate hair germ and interfollicular epidermis populations. B) Gene set enrichment analysis (GSEA) comparing sorted hair germ population transcriptome
with published E14.5 hair germ signature (microarray). C) Principle component analysis (PCA) on hair germ and interfollicular epidermis populations from control and Tg animals. D) GSEA comparing Tg interfollicular epidermis to published E14.5 hair germ signature (microarray).

**ii. miR-200 over-expression is not sufficient to induce hair germ fate**

Gene Set Enrichment Analysis (GSEA, Broad Institute) comparing the hair germ population with a published microarray for hair germs at E14.5 showed that the sorting did indeed enrich for hair germs (Fig. 4.7b). Because miR-200s are highly enriched in developing hair follicles, I wondered what the effect of over-expressing this miRNA family would be on the transcriptome of the interfollicular epidermis. If the miR-200 family is necessary to promote hair follicle fate, the transcriptome of the interfollicular epidermis with miR-200 family over-expression should be more similar to that of the control hair germ population. I performed principle component analysis on differentially-expressed genes to determine whether miR-200 over-expression caused the interfollicular epidermis to become more hair germ-like (Fig. 4.7c). The hair germ and interfollicular populations from control samples were distinct and well-separated. Interestingly, the two populations from Tg samples were more related to each other than to either of the two populations from control samples. Furthermore, GSEA only showed minor enrichment for hair germ genes within Tg interfollicular epidermis (Fig. 4.7d). Thus, miR-200 over-expression does not cause a fate-switch in the interfollicular epidermis, but does cause the distinction between hair germ and interfollicular populations to be less clear.

**iii. CLEAR-CLIP targets are down-regulated in RNA-seq**

Within the RNA-seq datasets from hair germ and interfollicular epidermis, thousands of genes were down-regulated with miR-200 over-expression. Moving forward, we wanted to
know whether the genes we identified with CLEAR-CLIP were more likely to be down-regulated, as we would expect from *bona fide* miR-200 targets. Importantly, the transcriptomes of cultured keratinocytes, hair germ and interfollicular epidermis largely overlap (Fig. 4.8a), indicating that our CLEAR-CLIP performed on cultured keratinocytes is likely to capture miR-200 targets that are relevant *in vivo*.

**Figure 4.8** Genes with miR-200 family CLEAR-CLIP reads are down-regulated with miR-200 over-expression.

A) Venn diagram comparing transcriptomes from cultured keratinocytes, hair germ, and interfollicular epidermis. B) CDF plot of RNA-seq on HG-enriched population comparing genes with a miR-200 CLEAR-CLIP read containing a seed (red line, $P<2.2^{-16}$ compared with
no CLEAR-CLIP read), a miR-200 CLEAR-CLIP containing no seed (blue line, \(P=3.26^{-5}\) compared with no CLEAR-CLIP read and \(P=0.005\) compared with genes with a seed), or no miR-200 CLEAR-CLIP read (black line). C) CDF plot of RNA-seq on IFE-enriched population comparing genes with a miR-200 CLEAR-CLIP read containing a seed (red line, \(P<2.2^{-16}\) compared with no CLEAR-CLIP read), a miR-200 CLEAR-CLIP read containing no seed (blue line, \(P=0.0003\) compared with no CLEAR-CLIP read or \(P=0.0009\) compared with genes with a seed), or no miR-200 CLEAR-CLIP read (black line).

Genes identified as miR-200 targets with CLEAR-CLIP were robustly down-regulated in RNA-seq from both hair germ and interfollicular epidermis (Fig. 4.8b-c). Interestingly, this down-regulation did not seem to be entirely dependent upon the seed: genes with a CLEAR-CLIP read but no seed match were still down-regulated, although to a more modest extent (Fig. 4.8b-c, blue lines). These results indicate that miR-200 target recognition generally leads to down-regulation at the mRNA level, and that seed-mediated target recognition by miR-200s leads to the most robust down-regulation.

**iv. miR-200s largely target the same processes in hair germ and interfollicular epidermis**

Although they are derived from the same lineage and share the majority of their transcriptomes, cells from the hair germ and the interfollicular epidermis represent distinct epidermal populations (Fig. 4.7). Thus, I wanted to know whether the miR-200 family has a distinct targeting profile in hair germs versus interfollicular epidermis. By combining CLEAR-CLIP with RNA-seq, I was able to determine which of the miR-200 family targets are relevant in each population. Because miRNAs repress each individual target to only a modest extent, I considered any gene with a CLEAR-CLIP read and down-regulated \(>10\%\) in one or both populations to be a miR-200 family target (Fig. 4.9a). In this way, I identified 141 miR-200 family targets specific to the hair germ, 224 specific to
the interfollicular epidermis, and 388 shared between the two populations, for a total of 753 high-confidence miR-200 family targets in the skin (Fig. 4.9a).

**Figure 4.9** Identification of high-confidence miR-200 family targets with CLEAR-CLIP and RNA-seq.

A) Number of genes identified in CLEAR-CLIP and down >10% in HG- or IFE-enriched populations. B) GO-term analysis on miR-200 family targets specific to the hair germ. C) GO-term analysis on miR-200 family targets specific to the interfollicular epidermis.

To determine whether there is any significant difference in the processes or pathways regulated by miR-200s in the hair germ and interfollicular epidermis, I performed GO-term analysis on genes enriched in each population. There was some enrichment for unique signatures between the two, such as membrane organization in the hair germ, and tissue and epithelium development in the interfollicular epidermis (Fig. 4.9b-c). Overall, however, the enrichment for these processes was small, and examination of the genes involved also did not reveal any strong repression of any particular type of gene in either population (such as transcription factors, signaling molecules, and others). Therefore,
although there are some distinct targets for the miR-200s in the hair germ and interferollicular epidermis, no strong targeting of a particular process or pathway occurs in only one population.

v. Identification of high-confidence miR-200 family targets with CLEAR-CLIP and RNA-seq

Moving forward, I chose to consider all 753 targets identified in either the hair germ, the interferollicular epidermis, or shared between the two as one group of skin-specific miR-200 family targets. Interestingly, neither CLEAR-CLIP nor RNA-seq revealed any evidence of miR-200 interaction with Zeb1 or Zeb2, the two most well-characterized targets of the miR-200 family (Gregory et al. 2008). This is likely due to the fact that Zeb1 and Zeb2 are mesenchymal genes not typically expressed in normal epidermis or keratinocytes—in fact, CLEAR-CLIP showed no reads for Zeb1 or Zeb2 at all. Importantly, I was able to recover several previously identified miR-200 family targets, such as Sec23a and Cfl2, which are expressed in the skin (Korpal et al. 2011; Bracken et al. 2014). Thus, the lack of any miR-200:Zeb1/2 interactions validates the tissue-specificity of my approach to target identification, as well as highlights the importance of understanding the functions of this miRNA family in endogenous contexts.

GO-term analysis on all 753 genes identified as miR-200 family targets revealed strong signatures for pathways in cancer, focal adhesion, actin cytoskeleton, and the Hippo signaling pathway (Fig. 4.10a). The regulation of genes involved with several of these signatures has been previously reported in cancer cell lines (Gregory et al. 2008; Bracken et al. 2014); however, the widespread and coordinated regulation of these processes by the miR-200 family in the skin is striking. Our analysis showed that many,
rather than a few genes within each pathway are targeted by miR-200s, highlighting the ability of this miRNA family to cooperate in target recognition. Furthermore, chimeric reads generated with CLEAR-CLIP allowed us to identify which miR-200 family member was interacting with a given transcript (Fig. 4.10b). By examining genes involved with different processes, such as focal adhesion, cell cycle and Hippo signaling, we were able to observed that not only are miR-200s cooperating to repress multiple genes within a given pathway, but that they further collaborate by interacting with many of the same actual targets (Fig. 4.10b).

vi. miR-200s cooperate in target repression but there is specificity in their targeting

In order to validate these miR-200 targets and to further study coordinated regulation by the miR-200 family, I selected a set of ten new or previously-validated targets for a heterologous luciferase assay (Fig. 4.10b).

<table>
<thead>
<tr>
<th>GO Biological Process Enriched in All miR-200 Targets</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathways in cancer</td>
<td>3.47e-8</td>
</tr>
<tr>
<td>Hippo signaling pathway</td>
<td>1.01e-7</td>
</tr>
<tr>
<td>Focal adhesion</td>
<td>1.98e-6</td>
</tr>
<tr>
<td>Regulation of actin cytoskeleton</td>
<td>3.64e-5</td>
</tr>
<tr>
<td>Cell cycle</td>
<td>1.06e-4</td>
</tr>
<tr>
<td>Adherens junction</td>
<td>1.10e-4</td>
</tr>
</tbody>
</table>

**Figure 4.10** Cooperative targeting by the miR-200 family.

A) GO biological process enriched in all miR-200 family targets identified with CLEAR-CLIP and RNA-seq. B) Selected miR-200 family targets and the miR-200 family member(s) they were found to be interacting with.
I intentionally chose targets that interact with the miR-200 family in a diverse manner. For example, there are targets that appear to be exclusive to one family member (Fat1), or to one seed (Ywhab, Lats1, Lats2, Fat2). Additionally, there are targets interacting exclusively with one cluster (Ccnd2), which inherently contains both seeds. Finally, several of the targets were found to be interacting with both clusters simultaneous, or even to be interacting with all five miR-200s (Ccng2, Snai2).

Using luciferase reporters with perfect binding sites for several of the miR-200 family members, I was able to determine that expression of either the miR-200b or the miR-200c cluster resulted in robust down-regulation of the miRNAs found within that cluster (Fig. 4.11). Interestingly, expression of either cluster alone also down-regulated luciferase expression for all miRNA-specific reporters, indicating that the seed is an important determining factor in some target repression.
**Figure 4.11** Heterologous luciferase assay.

Relative luciferase activity of 3’ UTRs for target genes with the addition of the miR-200b cluster, the miR-200c cluster, or both; n=6. *P<0.05, **P<0.01, ***P<0.001, ns=non-significant; error bars represent standard deviation.

I next assessed the ability of each cluster alone, or both in combination, to repress luciferase reporters specific to target 3’ UTRs. In some cases, addition of one cluster but not the other caused repression (Ywhab, Lats1, Lats2, Fati, Fat2, Ccnd2, Fig. 4.11). For example, only the miR-200c but not the miR-200 cluster repressed Fati, which was recognized only by miR-141. In the case of several other genes, which were found to be interacting with both clusters, addition of either cluster individually was able to induce repression (Ptpn14, Ccng2, Cfl2, Snai2). Of note, addition of both clusters caused stronger repression of Ywhab, Ccng2, Cfl2 and Snai2, all of which showed interactions with miR-200s from both clusters. In the case of Ccng2 and Snai2, all five miR-200s were found to
be interacting with their 3’ UTRs, further highlighting the ability of this miRNA family to cooperate in target repression. Overall, all of the selected targets were confirmed by the luciferase assay, supporting the ability to identify miR-200 family targets by combining CLEAR-CLIP and RNA-seq.

c. Chapter Four Summary

In order to assemble the most comprehensive list of high-confidence targets for the miR-200 family in the skin, I combined CLEAR-CLIP with RNA-seq. CLEAR-CLIP generates sequencing reads that are chimeras between miRNAs and their mRNA targets, allowing us to specifically dissect how miR-200 family members interact with their targets. About half of all of the genes identified in miR-200 family chimeric reads were only interacting with one family member, indicating that there is a certain level of specificity to the targets for each miR-200. A nearly equal number of genes, however, were found to be interacting with two or more family members, in many cases at the same site. Thus, miR-200s cooperate in target repression by concurrently binding to separate regions on the same 3’ UTR, as well as by targeting the same sequences. Furthermore, miR-200s were found to target many genes within the same pathways, indicating that their regulation extends beyond individual genes to impact entire regulatory networks.

To identify miR-200 family targets relevant in the skin, and to examine whether there is differential targeting by miR-200s in different populations, I performed RNA-seq on hair germ and interfollicular epidermis. This analysis showed that there are miR-200 family targets specific to either the hair germ or the interfollicular epidermis, but no strong enrichment for pathways or processes exclusive to just one population. Overall, I
identified 753 miR-200 family targets in the skin, which showed strong signatures involved with focal adhesions, actin cytoskeleton, cell cycle regulation, and the Hippo signaling pathway. A heterologous luciferase assay validated several of these targets. Furthermore, the assay confirmed the CLEAR-CLIP results, showing that individual miR-200 clusters are able to repress some targets but not others, despite the presence of both seeds in each cluster. In addition, genes which showed interaction with all five miR-200s in CLEAR-CLIP were most strongly repressed when both clusters were present. These results suggest that while there is specificity to miR-200 family targeting, different family members also cooperate to cause strong repression of some targets.

Together, these data represent the most comprehensive examination of miR-200 family targeting paradigms to date, and establish the largest list of high-confidence targets for these miRNAs in the skin. Furthermore, they identify the pathways and processes regulated by the miR-200 family, laying the groundwork for better definition of miR-200s’ functions in the skin.
CHAPTER FIVE

PHENOTYPES ASSOCIATED WITH MIR-200 FAMILY OVER-EXPRESSON

The identification of high-confidence miR-200 family targets described in Chapter 4 revealed that miR-200s regulate diverse targets with a broad array of functions. In order to better understand the role of the miR-200 family in skin, and how suppression of the identified targets may contribute to miR-200 function, I developed two transgenic mouse models for over-expression of miR-200b, -200a, and -429 (see Chapter 3 for details). Importantly, the over-expression of these three miRNAs includes both of the seed sequences present within the miR-200 family. CLEAR-CLIP data presented in Chapter 4 indicates that while there is some specificity to targeting by miR-200s, members of this family do share target sites; thus, over-expression of both seeds should allow for better analysis of the functions of this family as a whole.

This chapter details the use of the over-expression models to assess cellular and developmental phenotypes associated with miR-200 family over-expression. Guided by miR-200 family targets identified through CLEAR-CLIP and RNA-seq, the experiments detailed in this chapter reveal miR-200 family regulation of proliferation, migration, cell adhesion, and signaling pathways both in vitro and in vivo. As the experiments

8 Whereas the straight Tg (K14-Tg) consitutively and robustly over-expresses miR-200b, -200a and -429, the inducible Tg requires the addition of doxycycline to induce such over-expression. Feeding females doxycycline chow throughout pregnancy resulted in progeny with the most robust miR-200 family over-expression; thus all Tg animals used in experiments for this chapter were generated from continual doxycycline feeding throughout pregnancy, unless otherwise noted.
performed here show, modulation of such diverse processes ultimately allows the miR-200 family to regulate hair follicle development.

**a. miR-200 over-expression leads to neonatal lethality and decreased hair follicle number**

**ii. Neonatal lethality**

As is detailed in Chapter 3, I developed two separate mouse models for miR-200 over-expression: a K14-driven straight transgenic (K14-Tg) and a doxycycline-inducible transgenic (Tg). At birth, both K14-Tg and Tg animals were indistinguishable in size from littermate controls (Fig. 5.1a,c). However, animals with miR-200 over-expression noticeably lacked a milk-spot, an indication of healthy feeding. Over time, K14-Tg and Tg animals became smaller than littermate controls, and eventually succumbed to neonatal lethality before P10 (Fig. 5.1b,d). Although I did not specifically investigate the cause of lethality, given the lack of a milk spot in K14-Tg and Tg pups, they likely died from some form of malnutrition. This may have been due to competition from control littermates, or it may have been a primary defect associated with miR-200 over-expression— K14 is expressed not only in the skin but also many other stratified epithelia, so malnutrition may have arisen from improper esophageal or fore-stomach development, for example. To avoid possible secondary effects of malnutrition in observing phenotypes, the majority of the analyses on Tg animals were performed immediately after birth, at P0.5.
**Figure 5.1** miR-200 over-expression is neonatal lethal.

K14-Tg animals at A) P0.5 and B) P9. Inducible Tg animals at C) P0.5 and D) P4.5.

**ii. Hair follicle staging**

Despite the neonatal lethality observed with miR-200 family over-expression,\(^9\) epidermal stratification and differentiation remained intact (Fig. 5.2a-c). This, combined with the fact that the lethality did not occur until several days after birth, indicates that miR-200 family over-expression did not disrupt epidermal barrier formation. I next wanted to determine whether there were any defects in the progression of hair morphogenesis. At Po.5, back-skin hair follicles are typically found in a variety of different stages, reflecting the various waves in which hair follicles develop. When I examined the percentage of hair follicles at each developmental stage, I saw no difference between Tg and control animals (Fig. 5.2d). Interestingly, however, I observed a decrease in the number of hair follicles in Tg animals compared with controls (Fig. 5.2e). Therefore, miR-200 over-expression does

---

\(^9\) Because the K14-Tg line was derived from a single male founder, and the F1 generation with miR-200 family over-expression all died prematurely, the number of samples from that line was limited. Thus, the majority of the analyses done with miR-200 family over-expression were performed on animals from the inducible Tg line. Experiments performed with the K14-Tg are noted as such.
not affect the staging of hair follicles during development, but may decrease hair follicle specification.

**Figure 5.2 Epidermal stratification and hair follicle staging with miR-200 over-expression.**

A) Immunofluorescence for loricrin (red) and K5 (green) in Tg backskin at P0.5. B) Immunofluorescence for K1 (red) and K5 (green) in K14-Tg backskin at P0.5. C) H&E stain on Tg backskin at P0.5. D) Percentage of hair follicles at each stage of development in control and Tg animals at P0.5; n=3. E) Number of hair follicles per section from control and Tg animals at P0.5; n=3 animals. Error bars represent standard deviation; scale bars are 100 um.
**b. miR-200 over-expression reduces proliferation and migration in vitro**

i. Reduced proliferation

![Cell cycle table and colony formation assay](image)

**Figure 5.3 miR-200 over-expression reduces proliferation.**

A) miR-200 family targets identified with CLEAR-CLIP and RNA-seq involved with proliferation. B) Colony formation assay of inducible keratinocytes untreated or with doxycycline and signal intensity of colonies formed by inducible keratinocytes untreated or with doxycycline; n=3. C) Number of colonies formed from freshly-isolated K14-Tg and control keratinocytes; n=3. D) Representative flow plots of EdU incorporation in Tg keratinocytes untreated or treated with doxycycline for 48 hours; n=3. Error bars represent standard deviation. *P<0.05; **P<0.01; ***P<0.001.
Moving forward in assessing the effects of miR-200 family over-expression, I was guided by the miR-200 family targets identified through CLEAR-CLIP and RNA-seq. Among the strongest signatures within those targets was positive regulation of cell cycle. This group included many cyclins, as well as Cdk6, Mcm4 and Myc (Fig. 5.3a). Thus, miR-200 over-expression should cause more robust down-regulation of these targets, and I wanted to determine the effect of that over-expression on proliferation. To that end, I performed a colony formation assay on keratinocytes cultured from Tg animals, left untreated or treated with doxycycline to induce over-expression. miR-200 induction reduced colony formation (Fig. 5.3b). A colony formation assay on freshly-isolated keratinocytes from K14-Tg animals also revealed that those cells have reduced colony forming ability (Fig. 5.3c). Additionally, an EdU incorporation assay on cultured Tg keratinocytes with or without doxycycline showed that miR-200 over-expression reduced EdU incorporation (Fig. 5.3d). Taken together, these results indicate that miR-200 over-expression reduces proliferation and inhibits cell cycle progression in keratinocytes.

**ii. Reduced migration**

Another strong signature that arose within miR-200 family targets was regulation of cell migration, which included genes involved with both focal adhesion and actin cytoskeleton (Fig. 5.4a). To examine cell migration in vitro, I performed a scratch assay. miR-200 over-expression reduced the number of cells which migrated, indicating that miR-200 over-expression inhibits migration of cultured keratinocytes (Fig. 5.4b-c).
Figure 5.4 miR-200 over-expression reduces cell migration.

A) miR-200 family targets identified with CLEAR-CLIP and RNA-seq involved with the regulation of focal adhesions and actin cytoskeleton. B) Scratch assays on inducible keratinocytes untreated or with doxycycline 0 hours or 8 hours after scratching. C) Number of cells migrated in scratch assay; n=3. Error bars represent standard deviation; ***P<0.001.

c. miR-200 over-expression affects cell adhesion and differentiation

i. Disrupted focal adhesion formation

As mentioned above, a number of key regulators of focal adhesion and actin were identified as miR-200 family targets, including Ptk2 (focal adhesion kinase/FAK), Crk, Rock2 and Wasl (N-WASP, Fig. 5.4.a). Therefore, I wanted to know whether miR-200 family over-expression had any effect on the formation of focal adhesions. Upon miR-200 family over-expression in cultured keratinocytes grown in low calcium conditions, I observed an increased number of focal adhesions as evidenced by staining for vinculin and phalloidin (Fig. 5.5a). This result is similar to what was seen when miR-200s were added back to a breast cancer line lacking expression of the family (Bracken et al. 2014). I observed that when miR-200s were over-expressed in keratinocytes, the focal adhesions that formed were more numerous but much smaller than those seen in control cells (Fig.
Thus, the coordinated repression of multiple genes involved in focal adhesion appears to cause an increase in number, but a decrease in size, of such structures. Interestingly, miR-200 over-expression did not affect the ability of keratinocytes to adhere in culture, indicating that the observed disruptions in focal adhesion formation do not strongly affect cell adhesion.

**Figure 5.5 miR-200 over-expression disrupts focal adhesion formation.**

A) Immunofluorescence of Vinculin (red) and Phalloidin stain for actin (green) on inducible cells untreated or treated with doxycycline plated in low calcium; B) Number of focal adhesions per cell; n=76 untreated cells, n=77 dox-treated cells. Scale bars are 20 um. ***P<0.001.

**ii. Impaired adherens junction formation**

Actin cytoskeleton dynamics and adherens junction formation are intimately linked (Vasioukhin et al. 2000; Kanchanawong et al. 2010; Ratheesh & Yap 2012; Case & Waterman 2015). Because many miR-200 family targets such as Cdc42 and Actn1 (α-actinin) are involved with actin cytoskeleton dynamics and filopodia formation
(Krugmann et al. 2001), I wanted to examine the effect of miR-200 over-expression on adherens junction formation.

In culture, addition of calcium to keratinocytes quickly causes cytoskeletal rearrangements and filopodia formation, resulting in robust cell-cell junctions after less than 24 hours (Vasioukhin et al. 2000). To examine adherens junction formation with miR-200 over-expression, I induced cultured keratinocytes with calcium for three to seven hours. As expected, relocation of E-cadherin to the cell membrane and attachment of stress fibers to newly formed adherens junctions was readily detected in control cells three hours after calcium addition (Fig. 5.6a-b). Strikingly, although E-cadherin relocation was still visible with miR-200 over-expression, the attachment of actin stress fibers to this protein was strongly repressed (Fig. 5.6b, arrows). By seven hours post-calcium addition, control cells had largely sealed cell-cell contacts, and E-cad/stress fiber co-localization was clearly visible at the cell cortex. With miR-200 over-expression, however, the earlier-stage E-cadherin zippers were still present, and actin stress fibers were generally not localized to the cell cortex, reflecting impaired adherens junction formation (Fig. 5.6a-b). Thus, miR-200 over-expression impairs the process of adherens junction formation in cultured keratinocytes.
iii. Disrupted cell adhesion in vivo

Because I saw disrupted adherens junction formation in vitro, I wanted to know whether miR-200 over-expression disrupted cell-cell contacts in vivo as well. Additionally, many miR-200 family targets are specifically involved with adherens and tight junctions in vivo, such as Ocln, Tjp1, and several claudins (Fig. 5.7a). These structures are primarily important for cell-cell interactions in the suprabasal layers of the skin, which are undergoing keratinization, and therefore disruptions may lead to problems with barrier formation. However, as mentioned above, epidermal differentiation and barrier formation were intact in Tg animals, indicating no severe defects in initial cell-cell contact formation (Fig. 5.2a-b). On the other hand, staining for α-catenin did show more diffuse and scattered localization in Tg back-skin compared with controls (Fig. 5.7b). Additionally, as Tg animals aged, their skin appeared to thicken and crack in some places, taking on a scaly appearance and peeling away, particularly around the mouth, legs and belly (Fig. 5.7c-d, arrows). Thus, while miR-200 over-expression does not disrupt the initial formation of cell junctions, it appears to cause disruptions in cell-cell contacts over time, ultimately leading to problems with skin integrity.
Figure 5.7 miR-200 over-expression disrupts adherens junction formation in vivo.

A) miR-200 family targets identified with CLEAR-CLIP and RNA-seq involved with tight and adherens junctions. B) Immunofluorescence for α-catenin in control and Tg backskin at P0.5. C) Tg (left) and control (right) pups at P5.5. D) Transgenic pup at P5.5; arrows indicate areas where skin is breaking and peeling away from the animal. Scale bars are 25 um.

iv. miR-200s regulate components of the Notch signaling pathway

Among the list of miR-200 targets were genes involved with Notch signaling, including Notch2 and Jag1 (Fig. 5.8a). In the skin, Notch activation slows cell growth and promotes differentiation, and has also been suggested to play a role in hair morphogenesis (Kopan & Weintraub 1993; Vauclair et al. 2005; Okuyama et al. 2008). In addition, miR-200s have previously been shown to regulate Notch signaling, although this relationship has never been shown in the epidermis (Yang et al. 2011; Vallejo et al. 2011). Therefore, I wanted to determine whether miR-200s regulate Notch signaling in epidermal cells. As expected, a luciferase assay showed decreased Notch activity with miR-200 over-expression during keratinocyte differentiation (Fig. 5.8b). Thus, although the differentiation appears to
occur normally with miR-200 over-expression (Fig. 5.2a-b), Notch activity is reduced, and may contribute to some of the observed phenotypes.

Figure 5.8 miR-200 over-expression reduces Notch activity in vitro.

A) miR-200 family targets identified with CLEAR-CLIP and RNA-seq involved with Notch signaling. B) Relative luciferase activity of Notch-RBPj reporter in Tg cells induced to differentiate with 1.5μM calcium for 48 hours and untreated or treated with doxycycline for 48 hours; n=3. Error bars represent standard deviation. ***P<0.001.

d. miR-200s regulate components of the Hippo signaling pathway

i. Nuclear retention of Yap

In addition to the signatures mentioned above, the list of high-confidence miR-200 family targets included several genes that regulate the nuclear/cytoplasmic localization of Yap, the key transcription factor of the Hippo signaling pathway (Fig. 5.9a). This pathway is involved in regulating organ growth and mediating tumor suppression, and is intimately linked to cell adhesion and actin cytoskeletal dynamics (Matsui & Lai 2013). For example, when there are few cell-cell connections, Yap1 is translocated to the nucleus to promote cell proliferation and reduce apoptosis (Mo et al. 2014). Proper Hippo signaling has also been suggested to be important for hair follicle development (H. Zhang et al. 2011). Yap
localization is largely controlled by its phosphorylation state and its interaction with cell-adhesion proteins. Within the list of high-confidence miR-200 targets were genes whose expression positively enhances Yap1 phosphorylation including Lats1, Lats2, and Ptpn14. Additionally, there were several genes that promote cytoplasmic retention of phosphorylated Yap1, including several 14-3-3 proteins (Ywhab, Ywhag, and Ywhaz, Fig. 5.9a).

**Figure 5.9** miR-200 over-expression induces retention of nuclear Yap.
A) miR-200 targets identified with CLEAR-CLIP and RNA-seq involved with Yap nuclear localization. B) Immunofluorescence of K5 (green) and Yap1 (red) in control and Tg backskin at P0.5. Representative image from n=6 pairs of animals. C) Immunofluorescence of Yap1 and Hoescht in Tg keratinocytes untreated or treated with doxycycline for 48 hours. D) Fluorescence intensity of nucleus compared with cytoplasm from cells in C; n=3. Scale bars are 50 um. Error bars represent standard deviation. *P<0.05.

Thus, I first wanted to determine the nuclear/cytoplasmic localization of Yap in Tg and control skin. In wildtype animals, Yap displays nuclear localization throughout the developing epidermis early in development; however, by Po.5, Yap nuclear localization becomes restricted to developing hair follicles, and is excluded from the interfollicular epidermis (H. Zhang et al. 2011). In Tg animals, I observed widespread nuclear retention of Yap throughout the epidermis as well as the hair follicles at Po.5 (Fig. 5.9b). Inducing miR-200 over-expression in culture also resulted in increased nuclear Yap signal (Fig. 5.9c-d).

The animals with widespread nuclear Yap localization at Po.5 were transgenic animals whose mothers were fed doxycycline chow throughout pregnancy. Thus, to determine whether miR-200 over-expression is able to cause Yap nuclear localization later during development, I induced miR-200 over-expression postnatally with doxycycline injection. Interestingly, this later induction not cause any translocation of Yap to the nucleus, suggesting that miR-200 over-expression causes nuclear retention rather than translocation of Yap to the nucleus.

**ii. Reduced Yap phosphorylation and disrupted cell-cell contacts**

I next wanted to determine what may be causing nuclear retention of Yap with miR-200 over-expression. As mentioned, Yap localization in the cell is partially controlled by its
phosphorylation state. Reasoning that a decrease in genes which phosphorylate Yap (Lats1, Lats2, Fig. 5.9a) may lead to a decrease in Yap phosphorylation, thus leading to its nuclear retention, I examined the phosphorylation state of Yap in Tg and control epidermis. I did note a decrease in the amount of phosphorylated Yap in Tg animals (Fig. 5.10a); however, this decrease did not prove to be statistically significant, indicating that reduced phosphorylation is not the primary driver of Yap nuclear localization in Tg backskin. In addition, despite the increased nuclear Yap signal in cultured keratinocytes with miR-200 over-expression, I did not observe any decreased Yap1 phosphorylation in that system.

**Figure 5.10** miR-200 over-expression reduces Yap phosphorylation and disrupts cell junctions.

A) Immunoblot for Yap\(^{\text{Ser127}}\) on whole epidermis from Tg and control animals; n=4; p=0.074. B) Immunofluorescence of \(\beta\)-catenin (green) and Yap1 (red) in control and Tg backskin at P0.5. Scale bars are 20 um. Arrow indicates co-localization of nuclear Yap with disrupted \(\beta\)-catenin. Representative image from n=6 pairs of animals.

In addition to being affected by its phosphorylation state, Yap localization can also be regulated by cell-cell contact. Because I observed disruptions in cell adhesion with miR-200 family over-expression, I wanted to determine whether there was any link between those disruptions and the observed Yap nuclear localization. Indeed, disruptions in \(\beta\)-catenin, which showed the same diffuse and scattered localization as \(\alpha\)-catenin (Fig.
5.7b), correlated with Yap nuclear localization in the interfollicular epidermis (Fig. 5.10b). Thus, miR-200s regulate Yap nuclear localization by modestly regulating its phosphorylation state, and also by disrupting the cell adhesion structures required for its cytoplasmic retention.

iii. miR-200s target processes and genes downstream of Yap

Nuclear accumulation of Yap is usually associated with increased cell growth, hyperproliferation, and protection against apoptosis (Elbediwy et al. 2016). However, multiple positive regulators of cell cycle and proliferation were direct miR-200 family targets (Fig. 5.3a). In addition, miR-200 targets include genes such as Xiap which are involved with repressing apoptosis, and over-expression of this family has previously been shown to promote apoptosis in the pancreas (Belgardt et al. 2015). Consistent with the decreased colony formation and EdU incorporation in keratinocytes over-expressing miR-200s in vitro (Fig. 5.3b-d), I observed decreased EdU incorporation in Tg animals compared with controls (Fig. 5.11a-b). Additionally, the backskin of Tg animals had many more apoptotic cells at P0.5 as evidenced by activated caspase-3 staining (Fig. 5.11c). Finally, I noticed that genes previously shown to be up-regulated by Yap were in fact down-regulated with miR-200 over-expression (Fig. 5.11d). One of these, Ctgf, appears to be a direct miR-200 family target based off of CLEAR-CLIP and RNA-seq. A fluorescence reporter for Yap activity also did not show any response with miR-200 induction, further supporting the idea that the nuclear Yap observed with miR-200 over-expression does not lead to up-regulation of its targets. Thus, although miR-200 over-expression leads to
nuclear retention of Yap, the pro-growth and anti-apoptotic effects associated with this localization appear to be dampened by the miR-200s.

**Figure 5.11** miR-200 over-expression dampens the effects of nuclear Yap.

A) Immunofluorescence of EdU (green) and K5 (red) in control and Tg backskin at P0.5. B) Percentage of EdU+/K5+ cells in control and Tg epidermis at P0.5; n=6 pairs of animals. C) Immunofluorescence of activated caspase-3 (red) and K5 (green) in control and Tg backskin at P0.5. D) qRT-PCR for Ctgf and Cyr61 on whole epidermis of control and Tg animals at P0.5; n=3. Error bars represent standard deviation. *P<0.05; ***P<0.001. Scale bars are 50 um.

e. **miR-200 over-expression severely disrupts hair follicle development**

i. Inverted hair germs

Having observed that miR-200s regulate proliferation, migration, and cell adhesion, as well as components of the Notch and Hippo signaling pathways, I wanted to determine whether I could observe defects caused by this coordinated mis-regulation. During skin
morphogenesis, miR-200s are highly expressed in developing hair follicles (see Chapter 3). Interestingly, hair follicle development requires precise coordination of many of the processes and pathways controlled by miR-200s (see Chapter 1). In addition, active developmental processes such as hair morphogenesis provide a unique opportunity to capture dynamic miRNA regulation, and thus afford the chance to capture miRNA functions that may be missed during homeostasis.

Therefore, I wanted to determine the effect of miR-200 over-expression on hair follicle development. I observed a striking defect in early-stage hair follicles with miR-200 family over-expression: rather than invaginating down towards the dermis as in controls, Tg hair germs evaginate upwards toward the surface of the skin (Fig. 5.12a). I also noted inverted hair germs in K14-Tg animals, although I focused further analysis on inducible Tgs (Fig. 5.12c). About one third of hair germs in Tg animals showed this unique defect (Fig. 5.12b). Lef-1 staining and alkaline phosphatase activity detected in the underlying dermal papilla indicated that these structures are in fact bona fide hair germs (Fig. 5.12d). Importantly, inverted hair germs still down-regulated e-cadherin at the leading edge, a step that has been shown to be important for hair germs (Fig. 5.12e). This suggests that over-expression of miR-200s in the hair germ does not promote e-cad expression, as has been seen in cancer (Gregory et al. 2008; Park et al. 2008; Korpal et al. 2008).
**Figure 5.12** miR-200 over-expression causes inverted hair germs.

A) Immunofluorescence of Lef-1 (green) and β4-integrin (red) in control and Tg backskin at Po.5.  
B) Percentage of ctrl and Tg hair germs with inverted appearance; n=6.  
C) Immunofluorescence of Lef-1 (red) and β4-integrin (green) in control and K14Tg backskin at Po.5.  
D) Alkaline phosphatase (blue) and eosin (pink) staining on ctrl and Tg back-skin at Po.5.  
E) Immunofluorescence of E-cadherin (green) and β-catenin (red) in control and Tg backskin at Po.5. Scale bars are 50 um. Error bars represent standard deviation. ***P<0.001.

**ii. Imprecise cell fate specification**

During normal hair germ specification, cells at the leading edge receive strong Wnt signaling and express high levels of Lhx2 (Rhee et al. 2006). In contrast, cells one layer behind the leading edge typically express Sox9 as a result of assymetric cell division and
lower Wnt signaling (Ouspenskaia et al. 2016b). In Tg skin, the demarcation between Lhx2+ cells at the leading edge and Sox9+ cells in the suprabasal layer was compromised. Whereas control hair germs exclusively displayed Lhx2+ cells at the leading edge, inverted hair germs often had suprabasal expression of this marker (Fig. 5.13a). In addition, Sox9+ cells were seen at the leading edge of inverted hair germs (Fig. 5.13b).

**Figure 5.13** Inverted hair germs display imprecise cell fate specification.

A) Immunofluorescence of Lhx2 (red) and β4-integrin (green) in control and Tg backskin at Po.5. Arrows indicate suprabasal Lhx2+ cells. B) Immunofluorescence of Sox9 (red) and K5 (green) in control and Tg backskin at Po.5. Arrows indicate Sox9+ cells at the leading edge. C) Immunofluorescence of α-catenin (red) and staining with phalloidin (green) in control and Tg backskin at Po.5. Arrow indicates strong actin signal in centralized suprabasal cell. All representative images from n=6 pairs of animals. Scale bars are 25 um for A and 50 um for B and C.

Staining for phalloidin also revealed differences in actin cytoskeleton in cells within inverted hair germs. Typically, strong actin bundles appear surrounding the central Sox9+ cell of the hair germ (Fig. 5.13c, arrow). Additionally, the cells forming the dermal papilla appear to have strong actin bundles in control hair germs (Fig. 5.13c). In Tg animals,
however, such strong actin expression was lost in cells within the hair germ, and the cells of the dermal papilla also had reduced actin signal (Fig. 5.13c). Thus, the boundary between Lhx2+ and Sox9+ cells in inverted hair germs is disrupted, and the typical actin cytoskeletal signatures appear to be compromised.

iii. Hair follicle arrest

![Figure 5.14](image)

**Figure 5.14 Inverted hair germs arrest during development.**

A) Immunofluorescence of Lef-1 (green) and β4-integrin (red) in control and Tg backskin at P4.5. B) Alkaline phosphatase (blue) and eosin (pink) in control and Tg backskin at P4.5. C) Whole-mount of control and Tg backskin at P8. Scale bars are 100 um.

By P4.5, staining for Lef-1 and alkaline phosphatase indicated that inverted hair germs appear to be arrested, suggesting that such hair germs do not lead to productive hair follicles (Fig. 5.14a). By P8, a whole mount of back-skin from Tg and control animals revealed a marked decrease in the ability of Tg hair follicles to fully grow down to the bottom of the dermis, a hallmark of maturity at this stage (Fig. 5.14b). These phenotypes
may be a result of disrupted signaling observed at earlier stages of development. In addition, they are consistent with the inhibition of cell growth and migration observed with miR-200 over-expression (Fig. 5.3, 5.4, 5.11), which would lead to a failure in hair follicle downward growth.

iv. Disrupted hair follicle orientation

While doing whole mounts to assess defects in hair follicle development, I noticed that Tg back-skin appeared to have disruptions in hair follicle orientation. Typically, hair follicles are aligned whereby the hair bulb is oriented toward the anterior of the animal, and the shaft is oriented toward the posterior. With this orientation, hair emerging from the skin lays flat against the animal. Thus, when viewed from the dermal surface, hair follicles appear to have a 90° orientation relative to the anterior of the animal.

As expected, hair follicles in control animals displayed largely typical orientation (Fig. 5.15a-b). Hair follicles in Tg animals, however, had disrupted orientation (Fig. 5.15a,c). Tg hair germs were often not aligned with their neighbors, and generally showed a lack of coordination. Additionally, from the dermal surface, many Tg hair follicles appeared to have no distinct polarity at all. H&E staining revealed that these hair follicles were angled perpendicularly to the surface of the skin, rather than the typical 50-70 degrees observed in controls (Fig. 5.15b-d, “0 degrees” in radial plot) (Devenport & Fuchs 2008). Thus, miR-200 over-expression disrupts hair follicle orientation relative to the basement membrane.
**Figure 5.15** miR-200 over-expression disrupts hair follicle orientation.

A) Whole mount images of control and Tg backskin (viewed from dermal side) at P0.5 and P5.5. B) Radial plot of hair follicle orientation in control backskin; n=3. C) Radial plot of hair follicle orientation in Tg backskin; n=3. D) H&E staining in control and Tg backskin at P0.5. Arrows indicate typical hair follicle angle in control and atypical 90° angle in Tg. Scale bars are 200 um.

Interestingly, the planar cell polarity gene Celsr1 was identified within the list of high-confidence miR-200 family targets. This gene has previously been associated with hair follicle orientation—complete ablation results in hair “whorls” and severe disruptions in polarity (Devenport & Fuchs 2008; Ravni et al. 2009). Other genes involved with establishing cell polarity were also identified as miR-200 family targets; these genes and their functions will be further discussed in Chapter 6.

**f. Chapter 5 Summary**

In order to better understand miR-200 family function in the skin, as well as to better explore the relationship between this miRNA family and its targets (see Chapter 4), I utilized mouse models with robust over-expression of miR-200b, -200a, and -429. Both straight (K14-Tg) and inducible (Tg) transgenic animals became smaller than littermate controls over time, and succumbed to neonatal lethality. In both models epidermal differentiation, barrier formation, and hair follicle development were intact, although Tg animals appeared to have fewer hair follicles overall.

To guide the analysis of phenotypes associated with miR-200 over-expression, I turned to the targets identified in Chapter 4 by combining CLEAR-CLIP and RNA-seq. Among the strongest signatures in these targets were cell cycle progression and cell migration. In line with such targeting, miR-200 over-expression *in vitro* reduced cell
proliferation and impaired migration. Although miR-200s had previously been shown to repress both of these processes in cancer cells (see Chapter 1), this is the first record of miR-200-mediated repression of proliferation and migration in keratinocytes. In addition, our approach to target identification allowed us to relate many more miR-200 family targets to these processes.

Within miR-200 targets involved with cell migration, many had functions in focal adhesion and adherens junction formation. miR-200 over-expression disrupted focal adhesion formation \textit{in vitro}, and impaired adherens junction formation both \textit{in vitro} and \textit{in vivo}. In addition, miR-200 over-expression repressed Notch signaling, an important signaling pathway for differentiation. Although initial epidermal differentiation and barrier formation were intact early in development, the skin of Tg animals began to crack and show problems with integrity at later stages. Thus, miR-200s coordinate repression of genes involved with cell adhesion, junction formation, and Notch signaling to orchestrate the formation and maintenance of proper cell-cell junctions.

The most novel set of targets I identified for the miR-200 family were those involved with Hippo signaling, a signaling pathway that has recently been implicated in hair follicle development (H. Zhang et al. 2011). In particular, many miR-200 family targets are involved in the regulation of Yap localization, a process largely controlled by its phosphorylation state and association with cell-cell junctions. miR-200 over-expression induced nuclear Yap \textit{in vivo} and in tissue culture, and nuclear Yap in Tg backskin correlated with disrupted cell-cell junctions. Mechanistically, miR-200 over-expression modestly reduced Yap phosphorylation \textit{in vivo}. Thus, nuclear Yap with miR-200 over-
expression is likely caused by a combination of reduced phosphorylation and reduced integrity of cell-cell junctions. Although nuclear Yap has typically been correlated with hyperproliferation and protection against apoptosis, miR-200 over-expression caused a reduction in EdU incorporation and a dramatic increase in the number of apoptotic cells. In addition, direct Yap transcriptional targets also appear to be targeted by miR-200s and are in fact down-regulated in Tg backskin. Thus, miR-200s regulate targets both up- and down-stream of Yap, reflecting unique regulation of the Hippo signaling pathway.

Having determined that miR-200s regulate a wide variety of pathways involved in hair follicle development, I wanted to determine whether hair follicle development itself was affected by miR-200 over-expression. Strikingly, miR-200 over-expression strongly disrupted hair follicle development, leading to inverted hair germs early in development that eventually arrested and failed to reach maturity. The boundary between cell fates appeared to be disrupted in these inverted hair germs, as leading edge markers were found in the suprabasal layers, and suprabasal markers had extended to the leading edge. I also observed disruptions in the actin cytoskeleton of hair germs cells with miR-200 over-expression. Finally, miR-200 over-expression severely disrupted hair follicle orientation, leading to hair follicles being misaligned from their neighbors, and with unusual polarity.

Overall, these results show widespread and diverse roles for the miR-200 family in controlling processes and signaling pathways in the skin. While miR-200 family over-expression caused strong and easily-discernible phenotypes, such an approach does not resolve whether the entire miR-200 family is required for regulation, or how much
redundancy there is in miR-200 function. To answer these questions, a loss-of-function approach must be taken, and is described in Chapter 6.
CHAPTER SIX

PHENOTYPES ASSOCIATED WITH COMPLETE ABLATION OF THE MIR-200 FAMILY

Over-expression studies described in Chapter 5 were valuable in establishing miR-200s' ability to repress various cellular processes, modulate signaling pathways, and affect hair follicle development. However, the fact that mammalian genomes contain multiple copies of both miR-200 family seeds indicates that this family likely has functions which can only be revealed through combined ablation of all five of its members. Supporting this idea, nearly half of the high-confidence miR-200 family targets identified with CLEAR-CLIP and RNA-seq were found to be interacting with at least two miR-200 family members. To date, only one true loss-of-function study on entire the miR-200 family has been performed in mice (Belgardt et al. 2015), likely due to the complication of knocking out the two separate genomic loci from which miR-200s are expressed. In order to study miR-200 loss-of-function, I combined a global miR-200b cluster knockout with conditional ablation of the miR-200c cluster (see Chapter 3 for details). This allowed for the generation of mice with either global (gdKO) or skin-specific (dKO) loss of the entire miR-200 family.

This chapter details the effects of loss of miR-200s\textsuperscript{10} on similar cellular and developmental processes as those that are described in Chapter 5. Guided primarily by the targets identified with CLEAR-CLIP and RNA-seq, as well as the results of the

\textsuperscript{10} Because I was primarily interested in determining miR-200 functions within the epidermal populations of the skin, the majority of analyses were performed using skin-specific dKO animals, to avoid any confounding results from possible cell non-autonomous interactions. Experiments performed with gdKO animals are noted as such.
experiments performed in Chapter 5, I examined the effects of loss of the entire miR-200 family on proliferation, migration, and cell adhesion. I also assessed the effect of loss of these miRNAs on the process of hair follicle development. Together, the results presented in this chapter confirm miR-200 regulation of proliferation, migration and cell adhesion, also reveal a new layer of complexity to miR-200 family regulation, as loss of all five miR-200s increases variability during hair follicle development.

a. Global or skin-specific loss of the entire miR-200 family leads to neonatal lethality

i. Fertility of single knock-outs

Single knock-outs for either miR-200b/a/429 or miR-200c/141 (global or skin-specific) were indistinguishable from control littermates, and lacked any discernible skin defects. However, single knock-out animals for miR-200b/a/429 never successfully bred, supporting previous work showing that miR-200b and miR-429 are required for female fertility (Hasuwa et al. 2013). Interestingly, I also observed that single global knock-outs for the miR-200c cluster (miR-200c/141fl/fl/E2acre) did not breed, indicating that miR-200c and miR-141 may also be required for fertility.
ii. Neonatal lethality

Figure 6.1 Loss of miR-200s causes neonatal lethality.

A) miR-200 dKO (skin-specific) animals at P0.5 and B) P4.5. C) Survival curve for miR-200 gdKO (global); no pups survived past weaning. n=5 pairs of animals.

miR-200 dKO animals were born at the expected Mendelian ratios, indicating that loss of miR-200s does not cause embryonic lethality. At birth, both gdKO and dKO animals were indistinguishable from littermate controls (Fig 6.1a). Within several days, however, animals with global or skin-specific ablation of miR-200s became significantly smaller than littermates, and died neonatally (Fig. 6.1b-c). Lethality occurred earlier in gdKOs, indicating that global loss of miR-200s leads to more severe defects than skin-specific loss. As I saw with both K14-Tg and Tg animals, gdKO and dKO pups also often lacked a visible milk-spot, which is indicative of healthy feeding in neonates. Thus, although I did not specifically examine the cause of neonatal lethality, it may have been caused by nutritional deficits due to either competition with control pups or a primary defect due to
loss of miR-200s. Again, as with K14-Tg and Tg animals, I focused my phenotypic analysis on animals at P0.5 to avoid interference from possible secondary effects of malnutrition.

iii. Hair follicle staging

Figure 6.2 Epidermal differentiation and hair follicle staging in dKO animals.

11 The miR-200 family is also highly expressed in various epithelial tissues other than the epidermis, including the stomach, the intestine, and others (see Appendix 1 for miR-200 family expression in tissues other than the epidermis). Global loss would affect miR-200 family expression in these tissues, and the K14cre used to generate the skin-specific knockout is also expressed in various other epithelial lineages. Thus, the observed neonatal lethality could be due to possible disruptions in many tissues.
A) Immunofluorescence of loricrin (red) and K5 (green) in control and dKO backskin at P0.5. 
B) H&E of control and dKO backskin at P0.5. C) Percentage of hair follicles in each stage of development in control and dKO animals at P0.5; n=3. D) Number of hair follicles per section in backskin of control and dKO animals at P0.5; n=3. Scale bars are 100 um. Error bars represent standard deviation.

With ablation of the entire miR-200 family, epidermal stratification and differentiation remained intact (Fig. 6.2a-b). As with Tg animals, intact epidermal differentiation along with the observed neonatal, rather than perinatal, lethality indicates that the epidermal barrier formed properly in animals lacking miR-200 expression. Although I observed no defects in the progression of hair follicle stages with miR-200 over-expression, I wanted to determine whether loss of the family had any effect on that process. I observed no difference in the percentage of hair follicles at each developmental stage in dKO and control animals (Fig. 6.2c). In contrast with miR-200 over-expression, however, I also did not observe any difference in the number of hair follicles in dKO and control backskin (Fig. 6.2d). Thus, loss of miR-200s does not affect hair follicle developmental progression or specification.

b. Loss of the miR-200 family does not cause EMT in keratinocytes

   i. miR-200 dKO keratinocytes display typical cell morphology

The most well-characterized role for the miR-200 family is in promoting epithelial identity and preventing epithelial-to-mesenchymal transition (EMT). Indeed, it has been well-established that the loss of miR-200s is sufficient to induce EMT in some cell lines, including various cancer cells and MDCK cells\(^{12}\) (Gregory et al. 2008; Park et al. 2008).

\(^{12}\) Madin-Darby Canine Kidney epithelial cells, derived from a Cocker Spaniel. As with many formally immortalized cultured cell lines, this cell line is hyperdiploid, and its properties in tissue culture may not reflect those of its tissue of origin.
However, whether loss of miR-200s is sufficient to induce such a strong cell-fate switch in keratinocytes has not been addressed.

In culture, keratinocytes containing a TuD RNA (see Chapter 3) targeting both miR-200 seeds did not display any atypical morphology compared with controls (Fig. 6.3a). dKO keratinocytes also did not display any atypical morphology compared with cells expressing only K14cre (Fig. 6.3b). Thus, loss of miR-200s in culture does not induce mesenchymal morphology.

![Figure 6.3](image_url)

**Figure 6.3** Loss of all five miR-200s does not cause EMT in keratinocytes.

A) Brightfield images of cultured keratinocytes transfected with TuD-ctrl or TuD-200a/b. B) Brightfield images of K14cre or miR-200 dKO cultured keratinocytes. C) Immunofluorescence of E-cadherin (green) and β-catenin (red) in control and dKO backskin at P0.5. Scale bars are 100um.
ii. Normal E-cadherin expression in vivo

Loss of miR-200s has been shown to promote EMT through de-repression of ZEB1/2, which are transcriptional repressors of E-cadherin (Gregory et al. 2008; Park et al. 2008; Korpal et al. 2008). Thus, in other cell types assessed, miR-200 family down-regulation increases ZEB1/2 and decreases E-cadherin expression. I previously mentioned that CLEAR-CLIP revealed no Zeb1 or Zeb2 interactions with miR-200s, even in dKO keratinocytes (see Chapter 4). This suggests that the well-established miR-200 relationship to ZEB1/2 and E-cadherin is not relevant in keratinocytes or the skin. To confirm that loss of miR-200s does not result in a downregulation of E-cadherin, I examined the expression of that protein in the backskins of dKO and control animals at P0.5. This analysis revealed no decreased E-cadherin expression in dKO animals (Fig. 6.3c). dKO hair germs displayed normal down-regulation of E-cadherin at the leading edge, but that down-regulation did not extend into the suprabasal layers (Fig. 6.3c).

iii. No up-regulation of mesenchymal genes with loss of miR-200s

Many of the studies assessing EMT rely upon increased expression of mesenchymal genes, such as Zeb1/2 and Vim, as a read-out of mesenchymal fate. As part of another project to study the miR-200 family in the Yi lab, Dr. Glen Bjerke performed mRNA-seq on both control and dKO keratinocytes. He observed very little expression of Zeb1 or Zeb2 (0-3 normalized reads per library) in either control or dKO keratinocytes, and saw no evidence of Zeb1/2 upregulation or E-cadherin down-regulation with loss of the miR-200 family. In addition, as mentioned in Chapter 4, no interaction between any miR-200 family member and either Zeb1 or Zeb2 was observed with CLEAR-CLIP. Together with the lack of
changes in morphology or E-cadherin expression, these results show that loss of miR-200s in normal keratinocytes, either *in vitro* or *in vivo*, causes no discernible EMT.

**c. Loss of the miR-200 family enhances proliferation and migration**

i. Enhanced proliferation

As mentioned in Chapter 5, many miR-200 targets are involved with cell cycle regulation (Chapter 5 Fig. 5.3). Given the strong alterations in proliferation I observed with miR-200 over-expression (Chapter 5 Fig. 5.3 & 5.10) I wanted to examine how loss of miR-200s affected proliferation. In particular, because many of the identified targets were found to be interacting with multiple miR-200 family members, I wanted to determine the effect of combined loss of all five miR-200s. Loss of all five miR-200s should reveal whether coordinated up-regulation of miR-200 targets involved with cell cycle regulation is sufficient to promote proliferation. To address this question, I first performed a colony formation assay using keratinocytes containing a TuD RNA targeting both miR-200 seeds. In this assay, miR-200 family down-regulation enhanced colony formation ability (Fig. 6.4a). I then performed the same assay with dKO and control keratinocytes (Fig. 6.4b). This also showed increased colony forming ability; whereas the number of colonies that formed was less, the size of those colonies was larger, reflecting enhanced proliferative abilities (Fig. 6.4c-d).

Next, to determine whether loss of miR-200s is sufficient to increase the proliferative ability of cells in a tissue, I performed an EdU incorporation assay *in vivo*. Quantification of either the K5+ population or of cells specifically within the hair germ did not reveal any discernible differences in proliferation (Fig. 6.4e-g). Thus, loss of miR-200s in culture
enhances cell proliferation, but miR-200s are not required to suppress cell cycle progression at this early stage in vivo when cells are already highly proliferative.

**Figure 6.4 Loss of miR-200s enhances proliferation in vitro but not in vivo**

A) Colony formation assay of keratinocytes transfected with TuD-ctrl or TuD-200a/b. B) Colony formation assay of ctrl and dKO keratinocytes. C) number and D) size of colonies formed by ctrl and dKO keratinocytes; n=3. E) Immunofluorescence of EdU (green) and K5 (red) in control and dKO backskin at Po.5. F) Percentage of EdU+/K5+ cells in control and dKO epidermis at Po.5; n=3. G) Percentage of EdU+ cells within the hair germ of control and dKO animals at Po.5; n=4. Error bars represent standard deviation. *P<0.05. Scale bars are 100 um.

**ii. Enhanced migration**

Within the list of miR-200 family targets were many genes involved with cell migration, and miR-200 over-expression dramatically impaired that process in vitro (Chapter 5 Fig. 5.4). Again, many of the targets identified as involved in this process were found to be interacting with multiple miR-200 family members, suggesting coordinated regulation of cell migration by this family. To examine the effect of combined loss of all miR-200s on cell migration, I performed a scratch assay in control and dKO keratinocytes. As expected, complete loss of all five miR-200s resulted in a higher number of cells migrating into the
scratch, suggesting that loss off all five miR-200s enhances the ability of keratinocytes to migrate (Fig. 6.5a-b). I also observed a decreased ability of miR-200 dKO keratinocytes to adhere in a cell adhesion assay, suggesting that the increased migratory ability of these cells may be due in part to decreased cell adhesion (Fig. 6.5c).

**Figure 6.5 Loss of miR-200s enhances cell migration and reduces cell adhesion.**

A) Scratch assays on ctrl or dKO keratinocytes 0 hours or 8 hours after scratching. B) Number of migrated cells from ctrl and dKO scratch assays; n=3. C) Crystal violet absorbance of ctrl and dKO keratinocytes plated in cell adhesion assay; n=3. Error bars represent standard deviation. **P<0.01; ***P<0.001.

d. *Loss of the miR-200 family reduces cell adhesion*

   i. Disrupted focal adhesion formation

As is noted in Chapter 5, miR-200 family over-expression disrupted normal focal adhesion development. Again, many of the high-confidence miR-200 family targets identified as being involved with focal adhesions and actin cytoskeleton appear to be regulated by more than one member of the miR-200 family, suggesting coordinated regulation (Chapter 5 Fig. 5.5). Thus, I wanted to determine the effect of complete loss of the miR-200 family on focal adhesion, and to that end assessed focal adhesion formation in dKO and control keratinocytes. Whereas miR-200 over-expression caused the formation of
more, smaller focal adhesions in cultured keratinocytes (Chapter 5 Fig. 5.5), I observed fewer but much larger focal adhesions with miR-200 family ablation (Fig. 6.6a-b). This analysis also showed increased actin bundles localizing toward focal adhesions with miR-200 loss-of-function compared with controls. Furthermore, the actin bundles seen in dKO keratinocytes were often radially organized, in contrast with the circumferentially organized actin usually seen in control cells (Fig. 6.6a-b). Thus, loss of miR-200s disrupts focal adhesion formation and actin organization in cultured keratinocytes.

**Figure 6.6 Loss of miR-200s disrupts focal adhesion formation**

A) Immunofluorescence of vinculin (red) and phalloidin stain for actin (green). Arrows indicate examples of focal adhesions. B) Number of focal adhesions per cell; n=43 ctrl cells; n=40 dKO cells. Scale bars are 20 um.
ii. Disrupted adherens junction formation

A

3 hours calcium

Phalloidin  E-cadherin  Hoescht  Merge

Ctrl

dKO

7 hours calcium

Ctrl

dKO

B

3 hr Calcium

Phalloidin  E-cadherin  Merge

Ctrl

dKO

7 hr Calcium

Ctrl

dKO
**Figure 6.7 Loss of miR-200s disrupts cell junction formation**

A) Immunofluorescence of E-cadherin (red) and Phalloidin stain for actin (green) on control and dKO keratinocytes induced with calcium for 3 or 7 hours. B) Closer images of cell-cell junctions in A. Arrows indicate dispersed E-cadherin localization in dKO cells; representative images from n>20 cell junctions imaged for each treatment. Scale bars are 20 um.

Because miR-200 gain-of-function caused severe defects in cell adhesion (Chapter 5 Fig. 5.5, 5.6, 5.7 & 5.10), I wanted to determine whether loss of this family also affected cell-cell junction formation. To that end, I induced dKO and control cells with calcium for three or seven hours, as in Chapter 5 section c, to see the effect of loss of all five miR-200s on the cytoskeletal remodeling required for cell junction formation. Whereas miR-200 over-expression caused a reduction in actin bundle attachment to E-cadherin aggregates at the cell cortex (Chapter 5 Fig. 5.6), miR-200 loss-of-function seemed to reduce E-cadherin relocation to the cell cortex altogether (Fig. 6.7b, bottom arrows). This was accompanied by disorganized actin stress fibers in dKO keratinocytes. There were a few examples where E-cadherin localized properly to the cell cortex, and in those cases actin attachment was detected. However, up to seven hours after calcium addition, cell-cell contacts between dKO cells had failed to properly seal, at which point they were fully sealed in control cells (Fig. 6.7a). Thus, complete ablation of the miR-200 family disrupts cell junction formation *in vitro*.

**iii. No effect on cell adhesion *in vivo***

miR-200 over-expression impaired adherens junction formation both *in vitro* and *in vivo*, ultimately leading to scaly and breaking skin in Tg animals (Chapter 5 Fig. 5.5 & 5.6). However, miR-200 loss-of-function caused no discernible differences in adherens junctions *in vivo* (Fig. 6.3c). Additionally, I did not observe any defects in the surface of
dKO skin (data not shown). Therefore, although complete loss of miR-200s causes severe disruptions in cell-cell junction formation *in vitro*, the loss of this miRNA family is not sufficient to disrupt these processes *in vivo*.

**iv. No change in localization of Yap**

Analysis with CLEAR-CLIP and RNA-seq revealed many novel miR-200 family targets involved with regulating the Hippo signaling pathway, particularly genes involved with regulating Yap nuclear/cytoplasmic localization (Chapter 5 Fig. 5.9). Supporting a role for miR-200s in regulating Yap localization, miR-200 over-expression caused widespread nuclear retention of Yap in the interfollicular epidermis *in vivo*, as well as increased nuclear Yap signal in cultured keratinocytes (Chapter 5 Fig. 5.9). Therefore, I wanted to determine whether loss of miR-200s had any effect on the localization of this transcription factor. Staining both *in vivo* and in cultured keratinocytes revealed no difference in Yap nuclear localization in dKOs and controls (Fig. 6.8a-b). A lack of discernible regulation by loss of miR-200s may not be surprising, given that miR-200 over-expression appeared to induce nuclear retention rather than translocation of Yap, at least *in vivo* (Chapter 5 section d). In addition, it appears that at least some of the nuclear localization of Yap with miR-200 over-expression was due to disruptions in adherens junctions *in vivo*, a phenotype I did not observe with miR-200 loss-of-function (Chapter 5 Fig. 5.10). Thus, whereas miR-200 over-expression strongly influences Yap localization, loss of the miR-200 family is not sufficient to cause any change.
Figure 6.8 miR-200 ablation does not affect Yap localization

A) Immunofluorescence of Yap1 (red) and K5 (green) in control and dKO backskin at P0.5. Scale bar is 50 um. B) Immunofluorescence of Yap1 (red) in control and dKO cultured keratinocytes. Scale bar is 100 um.

e. Loss of the miR-200 family increases variability during cell-fate specification

i. Disrupted actin cytoskeleton

With miR-200 over-expression, I observed strong defects in actin cytoskeleton, proliferation, and reduced stringency in the molecular boundary between distinct progenitor cells in the hair germ (see Chapter 5). Most intriguingly, the targets involved in all of these processes were often found to be interacting with two or more miR-200
family members. Thus, complete ablation of the miR-200 family may be sufficient to disrupt these processes \textit{in vivo}.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure6.9.png}
\caption{Complete loss of miR-200s disrupts the actin cytoskeleton in hair germs.}
\end{figure}

\textbf{A)} miR-200 family targets identified with CLEAR-CLIP and RNA-seq involved with actin cytoskeleton. \textbf{B)} Immunofluorescence of \(\alpha\)-catenin (red) with Phalloidin stain for actin (green) on ctrl and dKO hair germs at P0.5; arrow indicates actin bundle. \textbf{C)} Percentage of hair germs from ctrl and dKO animals at P0.5 with centralized actin or lack of clear signal; \(n=6\). Scale bars are 20 um. Error bars represent standard deviation. **P\(<0.01.\)

To further understand the requirement for the miR-200 family during development, as well as to probe the function of such coordinated targeting \textit{in vivo}, I examined hair germ morphogenesis in miR-200 dKO skin. Many miR-200 targets are regulators of actin cytoskeleton and are coordinately regulated by multiple members of the family (Fig. 6.9a). In addition, actin was disrupted in both Tg and dKO keratinocytes; therefore, I first
visualized actin cytoskeleton in the hair germ. During hair follicle development, asymmetric divisions lead to a partitioning of different cell fates between the basal cells at the leading edge, which express Lef1 and Lhx2, and the suprabasal cells, which express Sox9 (see Chapter 1). Consistent with this partitioning, and as was also noted in Chapter 5, phalloidin staining revealed that the suprabasal, Sox9+ cells have distinctly strong actin signals (Fig. 6.9b). Statistical analysis showed that nearly 90% of the hair germs in control skin had the characteristic signals for strong actin bundles interacting with the centrally-located suprabasal hair germ cell (Fig. 6.9b, arrow). In contrast, such signals were largely absent from dKO skin, and only 30% of dKO hair germs were found to harbor such cells, with weaker actin signals (Fig. 6.9c). Thus, loss of miR-200s causes a reduction in actin signals within the developing hair germ.

ii. Disrupted cell polarity

The disruptions I observed with actin signals, combined with a number of miR-200 family targets involved with regulating cell polarity, such as Pard6b (Par6) and Celsr1, prompted me to examine the polarity of the basal and suprabasal cells of the hair germ (Fig. 6.10a). To that end, I stained for pericentrin, a component of the centrosome which can be used to determine cell polarity (Delaval & Doxsey 2010).
**Figure 6.10** Loss of miR-200s disrupts cell polarity.

A) miR-200 family targets identified with CLEAR-CLIP and RNA-seq involved with cell polarity. B) Immunofluorescence of Pericentrin (red) and β4-integrin (green) on ctrl and dKO hair germs at P0.5; arrow indicates mis-angled cell in dKO. C) and D) Angle of hair germ cells relative to the basement membrane from ctrl and dKO animals at P0.5; n=4 pairs of animals, n=16 ctrl cells, n=30 dKO cells. Scale bars are 50 um. Error bars represent standard deviation. **P<0.01.

As expected, basal cells in control hair germs occupied the basement membrane in a largely perpendicular orientation (Fig. 6.10b-d). This is consistent with the recently reported asymmetric cell division pattern of basal hair germ cells (Ouspenskaia et al. 2016b). In contrast, basal cells within dKO hair germs lost uniformly oriented cell polarity, and nearly half of dKO basal cells were more than 15 degrees off of the perpendicular axis (Fig. 6.10b-d). Therefore, loss of miR-200s in developing hair follicles disrupts cell polarity.
iii. Increased variability during cell fate specification

Because asymmetric cell division and perpendicular cell polarity are integral to cell fate specification with the hair germ (Ouspenskaia et al. 2016b), I next examined whether cell fate is altered in dKO hair germs. I first examined the expression of Lef1, a binding partner of β-catenin known to mark hair germ cells. Rather than remaining restricted only to the leading edge as in control placodes and hair germs, Lef1 was often detected throughout the entire hair germ in dKOs (Fig. 6.11a). Quantification revealed that such atypical hair germs do occur in controls, but they are a minor number compared with those that have Lef1 expression exclusively at the leading edge (Fig. 6.11b). Cell morphology was also mildly changed in dKO hair germs, from columnar to a more rounded shape, perhaps reflecting compromised cell polarity (Fig. 6.11a-b). In dKO animals, there was an increase in atypical hair germs at the expense of those developing normally.
**Figure 6.11** Loss of miR-200s increases variability in Lef1 specification

A) Immunofluorescence of Lef-1 (red) and β4-integrin (green) on ctrl and dKO hair germs at P0.5; arrows indicate supra-basal Lef-1+ cells in dKO. B) Percentage of hair germs from ctrl and dKO animals at P0.5 with typical appearance, extended Lef-1, or rounded cells at the leading edge (hair germs could have both rounded cells and extended Lef-1); n=6. Scale bars are 50 um. Error bars represent standard deviation. *P<0.05; **P<0.01.

In contrast, Sox9 expression was expanded from the suprabasal cells to the basal cells at the leading edge of dKO hair germs (Fig. 6.12a). Again, this abnormal expression pattern was occasionally observed in control hair germs; however, it was much more prevalent in dKO hair germs (Fig. 6.12b). Because of this extended Sox9 expression pattern, and because I did observe a single miR-141:Sox9 CLEAR-CLIP read in the data generated from cultured keratinocytes (Fig. 6.12c), I wanted to know whether this transcription factor may be a direct target of the miR-200 family. However, this single read did not pass the threshold we established for identifying *bona fide* miR-200 targets, potentially because Sox9 is lowly expressed in cultured keratinocytes (Adam et al. 2015). In addition, the 3’ UTR of Sox9 was not repressed by the miR-200s in a luciferase assay (Fig. 6.12d). Thus, Sox9 is not a direct target of the miR-200s, and its expanded expression likely reflects the altered actin cytoskeleton and cell polarity observed in dKO hair germs. Taken together, these results show that loss of miR-200s in the skin leads to increased variation in cell fate specification during hair follicle development, correlating with compromised actin cytoskeleton and cell polarization defects in developing hair germs.
Figure 6.12 miR-200 ablation increases variability in Sox9 specification

A) Immunofluorescence of Sox9 (red) and Keratin-5 (green) on ctrl and dKO hair germs at P0.5; arrows indicate leading edge Sox9+ cells in dKO. B) Percentage of hair germs from ctrl and dKO animals at P0.5 with typical Sox9 expression pattern and Sox9 at the leading edge; n=6. C) Single CLEAR-CLIP read for miR-141 ligated to the 3’ UTR of Sox9. D) Relative luciferase activity of Sox9 3’ UTR with the addition of the miR-200b cluster, the miR-200c cluster, or both; n=6. Scale bars are 50 um. **P<0.01.

f. Chapter Six Summary

Efforts to understand the effects of endogenous miR-200 loss-of-function have been complicated by the two separate genomic clusters in which these miRNAs are located. I
used mouse models with either global or skin-specific loss of the entire miR-200 family to examine the functions of these miRNAs during development and in the skin. Loss of either cluster alone caused no overt developmental phenotypes, although single cluster knockouts were unable to breed, suggesting that individual miR-200 family members play essential roles in regulating fertility. Surprisingly, global or skin-specific loss of the miR-200 family led to unexplained neonatal lethality, indicating that the redundant functions of this family are essential for development.

The most well established role for the miR-200 family is in promoting epithelial identity, and preventing EMT. miR-200 loss in keratinocytes, however, led to no discernible up-regulation of mesenchymal characteristics or genes. Thus, the miR-200 family is not required to maintain the epithelial identity of keratinocytes.

In vitro, miR-200 loss-of-function enhanced proliferation and cell migration, but reduced cell adhesion. Loss of miR-200s also disrupted the formation of focal adhesions, causing fewer but much larger focal adhesions to form in dKO cells compared with controls. In addition, loss of miR-200s disrupted cell-cell junction formation in vitro, but had no discernible effect on the same process in vivo.

Finally, miR-200 loss of function in developing skin disrupted the actin cytoskeleton and impaired cell polarity. Intriguingly, miR-200 loss-of-function also increased variability of cell fate specification during hair follicle morphogenesis, causing an expansion of Lef-1+ cells beyond the leading edge, and conversely causing Sox9 to be specified within leading edge cells. This was not due to direct targeting of either of those
transcription factors involved with cell fate specification, but was more likely a consequence of disrupted actin cytoskeleton and cell polarity.
CHAPTER SEVEN

DISCUSSION AND LOOKING TO THE FUTURE

The experiments and results presented in this dissertation provide a broad examination of miR-200 family targets and functions within the skin and hair follicle. In addition to providing the most comprehensive list of targets yet identified for the miR-200s, the results shown here begin to dissect the complex regulatory networks controlled by this unique family during skin and hair follicle development. In this chapter, I discuss and contextualize those results and briefly summarize remaining questions regarding miR-200 family regulation in the skin.

a. miR-200 family targeting paradigms

Over the past twenty-five years, our understanding of miRNA functions and mechanisms of action has greatly improved. Despite this, there are many remaining questions, partially due to the limitations of previous work. Because traditional methods for identifying miRNA targets have primarily relied upon seed sequence, a long-standing question in the field of miRNA biology has been the interplay between miRNAs with the same or similar seed sequences. The miR-200s present a unique opportunity to examine this question, as there are two nearly-identical seed sequences present within the five-member family, and each seed is found in at least two family members (Chapter 1). Novel techniques, such as CLEAR-CLIP, allow for identification of direct miRNA:mRNA interactions, and are crucial to answering such questions and dissecting miRNA family regulatory networks.
Here, I present multiple lines of evidence that miR-200 family members are able to form unique interactions with some mRNAs, but also that they are somewhat promiscuous in their targeting. First, over half of mRNAs found to be interacting with miR-200s in CLEAR-CLIP were interacting with only one family member, suggesting a certain level of specificity (Chapter 4). These analyses also revealed that the presence of a miR-200b seed, for example, does not guarantee interaction with all miR-200s with that seed (Chapter 4). In addition, a heterologous luciferase assay showed that some targets are repressed exclusively by one cluster, but not the other. This occurs despite the fact that each miR-200 cluster contains both seed sequences, suggesting that a region beyond the seed may provide some specificity to targeting (Chapter 4). On the other hand, nearly half of the targets identified with CLEAR-CLIP were found to be interacting with two or more miR-200 family members, indicating that this family often does bind the same mRNA targets (Chapter 4). Furthermore, a TuD RNA designed with perfect binding sites for miR-200a and miR-200b repressed expression of the entire family, providing additional support for the idea that miR-200s are able to share targets (Chapter 3).

Previously, the question of whether miR-200s with distinct seeds are able to interact with the same targets remained largely unaddressed. Data generated with CLEAR-CLIP also shows that cross-seed binding does occur between miR-200 family members. However, the endogenous occurrence of these interactions is infrequent, and the seed region is an important factor in determining targets (Chapter 4).

Importantly, cooperation appears to be paramount for miR-200 family targeting. Within miR-200 family targets identified with CLEAR-CLIP, up to 22% appear to be
interacting with four or even all five miR-200 family members (Chapter 4). The heterologous luciferase assay also showed that, in at least some cases, such targets are more strongly repressed in the presence of all five miR-200s (Chapter 4). Interestingly, several of the targets that appear to be most strongly regulated by this family have been previously identified and confirmed by other groups. This is perhaps because they show the strongest regulation by the miR-200s in comparison with other targets, and are thus easier to predict using traditional methods. Further exploration of these highly-targeted genes may reveal more about the importance of redundancy for miRNA families. Finally, targets presented throughout this dissertation show that not only are miR-200s cooperating to repress the same targets, they are also collaborating to target multiple genes within the same pathways. This will be discussed further in the following sections.

b. miR-200 expression in the skin

Using RNA-seq and in situ hybridization, I established that the miR-200 family is a highly-expressed miRNA family with the skin, with a unique expression pattern in developing hair follicles. As I showed with in situ hybridization, the miR-200 family becomes restricted to developing hair follicles only during the second and third waves of development. There is evidence that the expression of some regulatory molecules, such as Dkk4, is exclusive to certain waves of development, and that their disruption only affects the development of the other hair types (Fliniaux et al. 2008; Cui et al. 2010). Therefore, the fact that miR-200s are expressed throughout the developing skin during the first wave of hair follicle development, but become restricted to placodes during the second and third waves (Chapter 3), suggests that their function may be more specific to the
development of awl/auchene and zig-zag hairs, but not of guard hairs. With miR-200 over-expression, I observed a decrease in the number of backskin hair follicles compared with controls, but saw no difference in hair follicle staging (Chapter 5), suggesting perhaps that miR-200 over-expression prevents the specification of some hair follicles but does not affect the progression of others. The idea that miR-200s are involved with the morphogenesis of only some hair follicles is further supported by the fact that only ~30% of developing hair follicles are disrupted with miR-200 family over-expression (Chapter 5), and that there are still some undisrupted hair follicles with typical morphology even with ablation of the entire family (Chapter 6).

In addition, the expression of the miR-200 family at the leading edge of the hair follicle, and particularly within cells contacting the dermal papilla (Chapter 3), may suggest a function for these miRNAs in mediating epithelial-mesenchymal interactions during hair follicle development. Although I did not specifically explore such a function, I did observe disruptions in cell fate specification (see below), which could reflect disturbances in epithelial-mesenchymal signaling.

c. miR-200s perform distinct functions in the skin than in cancer cells

Given the lack of evidence for miR-200 interactions with Zeb1/2 in CLEAR-CLIP and RNA-seq, it is not surprising that this family does not appear to regulate epithelial identity in the same way in keratinocytes as it has been shown to do in other cell types. This difference is not due to the difference in knocking out versus knocking down these miRNAs, as keratinocytes with either full ablation of the miR-200 family or with the family down-regulated using a TuD showed no signs of mesenchymal characteristics,
including morphological changes or upregulation of mesenchymal genes (Chapter 6).

Therefore, the miR-200 family does not regulate EMT in cultured keratinocytes, although my results show that they play other critical functions.

*d. Coordinated regulation of pathways and processes by the miR-200 family influences hair follicle development*

Although miRNA-mediated regulation is generally mild for individual targets, my analyses revealed coordinated targeting of many components of focal adhesion, actin cytoskeleton, adherens junctions, and cell cycle by the miR-200 family. Genetic manipulation of miR-200 family levels both *in vitro* and *in vivo* supported functions for this miRNA family in regulating those processes. While the inhibitory effects of miR-200s on proliferation, focal adhesion formation, and migration have been previously shown during tumorigenesis (Korpal et al. 2008; Gregory et al. 2008; Shimono et al. 2009; Bracken et al. 2014), they have not been shown in a normal cell population in which this family is highly expressed.

Intriguingly, the prominent defects I observed in cell junction formation and cell polarity with miR-200 genetic manipulation are novel. Both miR-200 over-expression and loss-of-function disrupted the formation of cell junctions *in vitro*, highlighting a critical role for this family in that process. miR-200 over-expression also caused defects *in vivo*: rather than forming robust adherens junctions and establishing cell-cell adhesions along the lateral and apical membranes, punctate signals for both $\alpha$-catenin and $\beta$-catenin were observed in basal cells (Chapter 5). Additionally, miR-200 over-expression severely disrupted hair follicle orientation, indicating defects with polarity (Chapter 5). Supporting
a role for this family in mediating cell polarity, complete ablation of miR-200s in the skin caused defects in the angling of cells relative to the basement membrane (Chapter 6). These defects are likely caused by mis-regulation of a variety of miR-200 family targets, including regulators of focal adhesion and actin polymerization (Fak, Cdc42, α-actinin, N-WASP and others), tight and adherens junctions (Claudins 1, 12 and 23, Occludin, and others) and cell polarity (Celsr1, Pard6b, and others). Although future work will be required to reveal the exact mechanism mediating miR-200 regulation of cell junctions and polarity, these results clearly show that miR-200s are critical regulators of these processes.

In contrast with the direct regulation of the aforementioned processes, the interaction between the miR-200s and the Hippo signaling pathway is less clear. Although I identified a strong signature for targets that regulate the nuclear/cytoplasmic localization of Yap, and indeed detected increased nuclear Yap accumulation with miR-200 over-expression, I did not observe outcomes consistent with nuclear Yap activation such as downstream target activation and enhanced cell proliferation (Chapter 5). It is likely that direct miR-200 targeting of the cell cycle machinery, regulators of apoptosis, and Ctgf, one of the best characterized Yap1 targets, dampens the effects of nuclear Yap, although why this occurs is not clear. One possible explanation is that phosphorylated, cytoplasmic Yap may have unexplored functions. Notably, E-cadherin and α-catenin, two essential components of AJs, have both been shown to regulate Yap1 functions (Schlegelmilch et al. 2011; Kim et al. 2011). Thus, miR-200-mediated regulation of the
Hippo/Yap1 pathway, especially the role of pYap1 in the cytoplasm, warrants further investigation.

Over-expression of miR-200s severely disrupts hair follicle development, leading to inverted hair germs that evaginate away from the dermis in the opposite direction of controls (Chapter 5). This striking phenotype is likely due to coordinated down-regulation of cell proliferation and migration, and also to disruptions in cell-cell contacts, through miR-200 targets involved in those processes. Although I saw no evidence of Yap-regulated genes being upregulated, forced nuclear expression of Yap caused similar evaginated hair germs (H. Zhang et al. 2011), suggesting that the nuclear retention of Yap seen in inducible hair germs may be connected to the inverted hair germs. However, as mentioned previously, more work must be done to clearly delineate the link between miR-200s and the Hippo signaling pathway.

Whereas the experiments done with miR-200 over-expression clearly establish a role for this miRNA family in regulating various cellular processes and the Hippo signaling pathway, experiments performed with total loss of the entire family highlight the coordinated repression being performed by this miRNA family. Interestingly, the strongest phenotypes I observed with miR-200 family loss-of-function were primarily in cultured keratinocytes, and did not always carry over to the skin itself. For example, complete ablation of the miR-200 family enhanced cell proliferation in vitro, but had no discernible effect on this process in vivo (Chapter 6). This was also true for the disruptions in cell-cell junction formation caused by miR-200 family ablation, which were observed in vitro but not in vivo. These results show that loss of miR-200s alone is not sufficient to
disrupt these important processes within the developing skin, suggesting that they are primarily controlled by other, more robust, regulatory mechanisms. Since the phenotypes observed with loss of miR-200s and many other miRNAs are subtle, one way to determine more functions for these small RNAs would be to combine their ablation with either stress conditions or with ablation of other regulators, as is discussed in the Introduction.

Complete ablation of the miR-200 family was not sufficient to disrupt hair follicle development as severely as over-expression. However, I did observe disruptions in actin localization and in cell orientation. In addition, and most interestingly, I observed increased variability during hair germ development—hair germs in dKO animals were more likely to have Lef-1 extend into the suprabasal layers, and to have Sox9 extend to the leading edge. As both of these transcription factors have been shown to be involved with hair germ development and HFSC specification, these results likely reflect subtle defects in both of those processes. However, the neonatal lethality I observed in animals with skin-specific ablation of the miR-200 family prevented thorough examination of later stages of development, such as whether HFSCs are properly specified and exist in the same numbers in dKO hair follicles compared with controls. Thus, further study is required to better understand the outcomes of such variability during development.

It is especially intriguing that both over-expression and loss-of-function of this family was sufficient to disrupt normal hair follicle development in vivo, indicating that gene dosage control by the miR-200 family is critical for the regulation of this complex developmental process. While miR-200 over-expression did not appear to cause a complete shift in cell fate (Chapter 4), both miR-200 over-expression and loss-of-function
caused imprecise cell fate specification—within hair germs from both of these mouse models, Lhx2, Lef-1 and Sox9 were observed in populations where they are not normally seen (Chapters 5 & 6). Although I was not able to determine the specific mechanism by which miR-200 disruption leads to defects in the localization of these transcription factors, it is possible that the defects I observed in cell junction formation and polarity could be contributing to their mis-expression, as it has been suggested that those processes are involved with cell fate determination (Ouspenskaia et al. 2016a). Regardless of the precise mechanisms, these results clearly show that gene dosage control by the miR-200 family is critical for cell fate specification and proper hair follicle development (Figure 7.1).
**Figure 7.1** Model of miR-200 family function during hair follicle development

miR-200s regulate genes involved with cell cycle, actin cytoskeleton, cell junctions and polarity, and Yap localization to strongly influence cell fate specification and hair follicle development.

e. Unexplored miR-200 targets and functions

By combining CLEAR-CLIP with RNA-seq, I identified over 700 high-confidence miR-200 family targets in the skin—and this is not considering the additional 400 CLEAR-CLIP identified transcripts which would be included if I did not require any down-regulation at the mRNA level (for the full list of targets, see Appendix 4). Within the experiments performed for this dissertation, I only addressed the functions of a small portion of these targets. Thus, there are likely to be many more important functions for the miR-200 family than are accounted for within these pages. For example, my data clearly shows that miR-200s regulate hair follicle development, as their over-expression or loss-of-function disrupts this process. The results of my analysis, combined with many of the targets identified, suggest that such disruption is due to misregulation of a variety of cellular processes. Within the list of high-confidence miR-200 family targets, however, are several genes that I did not specifically explore which are involved with signaling pathways shown to be critical for hair follicle development, including Wnt (Lrp6, Nfat5, Fzd5, Gsk3b, Ctnnb1, and others), Tgf-β (BmprA1, Bmpr2, Acvr2a, Sp1, Fst, Smad2, Tgfb2, and others) and SHH (Sufu, Csnkia1, Csnkig1, Csnkig3). Therefore, miR-200 regulation of these pathways may also contribute to their influence over hair follicle development (Fig. 7.2). Further exploration of the relationship between the miR-200 family and these signaling pathways, as well as the hundreds of other unexplored targets, will be needed to fully define their function.
Figure 7.2 Coordinated targeting by miR-200s regulates hair follicle morphogenesis.

In addition to targets discussed within this dissertation, there are other unexplored targets for the miR-200s, such as those involved with various signaling pathways. Coordinated mis-regulation of all miR-200 family targets likely mediates development of the hair follicle phenotypes observed with miR-200 over-expression or loss-of-function.

f. Remaining questions

The data presented within this dissertation address many questions regarding miR-200 family targeting paradigms. However, further work needs to be done to address how miR-200s with the same seed are able to distinguish between their unique targets. In addition, evidence shown here suggests that the seed is the most important factor determining target interactions. However, nearly 30% of identified targets did not have a seed in or near the sequenced mRNA fragment (Chapter 4). While I did not validate any of these seedless interactions, they are more likely to be downregulated in RNA-seq compared with genes without any miR-200 interaction (Chapter 4). Therefore, more work is needed.
to specifically define non-canonical binding sites and determine the importance of seedless interactions.

In addition, the miR-200 family is highly expressed in a variety of tissues outside of the epidermis (see Appendix 1). The results presented in this dissertation clearly show that the miR-200 family performs a distinct function in the epidermis compared with what they have been shown to do in cancer. Thus, these miRNAs may have context-specific functions that would be revealed by further investigation of this unique family in other tissues.

**g. Conclusion**

As miRNA target prediction methods continue to become more sophisticated, the focus of miRNA studies will shift towards understanding whole regulatory networks influenced by these small RNAs, rather than individual target interactions. This dissertation presents the first example of such a study on a family of miRNAs, the miR-200s, in the skin and hair follicle. The results presented here clearly show that miR-200 family targeting paradigms are complex, and that they form part of a broad regulatory network mediating proliferation, migration, cell adhesion, and signaling within the skin and hair follicle. Importantly, this family of miRNAs cooperates in targeting, often regulating the same transcripts or coordinately repressing multiple genes within a pathway. Ultimately, coordinated regulation of a wide variety of cellular processes and signaling pathways allows this unique family of miRNAs to mediate the complex process of hair follicle morphogenesis.


Chauilk, S.G. et al., 2011. Role of pri-miRNA tertiary structure in miR-17–92 miRNA biogenesis. RNA Biology, 8(6), pp.1105–1114.


Dry, F., 1926. The coat of the mouse (Mus musculus). *J. Genet*.


Gregory, P. a et al., 2008. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. Nature cell biology, 10(5), pp.593–601.


Lee, Y. et al., 2004. MicroRNA genes are transcribed by RNA polymerase II. The EMBO Journal, 23(20), pp.4051–4060.


Morita, S. et al., 2007. One Argonaute family member, Eif2c2 (Ago2), is essential for development and appears not to be involved in DNA methylation. *Genomics*, 89(6), pp.687–696.


Renthal, N.E. et al., 2010. miR-200 family and targets, ZEB1 and ZEB2, modulate uterine quiescence and contractility during pregnancy and labor. *Proceedings of the National Academy of Sciences*, 107(48), pp.20828–20833.


Tao, C. et al., 2016. Adipocyte miR-200b/a/429 ablation in mice leads to high-fat-diet-induced obesity. Oncotarget, 7(42), pp.67796–67807.


Uhlmann, S. et al., 2010. miR-200bc/429 cluster targets PLCgamma1 and differentially regulates proliferation and EGF-driven invasion than miR-200a/141 in breast cancer. Oncogene, 29(30), pp.4297–4306.


Appendixes

Appendix 1: *miR-200 expression in tissues beyond the skin*

Introduction

Previous studies established that miR-200 expression within the skin is high (Yi et al. 2006). However, despite the fact that expression of this family is highly correlated with epithelial identity, there is little data on their expression in most epithelial tissues. In addition, several early studies examining the functions of this miRNA family established important roles for them in non-epithelial tissues, such as brain. Therefore, I wanted to determine whether, and at what relative levels, these miRNAs are expressed in various murine tissues.

Materials & Methods

Establishing miR-200 family expression levels in tissues other than the skin was done with qRT-PCR and *in situ* hybridization. For a detailed description of those methods, see Chapter 2. Tissue from brain, brown fat, epidermis, intestine, bladder, kidney, heart, tongue, liver, lung, spleen and stomach was collected from neonatal animals, RNA was extracted, and cDNA was made according to the methods in Chapter 2. *In situ* hybridization was performed as described in Chapter 2 on wildtype embryos at E18.5.

Results

To determine the levels of the miR-200 family, I performed qRT-PCR for miR-200b on RNA isolated from various tissues from neonatal mice. miR-200s have little to no
expression in brown fat, brain, spleen, and heart (Fig. A1a-b). They are expressed at low levels in the liver. Of all the tissues tested, their expression is highest in the stomach, followed by the intestine, kidney, epidermis, lung, bladder, and tongue (Fig. A1a-b). In \textit{situ} hybridization using a probe against miR-200b revealed high miR-200 expression in prenatal nasal epithelium, tongue and palate, and whisker buds (Fig. A1c-e).

\textbf{Figure A1} miR-200 expression in various neonatal tissues

A) Heat map of expression levels of miR-200b in various neonatal tissues; red indicates highest expression, green indicates lowest. Generated from qRT-PCR data. B) qRT-PCR of miR-200b in various neonatal tissues; \( n=1 \). C) \textit{In situ} hybridization on nasal epithelium, D) tongue and palate and E) whisker buds from wildtype animals at E18.5. Scale bars are 200 um.
Discussion

As expected, miR-200 family expression is highest in epithelial tissues, and they display little to no expression in tissues of other origins. However, there are several caveats to these results. First, all analyses were done examining only the expression of miR-200b, just one of the five miR-200s. While I showed that the expression levels of the miR-200 family are relatively equal within the epidermis, it is possible that miR-200b has differential expression than the rest of the family in other tissues. Therefore, future studies should use miRNA-seq to more specifically define expression of the entire family.

In addition, the results shown here were obtained from whole tissues, and not from any particular cell type. Thus, to better define the expression patterns of the miR-200s in each of the tissues examined, in situ hybridization and cell sorting must be used to examine the various cell types present.
Appendix 2: RNA-seq on whole epidermis at P4.5

Introduction

RNA sequencing can be critical to contextualizing miRNA targets and functions. Earlier in this dissertation, I described miR-200 target identification using a combination of CLEAR-CLIP and RNA-seq from two different epidermal populations at P0.5. I chose to focus my analyses on hair germs and interfollicular epidermis at P0.5 to most accurately capture targets related to the observed hair germ phenotypes, and to avoid possible complications from neonatal lethality. However, I also performed RNA-seq on whole epidermis from animals at P4.5, in a further effort to identify miR-200 family targets. Targets identified by combining CLEAR-CLIP with this dataset may assist in better understanding the miR-200 family’s role during later stages of development, as it may reveal targets that are more specific to hair follicles and interfollicular epidermis at P4.5 than those at P0.5. In addition, combining all sets of RNA-seq provides an even more stringent strategy for identifying targets, and may reveal miR-200 family targets with strong regulation and with critical functions that do not depend upon developmental stage.

Materials & Methods

RNA was isolated from whole epidermis of control and Tg animals at P4.5. 1000 ng starting material was used to create RNA-seq libraries using the NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (NEB, catalog #E7420L). Sequencing, mapping, and data analysis were performed as described in Chapter 2. Gene ontology analysis was done with the Molecular Signatures Database (mSigDb, Broad Institute).
**Results**

Genes with a miR-200 CLEAR-CLIP read were robustly downregulated with miR-200 over-expression at P4.5 (Fig. A2a). Interestingly, genes with a CLEAR-CLIP read but no seed, which were downregulated in P0.5 data, did not appear to respond as strongly to miR-200 over-expression at P4.5 (Fig. A2a, blue line). As with the data at P0.5, I chose to consider genes down >10% in RNA-seq and interacting with miR-200s in CLEAR-CLIP as miR-200 family targets. In this way, I identified 513 miR-200 family targets important for the skin at P4.5 (for a full list of targets see Table A5). GO term analysis on these targets revealed many of the same processes as I identified at P0.5, including tight junctions, focal adhesions, and cell cycle (Fig. A2b).

**Figure A2 Signatures of miR-200 targets identified at P4.5**

A) Cumulative distribution plot of RNA-seq on P4.5 Tg and control whole epidermis comparing genes with a miR-200 CLEAR-CLIP read containing a seed (red line), a miR-200...
CLEAR-CLIP containing no seed (blue line), or no miR-200 CLEAR-CLIP read (black line). B) Select KEGG pathways enriched in miR-200 targets found at P4.5. C) Select KEGG pathways enriched in miR-200 targets shared between P0.5 and P4.5.

I next wanted to determine the differences in targets between P0.5 and P4.5. There were 356 genes shared between CLEAR-CLIP, P0.5 RNA-seq (hair germ and interfollicular epidermis) and P4.5. Within these genes I found similar enrichment for the same processes as P4.5 alone (Fig. A2c). These pathways were also highly enriched in genes found at P0.5 in hair germ and interfollicular epidermis (Chapter 4). Thus, despite the slight difference in targets between P4.5 and P0.5, the miR-200 is largely regulating the same processes at both time points.

**Discussion**

I performed RNA-seq on whole epidermis at P4.5 to further characterize the role of the miR-200 family in the skin and hair follicle. Combining this data with CLEAR-CLIP revealed a set of genes with largely the same signatures as those identified at P0.5. While I did identify genes which appear to be unique to either P0.5 or P4.5, there was no strong enrichment for any processes or pathways in one or the other. Therefore, although the miR-200 family may be interacting with some different targets at different developmental timepoints, they appear to be largely regulating the same processes throughout development of the epidermis and hair follicle, at least at these neonatal stages. Further work will be needed to determine the role, if any, of this miRNA family in the adult hair follicle cycle. Indeed, unpublished data indicates that the miR-200 family is more highly expressed during anagen than telogen, suggesting that miR-200s may also play a role in that transition.
Appendix 3: miR-200 family role in wound healing in vivo

Introduction

Wound healing is a complex process involving not only proliferation and migration of epithelial cells, but also inflammation, immune response, and angiogenesis. This complicated process has been extensively studied in the skin, where there is high accessibility to and high visibility of wounds. Additionally, wounding creates a stress to the skin which is likely to reveal functions for miRNAs that would not be uncovered during normal homeostasis. Indeed, a recent study showed that there is a change in the expression of several miRNAs after dermal wounding when compared with healthy skin, indicating that miRNAs are potential regulators of wound repair (Jin et al. 2013).

With miR-200 family over-expression and loss-of-function, I observed strong regulation of both cell proliferation and migration, two important processes for wound healing. In fact, the scratch assays I performed to assess cell migration are often used as a proxy for a wound-healing assay. Thus, I wanted to determine whether the miR-200 family was involved with wound healing in vivo. One previous study has been done showing that miR-200c is involved in regulating wound repair in aged skin (Aunin et al. 2017). However, the involvement of the other members of this family in wound healing has not been addressed. In addition, the role of miR-200s in wound healing of non-aged skin is unknown.

Materials & Methods

Six Tg animals and four control animals (five males and five females) at P28 were used for the wound healing experiment. All animals were fed doxycycline chow for two days prior
to wounding. Mice were put under anesthesia and shaved, and wounds were inflicted on backskin area using an 8 mm punch biopsy. Each mouse was given two wounds. Pictures and measurements were taken daily for 10 days following wounding; mice were fed doxycycline chow throughout. All animals were housed and treated according to the IACUC at the University of Colorado Boulder; IACUC protocol number was Yi 1408.01.

Results

Figure A3 miR-200 over-expression slows wound healing four days post wound.

Percentage of wound closed 1, 2, 3, 4 and 10 days post wounding. n=12 transgenic wounds; n=8 control wounds. **P<0.01.

At 1, 2, and 3 days post-wounding (DPW), there was no significant difference in the percentage of wound healed between transgenic and control animals (Fig. A2). However, at 4 DPW, whereas control wounds appeared to quicken their healing, transgenic wounds did not experience a change (Fig. A2). There was a statistically significant difference at 4 DPW, but not at any other stage. After 10 days, the wounds of both control and transgenic animals were healed to a similar percentage of their original side.
Discussion

These preliminary results suggest that miR-200 over-expression represses wound healing 
*in vivo*, as wounds in transgenic animals appeared to stall in their healing 4 days after 
wounding. However, this difference was not maintained during healing. This supports the 
results of *in vitro* scratch assays, which showed that miR-200 over-expression reduced cell 
migration (Chapter 5). Further work will be needed to examine specific miR-200 family 
targets that may be involved in this phenotype, although it is likely that the same targets 
regulating migration *in vitro* are involved with this phenotype. More careful studies with 
more precisely determined miR-200 over-expression, as well as with miR-200 loss-of-
function, will also be needed to more specifically define the role of the miR-200 family 
members in regulating wound healing.
## Appendix 4: Supplementary information

### Table A1: Oligos used for cloning and library preparation

<table>
<thead>
<tr>
<th>Cloning primers</th>
<th>Primer name</th>
<th>Sequence</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cloning primers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>JH16_TuD1_commonfwd</td>
<td>TGACGGCGCTAGGATCATCAAC</td>
<td>Common TuD sequence; forward</td>
</tr>
<tr>
<td></td>
<td>JH17_TuD2_commonrev</td>
<td>ACTGCTCGGAAAAAGCGGCGCT AGGATCATCTTTG</td>
<td>Common TuD sequence; reverse</td>
</tr>
<tr>
<td></td>
<td>JH18_TuD200a_3</td>
<td>GTATTCTGTGACCCAGAATTACTTGTTA CACGTGACATTGGAAGATGTTGTT GATGATCTTAG</td>
<td>miR200a/200b TuD; forward</td>
</tr>
<tr>
<td></td>
<td>JH23_TuD200b_4_real</td>
<td>CAAGTATTCTGTCAGAGAATACAAC TACTACATCTAAATGCAAACAGGCGCT AAGATGATCTTAGGCC</td>
<td>miR200a/200b TuD; reverse</td>
</tr>
<tr>
<td></td>
<td>JH20_TuD203_3</td>
<td>GTATTCTGTGACCAGAATTACTTGTTG AAAAGTAGATAGGACCAGACGATGTT GATGATCTTAG</td>
<td>miR203 TuD; forward</td>
</tr>
<tr>
<td></td>
<td>JH22_TuD203_4_real</td>
<td>CAAGTATTCTGTCAGAGAATACAAC TACTACATCTAAATGCAAACAGGCGCT AAGATGATCTTAGGCC</td>
<td>miR203 TuD; reverse</td>
</tr>
<tr>
<td></td>
<td>JH38_TuDctrl_3</td>
<td>TTCTGTGACCCAGAATTACTTGTTGTTGACGATTAGATGTTGATTGATGATCCTAG</td>
<td>TuD control cloning (5')</td>
</tr>
<tr>
<td></td>
<td>JH39_TuDctrl_4</td>
<td>CAAGTATTCTGTCAGAGAATACAAC TACTACATCTAAATGCAAACAGGCGCT AAGATGATCTTAGGCC</td>
<td>TuD control cloning (3')</td>
</tr>
<tr>
<td></td>
<td>DW138_miR200b_2x_F</td>
<td>CTAGATCATATTACCGGGCAAGGCTATT ATGACATATTACCGGGCAAGGCTATT</td>
<td>200b luciferase reporter; forward</td>
</tr>
<tr>
<td></td>
<td>DW139_miR200b_2x_R</td>
<td>TCGAGTAATACCTGGATATAGTAGGAGATCTAATACCTGGATATAGTAGG</td>
<td>200b luciferase reporter; reverse</td>
</tr>
<tr>
<td></td>
<td>JH24_200a_2x_F</td>
<td>CTAGAACATCTTTACCAGACGTTAGT TTACGATCATATTACCGGGCAAGGCTATT</td>
<td>200a luciferase reporter; forward</td>
</tr>
<tr>
<td></td>
<td>JH25_200a_2x_R</td>
<td>TCGAGTAATACCTGGATATAGTAGGAGATCTAATACCTGGATATAGTAGG</td>
<td>200a luciferase reporter; reverse</td>
</tr>
<tr>
<td></td>
<td>JH26_141_2x_F</td>
<td>CTAGACCATTTCACCAGAGCAGTGTGTT AGATCCCCATTACCCAGACGAGTGTGTT</td>
<td>141 luciferase reporter; forward</td>
</tr>
<tr>
<td></td>
<td>JH27_141_2x_R</td>
<td>TCGAGTAATACCTGGATATAGTAGGAGATCTAATACCTGGATATAGTAGG</td>
<td>141 luciferase reporter; reverse</td>
</tr>
<tr>
<td></td>
<td>JH28_200c_2x_F</td>
<td>CTAGATCCATATTACCCGAGCAGTAT TAGATCTCCATATTACCCGAGCAGTA TTAC</td>
<td>200c luciferase reporter; forward</td>
</tr>
<tr>
<td>Gene_1</td>
<td>Gene_2</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>JH29_200c_2x_R</td>
<td>TCGAGTAATACTGCCGGGTAATGATGGAGATCTAATACTGCCGGGTAATGATGGGAT</td>
<td>200c luciferase reporter; reverse</td>
<td></td>
</tr>
<tr>
<td>JH30_429_2x_F</td>
<td>CTAGAACGGCATTACACGACTGAGATCTAATACTGCCGGGTAATGATTA</td>
<td>429 luciferase reporter; forward</td>
<td></td>
</tr>
<tr>
<td>JH31_429_2x_R</td>
<td>TCGAGTAATACTGTCTGGTAATGCCGTGATCTAATACTGTCTGGTAATGCCGT</td>
<td>429 luciferase reporter; reverse</td>
<td></td>
</tr>
<tr>
<td>various (restriction sites)</td>
<td>GTTGACCTCTCCACTACCTA</td>
<td>miR-200b cluster cloning; forward</td>
<td></td>
</tr>
<tr>
<td>various (restriction sites)</td>
<td>ACCAGTGGTATAGCCAGG</td>
<td>miR-200b cluster cloning; reverse</td>
<td></td>
</tr>
<tr>
<td>JH50_cg2_3'_Xb_F</td>
<td>ATGCTCTAGACAGACTGAGGAGAATTCACCTTC</td>
<td>cyclin G2 3' UTR; forward with Xba1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH51_cg2_3'_Xh_R</td>
<td>ATGCCTCGAGGCCAACAGGTAAACA</td>
<td>cyclin G2 3' UTR; reverse with Xho1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH93_Ptpn14_S_XbF</td>
<td>ATGCTCTAGATGCTCTGCTAGACAGTAAACA</td>
<td>Ptpn14 3' UTR; forward with Xba1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH94_Ptpn14_S_XhR</td>
<td>ATGCCTCGAGGAGGGAACAGTGCAAAGGAATA</td>
<td>Ptpn14 3' UTR; reverse with Xho1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH110_YwaB_XbF</td>
<td>ATGCTCTAGAGTTGACCTCTCCACTACCTA</td>
<td>Ywhab 3' UTR; forward with Xba1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH111_YwaB_XhR</td>
<td>ATGCCTCGAGCCACCACAGAAGCAAACATTAG</td>
<td>Ywhab 3' UTR; reverse with Xho1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH146_Egfr_XbF</td>
<td>ATGCTCTAGATGCTCTGCTAGACAGTAAACA</td>
<td>Egfr 3' UTR; forward with Xba1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH147_Egfr_XhR</td>
<td>ATGCCTCGAGGCCAACAGGTAAACA</td>
<td>Egfr 3' UTR; reverse with Xho1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH152_Snai2_XbF</td>
<td>ATGCTCTAGATGCTCTGCTAGACAGTAAACA</td>
<td>Snai2 3' UTR; forward with Xba1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH153_Snai2_XhR</td>
<td>ATGCCTCGAGGCCAACAGGTAAACA</td>
<td>Snai2 3' UTR; reverse with Xho1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH166_Cfl2_XbF</td>
<td>ATGCTCTAGAGTTGACCTCTCCACTACCTA</td>
<td>Cfl2 3' UTR; forward with Xba1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH167_Cfl2_XhR</td>
<td>ATGCCTCGAGGCCAACAGGTAAACA</td>
<td>Cfl2 3' UTR; reverse with Xho1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH170_Ccnd2_XbF</td>
<td>ATGCTCTAGAGTTGACCTCTCCACTACCTA</td>
<td>Ccnd2 3' UTR; forward with Xba1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH171_Ccnd2_XhR</td>
<td>ATGCCTCGAGGCCAACAGGTAAACA</td>
<td>Ccnd2 3' UTR; reverse with Xho1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH359_Lats1_XbF</td>
<td>ATGCTCTAGATGCTCTGCTAGACAGTAAACA</td>
<td>Lats1 3' UTR; forward with Xba1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH360_Lats1_XhR</td>
<td>ATGCCTCGAGGCCAACAGGTAAACA</td>
<td>Lats1 3' UTR; reverse with Xho1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH365_Lats2_XbF</td>
<td>ATGCTCTAGAGTTGACCTCTCCACTACCTA</td>
<td>Lats2 3' UTR; forward with Xba1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH366_Lats2_XhR</td>
<td>ATGCTCTAGAGTTGACCTCTCCACTACCTA</td>
<td>Lats2 3' UTR; reverse with Xho1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH371_Amotl2_XbF</td>
<td>ATGCTCTAGAGTTGACCTCTCCACTACCTA</td>
<td>Amotl2 3' UTR; forward with Xba1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH372_Amotl2_XhR</td>
<td>ATGCTCTAGAGTTGACCTCTCCACTACCTA</td>
<td>Amotl2 3' UTR; reverse with Xho1 restriction site</td>
<td></td>
</tr>
<tr>
<td>Sample ID</td>
<td>Sequence</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>JH376_Fat1_XbF</td>
<td>ATGCTCTAGACCATTTCCAGCGTCCTAACT</td>
<td>Fat1 3' UTR; reverse with Xho1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH377_Fat1_XhR</td>
<td>ATGCCTCGAGCTTACACAAGACCATTGCATCAC</td>
<td>Fat1 3' UTR; reverse with Xho1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH384_Fat2_XbF</td>
<td>ATGCTCTAGAGGTGAGAATAGGAATGGTAAT</td>
<td>Fat2 3' UTR; reverse with Xho1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH385_Fat2_XhR</td>
<td>ATGCCTCGAGGCAACAGCTTTGTGAATC</td>
<td>Fat2 3' UTR; reverse with Xho1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH388_Ccnd1_XbF</td>
<td>ATGCTCTAGACACACGCTGGGTAGGATGAAATAG</td>
<td>Ccnd1 3' UTR; reverse with Xho1 restriction site</td>
<td></td>
</tr>
<tr>
<td>JH389_Ccnd1_XhR</td>
<td>ATGCCTCGAGGACGGAACACCTAGACCTAACAG</td>
<td>Ccnd1 3' UTR; reverse with Xho1 restriction site</td>
<td></td>
</tr>
<tr>
<td>Primer name</td>
<td>Sequence</td>
<td>Use</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>----------</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td>RY189-TRE2-376F</td>
<td>CGCCTGGAGACGCCATCCACGCT</td>
<td>pTRE2-200bcl genotyping; forward</td>
<td></td>
</tr>
<tr>
<td>DW129-200 Cluster1_screening_R</td>
<td>ATGCTGCCCAGTAAGATG</td>
<td>pTRE2-200bcl genotyping; reverse</td>
<td></td>
</tr>
<tr>
<td>RY69-K14-2609F</td>
<td>ctcacagctcttggcaacgtg</td>
<td>K14 forward primer for K14rtTA and K14cre</td>
<td></td>
</tr>
<tr>
<td>DW143 rtTA_R</td>
<td>TGGGCGGGTGGCGAGAT</td>
<td>K14rtTA genotyping; reverse</td>
<td></td>
</tr>
<tr>
<td>DW118-CreF</td>
<td>tgctgtttactgttatgcgg</td>
<td>cre genotyping; forward</td>
<td></td>
</tr>
<tr>
<td>DW119-CreR</td>
<td>ttgccccctgctactatccag</td>
<td>cre genotyping; reverse for cre and K14-cre</td>
<td></td>
</tr>
<tr>
<td>JH14_miR200cko_F</td>
<td>CAACAGGCTCTGACTTTAACC</td>
<td>miR-200c cluster floxed allele; forward</td>
<td></td>
</tr>
<tr>
<td>JH15_miR200cko_R</td>
<td>CTTCTGGGCAGACAAGAAT</td>
<td>miR-200c cluster floxed allele; reverse</td>
<td></td>
</tr>
<tr>
<td>DW218_200b-a-429_GT_L1F</td>
<td>GCTGAGCATCCGGAAAGTA</td>
<td>miR-200bcl CRISPR knock-out genotyping, left side; forward; forward for knock-out band</td>
<td></td>
</tr>
<tr>
<td>DW219_200b-a-429_GT_L1R</td>
<td>CAGGCTAACCCTTTAGTGCTGT</td>
<td>miR-200bcl CRISPR knock-out genotyping, left side; reverse</td>
<td></td>
</tr>
<tr>
<td>DW222_200b-a-429_GT_RF</td>
<td>ATGGGGATTTAGTGGTGT</td>
<td>miR-200bcl CRISPR knock-out genotyping, right side; forward</td>
<td></td>
</tr>
<tr>
<td>DW223_200b-a-429_GT_RR</td>
<td>TGCAAGGCTCTTGAATG</td>
<td>miR-200bcl CRISPR knock-out genotyping, right side; reverse; reverse for knock-out band</td>
<td></td>
</tr>
<tr>
<td>DW183_miR200b-a-429_L1_F</td>
<td>GTGCCAGCCTCTGGCAGCG</td>
<td>L gRNA for miR-200b cluster</td>
<td></td>
</tr>
<tr>
<td>DW188_miR200b-a-429_R1_R</td>
<td>ATCGTGCCTCTATAATAGG</td>
<td>R gRNA for miR-200b cluster</td>
<td></td>
</tr>
<tr>
<td>Primer name</td>
<td>Sequence</td>
<td>Use</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-------------------------------</td>
<td>-----------------------------------</td>
<td></td>
</tr>
<tr>
<td>JH45_cycE2_q_F</td>
<td>ATGTAAGACGCAGCGGTTTA</td>
<td>qPCR for cyclin E2; forward</td>
<td></td>
</tr>
<tr>
<td>JH46_cycE2_q_R</td>
<td>GCTGATTTCTCCAGACAGTACA</td>
<td>qPCR for cyclin E2; reverse</td>
<td></td>
</tr>
<tr>
<td>JH47_cycG2_q_F</td>
<td>AGGGGGTTCAGCTTTCCGGATT</td>
<td>qPCR for cyclin G2</td>
<td></td>
</tr>
<tr>
<td>JH48_cycG2_q_R</td>
<td>AGTGTATATCTTCTCCGGGTAG</td>
<td>qPCR for cyclin G2; reverse</td>
<td></td>
</tr>
<tr>
<td>JH87_Ptpn14_qF</td>
<td>ACAAGAACAGACGGTGATGTT</td>
<td>qPCR for ptpn14; forward</td>
<td></td>
</tr>
<tr>
<td>JH88_Ptpn14_qR</td>
<td>CTGTGGACACCTTGCTCTAGG</td>
<td>qPCR for ptpn14; reverse</td>
<td></td>
</tr>
<tr>
<td>JH305_YWHAEqF</td>
<td>ACTGGCGAGTCCAAGGTTTTC</td>
<td>qPCR for ywhae; forward</td>
<td></td>
</tr>
<tr>
<td>JH306_YWHAEqR</td>
<td>TGATCATCACCATTCAGC</td>
<td>qPCR for ywhae; reverse</td>
<td></td>
</tr>
<tr>
<td>JH311_YWHAZqF</td>
<td>GAAAGTTTCTTGGATCACAATGC</td>
<td>qPCR for ywhaz; forward</td>
<td></td>
</tr>
<tr>
<td>JH312_YWHAZqR</td>
<td>TGATCTGGTCAACATTCCTT</td>
<td>qPCR for ywhaz; reverse</td>
<td></td>
</tr>
<tr>
<td>JH313_YWHAQf</td>
<td>GATGCTGGATGACCCCTC</td>
<td>qPCR for ywhab; forward</td>
<td></td>
</tr>
<tr>
<td>JH316_YWHAQr</td>
<td>CTCGTTGGGATCCTTGGTTACG</td>
<td>qPCR for ywhab; reverse</td>
<td></td>
</tr>
<tr>
<td>JH317_CCND1qF</td>
<td>GCCTACCTGACACCAATCTC</td>
<td>qPCR for ccnd1; forward</td>
<td></td>
</tr>
<tr>
<td>JH318_CCND1qR</td>
<td>CTCTCTTGGCATCTTCCG</td>
<td>qPCR for ccnd1; reverse</td>
<td></td>
</tr>
<tr>
<td>JH319_CCND2qF</td>
<td>GAGTCGGGATCCTTGTTG</td>
<td>qPCR for ccnd2; forward</td>
<td></td>
</tr>
<tr>
<td>JH320_CCND2qR</td>
<td>CGCACTTGGCAGGAAGGT</td>
<td>qPCR for ccnd2; reverse</td>
<td></td>
</tr>
<tr>
<td>JH321_CFL2qF</td>
<td>GCATCTGGGCTGCTAGGTAAAGTA</td>
<td>qPCR for cfl2; forward</td>
<td></td>
</tr>
<tr>
<td>JH322_CFL2qR</td>
<td>CACCAATGTCACCCACCAAGA</td>
<td>qPCR for cfl2; reverse</td>
<td></td>
</tr>
<tr>
<td>JH323_EGFrF</td>
<td>GCCATCTGGGCTGACCCCTC</td>
<td>qPCR for egfr; forward</td>
<td></td>
</tr>
<tr>
<td>JH324_EGFrR</td>
<td>GTCTTGGCATGAAATAGGCGAAT</td>
<td>qPCR for egfr; reverse</td>
<td></td>
</tr>
<tr>
<td>JH325_METqF</td>
<td>GTGGAATGATGGATGATGGG</td>
<td>qPCR for met; forward</td>
<td></td>
</tr>
<tr>
<td>JH326_METqR</td>
<td>TGTAGTGGGGATGAGCGAGG</td>
<td>qPCR for met; reverse</td>
<td></td>
</tr>
<tr>
<td>JH327_PVRL1qF</td>
<td>GACTCCATGTATGGGGTCACTC</td>
<td>qPCR for pvrl1; forward</td>
<td></td>
</tr>
<tr>
<td>JH330_PVRL1qR</td>
<td>CATGTCCTGGTAGAGGAGG</td>
<td>qPCR for pvrl1; reverse</td>
<td></td>
</tr>
<tr>
<td>JH331_THBS1qF</td>
<td>GGGGGAGATAACGGTGTTGG</td>
<td>qPCR for thbs1; forward</td>
<td></td>
</tr>
<tr>
<td>JH332_THBS1qR</td>
<td>CGGAGGATCATGGTGGCATT</td>
<td>qPCR for thbs1; reverse</td>
<td></td>
</tr>
<tr>
<td>JH333_LAMC2qF</td>
<td>CACACAGGGAGATGGCTACT</td>
<td>qPCR for lamc2; forward</td>
<td></td>
</tr>
<tr>
<td>JH334_LAMC2qR</td>
<td>CGACGTGCCACCTGCAGTG</td>
<td>qPCR for lamc2; reverse</td>
<td></td>
</tr>
<tr>
<td>JH335_LFNqF</td>
<td>CGAGGAATGATGGACTCCTCC</td>
<td>qPCR for lfng; forward</td>
<td></td>
</tr>
<tr>
<td>JH336_LFNqR</td>
<td>GCCAGGGCGACAGAACATTCG</td>
<td>qPCR for lfng; reverse</td>
<td></td>
</tr>
<tr>
<td>JH337_NOTCH1qF</td>
<td>CCGTCCTCTGCGTTAGC</td>
<td>qPCR for notch1; forward</td>
<td></td>
</tr>
<tr>
<td>JH338_NOTCH1qR</td>
<td>GCAAGTGACTGCTGCTTGG</td>
<td>qPCR for notch1; reverse</td>
<td></td>
</tr>
<tr>
<td>JH339_NOTCH2qF</td>
<td>ATGGTGAACTGCTGCTTGG</td>
<td>qPCR for notch2; forward</td>
<td></td>
</tr>
<tr>
<td>JH340_NOTCH2qR</td>
<td>GGAAGCAGAGACAGACGCAC</td>
<td>qPCR for notch2; reverse</td>
<td></td>
</tr>
<tr>
<td>JH341_JAG1qF</td>
<td>CCTGAGGCAACTTGGAAAGTA</td>
<td>qPCR for jag1; forward</td>
<td></td>
</tr>
<tr>
<td>JH342_JAG1qR</td>
<td>CCTGAGGCACTTGGAAAGTA</td>
<td>qPCR for jag1; reverse</td>
<td></td>
</tr>
<tr>
<td>JH343_DLL1qF</td>
<td>CAGGACCTCTCTTTCCGGTATG</td>
<td>qPCR for dll1; forward</td>
<td></td>
</tr>
<tr>
<td>Primer ID</td>
<td>Sequence</td>
<td>qPCR for</td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>---------------------------</td>
<td>-------------------</td>
<td></td>
</tr>
<tr>
<td>JH344_DLL1qR</td>
<td>AAGGGGAATCGGATGGGGTT</td>
<td>dll1; reverse</td>
<td></td>
</tr>
<tr>
<td>JH345_LATS2qF</td>
<td>GGACCCCAGGAATGAGCAG</td>
<td>lats2; forward</td>
<td></td>
</tr>
<tr>
<td>JH346_LATS2qR</td>
<td>CCCTCGTAGTTGCACCACC</td>
<td>lats2; reverse</td>
<td></td>
</tr>
<tr>
<td>JH349_AMOTL2qF</td>
<td>GGAGAAGAGTTGCCCACCTAT</td>
<td>amotl2; forward</td>
<td></td>
</tr>
<tr>
<td>JH350_AMOTL2qR</td>
<td>TCGAAGAGCTTCCATCTTCTGCG</td>
<td>amotl2; reverse</td>
<td></td>
</tr>
<tr>
<td>JH410_Cyr61_qF</td>
<td>CCCCCGGCTGGTGAAAGT</td>
<td>Cyr61; forward</td>
<td></td>
</tr>
<tr>
<td>JH411_Cyr61_qR</td>
<td>GCGGTTCGGTGCCAAAGA</td>
<td>Cyr61; reverse</td>
<td></td>
</tr>
<tr>
<td>JH412_Cgtf_qF</td>
<td>GGGCCTCTCTCGGATTTC</td>
<td>Ctgf; forward</td>
<td></td>
</tr>
<tr>
<td>JH413_Cgtf_qR</td>
<td>ATCCAGGCAAGTGCATTGGTA</td>
<td>Ctgf; reverse</td>
<td></td>
</tr>
</tbody>
</table>
Table A2: Genotypes, cell numbers, and mapping statistics for RNA-seq

**P0.5 RNA-seq**

<table>
<thead>
<tr>
<th>Animal/Genotype</th>
<th>Library</th>
<th>Raw reads</th>
<th>Mapped reads</th>
<th>Multiple alignment</th>
<th>% mapped</th>
</tr>
</thead>
<tbody>
<tr>
<td>K14rtTA 3</td>
<td>Hair germ 3</td>
<td>29745331</td>
<td>21551116</td>
<td>1016760 (4.7%)</td>
<td>72.50%</td>
</tr>
<tr>
<td></td>
<td>IFE 3</td>
<td>32835287</td>
<td>19547820</td>
<td>3248301 (16.6%)</td>
<td>59.50%</td>
</tr>
<tr>
<td>pTRE2-200bcl/K14rtTA 4</td>
<td>Hair germ 4</td>
<td>43802741</td>
<td>26382601</td>
<td>2110267 (8.0%)</td>
<td>60.20%</td>
</tr>
<tr>
<td></td>
<td>IFE 4</td>
<td>44921691</td>
<td>26087694</td>
<td>5249410 (20.1%)</td>
<td>58.10%</td>
</tr>
<tr>
<td>pTRE2-200bcl 5</td>
<td>Hair germ 5</td>
<td>35120986</td>
<td>32358824</td>
<td>1217179 (3.8%)</td>
<td>92.10%</td>
</tr>
<tr>
<td></td>
<td>IFE 5</td>
<td>29335695</td>
<td>26714563</td>
<td>1115661 (4.2%)</td>
<td>91.10%</td>
</tr>
<tr>
<td>pTRE2-200bcl/K14rtTA 6</td>
<td>Hair germ 6</td>
<td>25753274</td>
<td>17401894</td>
<td>905492 (5.2%)</td>
<td>67.60%</td>
</tr>
<tr>
<td></td>
<td>IFE 6</td>
<td>34269340</td>
<td>25307721</td>
<td>1291168 (5.1%)</td>
<td>73.80%</td>
</tr>
</tbody>
</table>

**P4.5 RNA-seq**

<table>
<thead>
<tr>
<th>Library</th>
<th>Genotype</th>
<th>Raw reads</th>
<th>Mapped reads</th>
<th>Multiple alignment</th>
<th>% mapped</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wildtype 4</td>
<td>Wildtype</td>
<td>47831347</td>
<td>39891233</td>
<td>6973423 (17.5%)</td>
<td>83.40%</td>
</tr>
<tr>
<td>Wildtype 7</td>
<td>Wildtype</td>
<td>43513730</td>
<td>39837856</td>
<td>4026375 (10.1%)</td>
<td>91.60%</td>
</tr>
<tr>
<td>Inducible 2</td>
<td>pTRE2-200bcl/K14rtTA</td>
<td>41760393</td>
<td>37072560</td>
<td>5118812 (13.8%)</td>
<td>88.80%</td>
</tr>
<tr>
<td>Inducible 3</td>
<td>pTRE2-200bcl/K14rtTA</td>
<td>47360811</td>
<td>39767680</td>
<td>6936834 (17.4%)</td>
<td>84.00%</td>
</tr>
</tbody>
</table>
Table A3: FIJI Macro for focal adhesion quantification

```java
imgName = getTitle();
run("Split Channels");
selectWindow("C2-" + imgName);
run("Enhance Contrast", "saturated=0.35");
setAutoThreshold("Mean dark");
run("Convert to Mask");
run("Dilate");
run("Close-");
run("Fill Holes");
run("Analyze Particles...", "size=1200-Infinity circularity=0.00-0.99");
run("Create Selection")
roiManager("ADD");
selectWindow("C1-" + imgName);
roiManager("AND");
selectWindow("C3-" + imgName);
roiManager("AND");
selectWindow("C1-" + imgName);
run("Enhance Local Contrast (CLAHE)", "blocksize=75 histogram=256 maximum=6 mask=*None* fast_(less_accurate)" );
run("Enhance Contrast", "saturated=0.35");
setAutoThreshold("Default dark");
run("Analyze Particles...", "size=25-Infinity circularity=0.00-1.00 display summarize");
selectWindow("C3-" + imgName);
run("Subtract Background...", "rolling=50 sliding");
run("Enhance Local Contrast (CLAHE)", "blocksize=19 histogram=256 maximum=6 mask=*None* fast_(less_accurate)" );
run("Exp");
run("Enhance Contrast", "saturated=0.35");
run("LoG 3D", "sigmax=5 sigmay=5");
run("8-bit");
setAutoThreshold("Default dark");
run("Analyze Particles...", "size=1-2000 circularity=0.00-0.99 show=Outlines display summarize exclude");
roiManager("Delete");
```
Table A4: Full list of miR-200 family targets identified with CLEAR-CLIP and P0.5 RNA-seq

Hair germ only

<table>
<thead>
<tr>
<th>A230046K03RIK</th>
<th>VPS36</th>
<th>HERPUD2</th>
<th>ATP2B1</th>
<th>ARF4</th>
</tr>
</thead>
<tbody>
<tr>
<td>241006H16RIK</td>
<td>PIP4K2B</td>
<td>ZC3H12C</td>
<td>XBP1</td>
<td>SOCS4</td>
</tr>
<tr>
<td>SDF2</td>
<td>TIMP2</td>
<td>TPP2</td>
<td>C1D</td>
<td>GPR180</td>
</tr>
<tr>
<td>SNAPC1</td>
<td>TGIF1</td>
<td>CAB39</td>
<td>PPP1R9B</td>
<td>OXR1</td>
</tr>
<tr>
<td>TRAF6</td>
<td>ZFP608</td>
<td>COPS8</td>
<td>MAP3K3</td>
<td>EP300</td>
</tr>
<tr>
<td>LIN7C</td>
<td>NFE2L2</td>
<td>PPP1R15B</td>
<td>ITSN2</td>
<td>ALG10B</td>
</tr>
<tr>
<td>SCMH1</td>
<td>MAP7D1</td>
<td>ZBTB18</td>
<td>MPLKIP</td>
<td>CCDC50</td>
</tr>
<tr>
<td>AACS</td>
<td>UBIAD1</td>
<td>LPGAT1</td>
<td>TMEM170B</td>
<td>OSBPL11</td>
</tr>
<tr>
<td>TSC22D4</td>
<td>WDFY3</td>
<td>AES</td>
<td>ABHD6</td>
<td>CPOX</td>
</tr>
<tr>
<td>TPRKB</td>
<td>FOXP1</td>
<td>CDK17</td>
<td>SLC4A7</td>
<td>MORC3</td>
</tr>
<tr>
<td>ATP6V0E</td>
<td>ZSWIM3</td>
<td>MSANTD4</td>
<td>STX11</td>
<td>PAXBP1</td>
</tr>
<tr>
<td>DMXL1</td>
<td>RAP2B</td>
<td>MSANTD2</td>
<td>FOXO3</td>
<td>STK38</td>
</tr>
<tr>
<td>TMX3</td>
<td>UBE2J1</td>
<td>SIAE</td>
<td>LPIN1</td>
<td>MCFD2</td>
</tr>
<tr>
<td>SLC29A2</td>
<td>TBC1D1</td>
<td>ADPGK</td>
<td>PQLC3</td>
<td>NPC1</td>
</tr>
<tr>
<td>CEP55</td>
<td>RABGGE1</td>
<td>GCLC</td>
<td>SEC23A</td>
<td>MYOF</td>
</tr>
<tr>
<td>ENTPD7</td>
<td>PLEKHA8</td>
<td>ATP7A</td>
<td>NEK9</td>
<td>SCD1</td>
</tr>
<tr>
<td>FAM160B1</td>
<td>AAK1</td>
<td>RPS6KA3</td>
<td>MAP3K1</td>
<td>ZBTB34</td>
</tr>
<tr>
<td>DYNC1I2</td>
<td>EDEM1</td>
<td>PFKFB2</td>
<td>SLC39A14</td>
<td>BAZ2B</td>
</tr>
<tr>
<td>AGPS</td>
<td>MYADM</td>
<td>CEP350</td>
<td>ETV5</td>
<td>ARHGAP11A</td>
</tr>
<tr>
<td>YWHAB</td>
<td>CLN8</td>
<td>SFT2D2</td>
<td>GTF2E1</td>
<td>ELMO2</td>
</tr>
<tr>
<td>CCNL1</td>
<td>SMURF1</td>
<td>IDS</td>
<td>MARCKS</td>
<td>MAGT1</td>
</tr>
<tr>
<td>HIPK1</td>
<td>MKRN1</td>
<td>TSC22D3</td>
<td>SERINC1</td>
<td></td>
</tr>
<tr>
<td>TET2</td>
<td>RAB7</td>
<td>SEC63</td>
<td>APAF1</td>
<td></td>
</tr>
<tr>
<td>CISD2</td>
<td>CECR5</td>
<td>FGFR1OP</td>
<td>RBMS1</td>
<td></td>
</tr>
<tr>
<td>CDC42</td>
<td>CHD2</td>
<td>PPP1R18</td>
<td>SPATA2</td>
<td></td>
</tr>
<tr>
<td>VAMP3</td>
<td>DCUN1D3</td>
<td>MAP3K2</td>
<td>PTPN12</td>
<td></td>
</tr>
<tr>
<td>RHEB</td>
<td>IST1</td>
<td>ITGAV</td>
<td>TMEM168</td>
<td></td>
</tr>
<tr>
<td>BLOC1S4</td>
<td>MAF</td>
<td>SPG20</td>
<td>WNK1</td>
<td></td>
</tr>
<tr>
<td>SDAD1</td>
<td>FAM63B</td>
<td>MSN</td>
<td>ZFP36</td>
<td></td>
</tr>
<tr>
<td>GLTP</td>
<td>ATP11C</td>
<td>STK17B</td>
<td>MAPK6</td>
<td></td>
</tr>
</tbody>
</table>
### Interfollicular epidermis only

<table>
<thead>
<tr>
<th>Gene 1</th>
<th>Gene 2</th>
<th>Gene 3</th>
<th>Gene 4</th>
<th>Gene 5</th>
<th>Gene 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTGF</td>
<td>DHCR7</td>
<td>GJA1</td>
<td>SERINC5</td>
<td>IKZF2</td>
<td>LRP6</td>
</tr>
<tr>
<td>MYH10</td>
<td>SUGP2</td>
<td>ZWINT</td>
<td>HOMER1</td>
<td>PKP1</td>
<td>MEX3B</td>
</tr>
<tr>
<td>UBE2G1</td>
<td>FZD5</td>
<td>FGD6</td>
<td>HMGCS1</td>
<td>RAPGEF2</td>
<td>LDLR</td>
</tr>
<tr>
<td>CRK</td>
<td>SSR1</td>
<td>TMEM194</td>
<td>PTPRG</td>
<td>MAN1A2</td>
<td>RCN2</td>
</tr>
<tr>
<td>MIPO1</td>
<td>BMPR1A</td>
<td>PIK3IP1</td>
<td>RICTOR</td>
<td>LRIG2</td>
<td>LARS2</td>
</tr>
<tr>
<td>ARRC3</td>
<td>TRIO</td>
<td>RNF145</td>
<td>TNRC6B</td>
<td>PDLIM5</td>
<td>LEMD3</td>
</tr>
<tr>
<td>MYO10</td>
<td>LNPEP</td>
<td>CNOT8</td>
<td>PPHLN1</td>
<td>PLAG1</td>
<td>EPHA2</td>
</tr>
<tr>
<td>CLIP4</td>
<td>PTPRS</td>
<td>NUFIP2</td>
<td>ACVR1B</td>
<td>RIPK2</td>
<td>MTF2</td>
</tr>
<tr>
<td>CDC42BP</td>
<td>METTL4</td>
<td>ACACA</td>
<td>SP1</td>
<td>ZFP292</td>
<td>F2RL1</td>
</tr>
<tr>
<td>AIF1L</td>
<td>TTF2</td>
<td>SOCS7</td>
<td>PPM1F</td>
<td>TMEM245</td>
<td>SIK1</td>
</tr>
<tr>
<td>USP33</td>
<td>RNF11</td>
<td>PSME3</td>
<td>CRKL</td>
<td>KLHL9</td>
<td>DSC3</td>
</tr>
<tr>
<td>SLC6A9</td>
<td>CYP26B1</td>
<td>NBR1</td>
<td>EIF4A2</td>
<td>XKR8</td>
<td>NR6A1</td>
</tr>
<tr>
<td>SLC2A1</td>
<td>TP1</td>
<td>SDC1</td>
<td>DLG1</td>
<td>CHKA</td>
<td>TACSTD2</td>
</tr>
<tr>
<td>SNHG12</td>
<td>CLCN3</td>
<td>TMX1</td>
<td>UHRF1BP1</td>
<td>VLDLR</td>
<td>MAT2A</td>
</tr>
<tr>
<td>SH2D5</td>
<td>PIKFYVE</td>
<td>HIF1A</td>
<td>CRIM1</td>
<td>ACVR2A</td>
<td>TIAM1</td>
</tr>
<tr>
<td>IFFO2</td>
<td>GIGYF2</td>
<td>CHURC1</td>
<td>EML4</td>
<td>TANC1</td>
<td>ZFP945</td>
</tr>
<tr>
<td>CERS3</td>
<td>SERPINB2</td>
<td>RREB1</td>
<td>SVIL</td>
<td>CERS6</td>
<td>SLC39A6</td>
</tr>
<tr>
<td>SEMA4B</td>
<td>TADA1</td>
<td>EDN1</td>
<td>SLC12A2</td>
<td>BTBD3</td>
<td>CEP76</td>
</tr>
<tr>
<td>LRRC32</td>
<td>PIGM</td>
<td>ELL2</td>
<td>CEP192</td>
<td>MTRF1</td>
<td>SEC61A2</td>
</tr>
<tr>
<td>CD81</td>
<td>BEND3</td>
<td>TMEM167</td>
<td>PARD6G</td>
<td>NPEPPS</td>
<td>ITGB6</td>
</tr>
<tr>
<td>BTAF1</td>
<td>EXTL2</td>
<td>FSCN1</td>
<td>ATP11A</td>
<td>CDC27</td>
<td>ZSCAN29</td>
</tr>
<tr>
<td>CCNJ</td>
<td>2610301B20RIK</td>
<td>SERBP1</td>
<td>WHSC1L1</td>
<td>STX12</td>
<td>SLC30A4</td>
</tr>
<tr>
<td>SUFU</td>
<td>SRM</td>
<td>FAM136A</td>
<td>FAT1</td>
<td>PDIK1L</td>
<td>PHC3</td>
</tr>
<tr>
<td>DLX1</td>
<td>SEMA3C</td>
<td>IL17RC</td>
<td>DDX39</td>
<td>STEAP2</td>
<td>GOLM4</td>
</tr>
<tr>
<td>FNBP4</td>
<td>EPT1</td>
<td>DMWD</td>
<td>CYLD</td>
<td>APBB2</td>
<td>SIDT2</td>
</tr>
<tr>
<td>KATNBL1</td>
<td>RGS12</td>
<td>ZFP568</td>
<td>NFAT5</td>
<td>SLC7A1</td>
<td>ARIH1</td>
</tr>
<tr>
<td>TFAP2C</td>
<td>AFAP1</td>
<td>KRTDAP</td>
<td>NPAT</td>
<td>WASL</td>
<td>MYZAP</td>
</tr>
<tr>
<td>YTHDF3</td>
<td>CPEB2</td>
<td>BCO30336</td>
<td>OA22</td>
<td>ZC3HC1</td>
<td>ARIH2</td>
</tr>
<tr>
<td>4932438A13RIK</td>
<td>SRP72</td>
<td>FBXL19</td>
<td>CSNK1G1</td>
<td>PCYOX1</td>
<td>MBTPS2</td>
</tr>
<tr>
<td>4930579G24RIK</td>
<td>PTPN13</td>
<td>COL4A2</td>
<td>FUND2</td>
<td>PLXNA1</td>
<td>ZADH2</td>
</tr>
<tr>
<td>DMD</td>
<td>BLZF1</td>
<td>EGLN3</td>
<td>ZFP26</td>
<td>MEX3C</td>
<td>CLDN12</td>
</tr>
<tr>
<td>ZXDB</td>
<td>SYDE1</td>
<td>E2F3</td>
<td>KLF6</td>
<td>TMEM97</td>
<td>6030458C11RIK</td>
</tr>
<tr>
<td>DLG3</td>
<td>DGKA</td>
<td>SERPINB6C</td>
<td>DEPDC1B</td>
<td>PRR11</td>
<td>UBR5</td>
</tr>
<tr>
<td>OGT</td>
<td>IGFBP3</td>
<td>OCLN</td>
<td>TRP63</td>
<td>SMG1</td>
<td>RBFOX2</td>
</tr>
<tr>
<td>ALG13</td>
<td>SPTBN1</td>
<td>ZFP131</td>
<td>REV1</td>
<td>TUBGCP3</td>
<td>DCL1</td>
</tr>
<tr>
<td>FAM135A</td>
<td>FAT2</td>
<td>NSUN3</td>
<td>NDUFS1</td>
<td>GOLGA7</td>
<td>AMFR</td>
</tr>
<tr>
<td>RPL13A</td>
<td>ZFP565</td>
<td>KRAS</td>
<td>ABCC5</td>
<td>CLDN23</td>
<td>CIAPIN1</td>
</tr>
<tr>
<td>FLRT3</td>
<td>APLP2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>197</td>
<td>NUCKS1</td>
<td>SEC24A</td>
<td>EGFR</td>
<td>PHF20L1</td>
<td>FUBP3</td>
</tr>
<tr>
<td>-----</td>
<td>--------</td>
<td>--------</td>
<td>------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>GAS5</td>
<td>UTP18</td>
<td>RTN4</td>
<td>LARP4</td>
<td>GPR21</td>
<td>DR1</td>
</tr>
<tr>
<td>LIN9</td>
<td>ACTN1</td>
<td>PWWPA2</td>
<td>RRN3</td>
<td>RIF1</td>
<td>DTX2</td>
</tr>
<tr>
<td>CD24A</td>
<td>BCL11B</td>
<td>PPP2CA</td>
<td>IL1RAP</td>
<td>PPIG</td>
<td>CYTH3</td>
</tr>
<tr>
<td>NSD1</td>
<td>TMEM65</td>
<td>ZCCHC10</td>
<td>ZFP148</td>
<td>CELF1</td>
<td>PAN3</td>
</tr>
<tr>
<td>MYC</td>
<td>CELSR1</td>
<td>MPRI P</td>
<td>NAA50</td>
<td>PAK6</td>
<td>PDS5B</td>
</tr>
<tr>
<td>GSK3B</td>
<td>GRPEL2</td>
<td>RHTO1</td>
<td>CBLB</td>
<td>MGA</td>
<td>SMARCAD1</td>
</tr>
<tr>
<td>CHD1</td>
<td>SGMS1</td>
<td>SRSF1</td>
<td>GABPA</td>
<td>CSRP2BP</td>
<td>0610030E20RIK</td>
</tr>
<tr>
<td>ZFP532</td>
<td>COL17A1</td>
<td>TOB1</td>
<td>SON</td>
<td>PLCG1</td>
<td>TMEM43</td>
</tr>
<tr>
<td>CHMP1B</td>
<td>CAPRIN1</td>
<td>CBX1</td>
<td>MRPS6</td>
<td>PARD6B</td>
<td>NR2C2</td>
</tr>
<tr>
<td>ATL3</td>
<td>KPN4</td>
<td>CDK12</td>
<td>TULP4</td>
<td>RAB22A</td>
<td>PPP4R2</td>
</tr>
<tr>
<td>EPC2</td>
<td>AQP3</td>
<td>MSL1</td>
<td>TBP</td>
<td>PROSER1</td>
<td>CHL1</td>
</tr>
<tr>
<td>PLEKH3</td>
<td>JAK1</td>
<td>PUM2</td>
<td>MAN2A1</td>
<td>FBXW7</td>
<td>PPFI B1</td>
</tr>
<tr>
<td>F3</td>
<td>FGFBP1</td>
<td>LAPT M4A</td>
<td>PPP4R1</td>
<td>MEX3A</td>
<td>ZC3H4</td>
</tr>
<tr>
<td>BCA R3</td>
<td>KBTBD2</td>
<td>ROCK2</td>
<td>EPAS1</td>
<td>ANP32E</td>
<td>ZFP60</td>
</tr>
<tr>
<td>WLS</td>
<td>USP5</td>
<td>RRM2</td>
<td>DSG3</td>
<td>IGSF3</td>
<td>RAB6A</td>
</tr>
<tr>
<td>MEG6</td>
<td>COL4A1</td>
<td>PXDN</td>
<td>AMMECR1L</td>
<td>SLC16A1</td>
<td>ADM</td>
</tr>
<tr>
<td>YES1</td>
<td>GNL3L</td>
<td>FOXC1</td>
<td>SLC35A4</td>
<td>ELOVL6</td>
<td>ACADS B</td>
</tr>
<tr>
<td>WHSC1</td>
<td>PCMTD1</td>
<td>SNRN P48</td>
<td>TGERG1</td>
<td>TG51</td>
<td>PRD X2</td>
</tr>
<tr>
<td>BAZ1B</td>
<td>BMPR2</td>
<td>SMIM13</td>
<td>CSNK1A1</td>
<td>CCNE2</td>
<td>FAM65A</td>
</tr>
<tr>
<td>CDKN1B</td>
<td>RQCD1</td>
<td>BICD2</td>
<td>SMAD2</td>
<td>USP45</td>
<td>ST3 GAL2</td>
</tr>
<tr>
<td>ZFP260</td>
<td>AGFG1</td>
<td>SMIM15</td>
<td>2700081O15RIK</td>
<td>SMC2</td>
<td>HMG20A</td>
</tr>
<tr>
<td>USP47</td>
<td>CCNT2</td>
<td>4833420G17RIK</td>
<td>KLF9</td>
<td>RAD23B</td>
<td>NEDD4</td>
</tr>
<tr>
<td>LP AR2</td>
<td>ZFP2B1</td>
<td>PKX</td>
<td>KANK1</td>
<td>SRSF10</td>
<td>ME4</td>
</tr>
<tr>
<td>GADD45GIP1</td>
<td>ABL2</td>
<td>DENND6A</td>
<td>TBC1D12</td>
<td>LUZP1</td>
<td>SH3BGRL2</td>
</tr>
<tr>
<td>MYO9A</td>
<td>CNIH4</td>
<td>BAP1</td>
<td>FAM178A</td>
<td>ERRF1</td>
<td>CDC25A</td>
</tr>
<tr>
<td>COL7A1</td>
<td>TRAF3 IP2</td>
<td>ZMYM2</td>
<td>SLK</td>
<td>CCNL2</td>
<td>CLASP2</td>
</tr>
<tr>
<td>SETD2</td>
<td>HSF2</td>
<td>REEP4</td>
<td>SEPHS1</td>
<td>CDK6</td>
<td>TCAIM</td>
</tr>
<tr>
<td>CTNNB1</td>
<td>PPP1R12A</td>
<td>NUP155</td>
<td>COMMD3</td>
<td>DCUN1D4</td>
<td>FMR1</td>
</tr>
<tr>
<td>DDX3X</td>
<td>PHLDA1</td>
<td>SQLE</td>
<td>SPTAN1</td>
<td>SLC4A4</td>
<td>ZFP185</td>
</tr>
<tr>
<td>VBP1</td>
<td>LGALS1</td>
<td>MAST4</td>
<td>ZFP871</td>
<td>CPNE1</td>
<td>YWHAG</td>
</tr>
<tr>
<td>FAM168B</td>
<td>NPM1</td>
<td>ERBB2IP</td>
<td>TUBB5</td>
<td>ELF2</td>
<td>NYAP1</td>
</tr>
<tr>
<td>KDEL C1</td>
<td>CNOT6</td>
<td>FST</td>
<td>NFYA</td>
<td>SSR3</td>
<td>N4BP2L2</td>
</tr>
<tr>
<td>SLC39A10</td>
<td>USP22</td>
<td>NR1D2</td>
<td>PJA2</td>
<td>EFNA1</td>
<td>ZC3HAV1L</td>
</tr>
<tr>
<td>GLS</td>
<td>NCOR1</td>
<td>GMFB</td>
<td>ATL2</td>
<td>SNX27</td>
<td>EPHA1</td>
</tr>
<tr>
<td>INO80D</td>
<td>TAOK1</td>
<td>SLITRK6</td>
<td>MAP4K3</td>
<td>PIP5K1A</td>
<td>HOXA1</td>
</tr>
<tr>
<td>CAMSAP2</td>
<td>MED13</td>
<td>DOCK9</td>
<td>NR3C1</td>
<td>OSTC</td>
<td>CNBP</td>
</tr>
<tr>
<td>TROVE2</td>
<td>ACLY</td>
<td>AW549877</td>
<td>YIPF5</td>
<td>IMPAD1</td>
<td>ZFP248</td>
</tr>
<tr>
<td>LAMC1</td>
<td>GPATCH8</td>
<td>YWHAZ</td>
<td>ADRB2</td>
<td>PNRC1</td>
<td>MICAL3</td>
</tr>
<tr>
<td>PBX1</td>
<td>CCDC47</td>
<td>ASAP1</td>
<td>PPP6R3</td>
<td>TOPORS</td>
<td>CD9</td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>----------</td>
</tr>
<tr>
<td>HSD17B7</td>
<td>SRSF2</td>
<td>PTK2</td>
<td>RAB11FIP2</td>
<td>ALG2</td>
<td>CCND2</td>
</tr>
<tr>
<td>HNRNPU</td>
<td>DNAJB9</td>
<td>TUBA1B</td>
<td>BRD3</td>
<td>CTNNAL1</td>
<td>ERGIC2</td>
</tr>
<tr>
<td>LBR</td>
<td>SPTSSA</td>
<td>FIGNL2</td>
<td>GTF3C4</td>
<td>NFIB</td>
<td>AU018091</td>
</tr>
<tr>
<td>MFSD7B</td>
<td>CFL2</td>
<td>GSPT1</td>
<td>GGTA1</td>
<td>CTP5</td>
<td>MBOAT7</td>
</tr>
<tr>
<td>UTRN</td>
<td>SOS2</td>
<td>MCM4</td>
<td>STRBP</td>
<td>C77080</td>
<td>PVRL2</td>
</tr>
<tr>
<td>MEX3D</td>
<td>PTPN21</td>
<td>PLCXD2</td>
<td>RND3</td>
<td>YTHDF2</td>
<td>MCTP2</td>
</tr>
<tr>
<td>CPSF6</td>
<td>TFAP2A</td>
<td>CRYBG3</td>
<td>ATF2</td>
<td>SEPN1</td>
<td>IDH2</td>
</tr>
<tr>
<td>NUP107</td>
<td>KIF13A</td>
<td>QK</td>
<td>LNP</td>
<td>UBA6</td>
<td>NUP98</td>
</tr>
<tr>
<td>XPOT</td>
<td>CLTB</td>
<td>MRPL18</td>
<td>COP52</td>
<td>ZFP644</td>
<td>ATE1</td>
</tr>
<tr>
<td>PURB</td>
<td>CAST</td>
<td>BRD4</td>
<td>JAG1</td>
<td>EP400</td>
<td>OAT</td>
</tr>
<tr>
<td>CCND1</td>
<td>USP27X</td>
<td>ANO6</td>
<td>CHORDC1</td>
<td>TGFB3</td>
<td>GNAI3</td>
</tr>
<tr>
<td>U8XN8</td>
<td>CUL4B</td>
<td>SNAI2</td>
<td>AMOTL2</td>
<td>UBA3</td>
<td>CPNE3</td>
</tr>
<tr>
<td>GAB1</td>
<td>RAP2C</td>
<td>LRRCS8</td>
<td>DUSP7</td>
<td>PRSS23</td>
<td>KLF4</td>
</tr>
<tr>
<td>DYNC1LI2</td>
<td>MBNL3</td>
<td>USP25</td>
<td>XIAP</td>
<td>SERPINH1</td>
<td>JUN</td>
</tr>
<tr>
<td>CFD1P1</td>
<td>MMGT1</td>
<td>CXADR</td>
<td>AHCTF1</td>
<td>EIF4G2</td>
<td>FAM126A</td>
</tr>
<tr>
<td>2310022B05RIK</td>
<td>PLS3</td>
<td>CREBKF</td>
<td>REEP3</td>
<td>FGFR2</td>
<td>ATAD2B</td>
</tr>
<tr>
<td>EGN1</td>
<td>TSPAN6</td>
<td>MATR3</td>
<td>LATS2</td>
<td>KCNQ1OT1</td>
<td>ARHGAP5</td>
</tr>
<tr>
<td>ARHGAP42</td>
<td>VAMP4</td>
<td>DCP2</td>
<td>SLC38A2</td>
<td>TGFB2</td>
<td>MGAT2</td>
</tr>
<tr>
<td>ZFP266</td>
<td>PTPN14</td>
<td>CSNK1G3</td>
<td>CLDN1</td>
<td>CCNG2</td>
<td>ESYT2</td>
</tr>
<tr>
<td>PAFAH1B2</td>
<td>LATS1</td>
<td>AHNAK</td>
<td>ATP13A3</td>
<td>MET</td>
<td>DSP</td>
</tr>
<tr>
<td>CUL5</td>
<td>PERP</td>
<td>ARL5B</td>
<td>BRWD1</td>
<td>TNRC6A</td>
<td>119002N15RIK</td>
</tr>
<tr>
<td>FEM1B</td>
<td>PANK3</td>
<td>TSC22D2</td>
<td>NDST1</td>
<td>OTUD4</td>
<td>PIK3CB</td>
</tr>
<tr>
<td>POLR2M</td>
<td>AFF4</td>
<td>MBNL1</td>
<td>RANBP6</td>
<td>TMEM123</td>
<td>RBM6</td>
</tr>
<tr>
<td>TMOD3</td>
<td>VEZF1</td>
<td>DENND2C</td>
<td>SH3PX2A</td>
<td>DAG1</td>
<td></td>
</tr>
<tr>
<td>TBC1D2B</td>
<td>SPAG9</td>
<td>PPP1CB</td>
<td>HIPK3</td>
<td>CDCP1</td>
<td></td>
</tr>
<tr>
<td>1700066M21RIK</td>
<td>TTC39B</td>
<td>E030024N20RIK</td>
<td>CAPNS2</td>
<td>APP</td>
<td>LEPROTL1</td>
</tr>
<tr>
<td>---------------</td>
<td>--------</td>
<td>--------------</td>
<td>--------</td>
<td>-----</td>
<td>----------</td>
</tr>
<tr>
<td>MAP2</td>
<td>DDI2</td>
<td>ABHD17B</td>
<td>ARL2BP</td>
<td>DUSP1</td>
<td>ASAH1</td>
</tr>
<tr>
<td>ACSL3</td>
<td>GATAD1</td>
<td>RIC1</td>
<td>CBFB</td>
<td>FEZ2</td>
<td>MSMO1</td>
</tr>
<tr>
<td>FAM129A</td>
<td>SERPIN1</td>
<td>TNKS2</td>
<td>ZFHX3</td>
<td>ZFP36L2</td>
<td>N4BP1</td>
</tr>
<tr>
<td>PSAP</td>
<td>CRTC3</td>
<td>MAR5</td>
<td>RPL13</td>
<td>CCNY</td>
<td>FHOD1</td>
</tr>
<tr>
<td>TMCC3</td>
<td>TSKU</td>
<td>WBP1L</td>
<td>AFG3L1</td>
<td>DSC2</td>
<td>KANK2</td>
</tr>
<tr>
<td>PITPNA</td>
<td>ITPRIPL2</td>
<td>GSTO1</td>
<td>NRP1</td>
<td>TMED7</td>
<td>1810064F22RIK</td>
</tr>
<tr>
<td>SYPL</td>
<td>CPE</td>
<td>USP6NL</td>
<td>ZBTB44</td>
<td>RTN3</td>
<td>DPY19L1</td>
</tr>
<tr>
<td>FOXG1</td>
<td>SLC7A5</td>
<td>TMEM203</td>
<td>31100390RIK</td>
<td>ERMP1</td>
<td>SC5D</td>
</tr>
<tr>
<td>FAM65B</td>
<td>SULF1</td>
<td>ANGPTL2</td>
<td>PLET1</td>
<td>ASAH2</td>
<td>CBL</td>
</tr>
<tr>
<td>FUT11</td>
<td>PDC3L</td>
<td>FMNL2</td>
<td>ULK3</td>
<td>ACTR1A</td>
<td>TSPAN3</td>
</tr>
<tr>
<td>MMP14</td>
<td>NRP2</td>
<td>DHR59</td>
<td>DPP8</td>
<td>ZDHHC12</td>
<td>PEAK1</td>
</tr>
<tr>
<td>MBNL2</td>
<td>TMEM185B</td>
<td>ZAK</td>
<td>NTSE</td>
<td>STAM2</td>
<td>ATP1B3</td>
</tr>
<tr>
<td>PIM3</td>
<td>KIF21B</td>
<td>SSFA2</td>
<td>PLOD2</td>
<td>CD44</td>
<td>DNAJC13</td>
</tr>
<tr>
<td>BACH1</td>
<td>ARPC5</td>
<td>SPRED1</td>
<td>CMTM6</td>
<td>IL1A</td>
<td>ACSL4</td>
</tr>
<tr>
<td>TNFRSF21</td>
<td>PVRL4</td>
<td>BCL2L11</td>
<td>GPKOW</td>
<td>4930402H24RIK</td>
<td>DDX3Y</td>
</tr>
<tr>
<td>TCF4</td>
<td>VSIG8</td>
<td>TMEM87B</td>
<td>SLC35A2</td>
<td>SLC23A2</td>
<td>PTGS2</td>
</tr>
<tr>
<td>MBD1</td>
<td>DES12</td>
<td>RPN2</td>
<td>ATP6AP2</td>
<td>TMX4</td>
<td>SLC19A2</td>
</tr>
<tr>
<td>CLCF1</td>
<td>MTHFD1L</td>
<td>SRC</td>
<td>DIAPH2</td>
<td>CST3</td>
<td>OSTM1</td>
</tr>
<tr>
<td>SNHG1</td>
<td>CITED2</td>
<td>RALGAPB</td>
<td>TMEM164</td>
<td>PLAGL2</td>
<td>PTTG1IP</td>
</tr>
<tr>
<td>ZFAND5</td>
<td>RUFY2</td>
<td>ARFGEF2</td>
<td>KDM5D</td>
<td>SULF2</td>
<td>UBE2G2</td>
</tr>
<tr>
<td>PGAM1</td>
<td>CCDC6</td>
<td>RBM38</td>
<td>ARFGEF1</td>
<td>TMEM189</td>
<td>MMD</td>
</tr>
<tr>
<td>MRPS52</td>
<td>IPMK</td>
<td>GID8</td>
<td>TRAK2</td>
<td>ZFP217</td>
<td>CTAGE5</td>
</tr>
<tr>
<td>KBTBD4</td>
<td>EEA1</td>
<td>TBL1XR1</td>
<td>IDH1</td>
<td>TAF4</td>
<td>WRNIP1</td>
</tr>
<tr>
<td>TRP53INP2</td>
<td>OSBPL8</td>
<td>JADE1</td>
<td>ABCA12</td>
<td>FNDC3B</td>
<td>PLAU</td>
</tr>
<tr>
<td>CEBBP</td>
<td>TMIM4</td>
<td>SELT</td>
<td>SRGAP2</td>
<td>SLC7A11</td>
<td>GOLPH3</td>
</tr>
<tr>
<td>NOTCH2</td>
<td>ANKR52D</td>
<td>GM21949</td>
<td>IER5</td>
<td>FNIP2</td>
<td>LRCH3</td>
</tr>
<tr>
<td>RECK</td>
<td>PYM1</td>
<td>SHC1</td>
<td>XPR1</td>
<td>ADGRL2</td>
<td>PKD1</td>
</tr>
<tr>
<td>TGFB1</td>
<td>PEL1</td>
<td>CERS2</td>
<td>TNFAIP3</td>
<td>RABG5GTB</td>
<td>CDKN1A</td>
</tr>
<tr>
<td>2610528J11RIK</td>
<td>FBXW11</td>
<td>TXNIP</td>
<td>PRDM1</td>
<td>ORC3</td>
<td>RN45S</td>
</tr>
<tr>
<td>RRAGC</td>
<td>ALKBH5</td>
<td>RBSN1</td>
<td>LIN28B</td>
<td>B4GALT1</td>
<td>IMPACT</td>
</tr>
<tr>
<td>CXCL5</td>
<td>ADORA2B</td>
<td>PNISR</td>
<td>MKNK2</td>
<td>EPB41L4B</td>
<td>SMAD7</td>
</tr>
<tr>
<td>CXCL3</td>
<td>PRKAR1A</td>
<td>DNAJ1</td>
<td>ELK3</td>
<td>CPT2</td>
<td>SCD2</td>
</tr>
<tr>
<td>MLXIP</td>
<td>MBOAT2</td>
<td>RUSC2</td>
<td>RAB21</td>
<td>ZMPSTE24</td>
<td>LRRCA8</td>
</tr>
<tr>
<td>TMEM106B</td>
<td>KIDINS220</td>
<td>SLC44A1</td>
<td>LRP1</td>
<td>RCAN3</td>
<td>CDS2</td>
</tr>
<tr>
<td>TES</td>
<td>DOCK4</td>
<td>TMEM38B</td>
<td>DRG1</td>
<td>KIF1B</td>
<td>LHFP</td>
</tr>
<tr>
<td>IRAK2</td>
<td>IMMP2L</td>
<td>MTAP</td>
<td>SLC35E4</td>
<td>CROT</td>
<td>TRP53INP1</td>
</tr>
<tr>
<td>LAMP1</td>
<td>FAM179B</td>
<td>DHCR24</td>
<td>B3GNT2</td>
<td>PHTF2</td>
<td>UGCg</td>
</tr>
<tr>
<td>ADAM10</td>
<td>ARF6</td>
<td>LRP8</td>
<td>JADE2</td>
<td>WDR4</td>
<td>GIGYF1</td>
</tr>
<tr>
<td>Gene 1</td>
<td>Gene 2</td>
<td>Gene 3</td>
<td>Gene 4</td>
<td>Gene 5</td>
<td>Gene 6</td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>GOLGA4</td>
<td>JKAMP</td>
<td>INSIG1</td>
<td>ULK2</td>
<td>GM3414</td>
<td>FUS</td>
</tr>
<tr>
<td>5730559C18RIK</td>
<td>FOS</td>
<td>SLIT2</td>
<td>CLTC</td>
<td>CLOCK</td>
<td>SEP7</td>
</tr>
<tr>
<td>UHMK1</td>
<td>GPATCH2L</td>
<td>PCDH7</td>
<td>NOG</td>
<td>HOPX</td>
<td>CADM1</td>
</tr>
<tr>
<td>FBXO28</td>
<td>CPSF2</td>
<td>REST</td>
<td>RAB5C</td>
<td>RSRC2</td>
<td>ATP6AP1</td>
</tr>
<tr>
<td>NUDT4</td>
<td>SFXN1</td>
<td>SEP11</td>
<td>VAT1</td>
<td>STX2</td>
<td>DDIT4</td>
</tr>
<tr>
<td>MDM2</td>
<td>IDNK</td>
<td>ATP6V0A2</td>
<td>DUSP3</td>
<td>CUX1</td>
<td>CKAP4</td>
</tr>
<tr>
<td>HMGA2</td>
<td>DAPK1</td>
<td>BRI3BP</td>
<td>CBX4</td>
<td>FLT1</td>
<td>SLC25A3</td>
</tr>
<tr>
<td>KREMEN1</td>
<td>PLK2</td>
<td>VKORC1L1</td>
<td>ADAM17</td>
<td>UBL3</td>
<td>RAP1B</td>
</tr>
<tr>
<td>TNS3</td>
<td>ARHGEF3</td>
<td>GATSL2</td>
<td>KLHL28</td>
<td>UBE2H</td>
<td>CANX</td>
</tr>
<tr>
<td>NFE2L1</td>
<td>FAM208A</td>
<td>CAV2</td>
<td>ATXN1</td>
<td>TMEM209</td>
<td>ARHGDOA</td>
</tr>
<tr>
<td>RHOB</td>
<td>WAPL</td>
<td>UBN2</td>
<td>ZFP729A</td>
<td>KDM7A</td>
<td>ZSCAN26</td>
</tr>
<tr>
<td>ELMSAN1</td>
<td>ABHD4</td>
<td>AVL9</td>
<td>SCAMP1</td>
<td>RPIA</td>
<td>RANBP9</td>
</tr>
<tr>
<td>ISCA1</td>
<td>CCDC25</td>
<td>EVA1A</td>
<td>HMGCR</td>
<td>RMND5A</td>
<td>AGO2</td>
</tr>
<tr>
<td>CTSL</td>
<td>TSC22D1</td>
<td>ARL8B</td>
<td>ERO1L</td>
<td>TGOLN1</td>
<td>GXYLT1</td>
</tr>
<tr>
<td>MAP1B</td>
<td>KBTBD7</td>
<td>BRK1</td>
<td>CNIH1</td>
<td>MXD1</td>
<td>PPT2</td>
</tr>
<tr>
<td>ZSWIM6</td>
<td>LAPT4B</td>
<td>MAR8</td>
<td>HMBOX1</td>
<td>ADIPO2</td>
<td>VEGFA</td>
</tr>
<tr>
<td>SUB1</td>
<td>DIP28</td>
<td>IL17RA</td>
<td>SKP2</td>
<td>ORL1</td>
<td>THBD</td>
</tr>
<tr>
<td>REEP5</td>
<td>MFSD5</td>
<td>AEBP2</td>
<td>BASP1</td>
<td>PTHLH</td>
<td>BLCAP</td>
</tr>
<tr>
<td>HBE GF</td>
<td>COX17</td>
<td>ETNK1</td>
<td>MAR6</td>
<td>AI987944</td>
<td>ZDHHC21</td>
</tr>
<tr>
<td>NARS</td>
<td>MAPK14</td>
<td>KCNC3</td>
<td>SBF1</td>
<td>FANCF</td>
<td>TMED5</td>
</tr>
<tr>
<td>PLGRKT</td>
<td>FOXN2</td>
<td>IQCK</td>
<td>SLC38A1</td>
<td>CEMIP</td>
<td>CLDN4</td>
</tr>
<tr>
<td>ATAD1</td>
<td>BIN1</td>
<td>PLEKH1</td>
<td>KRT6A</td>
<td>ZFAND6</td>
<td>B630005N14RIK</td>
</tr>
<tr>
<td>NF5</td>
<td>KDM3B</td>
<td>AGPAT5</td>
<td>ZFP597</td>
<td>RNF169</td>
<td>ELMOD2</td>
</tr>
<tr>
<td>PMEPA1</td>
<td>SPINK5</td>
<td>ATP6V1B2</td>
<td>ACAP2</td>
<td>TPP1</td>
<td>IRF2BP2</td>
</tr>
<tr>
<td>MFSD14A</td>
<td>NAPG</td>
<td>MRPL34</td>
<td>ATP6V1A</td>
<td>CTBP2</td>
<td>KMT2A</td>
</tr>
<tr>
<td>METAP1</td>
<td>SUV420H1</td>
<td>CNEP1R1</td>
<td>SLC35A5</td>
<td>ADAM12</td>
<td>RASA2</td>
</tr>
<tr>
<td>TXN1</td>
<td>B4GAT1</td>
<td>CHD9</td>
<td>BBX</td>
<td>ELAVL1</td>
<td>PDCD6IP</td>
</tr>
</tbody>
</table>
Table A5: List of miR-200 family targets identified with CLEAR-CLIP and P4.5 RNA-seq

<table>
<thead>
<tr>
<th>SPTSSA</th>
<th>HOPX</th>
<th>TMEM167</th>
<th>MMGT1</th>
<th>STX12</th>
<th>ATAD2B</th>
</tr>
</thead>
<tbody>
<tr>
<td>OAT</td>
<td>PCMTD1</td>
<td>ATP6V1A</td>
<td>MSN</td>
<td>TMEM245</td>
<td>TMBIM4</td>
</tr>
<tr>
<td>SCD1</td>
<td>PDLIM5</td>
<td>ESYT2</td>
<td>RSRC2</td>
<td>SNHG1</td>
<td>RTN3</td>
</tr>
<tr>
<td>UBA6</td>
<td>ADAM17</td>
<td>CCDC47</td>
<td>ARL2BP</td>
<td>ASAH1</td>
<td>ATP6AP2</td>
</tr>
<tr>
<td>ELOVL6</td>
<td>ACY</td>
<td>FAM135A</td>
<td>PGAM1</td>
<td>TMEM43</td>
<td>GPKOW</td>
</tr>
<tr>
<td>CNIH4</td>
<td>ZFP185</td>
<td>CLTB</td>
<td>CERS6</td>
<td>FLT1</td>
<td>CLDN12</td>
</tr>
<tr>
<td>SMIM15</td>
<td>RNF11</td>
<td>CAMSAP2</td>
<td>GABPA</td>
<td>RUFY2</td>
<td>CIAPIN1</td>
</tr>
<tr>
<td>SERPINB2</td>
<td>CRYBG3</td>
<td>YES1</td>
<td>ROCK2</td>
<td>PFKFB2</td>
<td>TCF4</td>
</tr>
<tr>
<td>GNAI3</td>
<td>DEPDC1B</td>
<td>MARCKS</td>
<td>TRP53INP2</td>
<td>GSPT1</td>
<td>SSR3</td>
</tr>
<tr>
<td>LHFP</td>
<td>ZFP608</td>
<td>CST3</td>
<td>SNRN4P8</td>
<td>IMPAD1</td>
<td>ULK3</td>
</tr>
<tr>
<td>IMMPL2</td>
<td>MGAT2</td>
<td>AMOTL2</td>
<td>LAPT4M4B</td>
<td>FBXW7</td>
<td>ARIH2</td>
</tr>
<tr>
<td>FAM136A</td>
<td>ADM</td>
<td>FST</td>
<td>LAPT4M4A</td>
<td>CCND1</td>
<td>GAS5</td>
</tr>
<tr>
<td>CFL2</td>
<td>MAF</td>
<td>PITPNA</td>
<td>CPNE3</td>
<td>ACVR2A</td>
<td>USP47</td>
</tr>
<tr>
<td>CAST</td>
<td>AGFG1</td>
<td>ASAP1</td>
<td>RAB6A</td>
<td>FANCF</td>
<td>SRGAP2</td>
</tr>
<tr>
<td>FOXC1</td>
<td>TMEM38B</td>
<td>SLC39A10</td>
<td>ZCCHC10</td>
<td>LBR</td>
<td>PLEKHA3</td>
</tr>
<tr>
<td>UBE2G2</td>
<td>BRK1</td>
<td>PRDM1</td>
<td>CBFB</td>
<td>2610301B2ORIK</td>
<td>CAPRIN1</td>
</tr>
<tr>
<td>PRDX2</td>
<td>ATL2</td>
<td>TOB1</td>
<td>DENND6A</td>
<td>CPE</td>
<td>TSC22D1</td>
</tr>
<tr>
<td>ATP6V0E</td>
<td>BLCAP</td>
<td>UGCG</td>
<td>SDF2</td>
<td>DUSP3</td>
<td>UBE2H</td>
</tr>
<tr>
<td>PTPN21</td>
<td>MBOAT2</td>
<td>AACS</td>
<td>LARS2</td>
<td>B3GNT2</td>
<td>DSC3</td>
</tr>
<tr>
<td>HIPK3</td>
<td>BICD2</td>
<td>MEI4</td>
<td>SERINC1</td>
<td>NR2C2</td>
<td>KANBN1</td>
</tr>
<tr>
<td>MRPL34</td>
<td>PPFIBP1</td>
<td>HMGC</td>
<td>CCNE2</td>
<td>TMEM194</td>
<td>VAT1</td>
</tr>
<tr>
<td>EGLN3</td>
<td>DHCR24</td>
<td>PLGRKT</td>
<td>BIN1</td>
<td>IRF2BP2</td>
<td>TANC1</td>
</tr>
<tr>
<td>ZFAND6</td>
<td>ATPIB3</td>
<td>RAB7</td>
<td>RANBP6</td>
<td>MFS07B</td>
<td>PLAU</td>
</tr>
<tr>
<td>MMD</td>
<td>MRPL18</td>
<td>ZFP266</td>
<td>YIF5</td>
<td>TRP53INP1</td>
<td>PVR2</td>
</tr>
<tr>
<td>RAP2C</td>
<td>SEC24A</td>
<td>COPS2</td>
<td>PDIK1L</td>
<td>EPT1</td>
<td>ITSN2</td>
</tr>
<tr>
<td>AGPAT5</td>
<td>CUL5</td>
<td>BLOC1S4</td>
<td>MPLKIP</td>
<td>DPP57L1</td>
<td>ZFPB71</td>
</tr>
<tr>
<td>OLR1</td>
<td>SMC2</td>
<td>EGFR</td>
<td>HMGC</td>
<td>GCLC</td>
<td>ARFGEF2</td>
</tr>
<tr>
<td>DR1</td>
<td>KRTDAP</td>
<td>VAMP4</td>
<td>CEP192</td>
<td>ANP32E</td>
<td>BAP1</td>
</tr>
<tr>
<td>SMIM13</td>
<td>FGFBP1</td>
<td>EIF4G2</td>
<td>UBL3</td>
<td>CUX1</td>
<td>NDUFS1</td>
</tr>
<tr>
<td>PLCXD2</td>
<td>RAP2B</td>
<td>CNOT6</td>
<td>CPEB2</td>
<td>EML4</td>
<td>NUP107</td>
</tr>
<tr>
<td>CTPS</td>
<td>YWHAB</td>
<td>ITGB6</td>
<td>FLRT3</td>
<td>ZFP655</td>
<td>LIN7C</td>
</tr>
<tr>
<td>NAA50</td>
<td>MTFR1</td>
<td>VSG8</td>
<td>CERS3</td>
<td>LNP</td>
<td>FMR1</td>
</tr>
<tr>
<td>RRM2</td>
<td>TNFAIP3</td>
<td>IDS</td>
<td>CLCF1</td>
<td>REEP5</td>
<td>DPP8</td>
</tr>
<tr>
<td>SEC23A</td>
<td>GID8</td>
<td>CNBP</td>
<td>PPP1CB</td>
<td>ATP7A</td>
<td>BRWD1</td>
</tr>
<tr>
<td>TMEM185B</td>
<td>XPOT</td>
<td>DENND2C</td>
<td>MBTPS2</td>
<td>YWHA2</td>
<td>KIF1B</td>
</tr>
<tr>
<td>COMMD3</td>
<td>TSC22D2</td>
<td>MBNL3</td>
<td>MTF2</td>
<td>AIF1</td>
<td>ATP11C</td>
</tr>
<tr>
<td>SLC16A1</td>
<td>TMEM97</td>
<td>TADA1</td>
<td>PPT2</td>
<td>MAP3K1</td>
<td>BMPR1A</td>
</tr>
<tr>
<td>TMED5</td>
<td>IL17RC</td>
<td>DESI2</td>
<td>EDN1</td>
<td>GRPHEL2</td>
<td>PANK3</td>
</tr>
<tr>
<td>TMEM164</td>
<td>CDC42</td>
<td>CLDN1</td>
<td>SLK</td>
<td>RRAGC</td>
<td>DOCK4</td>
</tr>
<tr>
<td>PTGS2</td>
<td>MBNL1</td>
<td>TMEM65</td>
<td>SH3BGRL2</td>
<td>NRP2</td>
<td>RND3</td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>--------</td>
<td>----------</td>
<td>------</td>
<td>---------</td>
</tr>
<tr>
<td>BLZF1</td>
<td>ZDHHC21</td>
<td>ARF6</td>
<td>KPNA4</td>
<td>SYPL</td>
<td>A230046K03Rik</td>
</tr>
<tr>
<td>MAP2</td>
<td>TMEM168</td>
<td>STEAP2</td>
<td>FEZ2</td>
<td>NR3C1</td>
<td>LIN28B</td>
</tr>
<tr>
<td>CHURC1</td>
<td>PLEKHA8</td>
<td>AW549877</td>
<td>IST1</td>
<td>SLC39A14</td>
<td>N4BP2L2</td>
</tr>
<tr>
<td>IMPACT</td>
<td>COPS8</td>
<td>SLC35A4</td>
<td>LATS2</td>
<td>PLK2</td>
<td>KBTBD2</td>
</tr>
<tr>
<td>ADIPOR2</td>
<td>AFF4</td>
<td>PUM2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>