UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

Gene Regulatory Mechanisms in Epithelial Specification and Function

Permalink

https://escholarship.org/uc/item/3wb6v4pd

Author

Klein, Rachel Herndon

Publication Date

2015

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NoDerivatives License, availalbe at https://creativecommons.org/licenses/by-nd/4.0/

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA,
IRVINE

Gene Regulatory Mechanisms in Epithelial Specification and Function

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Biomedical Sciences

by

Rachel Herndon Klein

Dissertation Committee:
Professor Bogi Andersen, M.D., Chair
Professor Xing Dai, Ph.D.
Professor Anand Ganesan, M.D.
Professor Ali Mortazavi, Ph.D
Professor Kyoko Yokomori, Ph.D

2015

© 2015 Rachel Herndon Klein

DEDICATION

To

My parents, my sisters, my husband, and my friends

for your love and support,

and to Ben

with all my love.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I have so many people to thank for their support during my graduate school career. First, and foremost, I would like to thank Dr. Andersen for his mentorship, his patience, his invaluable insights and guidance while developing my project, and his willingness to let me determine the direction of my research based on the observations I found most interesting. Thanks to his mentorship, I feel like I have grown tremendously as a scientist. I would also like to thank my committee for their guidance and suggestions, which I feel have greatly improved my project.

I would like to say thank you my parents for their unconditional love and support in everything I do. An entire thesis devoted to the subject would not have space to explain how great they are. Finishing this project would not have been possible without their encouragement and support.

I am also forever grateful for the support day-in and day-out from my husband, who has always believed in me, and has been with me through the full grad-school experience. I am so lucky to have you.

I am also thankful for all my graduate school friends and lab-mates for their support and camaraderie. For Amy, who got me started in the Andersen Lab, and gave me so much advice and guidance over the years. For Michael, Will, and Hsiang for their valuable suggestions for improving my research. For Eric, Ghaidaa, Elyse, Jeff, and Bob for their friendship and assistance with my project.

CURRICULUM VITAE

Rachel Herndon Klein

EDUCATION:

University of California, Irvine
PhD in Biomedical Sciences
2015

University of California, Davis
Bachelor of Science in Genetics**,** with Honors
2008
GPA: 3.77/ 4.00

PUBLICATIONS:

Klein RH, Andersen B. Dynamic networking for epidermal differentiation. (2015) Invited Preview. Dev Cell. 32(6):661-2.

Klein RH, Stephens DN, Salmans ML, Andersen B. (manuscript in preparation) Cofactors of LIM domain (CLIM) proteins regulate corneal epithelial progenitor cell function through noncoding RNA H19.

Klein RH, Ceglia N, Lin Z, Chen J, Selwan L, Baldi P, Andersen B (manuscript in preparation) GRHL3 acts through super enhancers to regulate epidermal keratinocytes in migration and differentiation.

Klein RH, Ceglia N, Hu WK, Nguyen T, Ngo J, Baldi P, Andersen B (manuscript in preparation). KLF7regulates the corneal epithelial progenitor cell state acting antagonistically to KLF4.

Hopkin A, Gordon W, **Klein RH**, Espitia F, Daily K, Zeller M, Baldi P, and Andersen B. (2012) GRHL3/GET1 and Trithorax group members collaborate to activate the epidermal progenitor differentiation program. PLoS Genetics. 8(7):e1002829. PMCID: PMC3400561

Stephens DN, **Klein RH,** Salmans ML, Gordon W, Ho H, and Andersen B. (2013) A transcriptome profile of cornea over the lifetime of the mouse: Identification of EHF as a transcriptional regulator of the cornea epithelium differentiation program. J Biol Chem. 29;288(48):34304-24. PMCID: PMC3843046

Gordon W, Zeller M, **Klein RH**, Ho H, Baldi P, Andersen B. (2014) A GRHL3-regulated repair pathway suppresses immune-mediated epidermal hyperplasia. J Clin Invest. 124(12):5205-18. PMID:25347468

TEACHING EXPERIENCE:

Teaching Assistant- Cell Biology, UC Irvine, 2013-2014
Guest Lecturer- Cell Biology, UC Irvine, September 2013

PRESENTATIONS:

Society of Investigative Dermatology Annual Meeting – Phoenix, May 2011, Epidermal Structure and Function Session Speaker
"Combined biological and computational approaches to understand the role of Get1/Grhl3 in epidermal differentiation"

Gordon Conference in Cornea Biology- Ventura, March 2012, Poster Presentation
"Genome-wide identification of enhancers in human corneal epithelial cells provides insights into coordinated gene regulation during corneal development"

Society of Investigative Dermatology Annual Meeting – Raleigh, May 2012, Epidermal Structure and Function Session Speaker
"The Trithorax complex promotes epidermal keratinocyte differentiation in part through GRHL3/GET1- mediated recruitment"

AWARDS AND HONORS:
UC Davis Departmental Citation for Outstanding Performance, Genetics, 2008
UC Davis Dean's List 2005-2008
Phi Beta Kappa

FELLOWSHIPS:

NLM Biomedical Informatics Training Grant, UC Irvine (2011- 2013)
NLM NRSA T15LM00744

RESEARCH EXPERIENCE:

Graduate Student in the Department of Biological Chemistry, UC Irvine School of Medicine 2009-2015
Studied mechanisms of gene regulation during epithelial differentiation using ChIP-Seq.

Lab Manager, Pulmonary and Critical Care, Department of Internal Medicine, UC Davis School of Medicine, 2008-2009
Characterized the changes in signal transduction that occur in human lung epithelial cells under oxidative stress.

ABSTRACT OF THE DISSERTATION

A Genomics Approach to Studying Transcriptional Regulation in Epithelial Development

By

Rachel Herndon Klein

Doctor of Philosophy in Biomedical Sciences

University of California, Irvine, 2015

Professor Bogi Andersen, Chair

As epithelial tissues, the cornea epithelium and epidermis share many functional, morphological and regulatory similarities. They both serve important barrier functions, protecting the organism and underlying tissues from the external environment, and undergo constant cell turnover, requiring a precise balance of proliferation and differentiation to maintain tissue integrity. To accomplish this balancing act, they make use of similar mechanisms, and integrate information from pre-existing regulatory domains with the expression and actions of lineage-specific transcriptional regulators. The importance of transcription factor binding to distal regulatory regions is demonstrated by our finding that SNPs linked to disease in cornea or epidermis cause alterations in enhancing ability, and disrupt transcription factor motifs, a potential mechanism underlying the observed increased risk for disease. GRHL3 is an important regulator of both epidermal keratinocyte differentiation and migration; we have identified unique mechanisms of action for this single factor under the two different physiological conditions: GRHL3 associates with enhancers and super enhancers to activate structural

and barrier genes in differentiating keratinocytes, and with promoters to repress genes involved in inhibition of migration. Selectivity may come about through differential association with co-factors during migration and differentiation. Based on opposing expression patterns for KLF4 and KLF7 in cornea epithelium, we characterize and define a role for KLF7 in maintaining corneal epithelial progenitor cells in an undifferentiated state, acting antagonistically to KLF4, a pro-differentiation factor. Our work with CLIMs points to the importance of these factors in corneal epithelium development and maintenance, in part through regulation of noncoding RNA H19.

Chapter1. Introduction

Epithelial Tissues

Epithelial tissues play crucial roles in modulating the interactions of an organism with the external environment. Based on the requirements of a given tissue, this diverse group of cells can: absorb and transport nutrients to underlying tissues; detect stimuli in the environment and transmit this information internally; act as a structural, mechanical, and permeability barrier; and secrete mucous, enzymes, and other molecules into the ducts of epithelial glands. Examples of such tissues include the epidermis, cornea, intestine and lung epithelium, as well as the mammary gland.

Cells in epithelial tissues are organized into one of several patterns to meet the unique functional requirements of a given tissue. Simple epithelia, which are a single cell layer thick, are commonly found in tissues that specialize in absorption of gases or nutrients, such as the lung and intestinal epithelium. Epithelial tissues that must withstand large amounts of mechanical stress are often structured as stratified squamous epithelium. Both cornea epithelium and epidermis, which serve crucial barrier functions, protecting underlying tissues from UV damage, toxins, invasion of pathogens, and dehydration, are examples of such stratified squamous epithelial tissues. In this organization, cells are arranged in layers, becoming progressively flattened as they move outward in the tissue.

Despite their diverse functions and structural organization, epithelial tissues share a number of features related to their common roles. Cell adhesion is a defining and crucial aspect of epithelial cells, which must adhere to each other to serve as barriers. A number of junctional complexes including desmosomes, hemidesmosomes, adherens junctions, and gap junctions are produced by epithelial cells, and allow different epithelial tissues to customize their permeability

to suit their specific combination of absorption, and barrier functions. Many epithelial tissues also produce large amounts of intermediate filaments, which can connect to the nucleus as well as the plasma membrane at desmosomal cell junctions, giving the cell structure and providing strength to withstand high levels of mechanical stress. Additionally, due to their extensive contact with the external environment and potential for damage that results, most epithelial tissues must continually replenish themselves throughout the lifetime of the organism, and must be able to rapidly heal wounds. Such tissue homeostasis requires a precise balance between proliferation and differentiation, suggesting a high level of transcriptional control involved in this process.

The epidermis and cornea epithelium are two examples of stratified squamous epithelial tissues that both arise from the embryonic ectoderm. However, their execution of epithelial functions are crucially distinct due to the unique requirements and physiologies of each tissue. While much work has been done to morphologically characterize corneal and epidermal development and tissue homeostasis, mechanisms of transcriptional regulation in effect in each tissue are not well defined.

Epidermal Structure

The primary function of the epidermis is to provide a rigid barrier between the organism and the external environment; it prevents pathogens, toxins, and UV rays from penetrating and reaching underlying tissues, and also prevents deydration by blocking water loss from underlying tissues to the external environment. The epidermis is primarily composed of layers of keratinocytes; each layer plays a vital role in the formation of the epidermal barrier. Cells in the basal layer are proliferative and serve to replenish the tissue through horizontal divisions, giving rise to

additional basal keratinocytes [1]. Asymmetrical divisions of basal cells produces one basal progenitor cell, and one suprabasal cell, which will go on to differentiate as it moves upward in the epidermis [1]. Differentiation involves the progressive activation of genes required for cell adhesion (including claudins and occludins), genes involved in forming the rigid structure of the cornified envelope, a crucial barrier component (including transglutaminases, loricrin, and late cornified envelope proteins), as well as genes encoding enzymes involved in lipid synthesis, which help to seal the barrier [2].

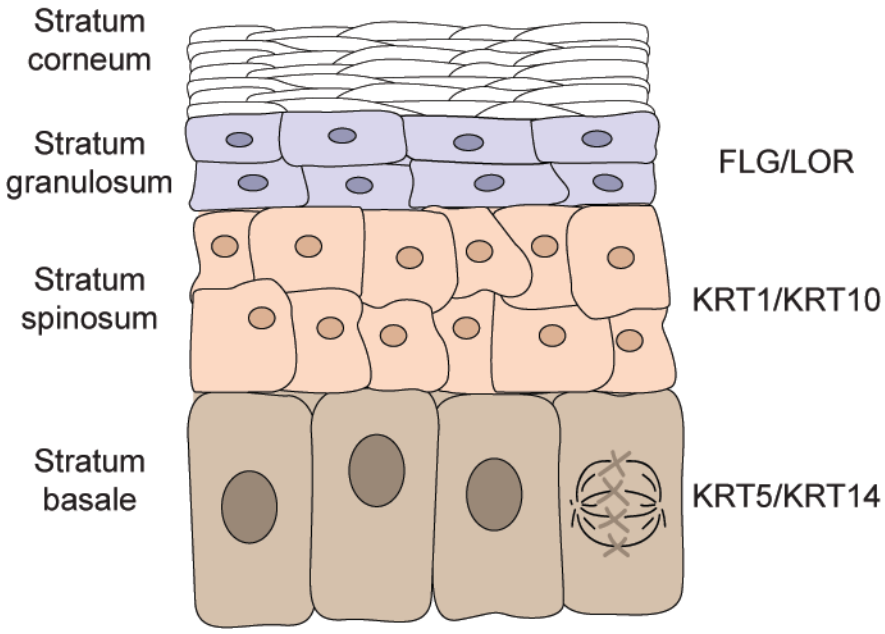


Figure 1-1. Structure and organization of the epidermis.

Corneal Structure

The cornea is the outermost layer of the eye, and is composed of three layers: endothelium, stroma, and epithelium. The cornea serves a barrier function, protecting the underlying ocular structures from UV damage, dehydration, and infection, while also refracting light to aid in vision; approximately 70 percent of the refractive power of the eye is directly attributable to the

cornea [3]. Due to this function in vision, the cornea must be completely clear; unlike other systems, this means the corneal stroma is avascular, and the cornea epithelium is not keratinized [4]. Rather than receiving oxygen and nutrients from the blood vessels, if must rely on the tear film and the aqueous humor [4].

In the multi-layered cornea epithelium, cells are continually sloughed off from the surface and are replenished by proliferation of limbally-located corneal epithelial stem cells [5, 6]. These cells divide infrequently and give rise to transient amplifying cells that move toward the center of the cornea and upward in the tissue, activating a terminal differentiation program in the process. While transcription factors [7-13] and signaling pathways [14-16] maintain limbal stem cells, how these multiple transcriptional inputs are integrated remains poorly understood.

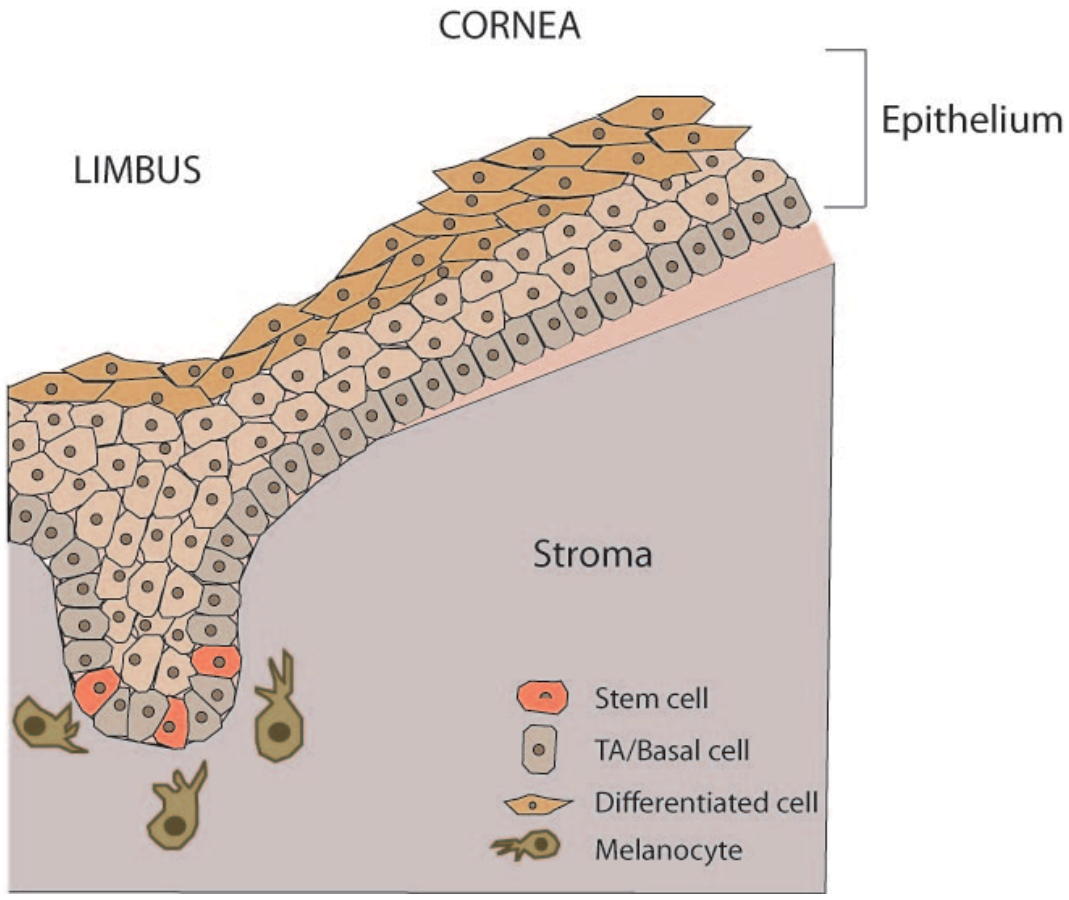


Figure 1-2. Structure of the limbus, and cornea epithelium.

Development of Ectoderm

In the early embryo, the ectoderm arises from the outermost cell layer; among the many tissues derived from this layer are the epidermis, the cornea, mouth and olfactory epithelia, and the central and peripheral nervous system. Cells from the surface ectoderm that receive BMP and WNT signals will go on to become the epidermis [17]. In the mouse, epidermal commitment of the primitive ectoderm occurs around (embryonic) day E9.5, when simple epithelial keratins 8 and 18 are replaced by expression of Keratin 14 and 5, markers of basal epidermal cells [18]. Around E15.5, the basal cells begins to stratify, and those in the upper layers acquire Krt1 and Krt10 expression, eventually resulting in a stratified squamous epithelium with a functional barrier soon after birth [18].

The neuroectodermal lineage also derives from the surface ectoderm; it diverges from the epidermal lineage by receiving Bone Morphogenic Protein (BMP) and WNT inhibitor signals [17]. This neuroectoderm will form the neural crest and neural tube; the neural crest will later produce melanocytes, cartilage, connective tissue, and a number of other cell lineages, while the neural tube will give rise to the brain and central nervous system.

The corneal epithelium also derives from the surface ectoderm; prior to specification, signals from the optic vesicle, a neuronal outgrowth, contact the surface ectoderm directly above and guide these ectodermal cells along an ocular fate, forming lens and corneal tissues [19]. Expression of master transcriptional regulator PAX6 is essential to this process, and to corneal epithelial fate specification [20]. Additional BMP inhibiting signals from migrating neural crest cells, which will eventually become stromal keratocytes, help specify the corneal epithelial fate [21]. By mouse embryonic day E14.5, all three cellular layers of the cornea are present, and the cornea epithelium is a single layer [22]. Prior to birth, proliferation will result in two layers, and

the formation of tight junctions, desmosomes, and hemidesmosomes helps to seal the epithelial barrier [23]. Upon eyelid opening around postnatal days 12-14, the epithelium continues to differentiate, acquiring corneal specific keratin 12 expression, and further stratifying to 5-6 cell layers thick [24].

Transcriptional Regulation

In a multicellular system a cell must integrate signals from itself, its neighbors, and its environment to produce the proper expression pattern for the given time and location. Such integration of diverse signals is most easily accomplished by a network of interconnected transcriptional regulators that allow input from differing sources, and provide many opportunities for regulation and fine-tuning of the transcriptional response. An additional layer of complexity added to this network is the idea that different networks act in different cells within a single tissue, and that a single transcription factor can be part of several unique networks depending on the spatial and temporal characteristics of the cell, as well as external stimuli. Under such circumstances, a transcription factor can have different targets in different cells based on the co-factors and co-regulators they interact with and the permissibility of the epigenetic landscape.

Grainyhead-like epithelial transactivator (Get1/Grhl3) is a transcription factor expressed in numerous epithelial tissues, and has been identified as a key regulator of epidermal terminal differentiation and barrier formation, a function that is evolutionarily conserved from flies to mammals. Grhl3 knockout mice die at birth, with numerous epidermal abnormalities including altered expression of adhesion proteins and changes in lipid composition, resulting in increased barrier permeability [25]. Gene expression profiling of Grhl3-/- mouse embryonic skin revealed over 200 genes aberrantly regulated in the absence of Grhl3, including many genes localized to

the Epidermal Differentiation Complex (EDC), a chromosome region containing many genes involved in barrier formation [26].

Intriguingly, GRHL3, also plays a major role in the regulation of wound healing, an epidermal function distinct from barrier formation, which requires activation of coordinated movement of keratinocytes and suppression of the differentiation program. In this process, GRHL3 is localized to keratinocytes at the leading edge of migrating cells in both embryonic and adult wounds [27], where it was shown to regulate genes involved in actin polymerization, possibly by affecting planar cell polarity pathways [28]. The role of GRHL3 in coordinating and regulating epidermal differentiation and wound healing offers a biologically and clinically relevant opportunity to understand a fundamental aspect of transcription factor function, namely their ability to regulate distinct genetic programs in the same cell type under different conditions.

Among the identified factors that have been linked to regulation of epithelial progenitor cell maintenance and differentiation are the Kruppel-like family of transcription factors (KLF). Members of this family of seventeen zinc finger transcription factors, which share a common DNA binding motif, have been implicated in the regulation of a diverse array of cellular functions; in many cases times the same factor can promote differentiation in certain systems, while maintaining the stem cell population in others [29]. For example, a role for KLF4 has been well established in embryonic stem cell maintenance [30], and KLF4 was among the original factors used to create induced pluripotent stem cells [31]. Intriguingly, this factor also plays a key role in promoting differentiation of epithelial tissues, including the intestine, epidermis, and cornea epithelium [8, 32, 33]. Another Kruppel-like family member, KLF7 appears to primarily play a role in maintaining progenitor cell state in systems like adipocytes [34], but is required for certain stages of neuronal differentiation [35].

Epigenetic Regulation

Epigenetic mechanisms of gene regulation, including DNA methylation and histone modification have been shown to play a major role in tissue development, homeostasis and repair after injury. In eukaryotes, DNA is wrapped around an octamer of histone proteins, forming a structure known as the nucleosome. Each nucleosome contains two copies each of histones H2A, H2B, H3, and H4. The tails of these histones extend out from the nucleosome core, and can be modified by the addition of various functional groups, including methyl, acetyl, and phosphate groups, among others. These modifications demarcate important regulatory domains, and have been shown to play a role in recruiting the enzymes and transcription factors necessary to effect gene expression or repression. For example, the H3 histone tail is marked by three methyl groups on lysine 4 (H3K4me3) at active promoters and is bound by factors involved in nucleosome remodeling as well as by a subunit of the TFIID complex [36]. In contrast, the repressive mark H3K27me3 is created by the histone methyltransferase Ezh2, a subunit of the Polycomb Repressive Complex 2 (PRC2). This mark is bound by the PRC1 complex, which then promotes ubiquitination of histone H2B, an example of the step-wise progression of gene activation or silencing [37]. There are many possible histone modifications, and it appears that a specific combination of modifications identifies the type of regulatory region; for example, the presence of H3K4me1 and H3K27ac in the absence of high levels of H3K4me3 marks active enhancers [38].

An additional level of regulation comes from the multiple copies of each histone, with the exception of histone H4, encoded in the genome. Some of these variants can confer unique architectures to the nucleosome structure and are associated with specific functions like

transcription, replication, recombination, and gene silencing. For example, the histone H3 variant known as CENP-A is found at active centromeres, and the combination of histone variants H3.3 and H2A.Z in a nucleosome causes instability and more rapid turnover, potentially acting to facilitate transcription factor access to regulatory regions [39].

In addition to histones and histone tail modifications, nucleosomes as an entity also play a role in the regulation of transcription. Active transcription is associated with regions of relatively open chromatin, and areas where histone modifications cause associations between histone protein and DNA helix to be much more relaxed. In fact, nucleosomes have been shown to act as a barrier to transcription factor binding in many situations. It has been suggested that a significant amount of nucleosome remodeling must occur both at promoters and at distal enhancers during the process of transcriptional activation [40]. This is reflected in the characterization of pioneering factors like Foxa1, which have been shown to bind first to nucleosome occupied sites, and recruit chromatin remodeling enzymes and other factors to alter the nucleosome landscape and allow for subsequent binding of other transcription factors to the now open chromatin [41].

Trithorax complex members regulate methylation of lysine 4 on histone H3

Lysine 4 on the tail of histone H3 can be mono-, di-, or tri-methylated. All three modifications can be found at active promoters, however, trimethylation (H3K4me3) localizes directly to the proximal promoter and has been most strongly associated with gene transcription. H3K4 dimethylation is found at active promoters, enhancers, and within the body of genes undergoing transcription [42]. H3K4 monomethylation is found at distal regulatory regions, including active and poised enhancers, and has also been identified more recently at the 3 prime ends of actively

transcribed genes [43]. Recent research points to a role of H3K4me1, which often encompasses broad domains that fully contain other marks, in restricting readers of other chromatin marks to specific domains, for example, containing readers of H3K4me3 like ING proteins in proximal promoter regions of active genes.

There are many different enzymes that can catalyze the transfer of a methyl group to lysine residues on a protein; however, all Trithorax family methyl transferases are characterized by the specific SET methyltransferase domain. The SET1 family of histone lysine methyltransferases catalyze the transfer of methyl groups to lysine 4 of histone H3. In humans, there are six such enzymes: MLL1-4, and SETd1A/B. MLL1 and MLL4 are most closely related to *Drosophila* Trithorax, while MLL2 and MLL3 are homologous to *Drosophila* Trithorax-related [44]. Each SET1 family enzyme interacts with the WRAD complex of proteins, so named because it comprises WD-40 repeat protein 5 (WDR5), Retinoblastoma-Binding Protein 5 (RBBP5), Absent Small Homeotic-2-like (ASH2L), and Dumpy-30 (DPY-30) [45, 46]. The exact function of each of these WRAD complex members is unclear, however, their association with a SET1 family member dramatically increases the enzyme's methyltransferase activity [45].

Drosophila Set1 and it's human counterparts SETD1A and SETD1B are thought to be the primary H3K4 -trimethylating SET1 enzymes, while the Trithorax-related homologs MLL4 and MLL3 have recently been shown to be responsible for H3K4 monomethylation at enhancers [47]. Trithorax homologs MLL1 and MLL4 have both been identified as mediators of H3K4 trimethylation at select loci, including bivalently-marked gene promoters in embryonic stem cells.

Chromatin regulation in epidermal development

The importance of chromatin regulation in epidermal differentiation dynamics has been well established through studies that show roles for a varied list of chromatin enzymes in the epidermis. Many of these epigenetic regulators have dynamic expression patterns during epidermal differentiation. For example, the Polycomb complex was found to be a repressor of the epidermal differentiation program, and is required to maintain keratinocytes in the basal, progenitor state [48]. Upon differentiation, levels of Ezh2, the catalytic component of Polycomb decrease in keratinocytes, concomitant with a release of epidermal differentiation genes from repression. Chromatin remodeling enzymes SATB1 and Brg1 are direct targets of master epidermal regulator p63 and are required for proper remodeling of the epidermal differentiation complex, a cluster of epidermal genes activated coordinately during differentiation [49, 50]. The downregulation of ACTL6A during differentiation is also required for Brg1-containing SWI/SNF complexes to remodel chromatin and activate differentiation genes [51].

Trithorax proteins regulate the epidermal differentiation program through interaction with GRHL3

In contrast to other chromatin regulators, whose expression levels change within a tissue in response to different physiological conditions, SET1 family members are highly expressed in the epidermis throughout development, homeostasis, and during wound healing [52]. This indicates that presence of Trithorax complexes in the cell is not enough to activate the epidermal differentiation program, and that targeting of Trithorax to the proper gene loci requires additional factors. While levels of each Trithorax complex member remain relatively constant across the differentiation process, there is great variation in the levels of different Trithorax complex

proteins within the epidermis. In differentiating keratinocytes, ASH2L is the most highly expressed Trithorax-associated protein, followed by WDR5 and MLL1 [52]. High expression of MLL1 and WDR5 is also observed in human whole epidermis samples; however, in mouse, when dermis and epidermis are separated, MLL2 is more highly expressed and is enriched in the epidermal fraction [52]. It is likely that unique combinations of Trithorax complexes are at work at the same time and in the same tissue, allowing for regulation of unique subsets of genes by these different Trithorax complexes. The composition of the Trithorax complex can also influence the enzymatic activity of the SET1 protein towards H3K4 mono-, di- or tri-methylation, therefore it is possible that multiple Trithorax complexes catalyze different methylations of lysine 4 at the same promoter, adding additional layers of regulation of epidermal gene expression.

There is strong evidence that multiple Trithorax complexes are involved in epidermal differentiation, as knockdown of SET1 enzymes MLL1, MLL2, MLL4 or Trithorax core component WDR5 causes a reduction in expression of epidermal differentiation marker gene TGM1 [52]. Of these Trithorax complex members, only MLL2 and WDR5 bind directly to the TGM1 promoter, indicating that MLL2-containing Trithorax complexes are direct regulators of TGM1, and suggesting that other complexes regulate TGM1 indirectly through promotion of epidermal differentiation at other targets. It is also possible that other Trithorax complexes regulate TGM1 through catalyzing H3K4 monomethylation at distal enhancer regions, rather than regulating H3K4 methylation in the proximal promoter.

TGM1 is also directly regulated by the epidermal differentiation-promoting transcription factor Grainyhead-like 3 (GRHL3), pointing to a potential interaction between MLL2 and GRHL3. GRHL3 knockout mice die at birth with numerous abnormalities including defects in

the epidermal barrier [26]. Microarray and ChIP-Seq experiments in mouse and human revealed that GRHL3 is important for coordinating the regulation of epidermal differentiation genes, as well as the activation of other transcriptional regulators. The role of GRHL3 extends beyond initial creation of the barrier; it is also required for repair after injury and in response to immune-mediated hyperplasia [53], it is required for keratinocyte migration [27], to close both embryonic and adult wounds [28], and it also regulates the barrier function of other tissues including bladder epithelium [54]. Depletion of GRHL3 during epidermal differentiation causes a reduction of MLL2 localization and H3K4 methylation at the promoters of genes it regulates [52]. Additional support for GRHL3 recruitment of MLL2-containing Trithorax complexes can be found in the observation that knockdown of GRHL3 or MLL2 affects a significant shared set of genes in epidermal keratinocytes; the most significant overlap is in genes that are classified as "late" differentiation genes, showing an increase in expression during the later stages of differentiation [52]. This is consistent with the defined role of GRHL3 in promoting terminal differentiation in the suprabasal layers of the epidermis, and indicates GRHL3 effects gene activation through recruitment of Trithorax complexes containing MLL2.

In contrast to GRHL3, MLL2 regulates a large fraction of genes expressed in epidermal keratinocytes, indicating a more general role in transcriptional activation, rather than being a selective regulator of the epidermal differentiation program. Indeed, siRNA knockdown of MLL2 affected a large swath of epidermal genes regardless of their specific expression pattern, suggesting it works with a number of transcriptional regulators that may provide additional temporal and spatial specificity to gene activation [52]. The large number of genes that are affected by MLL2 in epidermal keratinocyte differentiation but that do not appear to be regulated by GRHL3 suggests that other epidermal transcription factors also may promote differentiation

through recruitment of MLL2 to their target genes. Conversely, GRHL3 may recruit other Trithorax complexes to the genes it regulates independently of MLL2, or it may act through alternative mechanisms to activate these target genes.

Of note, GRHL3 is also downregualted by MLL2 knockdown [52]; whether this represents a feed-forward positive regulatory loop, whereby GRHL3 activates it's own expression through recruitment of MLL2, or whether this is an example of an alternative transcription factor working with MLL2-containing complexes to activate GRHL3 is not known. It is possible that both explanations are valid. During differentiation GRHL3 was not detected binding to its own proximal promoter, however, it does bind to enhancers within the *Grhl3* gene body, suggesting positive auto-regulation is possible. WDR5 localizes with GRHL3 to these enhancer regions, and is also found in the proximal promoter of GRHL3 indicating other transcriptional regulators may recruit Trithorax complexes to the *Grhl3* promoter to activate expression [52].

Further evidence comes from the observation that GRHL3 and core Trithorax member WDR5 interact directly, and GRHL3 also appears to weakly interact with MLL2, but not MLL1 or SETD1A [52]. While not conclusive, this suggests that MLL1 and SETD1A are recruited by GRHL3-independent mechanisms to gene targets they share with GRHL3 in epidermal keratinocytes. The interaction between GRHL3 and certain Trithorax complexes occurs genome-wide, as ChIP-Seq experiments revealed that WDR5 co-localizes with 88 percent of GRHL3 bound regions [52]. Forty-three percent of genes with a keratinocyte-differentiation expression pattern are bound by GRHL3 and WDR5, indicating Grhl3 provides some of the specificity for Trithorax targeting in the epidermis. This mechanism of Trithorax recruitment by GRHL3 may

also extend to other tissues where GRHL3 is expressed, and likely reflects a more general paradigm, whereby a tissue-specific transcriptional regulator recruits Trithorax to gene targets.

There are multiple mechanisms of epidermal differentiation gene activation

Initial studies focused on the antagonistic relationship between Polycomb and Trithorax at shared gene targets, however, work in the epidermis points to an independent role for Trithorax complexes at a subset of genes. While the repressive H3K27me3 modification is reduced concurrent with an increase in H3K4me3 during the process of differentiation at many genes, a number of epidermal differentiation genes lack any H3K27me3, even in the undifferentiated state, when the gene is not expressed. These genes show a strong increase in H3K4me3 upon epidermal differentiation [52]. It is possible that these genes were marked by repressive H3K27me3 at one point along their lineage trajectory, but that this mark was removed before acquisition of keratinocyte cell fate. The uncoupling of Polycomb repression and Trithorax activation may allow such genes to respond rapidly to signals for keratinocyte differentiation, as they do not need to recruit H3K27 demethylase enzymes to remove repressive marks before gene activation can occur. Alternatively, such genes may be regulated by transcription factors like GRHL3 that are only expressed in differentiating keratinocytes, and therefore do not need to be actively repressed in undifferentiated cells, where they are unable to be activated due to absence of transcription factor. This separation of Polycomb and Trithorax regulatory functions at certain gene promoters may also be a more general mechanism active in a number of tissue types.

Recent work has identified regulatory domains, termed super enhancers, which are characterized by the presence of clusters of enhancers that result in long stretches of active regulatory domains, and very high levels of histone modifications and transcription factor

binding [55, 56]. In accordance with their high levels of H3K27ac and other enhancer associated histone marks, these domains often overlap highly expressed genes. Intriguingly, these super enhancer-linked genes are not housekeeping genes, which are expressed highly in a number of cell types, but rather are tissue-specific genes that play a role in cell identity and carry out the functions of the particular cell type. The core Trithorax complex member WDR5 has recently been shown to localize to super enhancers; this is perhaps not surprising as Trithorax complexes are responsible for the deposition of the H3K4me1 histone mark found at enhancer domains. Indeed, in epidermal keratinocytes, many WDR5 binding events occur outside of proximal promoter regions [52]. GRHL3 and WDR5 co-binding events also occur in many distal regulatory regions, it is possible that GRHL3 also regulates epidermal differentiation through targeting of Trithorax to super enhancers.

Diseases of Trithorax complexes

Aberrant epigenetic regulation has been linked to many diseases, including cancer and developmental defects; due to their roles as crucial regulators of gene expression in many cell types, Trithorax and Polycomb have been found to be mutated in a variety of cancers. MLL proteins are so named because fusion proteins placing the N terminus of MLL1 together with another protein are common in mixed lineage leukemias. Intriguingly, the SET domain of MLL1 is located near the C terminus, and is not included in these fusion proteins, indicating that MLL1 promotes leukemiagenesis through mechanisms other than H3K4 methylation [57]. One common fusion is between MLL1 and the C terminus of DOT1, another methyltransferase that creates the active H3K79 methylation mark [58]. Because the MLL1 fragment can still bind DNA, this

fusion results in the incorrect targeting of H3K79me3 to domains that should have H3K4 methylation, an error that may promote the progression of leukemia.

In addition to roles in the progression of cancers of the hematopoietic system, MLLs have been found to be mutated in a number of solid tumors, including those of epithelial tissues. While MLL4 has been found to be mutated in a significant proportion of hepatocellular carcinomas [59], MLL3 is mutated in a wide range of cancers, including aggressive squamous cell carcinoma [60]. Furthermore, MLL2, MLL3, and the Polycomb enzyme EZH2 have been identified in a significant number of head and neck squamous carcinomas, a cancer that develops from epithelial mucosa in the upper aerodigestive tract, and is linked to tobacco use [61]. In contrast to the gene fusions that occur in leukemia, MLL mutations found in epithelial cancers are inactivating, pointing to a tumor suppressive role for MLL2 and MLL3 in epithelia, possibly through promotion of terminal differentiation.

Due to the role of Trithorax in regulation of HOX and other developmental clusters of genes, mutations in MLLs are also the cause of several developmental syndromes that share similar phenotypes; syndromes caused by mutation in MLL proteins are generally characterized by growth restriction leading to short stature and low weight, intellectual disability and facial structure abnormalities [62, 63]. Patients with these syndromes also have abnormal epidermal ridges on palmoplantar epidermis and abnormal hair growth, including long eyelashes, and excessive hair on forearms and elbows [63]. These findings suggest a role for MLL proteins in the establishment of epidermal structures during development, and in the hair follicle, an epidermal appendage. These syndromes occur with the mutation of a single copy of the MLL gene in question, leading to haploinsufficiency.

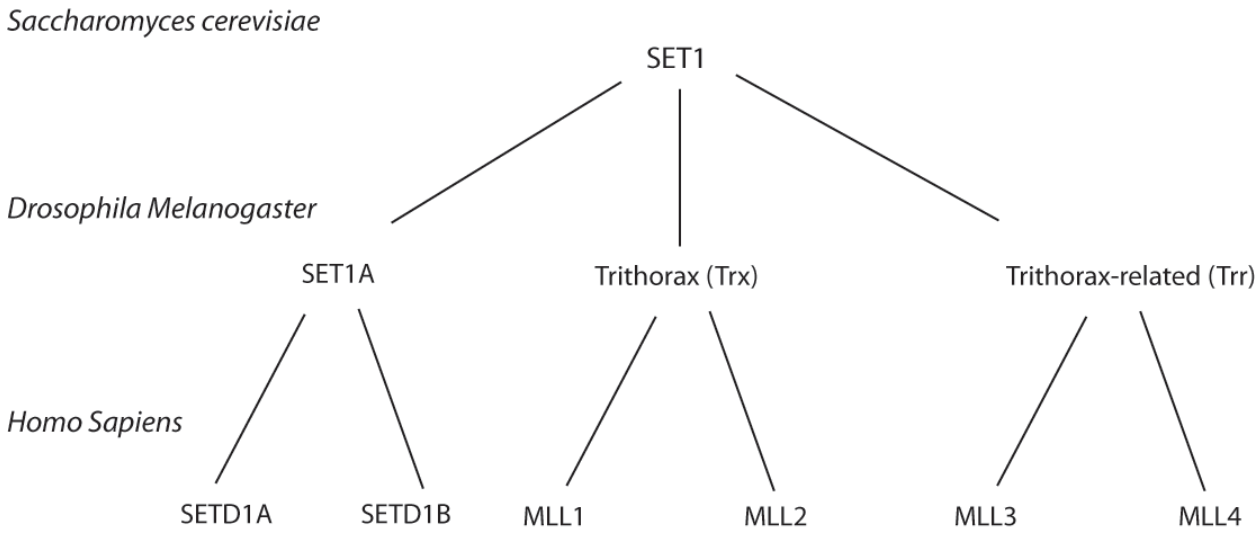


Figure 1-3. Comparison of H3K4 methylating SET enzymes between *S. cerevisiae*, *D. melanogaster*, and *H. sapiens*.

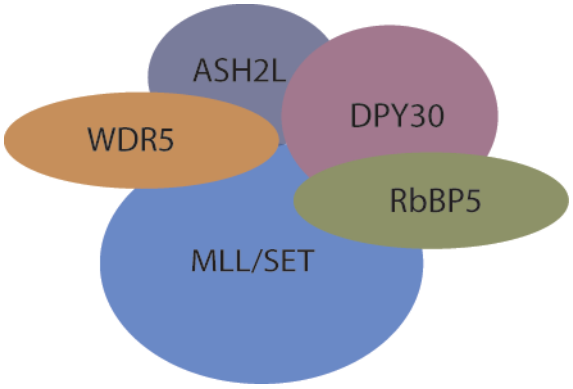


Figure 1-4. The WRAD complex associates with Trithorax SET enzymes.

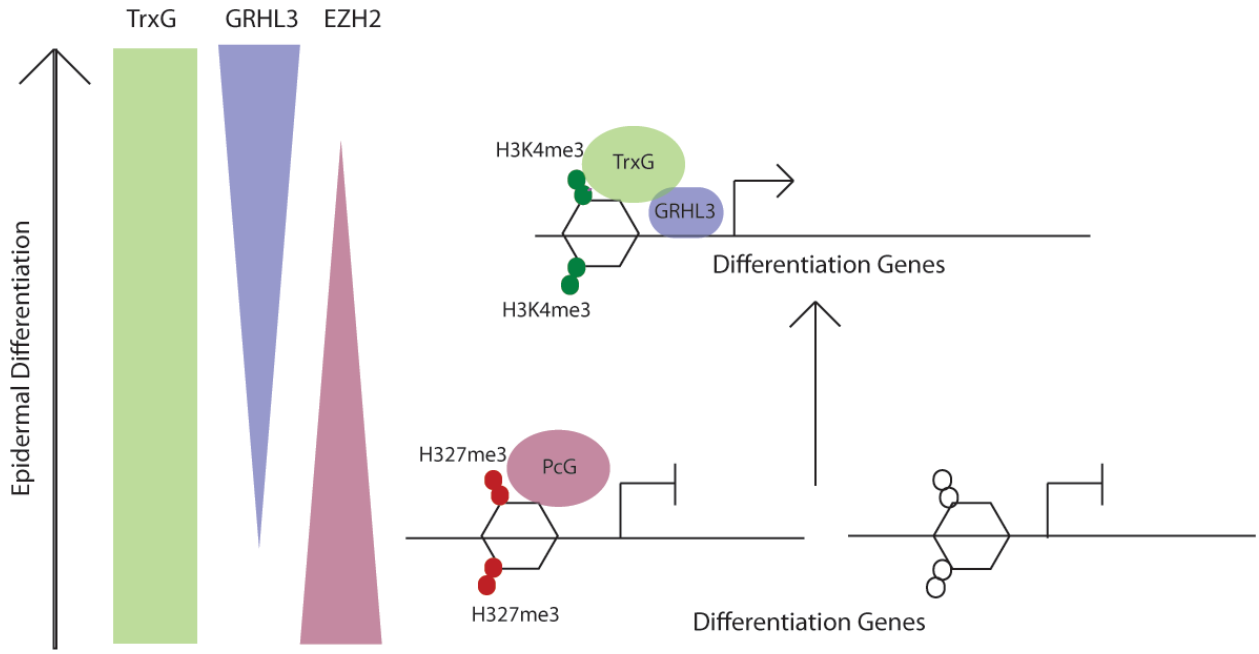


Figure 1-5. Model for GRHL3, Polycomb, and Trithorax –mediated regulation of epidermal differentiation genes.

Distal Regulatory Regions

During tissue development and homeostasis, distal regulatory regions called enhancers, which contain binding sites for numerous transcription factors, help to control and coordinate gene expression in a temporal and spatial specific manner. Unlike gene regulation that occurs at the proximal promoter, these enhancer regions can be situated at long distances from the transcription start site of the gene they regulate, and are thought to physically interact with the gene through chromatin looping [64]. Overcoming the inhibitory forces of RNA Pol II pausing to achieve transcriptional activation is thought to require the concerted action of many transcription factors and cofactors. In this context, with their increased density of transcription factor motifs, enhancers act to concentrate the factors necessary to produce transcription.

Prior to their time of action, enhancers can be poised for action, displaying both H3K4me1 and H3K27me3; in order to become fully active regulatory regions, they must

exchange H3K27me3 for H3K27ac, paving the way for transcription factor binding and gene regulation [38, 65]. In addition to regulating distal gene expression, enhancers have also been shown to be transcribed into short-lived eRNAs [66, 67]. This transcription may be partially a consequence of the proximity of enhancers to the transcriptional activity of the promoters they activate; however, these eRNAs have also been shown to affect gene expression.

The active enhancers within a given tissue show variation in intensity, both of histone modification and of transcription factor binding. Some enhancers, recently termed super enhancers, stretch for kilo bases and are extremely highly enriched in enhancer histone modifications as well as mediator and transcription factor binding [55]. These super enhancers are frequently tissue specific and thought to regulate the unique gene expression programs in each tissue [55].

Genomics Techniques

Rapid advances in DNA and RNA sequencing technologies have revolutionized approaches for studying transcriptional regulation. Assays that had until recently been directed toward a small number of genomic regions, can now be performed on a global, genome-wide scale; this includes the technique chromatin immunoprecipitation, which identifies protein-DNA interactions, and is a crucial aspect of transcriptional regulation studies. Chromatin immunoprecipitation sequencing (ChIP-Seq), can show all the map-able sites of transcription factor binding across the genome. It can also be used to define genomic regulatory regions and chromatin states based on histone modification.

These advances in sequencing technologies have also resulted in high throughput forms of numerous other transcriptional regulation techniques, including DNAse-Seq to identify open

chromatin domains, and 4C and 5C, which look at chromatin conformation and dynamics. A related technique is the newly developed ChIA-PET, which also looks at three dimensional chromatin, but focuses on the structure around the DNA binding sites of a specific transcription factor [68]. This is accomplished through an initial chromatin immunoprecipitation step; the technique is made high-throughput by paired-end sequencing of the ligated DNA fragments. Using RNA Pol II for the immunoprecipitation, Ruan et. al. showed that chromatin looping is very common throughout the genome, and that many regions containing gene clusters have large numbers of interacting domains, including multiple enhancers and promoters [69].

Chapter 2. Cofactors of LIM domain (CLIM) proteins regulate corneal epithelial progenitor cell function through noncoding RNA H19

Abstract

Cofactors of LIM domain proteins, CLIM1 and CLIM2, are widely expressed transcriptional cofactors that are recruited to gene regulatory regions by DNA-binding proteins, including LIM domain transcription factors. In the cornea, epithelial specific expression of a dominant negative (DN) CLIM under the *Keratin 14* (K14) promoter causes blistering, wounding, inflammation, epithelial hyperplasia and neovascularization, followed by epithelial thinning and subsequent epidermal-like differentiation of the cornea epithelium. The defects in corneal epithelial differentiation and cell fate determination, suggest that CLIM proteins may be involved in regulating corneal progenitor cells and the transition to differentiation. Consistent with this notion, the *K14-DN-Clim* cornea epithelium has fewer progenitor cells, and altered proliferation dynamics during epithelial development. Differentially regulated genes in DN-CLIM corneas include those in pathways crucial for progenitor cell maintenance, regulation of proliferation, and cornea differentiation. *In vivo* ChIP-Seq experiments in the cornea epithelium show that CLIM associates with DNA regulatory elements containing binding sites for HLH and other non-LIM homeodomain factors, as well as the estrogen receptor. Many genes that are highly expressed in the limbal epithelium, as well as genes with known roles in progenitor maintenance, are directly bound by CLIM cofactors. DN-CLIM binds to and decreases the expression of noncoding RNA H19, consistent with alterations in proliferation dynamics of DN-CLIM corneal cells during development. Together our results suggest an important role for H19 in corneal epithelial progenitor maintenance, and identify CLIMs as upstream regulators of H19 in the cornea.

Introduction

Epithelial tissues make an excellent model to study the regulation of the balance between progenitor state maintenance, proliferation, and differentiation. In the multi-layered cornea epithelium, cells are continually sloughed off from the surface as they are replenished by proliferation of limbally located corneal epithelial stem cells [5, 6]. While transcription factors [7-13] and signaling pathways [14-16] balance limbal stem cell maintenance with differentiation, how these multiple transcriptional inputs are integrated remains poorly understood.

CLIMs are a family (CLIM1/LDB2 and CLIM2/NLI/LDB1) of ubiquitously expressed co-factors [70-72] that form homo- and possibly heterodimers [70, 73] to mediate activation of genes in embryogenesis. CLIM2 is required for maintenance of fetal and adult hematopoietic stem cells [74], and epithelial stem cells in the crypts of the small intestine [75], in the bulge of hair follicles [10], and in the basal layer of the mammary gland [76]. Intriguingly, CLIM2 has also been shown to regulate the final stages of erythroid differentiation [77], making it likely that this factor acts at multiple stages of differentiation, maintaining progenitor cells in the stem cell niche while also promoting cell differentiation at later stages outside of the niche.

Characterized by their N-terminally located dimerization domain and C-terminally located LIM interaction domain, CLIM proteins lack DNA binding capacity and rely on interactions with other transcription factors to regulate gene expression. Through homodimerization, CLIMs have the potential to bring together multiple interacting proteins to form large multiprotein complexes that coordinate and integrate different input signals to regulate transcription [78]. It has also been suggested that the dimerization of CLIM proteins can mediate chromatin looping between promoters and enhancers [79]. While numerous interacting partners of CLIMs have been identified in various systems, the transcriptional regulators that

coordinate with CLIMs in the cornea remain undefined, as does the precise regulatory role of CLIM in this system.

Noncoding RNAs, an integral component of gene regulation during development, provide an additional layer of complexity to transcriptional regulation. These genes include both short RNA species, such as microRNAs, and long noncoding RNAs, which have been implicated in a wide range of transcriptional processes, including X chromosome inactivation and Hox locus regulation. H19, one of the few lincRNAs conserved between mouse and human, is a key regulator of proliferation during embryonic development, acting to antagonize the growth promoting effects of IGF2 expression. H19 and IGF2 are located adjacent to each other and are imprinted (IGF2 is expressed from the paternal allele, and H19 from the maternal allele) [80]. Differential methylation of an insulator element between the two determines which promoter can make contact with downstream enhancers, highlighting the importance of chromatin structure to gene regulation [80]. Both IGF2 and H19 are highly expressed in fetal tissues, underscoring the importance of balanced proliferation during development, and both are downregulated in most tissues soon after birth. While much work has been done to establish the mechanism of imprinting gene activation in this system, many questions remain about the tissue- and temporal-specific regulation of this locus during development, as well as the tissue-specific funtions of this noncoding RNA. Interestingly, H19 is expressed at a higher level in female than male corneas [81]; however, whether H19 plays a role in corneal epithelial cells is unknown.

Clim2 knockout mice die at E9.5 with severe patterning defects [78], whereas *Clim1* knockout mice are phenotypically normal, suggesting redundancy in function between the two genes. To avoid the embryonic lethality and the complex issue of redundancy, we developed a mouse expressing a dominant negative (DN) CLIM under the Krt14 promoter (*K14-DN-Clim*),

targeting DN-CLIM to the basal layer of stratified epithelial tissues including epidermis, hair follicles, and cornea [10]. The *K14-DN-Clim* mice exhibit decreased numbers of hair follicle stem cells, resulting in hair loss, and have abnormal corneas.

Right after birth, the *K14-DN-Clim* mice develop epithelial hyperplasia and corneal opacity [10]. Due to defective cell adhesion, caused at least in part by decreased expression of the hemidesmosome gene *Bp180*, stromal edema and blistering frequently occur, resulting in a strong inflammatory reaction and neovascularization. Recurrent wounding and inflammation persist, dramatically perturbing the ability of the cornea epithelium to maintain homeostasis. After this period of hyperplasia, the epithelium of *K14-DN-Clim* mice undergoes thinning around postnatal days 11 and 16, which persists up to 5 months of age, at which point the corneal epithelium begins to develop abnormal characteristics that mimic epidermal differentiation, including cornification in the upper layers of the epithelium; occasional terminal-phenotype mice also develop sebaceous- or goblet-like cells.

While we have shown that cell adhesion defects cause blistering and wounding [10], they alone cannot fully explain the defects in DN-CLIM cornea, suggesting additional regulatory roles for CLIMs. The epithelial thinning and later epidermal-like differentiation seen in *K14-DN-Clim* mice share similarities with some symptoms of corneal diseases in humans, including inflammatory human keratopathy, corneal epithelial squamous metaplasia, and limbal stem cell deficiency (LSCD), where inflammation, neovascularization, recurrent erosions of the epithelium and a period of epithelial thinning are characteristic [82-84]. Some of these diseases also result in abnormal keratinization of the cornea epithelium.

In this study we examined the early cellular and molecular cornea epithelial abnormalities, testing whether CLIMs have a role in corneal epithelial progenitor cells. We show

that when the function of CLIM is disrupted in the cornea epithelium, there is a reduction in self-renewal capacity of epithelial progenitor cells. Through gene expression profiling of *K14-DN-Clim* corneas as well as *in vivo* ChIP-Seq, we identify a number of genes known to be involved in epithelial progenitor cell function that are directly regulated by CLIM factors, including regulators of cell proliferation. Furthermore, we show that CLIMs bind to and regulate the noncoding RNA H19, pointing to an important and previously undefined role for this RNA as a regulator of proliferation in corneal progenitor cells. Together, these studies identify direct targets of CLIM in cornea epithelium indicate and a role for CLIMs in cornea epithelial maintenance and regulation of cell proliferation dynamics.

Results

Genes important for progenitor cell maintenance and tissue homeostasis are misregulated in *K14-DN-Clim* corneas

To identify early transcriptional changes that could underlie the striking terminal phenotype of the *K14-DN-Clim* corneas, we characterized the gene expression differences in whole corneas from WT and *K14-DN-Clim* mice at postnatal day 3 (P3). We identified 1099 differentially expressed genes with a significance of $p<0.05$ and a fold cutoff of 1.5; 361 genes are up regulated, while 738 genes are down regulated. Up regulated genes are overrepresented in eye development-related gene ontology (GO) categories [85, 86] (Figure 2-1A). These include retinal/neural crest genes, such as *Opn1sw*, and *Rcvrn*, which are not normally expressed in the cornea. The enriched GO categories for the down-regulated genes included blood vessel development, cell adhesion and extracellular matrix organization (linking gene expression changes to the loss of epithelial adhesion and stromal edema in the *K14-DN-Clim* corneas), as

well as regulation of proliferation (Figure 2-1A). The cell adhesion category of CLIM regulated genes contains a diverse group of proteins, including numerous laminins, integrins, and claudins. Among the cell proliferation genes down regulated by *K14-DN-Clim* in the cornea are cell cycle inhibitors *Cdkn1a/p21* and *Cdkn1b/p27*, consistent with the hyperproliferation observed in the early time points in DN-CLIM corneas [10]. We also found enrichment for components of several signaling pathways among the downregulated genes (Figure 2-1B). These include focal adhesion, TGF-beta, MAP kinase, and cytokine signaling pathways.

CLIM targets share similar expression dynamics across corneal development

We previously profiled gene expression changes in the cornea over the lifetime of the mouse (from embryonic day E14 through 2 years), identifying genes with similar spatial and temporal expression patterns [7]. By overlapping these data with genes differentially regulated in *K14-DN-Clim* mice, we found that several of the previously described developmental time course clusters harbor an overrepresentation of genes affected by DN-CLIM (Figure 2-1C). Of the genes upregulated by DN-CLIM, many fall in supercluster G, which contains genes involved in eye development and nervous system function. Many *K14-DN-Clim* downregulated genes are found in Cluster B, which is enriched for genes most highly expressed in the stroma, containing many genes with functions in extracellular matrix organization. Response to wounding is among the enriched GO categories for DN-CLIM affected genes in this cluster; the majority of genes that fall into this category, like Tgfbr2, Igfbp4, and Pros1, are highly stromally enriched and likely represent secondary effects of DN-CLIM expression in the cornea epithelium. Downregulated genes are also highly enriched in cluster D. Both clusters B and D contain many genes that are involved in the processes of cell adhesion and extracellular matrix organization, such as *Aebp1*,

Thbs1, and *Spon1* (Figure 2-1D), correlating well with the observed adhesion defect in the DN-CLIM corneas. These results suggest that CLIMs coordinately regulate groups of genes that are co-expressed across development and have similar functions. It also points to a role for CLIMs in higher level transcriptional organization of processes like cell proliferation, cell adhesion and cell signaling; which are crucial both to progenitor cell maintenance as well as the proper differentiation required to form a functional tissue.

K14-DN-Clim cornea epithelium exhibits altered proliferation dynamics

The enrichment of cell proliferation regulators among the genes affected by *K14-DN-Clim* in the cornea at P3, along with previous observations that the *K14-DN-Clim* phenotype includes a period of epithelial thinning, suggest that CLIM may play a role in regulating the balance between cell proliferation and differentiation, and in helping to maintain the progenitor population. To study the proliferation dynamics of the developing *K14-DN-Clim* corneal epithelium, we evaluated BrdU incorporation at P3, P8, and P10, in both the peripheral/limbal (Figure2-2A, B) and the central cornea (Figure 2-2A, C), corresponding to limbal stem cells/TA cells and TA cells/basal cells respectively (Figure 2-2A). At P3, the number of proliferating cells was significantly higher in *K14-DN-Clim* mice in both the central and peripheral cornea. A similar trend was observed at P8, although the difference did not reach statistical significance. By P10, BrdU labeling decreased in *K14-DN-Clim* mice, with significantly fewer proliferating cells in the peripheral cornea. Therefore, the corneal epithelium in *K14-DN-Clim* mice proliferates more actively than wild type cornea in the early days after birth, but as the phenotype progresses, as early as P10, the *K14-DN-Clim* limbally located cornea epithelial cells are less proliferative.

Progenitor cells are reduced and have less proliferative potential in *K14 DN-Clim* mice

To test whether the epithelial progenitor cells are perturbed in young *K14-DN-Clim* mice we adapted a method of culturing primary corneal epithelial cells for use in colony forming assays (Figure 2-3A). After isolation, the epithelial cells have high expression of progenitor cell markers Abcg2 and Krt19, and low expression of corneal epithelial differentiation markers like Krt12. After two weeks in culture, the expression of these progenitor markers still remains high compared to differentiation markers. Additionally, loricrin, a marker of epidermal differentiation is low at all timepoints. These results suggest the cells remain progenitor-like after two weeks in culture.

Corneal epithelial cells from *K14-DN-Clim* mice formed significantly fewer colonies than those from wild type littermates (Figure 2-4A, B). In addition, the size of *K14-DN-Clim* corneal epithelial cell colonies was smaller than controls, with the ratio of all *K14-DN-Clim* colonies skewed towards smaller sized colonies and away from large colonies, which comprised up to 10% of wild type colonies but less than 2% of *K14-DN-Clim* colonies (Figure 2-4C). Taken together, these data suggest that *K14-DN-Clim* mice have fewer corneal epithelial progenitor cells with decreased proliferative potential compared to wild type mice.

Limbal-enriched genes are regulated by CLIM

To gain insights into the effect of DN-CLIM on genes characteristic for limbal epithelial cells, we took advantage of previously published data comparing gene expression in central corneal and limbal cell populations captured by laser microdissection in the mouse [87] (Figure 2-5). There is a statistically significant overlap of genes differentially regulated by DN-CLIM at P3

and the limbal enriched genes determined by the aforementioned study [87]. Interestingly, all of the limbal epithelial enriched genes that are affected by DN-CLIM in the mouse cornea are down regulated, further supporting a role for CLIMs as transcriptional activators of a gene program regulating corneal progenitor cells (Figure 2-5). Among these are genes involved in cell growth, cell adhesion and in extracellular matrix deposition (Figure 2-5). These categories are highly similar to the GO categories found in the overall analysis for genes affected by DN-CLIM in the cornea. Several of the limbal enriched genes that are downregulated in the *K14-DN-Clim* cornea, include *Thbs2,* a cell adhesion gene. *Wnt5a,* which can act as an inhibitor of the canonical Wnt pathway in some systems, is also significantly downregulated in the aging cornea [7], further supporting the link between CLIM and regulation of the regenerative potential of the cornea.

Chromatin immunoprecipitation-sequencing (ChIP-Seq) identifies direct targets of CLIMs

To identify which of the genes affected by *K14-DN-Clim* are direct targets of CLIMs, we performed ChIP-Seq on corneal epithelium isolated from P7 mouse corneas, using an antibody to the myc tag of DN-CLIM. The P7 timepoint was selected as the earliest age from which sufficient chromatin could be obtained from corneal epithelia for ChIP-Seq. ChIP experiments were performed in duplicate; at a 5% FDR cutoff, approximately 60 percent of the peaks overlapped between the two replicates. These 27616 overlapping peaks were used for all further analyses.

Whereas 10 percent of CLIM peaks were found in proximal promoter regions, the great majority were found in distal regions, defined as 50kb upstream through 5kb downstream of a TSS, excluding proximal promoters (0-2kb upstream of TSS); these distal binding sites are candidate locations for regulatory regions like enhancers (Figure 2-6A). This genomic

distribution of CLIM peaks represents a strong deviation from the background genomic distribution, consistent with previous ChIP-Seq studies for CLIM in erythroid cells [77], and for the CLIM-interacting factor LHX2 in hair follicle stem cells [88]. Using a 20kb window around transcription start sites, genes with a nearby CLIM peak were enriched for gene ontology categories that included transcriptional regulation, proliferation, cell signaling, cell fate commitment, embryonic morphogenesis, and negative regulation of cell differentiation (Figure 2-7).

CLIM binding sites are enriched for selective transcription factor motifs

Using the 250 bases of flanking DNA from each side of the center summits of the ChIP-seq peaks (Figure 2-7A), we found over-represented DNA-binding motifs that were highly similar to the GATA, RUNX1, and the Estrogen receptor motif (Figure 2-6E). Directed motif searches on the CLIM binding regions revealed an enrichment of motifs for additional factors known to interact with CLIM in other systems, including GATA factors, TAL1, LMO2, GFI1, and FOXO1 (Figure 2-6F). Interestingly, LHX motifs were not enriched in CLIM binding regions. To determine if CLIM associates with multiple DNA-binding proteins, we looked for co-occurrence of motifs within the ChIP-Seq peaks (Figure 2-6G). Nearly all GATA3 motifs had nearby TAL1 motifs, indicating these two factors bind together. Additionally, we found an increase in the co-occurrence of GATA3 and RUNX1 motifs, as well as an increase in the number of ESR1 and LMO2 motifs found together in peaks, suggesting that biases exist in the grouping of CLIM-associated transcription factors at different locations.

CLIM target genes are involved in both stem cell maintenance and epithelial differentiation

We next overlapped the DN-CLIM ChIP-Seq peaks in corneal epithelium with genes whose transcripts were affected by DN-CLIM in P3 corneas. Four hundred sixty-nine genes, representing approximately 40 percent of differentially regulated transcripts, also had a nearby CLIM ChIP-Seq peak (Figure 2-6B). These genes were enriched for a subset of GO categories from the analysis of all genes bound by CLIM, including cell proliferation, blood vessel development, eye development and cytoskeletal development (Figure 2-6C), categories associated with the phenotype of the DN-CLIM cornea. Many proteins involved in ECM interaction and focal adhesion signaling are directly bound by CLIM, including thrombospondins, integrins, collagens and filamin. The Wnt signaling pathway is also overrepresented among CLIM targets. *Wnt5a, Lrp1, Lrp1b, Ctnnbl1*, and *Nkd2* are all directly bound by CLIM and misregulated in the *K14-DN-Clim* cornea epithelium. Additionally, several of the limbal-enriched genes that are affected by DN-CLIM at P3 also have CLIM ChIP-Seq peaks in the corneal epithelium at P7; these include *Wnt5a, Fgfr1*, and *Thbs2* (Figure 2-5).

CLIM targets h19 to regulate corneal epithelial proliferation dynamics

Among the genes bound by DN-CLIM with expression changes in the microarray is the noncoding RNA H19, a factor with well defined roles in regulation of proliferation and growth in other systems (Figure 2-8A). During development, H19 counteracts the growth-promoting role of IGF2, thereby reducing proliferation in progenitor cells [80]. DN-CLIM binds to two sites overlapping the H19 gene body, and to additional upstream and downstream sites. Fitting with the higher expression of H19 in the female cornea, we find a high scoring estrogen receptor motif within a DN-CLIM peak at the H19 locus, suggesting CLIM may bind with ER to regulate

H19 expression. We validated that H19 expression levels are decreased in DN-CLIM mice compared to WT littermate controls; H19 is also downregulated in primary human corneal epithelial cells (HCE) in the presence of DN-CLIM (Figure 2-8B). To determine the effect of DN-CLIM and H19 on cell proliferation, we performed MTT assays; cells transfected with DN-CLIM alone showed an increase in proliferation, consistent both with the reduction in H19 gene expression, and with the observed initial increase in proliferation in the *in vivo* BrDU assays. When H19 and DN-CLIM were co-transfected, the increase in proliferation was not observed, These results identify a role for H19 in modulating corneal epithelial proliferation, and indicate that CLIMs are important regulators of H19, potentially through interactions with ER.

Discussion

In this study we provide evidence that CLIM regulates the maintenance and behavior of cornea epithelial progenitor cells. Using a combined gene expression and ChIP-Seq approach, we identify direct targets of CLIM that are crucial to maintaining the population of corneal progenitor cells. One such target is H19, a noncoding RNA that acts to reduce proliferation in corneal epithelial progenitor cells, and maintain the proper balance between differentiation and proliferation required for tissue function.

CLIM regulates cornea progenitor cell function

The similarities in morphological changes between numerous corneal diseases, including LSCD [82-84] and metaplasia, and aspects of the DN-CLIM cornea phenotype suggest that DN-CLIM mice have the potential to be used as a model to gain insights into these diseases. Our data identify genes misregulated by DN-CLIM, as well as CLIM binding events on a genome wide

scale. Not only are these genes likely responsible for certain aspects of the DN-CLIM phenotype, they are also likely candidates to have altered expression levels in corneal diseases; such targets could be explored in the search for novel therapeutic approaches to these conditions.

Overlap of the *K14-DN-Clim* affected genes with the list of enriched limbal genes [87] revealed that a significant number of limbal enriched genes are downregulated by DN-CLIM in the cornea, including *Thrombospondin 2*, and cell signaling molecules *Fgfr1* and *Wnt5a*. WNT5A, which promotes quiescence in hematopoietic stem cells [89], can act as an inhibitor of canonical Wnt signaling [90].

CLIMs regulate noncoding RNA H19.

We identified the noncoding RNA H19 as a CLIM target in cornea epithelium, adding to the mechanistic complexity of CLIM-mediated transcriptional regulation, and defining for the first time a role for H19 as a negative regulator of proliferation in cornea epithelium. Two genes, H19 and IGF2, have opposing roles in regulating proliferation and are expressed from the same locus. Expression is determined by imprinted methylation that directs shared enhancer elements to the proper promoter (IGF2 on paternal allele, H19 on maternal allele) [91]; because of this, the structure and chromatin looping formed between enhancer and promoter is crucially important. Due in part to it's ability to homodimerize, CLIMs have been shown to play a major role in forming and maintaining chromatin loops at the beta globin locus during hematopoesis [79]; thus it is likely that CLIM functions at the H19/IGF2 locus to regulate chromatin structure and interactions between enhancer and promoter. Both IGF2 and H19 are expressed at higher levels during embryonic and early postnatal development than in mature tissue [80], suggesting they are particularly important in regulating progenitor cells.

Two developmental syndromes are linked to misregulation of this locus; Silver-Russell syndrome causes growth defects and is associated with loss of paternal IGF2 expression. Opthamologic abnormalities, including refractive errors have been linked to this syndrome, suggesting proper regulation of the IGF2/H19 locus is important for the development of occular structures [92]. The converse syndrome, Beckwith-Wiedemann, associated with hypermethylation of the locus, causes excessive growth, and predisposes patients toward developing tumors [93]. Additionally, sex-specific expression of H19 and IGF2 has been documented; both are significantly more highly expressed in the iris, retina, ciliary bodies, eyecup and cornea of the female mouse eye compared to the male [81]. This is intriguing, as DN-CLIM binding at H19 overlaps ER motifs, suggesting CLIMs may interact with ER to upregulate expression of this locus, leading to higher expression in the female eyes.

CLIMs help to regulate proliferation and repopulation dynamics in the cornea epithelium. The DN-CLIM corneas exhibit a striking misregulation of genes involved in cell proliferation, another aspect crucial for progenitor maintenance and tissue homeostasis. The BrdU incorporation data and colony forming assays together suggest that the corneal epithelial progenitor cell population in *K14-DN-Clim* mice is altered, with an initial increase in proliferation followed by a decrease by P10. The DN-CLIM cornea epithelial cells also have less repopulating potential compared to WT. This initial increase in proliferation, characterized by a reduction in cell cycle inhibitor expression, may cause the stem cell compartment to exhaust its potential early, resulting in later reductions in proliferation and repopulating capacity. A number of cell cycle regulators, including Cdkn1b, are bound directly by CLIM and misregulated in *K14-DN-Clim* corneas, suggesting they are direct targets of CLIM. Based on the large number of

potential direct CLIM targets within the Wnt signaling pathway, CLIM likely further influences the cell cycle through its regulation of Wnt signaling.

CLIM interacts with distinct cohorts of transcription factors to regulate gene expression in the cornea

Our ChIP-Seq experiments, the first *in vivo* ChIP-Seq experiments in the developing cornea epithelium, establish the feasibility of using such samples to study chromatin and transcriptional regulation in the cornea epithelium *in vivo*. This approach could be applied in the future to chromatin samples isolated from the epithelium of diseased corneas to identify the mechanisms of gene misregulation and potential gene targets for drug therapy.

Identifying genome-wide CLIM binding sites in the cornea provided insights into the mechanisms of CLIM function in cornea, including information on potential binding partners responsible for recruiting CLIM to chromatin. Thus we found that several DNA-binding motifs were enriched in CLIM binding regions, including motifs for TAL1, GATA, RUNX1, and others; all are factors known to interact with CLIM in other systems. Intriguingly, among the enriched motifs flanking CLIM peaks was the motif for estrogen receptor alpha (ERα). CLIM has previously been shown to be a cofactor for ERα in mammary epithelia, but not stratified epithelia; in mammary epithelial cells it functions to enhance ERα mediated gene expression [94]. Estrogen receptors are expressed in all ocular surface tissues, including the cornea, and a recent study highlighted the sex-specific differences in corneal wound healing [95]. These results suggest that the ERα-CLIM interaction also occurs outside of reproductive tissues.

CLIM was originally identified through its interaction with LHX proteins [70-72]. Therefore, it is interesting to note that the motif for Lim homeodomain proteins (LHX) was not

enriched within the flanking regions of Clim ChIP-Seq peaks. This is in contrast to hair follicle stem cells where CLIM seems to function through interactions with LHX2 [96]. LHX proteins are not highly expressed in the cornea [7]; this, combined with the lack of enriched LHX motifs, suggests that in the cornea epithelium, CLIM primarily interacts with other factors to effect its regulation of gene targets, and that the exact composition of the CLIM-mediated multiprotein complexes varies with tissue type. LMO proteins are expressed in the cornea epithelium and might play a role in recruiting CLIM to other non-LHX complexes.

The biases we discovered in co-occurrence rates for motifs of different CLIM interacting partners indicate that CLIM can act through different protein complexes in the same cell type to regulate gene expression, and suggests it plays a role in bringing together different transcription factors to function synergistically at specific regulatory locations. It is tempting to speculate that that these preferences for specific combinations of CLIM co-binding transcription factors may be a mechanism of regulation that is also active in other CLIM-expressing tissues.

CLIMs act through unique mechanisms to modulate progenitor cells in different tissues.

CLIMs play a number of roles in progenitor cell maintenance: they promote quiescence in hematopoietic stem cells [89], maintain hair follicle stem cells in the epidermis through interaction with Lim homeodomain protein LHX2 [10, 88, 96, 97], and maintain the proliferative potential of basal stem cells, acting upstream of FGFR2 in mammary epithelium [76]. In our studies characterizing CLIM regulation of corneal epithelial progenitor cell maintenance, we do not find evidence for CLIM binding to the proximal *Fgfr2* promoter, and motif analysis suggests that Clim acts independently of Lhx factors in the cornea. Our results therefore indicate that,

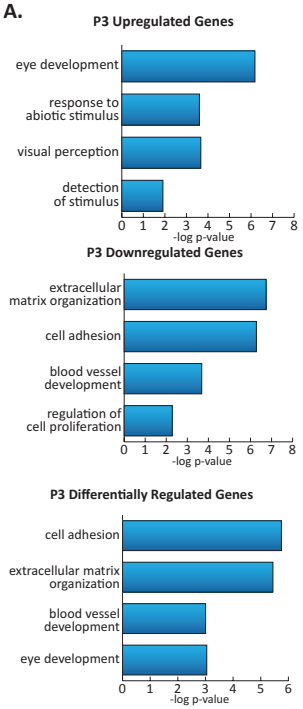
while CLIMs have a broad role in regulating progenitor cell dynamics in many tissues, they do so through unique mechanisms in different tissues.

Conclusion:

Interference with CLIM function in corneal epithelial cells causes a dramatic phenotype in mice, some aspects of which mirror symptoms of corneal diseases in humans. In this study, we show that CLIMs are essential for proper corneal epithelial progenitor maintenance, and suggest that they carry out this role by acting at a higher level to coordinate expression patterns for genes involved in cell adhesion, regulation of the extracellular matrix environment, and cell proliferation. We provide data indicating CLIMs interact with specific combinations of transcription factors to effect this regulation. Through identification of direct CLIM targets, we identify a role for CLIMs in regulation of H19; balanced proliferation is essential for corneal epithelial progenitor cell function.

Figure 2-1. Microarray gene expression analysis of postnatal day 3 (P3) whole mouse corneas reveals genes and pathways with altered expression in *K14-DN-Clim* mice. A) Gene ontology and pathway analysis for differentially expressed genes in P3 DN-CLIM corneas. Separate GO Analysis of up regulated genes and down regulated genes in P3 DN-CLIM corneas is also shown. B) Significant enrichment of signaling pathways among downregulated genes in P3 DN-CLIM corneas as determined by KEGG pathway analysis. A list of specific genes is shown. C) Overlap of genes affected by DN-CLIM with previously defined [7] clusters of genes with similar expression patterns and functional classifications across development and aging of the cornea. The graph shows the fraction of genes in each cluster affected by DN-CLIM as well as the fraction expected based on a random overlap. D) Heatmaps showing previously described gene expression changes of DN-CLIM affected genes in each developmental cluster. For each cluster, selected GO categories and genes are shown.

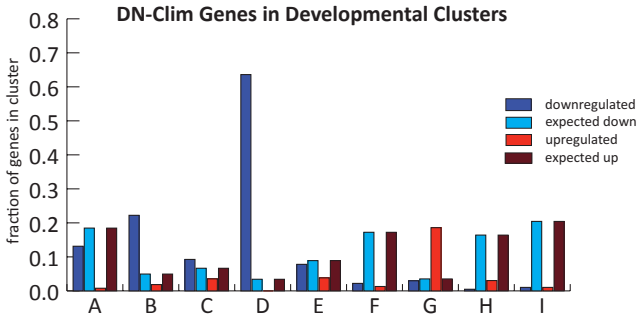
A.



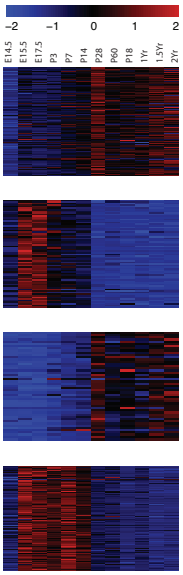
B.

Signaling Pathway	p-value	Genes
TGF-beta	3.70E-03	Acvr1a, Dcn, Fst, Ltbp1, Pitx2, Thbs1, Thbs2, Thbs3, Thbs4, Tgfb3, Tgfb2
Focal Adhesion	7.80E-09	Cav1, Ccnd2, Collagen type II, Collagen type IV, Collagen type XI, Flnc, Flna, Flnb, Igf1, Itga1, Itgb1, Lama2, Lama4, Myl9, Pdgfra, Thbs1, Thbs1
MAPK Signaling	1.20E-02	Flna, Flnb, Mapk13, Map3k6, PDGF, Tgfb3, Tgfbr2, Tnfr, Ntrk2, Nfatc4, Mapkapk3, Fgfr1, Fgf2
Cytokine- receptor interaction	4.70E-03	Ccl12, Cxcl12, Ifnar2, Il11ra1, Il13ra1, Il15ra, Kitl, Lepr, Pdgfra, Pdgfrb, Pf4, Tgfb3, Tgfbr2, Tnsf10, Tnsf1a, Osmr, Ghr, Avcr1

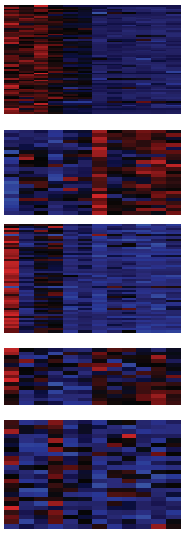
C.



D.



GO Terms	Example Genes
Cell proliferation	Cdkn1a, Cdkn1b, Alox12
Regulation of kinase activity	Cdkn1a, Cdkn1b, Cav1, Tlr4, Trib1, Pde6g
VEGF signaling	Pdgfra, Pdgfrb
Cell adhesion	Egfl6, Fbln2, Smoc2
Protein localization	Arntl, Rasef
Eye development	Cep290, Crygs, Gngt1
Cell adhesion	Aebp1, Spon1, Thbs2, Thbs4
Extracellular matrix	Col11a1, Fbln5, Ecm2, Nid1



GO Terms	Example Genes
Locomotion	Sema3a, Gnao1, Npas3
Cell adhesion	Lama4, Itga8, Vcan, Pcdh18
Dephosphorylation	Cdc14b, Ptpn21, Synj2
Sensory organ development	Mab21l2, Fzd3, Lim2, Gpr98, Sox2
Cell adhesion	Nrcam, Nlgn1, Pcp4
NA	NA
Nervous system	Pou3f2, Cplx3, Trim9

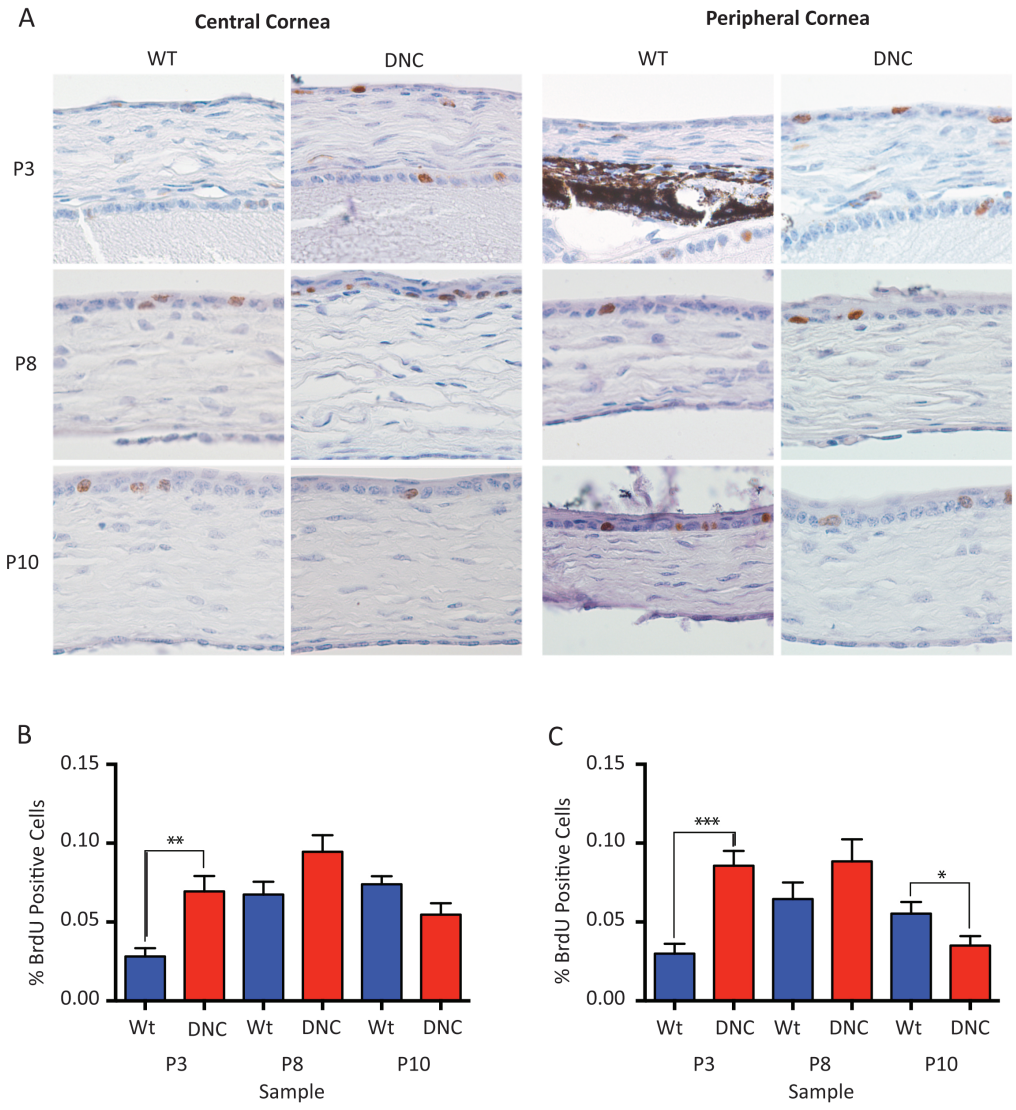


Figure 2-2. Cornea epithelial cell proliferation dynamics are altered in *K14-DN-Clim* mice. A) Representative images of BrdU labeling in the cornea at the indicated times (40X). B) Quantification of the fraction of BrdU labeled cells in central cornea at the indicated times (n= 3 WT and 3 DN-CLIM mice). C) Quantification of the fraction of BrdU labeled cells in peripheral/limbal cornea at the indicated times (n= 3 WT and 3 DN-CLIM mice).

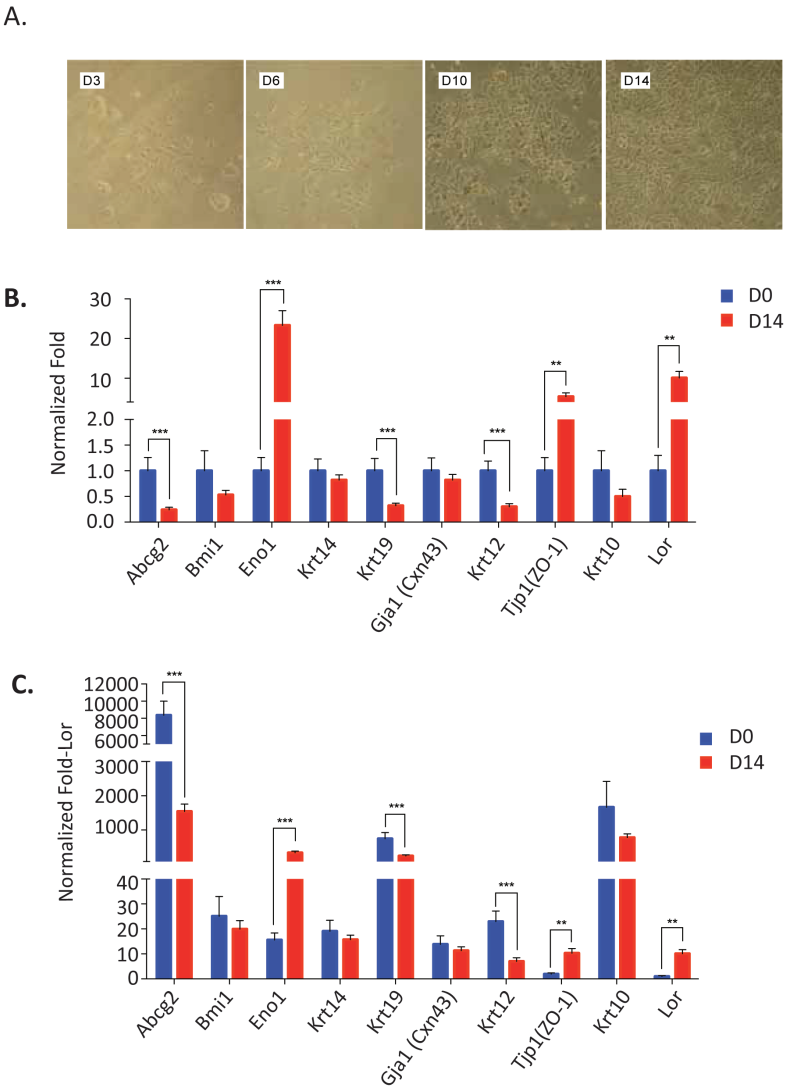
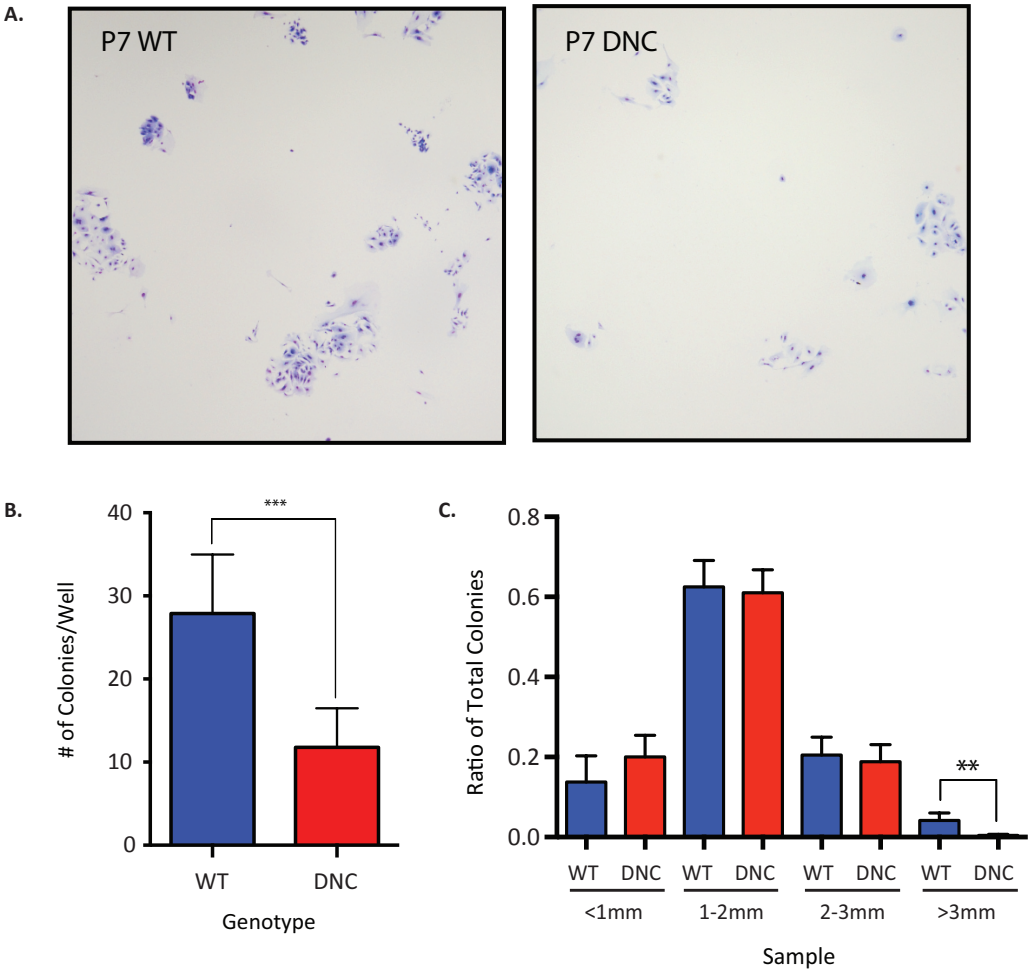
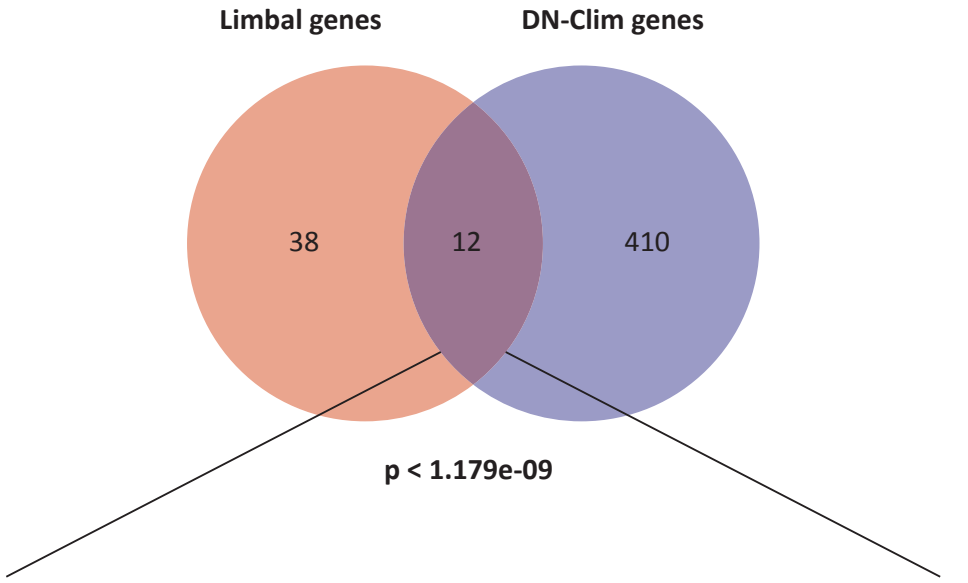


Figure 2-3. Characterization of primary mouse corneal epithelial cell culture. A) Growth of cells in culture over two weeks. B) Expression of markers at day 0. C) Expression of markers at day 14. All data were normalized to 18SRNA.

Figure 2-4. Cornea epithelial progenitor cells are decreased in *K14-DN-Clim* mice at P7. A) Giemsa stained representative images of CFU assays from P7 Wt and DN-CLIM primary cornea epithelial colonies (CECs) (4X). B) Quantification of number of colonies/well in WT and DN-CLIM cells (n= 7 WT and 7 DN-CLIM mice). C) The ratio of colonies (each genotype totals 1)


A.
P7 WT
P7 DNC
B.
# of Colonies/Well
40
30
20
10
0
***
WT
DNC
Genotype
C.
Ratio of Total Colonies
0.8
0.6
0.4
0.2
0.0
**
WT
DNC
<1mm
1-2mm
2-3mm
>3mm
Sample



Gene Symbol	Gene name	Role	Fold	p-value
Thbs2	Thrombospondin2	Cell Adhesion	-1.88	0.0031
Aebp1	AE-binding protein 1	Cell Adhesion	-1.86	0.0068
Col6a2	Procollagen, type VI α 2	Cell Adhesion	-1.48	0.0452
Spon1	Spondin 1	Cell Adhesion	-1.5	0.0346
Dab2	Disabled homolog 2	Differentiation	-1.87	0.0040
Ereg	Epiregulin	Cell Growth/ Development	-1.77	0.0105
Fgfr1	Fibroblast growth factor receptor 1	Cell Growth/ Development	-1.57	0.0151
Krt7	Keratin complex 2 basic gene 7	Cytoskeleton	-1.52	0.0179
Timp2	Tissue inhibitor of metalloproteinase 2	Extracellular Matrix Protein	-1.74	0.0094
Pdgfra	Platelet derived growth factor receptor α polypeptide	Extracellular Matrix Protein	-2.06	0.0016
Calcrl	Calcitonin receptor-like	Signal Transduction	-2.72	0.0002
Wnt5a	Wingless related MMTV integration site 5A	Signal Transduction	-2.05	0.0002

Figure 2-5. Previously defined limbal enriched genes [87] significantly overlap with differentially regulated genes in P3 *K14-DN-Clim* corneas. The overlapping genes are listed with the fold-decrease in DN-CLIM corneas as well as p-value for the differential regulation.

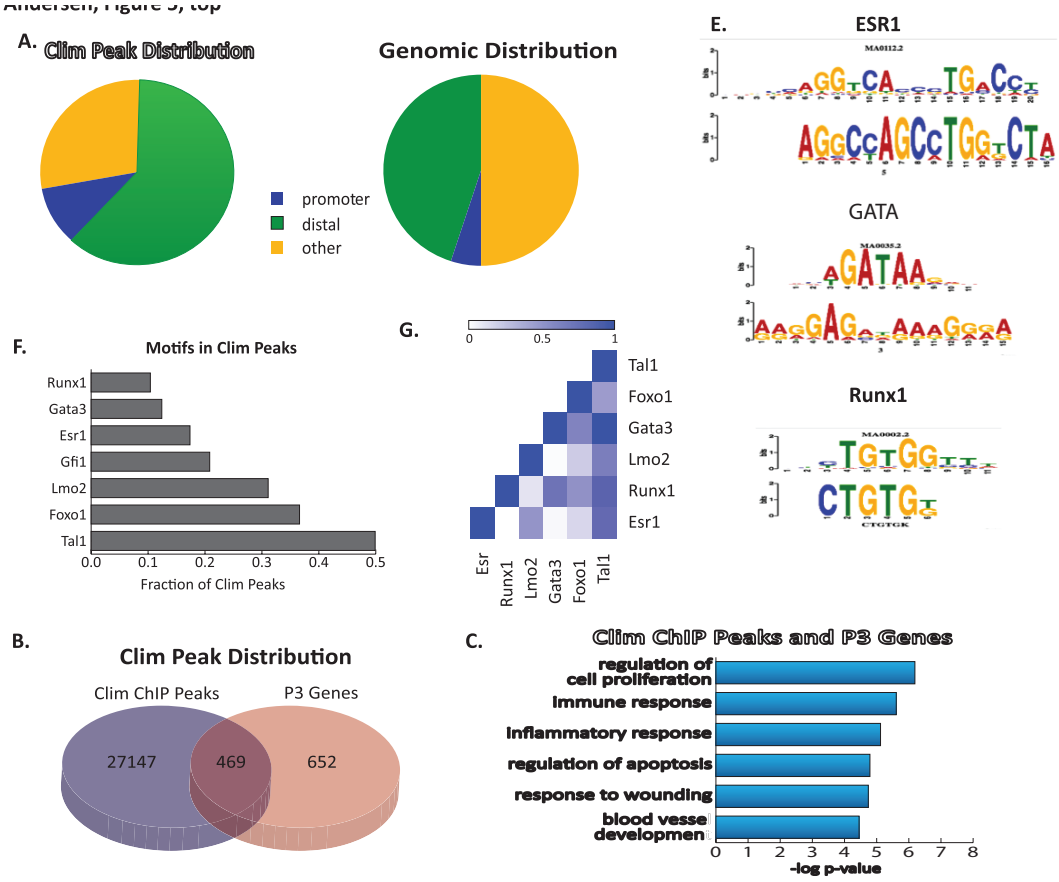


Figure 2-6. ChIP-Seq identifies direct targets of CLIMs in cornea epithelium. A) Distribution of ChIP-Seq peaks across genomic features: promoter= 0-2kb window upstream of TSS, distal = 50kb upstream of TSS through 5kb downstream of TSS. B) Enriched motifs found by MEME in regions flanking peaks. C) Enriched motifs found in flanking 250bp of ChIP-Seq peaks by directed motif searches. D) Co-occurrence of motifs found in CLIM peaks. E) Venn diagram showing overlap of ChIP-Seq peaks and genes affected by DN-CLIM in P3 corneas. F) GO Analysis of genes affected by DN-CLIM at P3 that have a Clim peak within a 40kb window of the TSS (20kb upstream through 20kb downstream).

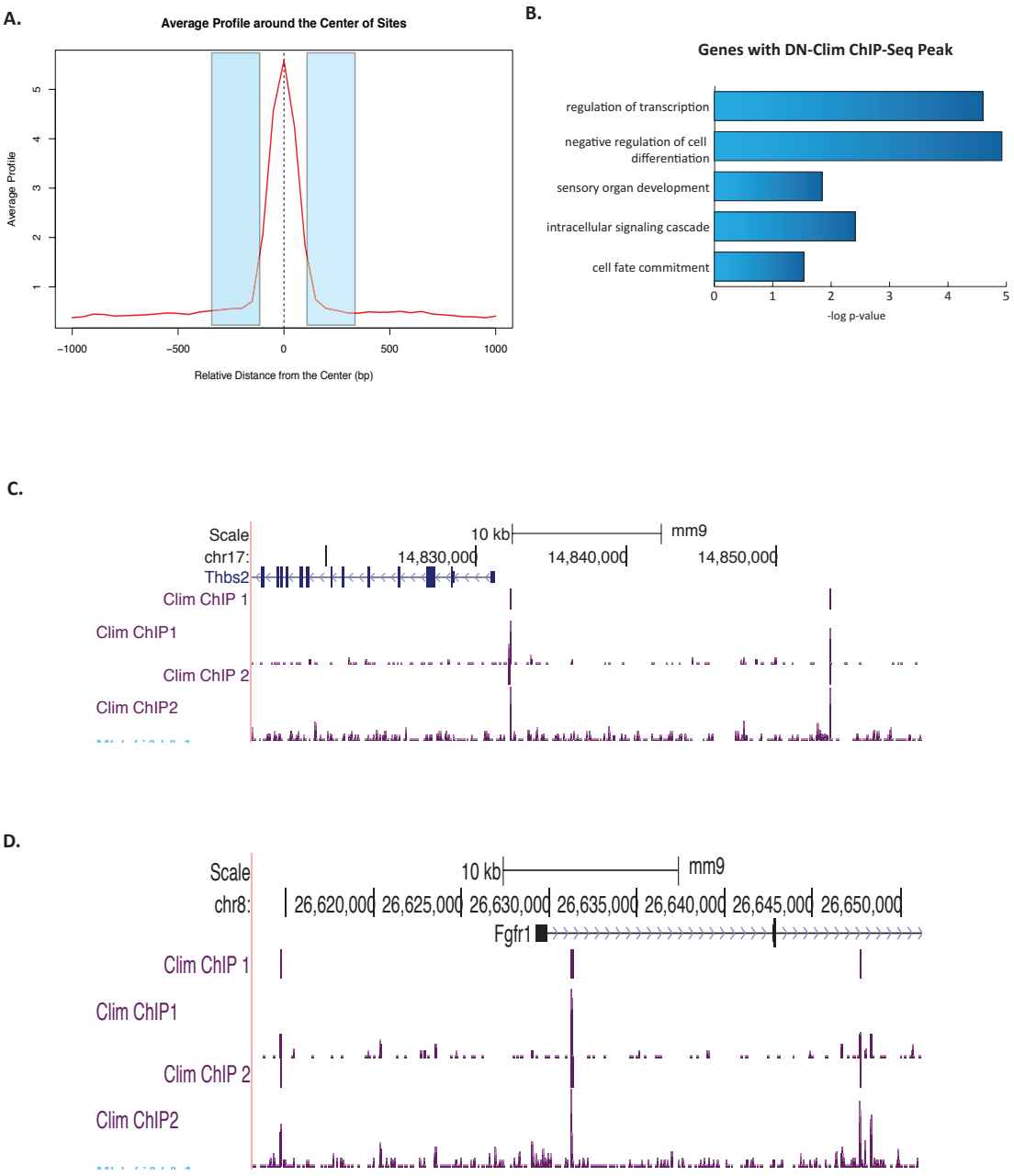


Figure 2-7. ChIP Seq identified CLIM targets in cornea epithelium. A) GO analysis for ChIP-Seq peaks within 40kb of TSS of genes. B) GO analysis of genes with CLIM ChIP-Seq peak in 40kb window around TSS. C) ChIP-Seq peaks near CLIM-regulated gene *Fgfr1*. D) ChIP-Seq peak at CLIM-regulated gene *Runx1* overlapping GATA binding in MEL cells.

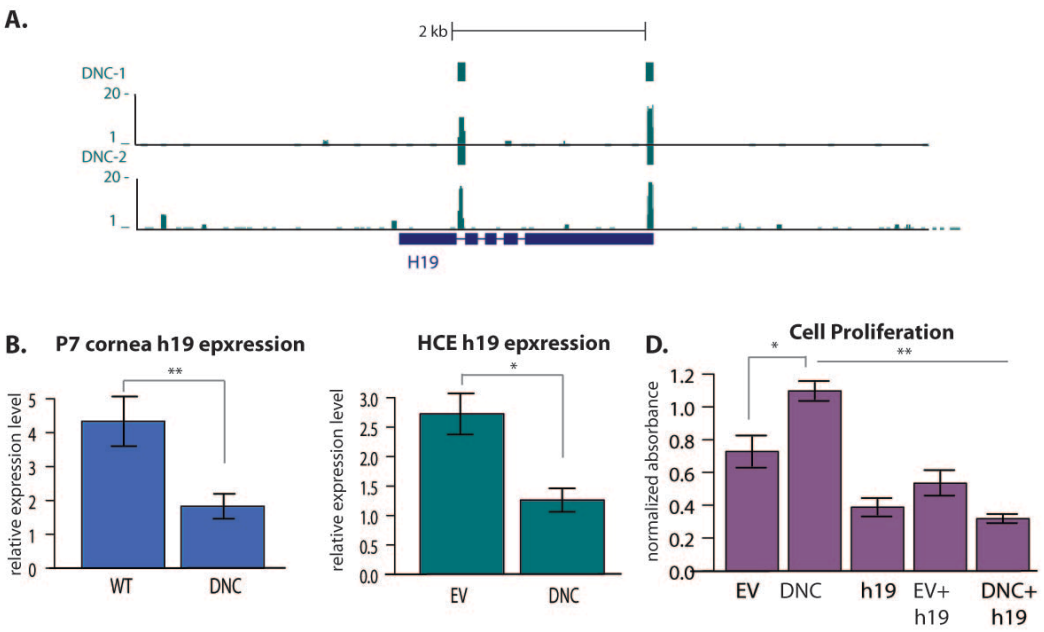


Figure 2-8. DN-CLIM regulates noncoding RNA H19. A) View of DN-CLIM ChIP-Seq peaks and wig files at H19 gene. B) Expression of H19 in WT and DNC corneas at P7. n=4 WT, 6 DNC. C) Expression of H19 in empty vector (EV) and DN-CLIM (DNC)- transfected human corneal epithelial cells (HCE), n=6 EV, 4 DNC. D) MTT cell proliferation assay for cells transfected with indicated constructs, n=6 per condition.
* $p < .05$, ** $p < .01$.

Chapter 3. KLF7 regulates the corneal epithelial progenitor cell state acting antagonistically to KLF4

Introduction:

During both tissue development and homeostasis, distal regulatory regions called enhancers, which contain binding sites for numerous transcription factors, help to control and coordinate gene expression in a temporal and spatial specific manner. Unlike the gene regulation that occurs at proximal promoters, these enhancer regions can be situated at long distances from the transcriptional start site of the gene they regulate, and are thought to physically interact with the gene through chromatin looping. Active enhancers can be identified through chromatin signatures correlated with enhancer function. Whereas active histone modifications like H3K4me3, H3K9ac, and H3K27ac mark promoters, enhancers have high levels of H3K4me1 and lower levels of H3K4me3 [43]. More recently, active enhancers have also been shown to be marked by high levels of H3K27ac, and inactive or repressed enhancers are enriched for H3K27me3 [38].

Within the library of enhancers in each cell type, a small group of enhancers exist that are unusually long (more than 12.5kb) and show extremely high intensity of transcription factor binding and chromatin enhancer signature marks. These regions, termed super enhancers, are frequently tissue specific and are thought to regulate the unique gene expression programs in each tissue. Within an organism, many different cell types share function and lineage history; epithelial cells are one such example. The extent of shared super enhancer usage in these cell types is not well understood. Additionally, the mechanisms by which these super enhancers integrate with typical enhancers in a cell to effect gene regulation remain to be defined.

The cornea epithelium provides an excellent system to study the process of enhancer mediated gene regulation during tissue development and homeostasis; it is a stratified squamous epithelium that undergoes constant cell turnover through the sloughing-off of cells from the surface and replenishment by the proliferating limbally-located cells [98, 99]. As these cells divide and move toward the center of the cornea and up to the surface of the tissue, they activate a terminal differentiation gene expression program. Such a process requires very precise regulation of gene expression to balance proliferation and differentiation and maintain a functioning epithelium.

To study the regulatory mechanisms of this balancing act, we identified enhancers in primary human corneal epithelial cells, through histone modification ChIP-Seq. These data provide a view of the complex regulatory dynamics active in proliferating cells; enhancer regions point to shared mechanisms of regulation among different epithelial cell types, while also highlighting cell type specific regulatory regions that control uniquely corneal functions. In addition, these data provide insight into corneal disease, as GWAS-associated SNPs lie within corneal enhancers, and cause alterations in the functionality of these regions. Motif analysis highlights the important role of KLF transcriptional regulators in cornea epithelium, and led to the discovery of a novel and previously undescribed role for KLF7 in cornea epithelial progenitor cells. We also defined a novel mechanism of KLF regulation: KLF7 acts to prevent corneal epithelial differentiation through competition for binding at epithelial differentiation genes and enhancers with KLF4, a pro-differentiation factor. The expression levels of these factors and their competing functions are crucial to the balancing of proliferation and differentiation during cornea epithelial development and homeostasis.

Results:

ChIP-Seq with histone modifications H3K4me1, H3K27ac and H3K4me3 identifies active enhancers in human corneal epithelial cells.

To gain insight into the regulatory landscape in corneal epithelial cells, we performed ChIP-Seq with antibodies to the histone marks H3K4me3, H3K4me1, and H3K27ac in primary human corneal epithelial cells (HCE). Using previously defined parameters; enhancers were identified as having high levels of H3K27Ac and H3K4me1 with relatively low levels of H3K4me3 [38]. Due to their importance in cell type specific functions, we also identified super enhancers using the H3K27ac mark. We identified 1,154 super enhancers (SE) in corneal epithelial cells, and 12,424 typical enhancers (TE), as well as 76,205 inactive regulatory regions, marked by H3K4me1 alone (Figure 3-1A).

As is characteristic of enhancers, the majority of these regions were found in intergenic and intronic regions of the genome rather than in proximal promoters (Figure 3-1B). While typical enhancers show a trend toward being localized near genes, super enhancers are even more strongly enriched near genes; most are found within 50kb of a transcriptional start site (TSS).

Both typical (TEs) and super enhancers (SEs) were located near key corneal epithelial identity genes, including transcription factor PAX6, a master regulator of sensory organ development. A super enhancer is located directly over the PAX6 gene body, and four typical enhancers are found within the upstream and downstream regions that have been linked to regulation of PAX6 expression [100], suggesting that both TEs and an SE regulate this gene (Figure 3-1C).

SE –linked genes are involved in regulation of apoptosis, epidermal development (encompassing many epithelial factors), hemidesmosome assembly, and regulation of fibroblast proliferation (Figure 3-1D). TE linked genes were enriched for similar categories (Figure 3-1D). Both TEs and SEs fall near highly expressed corneal epithelial genes, SEs more significantly so than TEs (Figure 3-1E), again highlighting the association of SEs with high levels of gene expression. While TEs may be more general regulators of genes expressed in numerous tissues, the genes linked to TEs in HCE are still very important to corneal and epithelial function, as evidenced by the fact that mouse phenotypes linked to these genes include abnormal corneal morphology, corneal opacity, and fused cornea and lens, in addition to numerous epidermal abnormalities (Figure 3-2).

Motif analysis identified many similar transcriptional regulators predicted to bind the two types of enhancers, with only subtle differences between the two. Motifs enriched in both SEs and TEs include AP1 family members, KLF, and ETS factors, all known to have important roles in corneal epithelial development (Figure 3-1F). Interestingly, TEs showed a low but significant enrichment for FOX family motifs that was not found in SEs (Figure 3-1F).

We next compared our enhancer data for corneal epithelial cells with the publicly available ChIP-Seq data from the ENCODE project for histone modifications in four additional primary cell types: mammary epithelium (HMEC), epidermal keratinocytes (NHEK), lung fibroblasts (NHLF), and skeletal muscle (HSMM) [101]. When clustering the TEs, we found that the epithelial cell types clustered together, indicative of their related functions and gene expression; within the epithelia, NHEK and HMEC cells clustered most closely (Figure 3-3A).

Analysis of super enhancers revealed similar relationships among the tissue types, with HMEC and NHEK the most similar. Interestingly, there was less similarity overall between cell

types, possibly indicative of the unique locations and roles of SEs in each tissue due to their association with tissue-specific functions (Figure 3-3A). Additionally, HCE no longer clustered with HMEC and NHEK, suggesting SEs in cornea are more distinct from SEs in HMEC and NHEK (Figure 3-3A). Comparing SEs between different epithelial cell types also shows that HMEC and NHEK share more overlapping SEs than either one does with HCE (Figure 3-3C). We next examined the relationship between super enhancers and typical enhancers among the different epithelial cell types and found that only about a third of super enhancers are shared between two or more epithelial cell types. However, SEs in one cell type are frequently TEs in another, suggesting that the majority of epithelial regulatory regions are common to multiple epithelial cell types, differing only in the levels of enhancer activity (Figure 3-3D).

To explore the active epithelial enhancers further, we identified overlapping and unique typical enhancers for each epithelial cell type. Gene ontology (GO) analysis for the genes near the TEs unique to each epithelial cell type revealed potential functional differences. While GO categories related to epithelial or ectodermal development appeared in many cases, genes involved in the cellular response to hypoxia were only enriched around uniquely corneal epithelial enhancers, perhaps reflective of the importance of dealing with hypoxic stress in this avascular tissue (Figure 3-3E). GO analysis for genes found near only NHEK enhancers revealed a group of actin cytoskeleton genes, which may reflect the importance of resistance to mechanical stress in the function of the skin (Figure 3-4A). The genes near HMEC unique enhancers were enriched for the GO category steroid metabolic process, which could be related to the hormone signaling that occurs in this reproductive tissue (Figure 3-4A).

Because SEs are thought to be more cell-type specific than TEs, we next looked at the unique SEs for each cell type. Using the SEs that did not overlap SEs or TEs in another epithelial

cell type, we found that unique HCE SEs were located near genes involved in sensory organ development and cell signaling (Figure 3-3F). SEs unique to NHEK were found near clusters of histone genes, as well as near the HOXA and HOXC gene clusters (Figure 3-4B), suggesting these SEs act as sites for coordinated regulation of genes within the cluster. Cell adhesion was the most enriched category for genes near HMEC-unique SEs; HMEC SE-linked genes were also found near histone clusters (Figure 3-4B). *De novo* motif analysis on the unique corneal epithelial SEs and TEs revealed that the motif for the ETS family member EHF was enriched in both unique SEs and TEs (Figure 3-3G, 3H). TEs unique to HCE also were enriched for the retinoic acid receptor alpha (RXRA), whereas SEs unique to HCE were enriched in AP1 and KLF motifs, further underscoring the important role for ETS and KLF family members in cornea epithelium.

The identification of KLF family motifs in both typical and super enhancers in cornea epithelium is not surprising, given the well defined roles of KLF4 and KLF5 in the cornea [8, 9, 102, 103]. However, the KLFs are a large family of transcription factors, many of which are highly expressed in both mouse and human cornea epithelium, and we wanted to understand the functions of these additional KLF family members, including how multiple KLFs could potentially interact to regulate gene expression within the same cell type. Using a previously published cornea transcriptome profiling gene expression changes over the lifetime of the mouse [7], we identified two significant patterns of KLF family member expression in the cornea: the majority of KLF factors, including KLF4, KLF5, and KLF6 showed a strong increase in expression across cornea development. In contrast, KLF7 and KLF12 were initially highly expressed, but decreased significantly during cornea development (Figure 3-5A).

Based on the opposing expression patterns, we hypothesized that KLF7 and KLF12 could be working in cornea development to maintain the progenitor state, acting antagonistically to pro-differentiation factors like KLF4, whose increase in expression parallels cornea terminal differentiation. We first established that a similar expression pattern was observed in human corneal epithelial cells in culture: KLF7 was more highly expressed in proliferating cells than differentiating, and KLF4 was more highly expressed in differentiating cells (Figure 3-7B). To test the role of KLF4 and KLF7 in corneal epithelium, we performed siRNA knockdown experiments in proliferating HCE with KLF7 or KLF4 and assessed gene expression changes by microarray. We found a significant difference in the proportion of upregulated and downregulated genes between siKLF7 and siKLF4: a significantly larger proportion of genes are upregulated when KLF7 is knocked down, indicating that KLF7 acts to repress a large number of genes in proliferating cornea epithelium; in contrast, a significantly larger proportion of genes are downregulated by siKLF4, pointing to an important role for KLF4 in gene activation (Figure 3-7C). Genes upregulated by knockdown of KLF7 include cornea differentiation genes *Pax6* and *Aldh3a1* (Figure 3-7D), further supporting the role of KLF7 as a progenitor factor that acts to repress the terminal differentiation program in cornea epithelium. We also identified many cell cycle regulators that are misexpressed in siKLF7 cells (Figure 3-7D), indicating KLF7 may also modulate the proliferation dynamics of progenitor cells. In comparing the two knockdown experiments, we found that KLF7 and KLF4 regulate similar functional categories of genes, but in the opposite direction; genes that are downregulated by siKLF7 fall into categories that contain genes upregulated by siKLF4, and vice versa (Figure 3-5B, 5C). Genes upregulated by siKLF7 include many epithelial differentiation factors, and epithelial differentiation is also a functional category for genes downregulated by siKLF4. Immune response and wounding

response were common categories for genes upregulated by siKLF4 and downregulated by siKLF7. We also found a significant overlap in specific genes regulated in opposite directions by KLF7 and KLF4 (Figure 3-5D, 5E).

We next performed the same experiment in cells that had been induced to differentiate. We did not observe an opposing expression pattern for genes regulated by KLF7 and KLF4 in differentiating corneal epithelial cells; however, such a relationship existed between proliferating and differentiating corneal epithelial cells. There was a significant overlap in genes downregulated by siKLF4 in differentiation that were upregulated by siKLF7 in proliferating cells, and genes upregulated by siKLF4 in differentiated and downregulated by siKLF7 in proliferating (Figure 3-6A). In differentiating cornea epithelial cells, we also observed a difference in the proportion of upregulated and downregulated genes in each siRNA experiment. In contrast to differentiation, siKLF7 was associated with significantly more downregulated genes, and siKLF4 with more upregulated genes, suggesting a shift in regulatory mechanisms for the two factors (Figure 3-7C)**.** Interestingly, many other KLF family members are affected by knockdown of KLF4 or KLF7: KLF12 is downregulated by siKLF7, and KLF10 and KLF5 are upregulated by siKLF4, suggesting auto regulation of family members is an important part of KLF function in cornea epithelium.

Due to the strong enrichment for genes in functional categories related to cell cycle in the KLF7 knockdown experiments in proliferating cells, we next wanted to understand the role of KLF7 in regulation of cell proliferation in cornea epithelium. MTT assays revealed a significant increase in proliferation in undifferentiated HCE when KLF7 is knocked down, suggesting KLF7 may act to temper proliferation and maintain progenitor cell quiescence (Figure 3-6B). This is

consistent with CFSE analysis, which showed that siKLF7 cells underwent more divisions than control cells (Figure 3-6C).

To investigate the functional role of enhancers in disease in corneal epithelium; we overlapped our active enhancer data for corneal epithelial cells with single nucleotide polymorphisms (SNPs) that have been linked to diseases through GWAS studies. Two cornea-specific enhancers overlap SNPs associated with disease phenotypes in cornea: rs3815087, a SNP on chromosome 6 associated with Stevens-Johnson syndrome [104], an immune-mediated condition that affects both skin and cornea, and rs6758183, a SNP on chromosome 2 linked to corneal astigmatism [105](Figure 3-8A). rs3815087 is located in an exon of *Psors1c1* (Figure 3-8A, 3-9A)**,** a gene identified as a psoriasis susceptibility locus, however, evidence points to this SNP being non-coding; the exon in which it is located is not translated in detected isoforms. Additionally, the enhancer histone modification pattern suggests that this SNP may affect disease risk through non-coding mechanisms. Even though enhancers are often thought of as distal regulatory regions, many have been found to directly overlap gene exons. rs6758183, linked to corneal astigmatism, is located in a gene desert, approximately 500kb from the nearest coding gene (Figure 3-8B). We cloned these regions into luciferase reporter vectors, and validated that both regions showed enhancer function (Figure 3-8C). We next mutated the SNP nucleotide to the disease-linked variant and compared enhancer activity to the WT variant. The region containing the disease-linked allele of rs3815087 showed significantly reduced enhancer activity, compared to the WT allele. In contrast, the region containing the disease-linked variant of rs6758183 showed significantly increased enhancer activity compared to the WT allele (Figure 3-8C). These results suggest that alterations in cornea enhancer function caused by the SNP variants could contribute to the risk of disease at these loci.

We hypothesized that SNP variants could disrupt motif recognition and transcription factor binding, leading to differences in enhancer activity for each allele. To study this, we identified transcription factor binding motifs overlapping each SNP that show significant differences in motif scores between WT and disease-linked alleles. At rs6758183, the disease associated A allele increased the strength of a SMAD motif, and decreased the strength of an ETS motif (Figure 3-8D), the disease associated allele of rs3815087 strengthened a CEBPB motif, among others (Figure 3-9B). The Ets factor EHF is highly expressed in cornea epithelium and has been shown to be an important regulator of this tissue [7], making it a candidate for binding to the disrupted Ets motif. Using ChIP, we found that EHF binds to rs6758183, suggesting disruption of EHF binding by the SNP could result in the observed differences in enhancer activity (Figure 3-8C). Intriguingly, the WT version of the SNP shows less enhancer activity than the mutant, suggesting EHF binding acts to repress the enhancer region. When this binding is lost, the enhancer becomes overactive, and could result in growth imbalances that cause alterations in the curvature and refractive power of the cornea.

Since rs6758183 is located in a gene desert, we looked for studies mentioning the surrounding area to this SNP. We found that a number of SNPs in the vicinity of rs6758183 have been linked to type II diabetes. Additionally, ChIA-PET studies in MCF7 cells found that this region (and the diabetes-linked SNPs) looped to form contact with the IRS1 promoter, 500kb away [69]. Insulin signaling is also very important for growth and development of the cornea, both epithelium and stroma, making IRS1 a strong candidate target for the enhancer containing rs6758183. Furthermore, in previously published siRNA data for EHF in corneal epithelial cells, we find that knockdown of EHF causes a small but significant increase in IRS1 gene expression (Figure 3-9C) [7]. Combined, these data suggest EHF may regulate IRS1 through binding to the

rs6758183 enhancer; disruption of the EHF motif causes a reduced affinity of EHF for this site, likely resulting in reduced EHF binding, and increased enhancer activity, potentially causing over expression of IRS1, which could disrupt the precise balance required for corneal growth that results in proper light refraction.

Discussion:

Our identification of enhancers in cornea epithelium has provided important insights into the mechanisms and factors involved in the dynamic regulation of epithelial tissue development and homeostasis. We have defined roles for both SEs and TEs in these processes, and characterized the regulatory relationships between epithelial cell types, helping further define the factors involved in cell identity. In combination with data from other sources, we have characterized GWAS-associated SNPs that lie within corneal enhancers, identifying factors with a potential role in disease risk and defining previously uncharacterized regulatory relationships. We have also defined a role for a novel KLF factor in cornea epithelium, and characterized the regulatory dynamics between KLF7 and pro-differentiation factor KLF4. Our work highlights the importance of regulating the balance between proliferation and differentiation, both for proper tissue development and for homeostasis.

Our identification and analysis of corneal epithelial enhancers has led to a number of interesting findings. In comparing SEs to TEs, we have confirmed that, as in other tissues, HCE SEs tend to show more tissue-specific function than TEs. Consistent with results for SEs in other cell types, we find that SEs are more positively correlated with gene expression, and are located near genes more strongly enriched in corneal epithelial functional categories. However, there is also a great amount of evidence that TEs are very important for gene regulation in cornea

epithelium. Genes near TEs are more strongly linked to corneal phenotypes in mice than genes near SEs. We also found both an SE and TEs in the more than 200kb upstream and downstream of the PAX6 TSS that is required for proper expression of this gene, suggesting that both types of enhancers can act together to regulate the same gene. It is possible that SEs, with their highly enriched transcription factor binding, provide the largest component of transcriptional activation, with TEs providing the fine tuning mechanism, allowing for more precise and controlled gene expression.

While SEs seem to be more involved in tissue specific regulation, the vast majority of SEs in HCE are either TEs in other epithelial cell types or show marks of regulatory potential. Epithelial tissues express many of the same genes, and appear to have very similar regulatory landscapes. The main differences are in the level of intensity of each enhancer; this suggests that rather than a binary ON-OFF determination, these regulatory regions that can be TEs or SEs depending on the cell type may serve to modulate the levels of expression of epithelial genes; providing the right dose of each for a given epithelial tissue.

Evidence for functionality of our identified enhancers comes in part from the finding that two GWAS-linked SNPs with corneal disease phenotypes overlap HCE TEs, and that the single base substitution is sufficient to cause changes in enhancer activity at these loci. It is interesting that GWAS SNPs linked to corneal disease are located in TEs instead of SEs; this supports the idea that TEs also have an important role in gene regulation. It is also possible that GWAS SNPs in cornea epithelium are found TEs but not in SEs because mutations in super enhancers have a higher chance of being very deleterious to the organism and are quickly selected against.

The finding that many GWAS-linked SNPs occur outside of coding regions led to the hypothesis that these variants fall within enhancers and other key regulatory regions, affecting

gene expression levels, rather than causing mutations in protein sequence. Our data supports this idea, as we found SNPs linked to corneal diseases overlap active enhancers in HCE, and that the single base difference causes a change in enhancer activity. A challenge to understanding the disease promoting mechanism for these noncoding SNPs, is that enhancers can be located at a great distance from the genes they regulate, making the link between gene expression and SNP variant difficult to decipher. Using a combination of data from published ChIA-PET studies, motif analysis, and siRNA experiments, we have been able to establish a role for ETS factor EHF in binding to the SNP region and regulating nearby gene IRS1. This is the first such study for a corneal-disease linked SNP, and it demonstrates the importance of integrating data from multiple sources in elucidating mechanisms of disease risk and progression.

Both cornea epithelium and stroma contribute to the refractive power of the eye and the stroma is primarily implicated in corneal astigmatism. While our studies were carried out in epithelial cells, this enhancer may also be active in stromal keratinocytes, leading to alterations in insulin signaling in both tissues that could underlie corneal astigmatism risk. Additionally, extensive signaling occurs between epithelium and stroma, and alterations in IRS1 expression in one tissue could lead to downstream growth effects in both.

Motif analysis highlighted the important role of KLF transcriptional regulators in cornea epithelium, and led to the discovery of a novel and previously undescribed role for KLF7 in cornea epithelial progenitor cells. KLF7 regulates neuronal and olfactory development [35], and has been characterized as an inhibitor of adipocyte differentiation [34, 35]. We found that KLF7 works to repress the corneal epithelial differentiation program and also reduces proliferation, potentially acting to maintain progenitor cells in a quiescent state. We also defined a novel mechanism of KLF7 regulation in corneal epithelium, which may apply to KLF7 regulation in

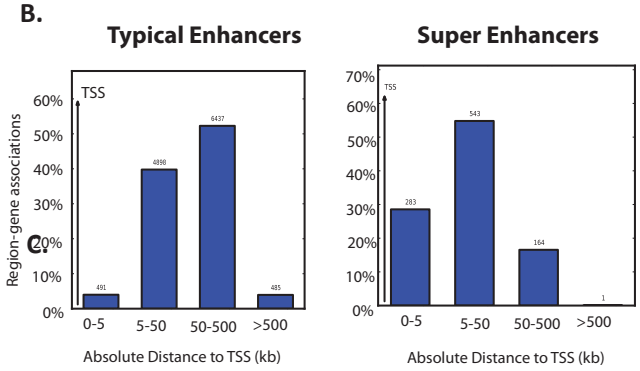
other tissues as well. We found that KLF7 acts to maintain the corneal progenitor state in part through antagonistic regulation of KLF4 gene targets. The expression levels of each KLF determine the transcriptional state of shared target genes; as KLF7 levels drop and KLF4 expression increases during differentiation, it is likely that KLF4 is able to out-compete KLF7, and change the response of shared targets. It appears that the expression levels of these factors and their competing functions are crucial to the balancing of proliferation and differentiation during cornea epithelial development and homeostasis. KLF4 and KLF7 are only two of the many KLF family members expressed in cornea epithelium. In the future, it will be interesting to understand both the unique roles of each factor, and how they work together to accomplish gene regulation.

Figure 3-1. Identification of Typical (TE) and Super Enhancers (SE) in primary corneal epithelial cells (HCE). A) Quantification of regulatory domains in HCE. B) Distribution of TE and SE around TSS. C). TE and SE near corneal identity regulator PAX6. D). Gene ontology (GO) for genes linked to TE and SE. E) GSEA analysis of TE and SE across HCE expressed genes (ranked high to low). F) Top enriched motifs in TE and SE.

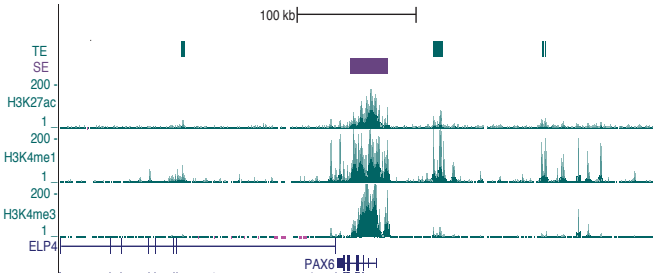
A.

Chromatin Domain	Typical Enhancer (TE)	Super Enhancer (SE)	H3K4me1
Number of Peaks FDR <.05	12424	1154	76,205

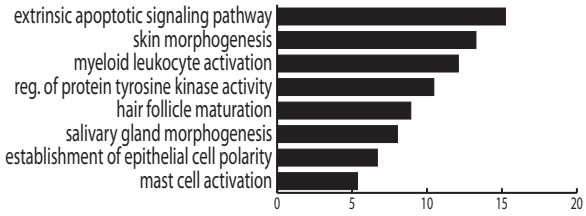
B.

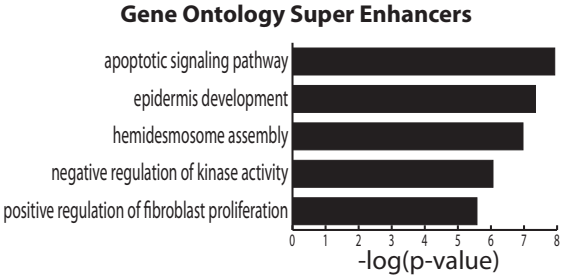


C.

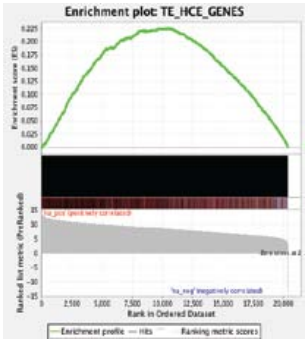


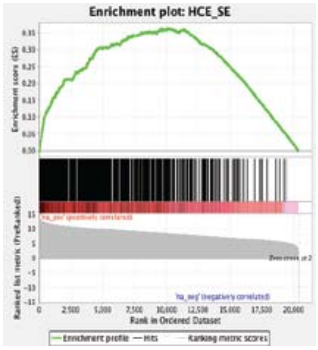
D.





E.





F.

Super Enhancer	Typical Enhancer
ATF3	FOS/AP1
ETS	ETS
KLF	KLF
FOS/AP1	FOXJ3
SRF	SRF

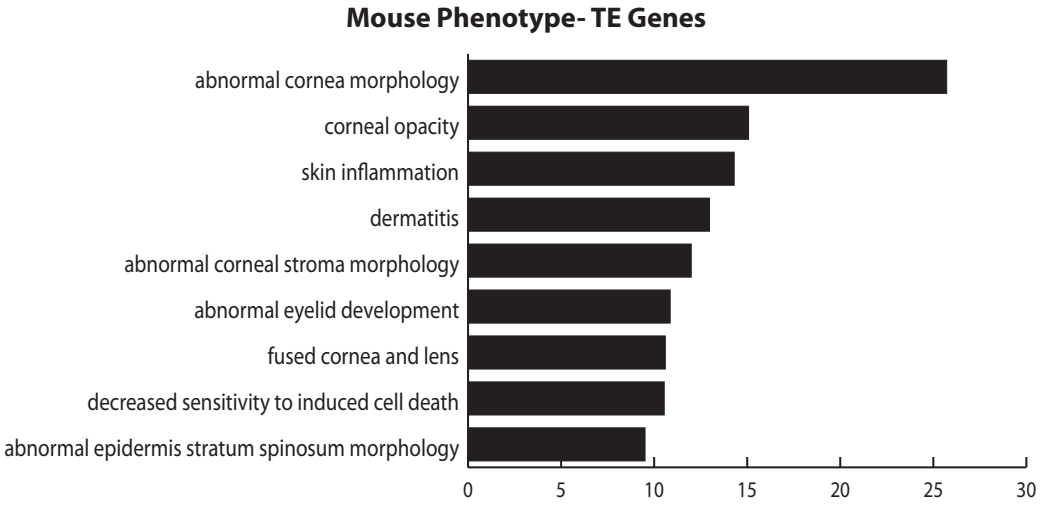
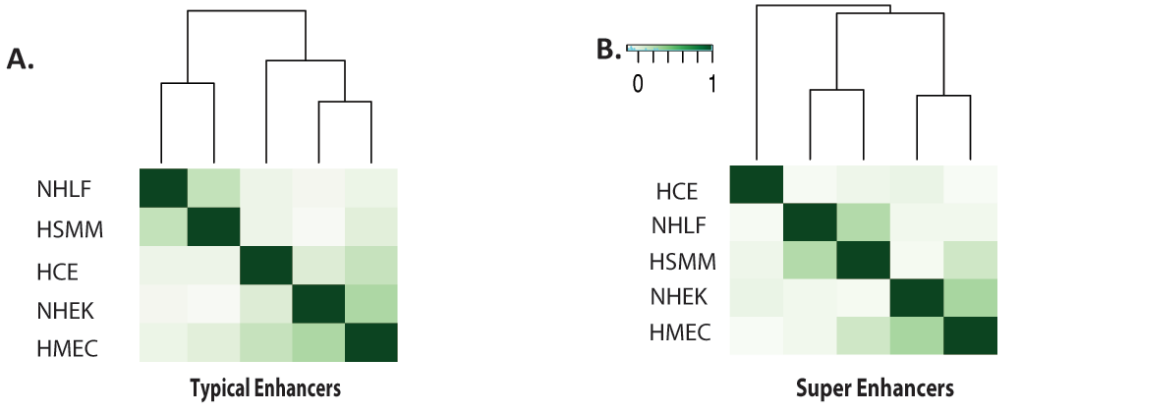
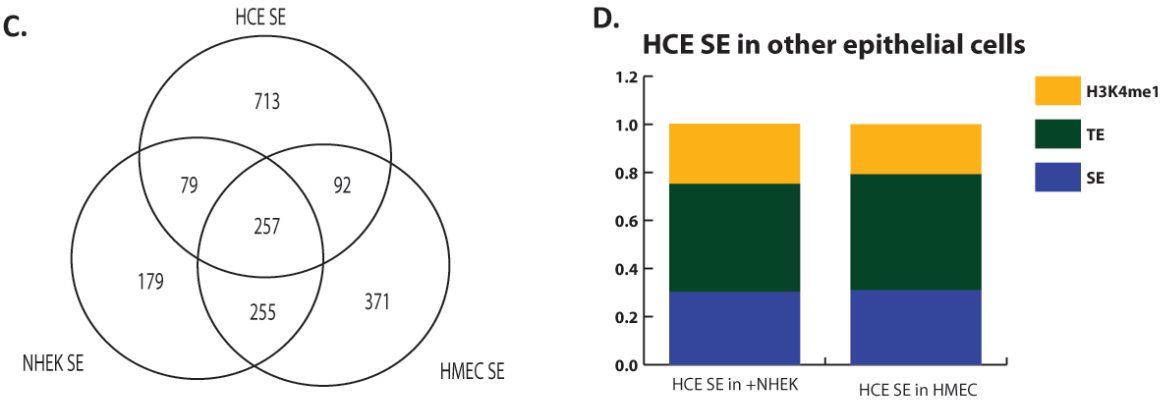
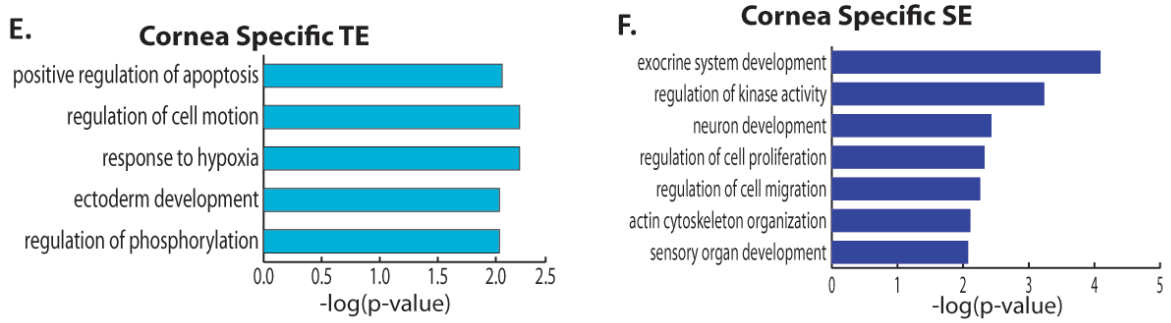


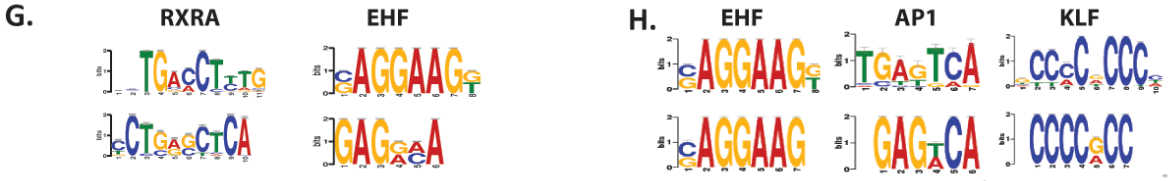
Figure 3-2. Corneal epithelial TE are linked to genes associated with corneal and epithelial phenotypes in mice. Mouse phenotype gene ontology analysis for nearest genes to HCE TE.

Figure 3-3. Comparison of regulatory regions between HCE and ENCODE cell types. A) Clustering of primary cells based on similarities in TE. B) Clustering of primary cells based on similarities in SE. C) Assessment of regulatory cell state of HCE SE regions in other epithelial cells (HMEC, NHEK). D) Overlap of SE between different epithelial cell types. E) Gene ontology for genes linked to cornea-specific TE. F) Gene ontology for genes linked to cornea-specific SE. G). Enriched motifs in cornea-specific TE. G) Enriched motifs in cornea-specific SE.


A.
NHLF
HSMM
HCE
NHEK
HMEC
Typical Enhancers
B.
0
1
HCE
NHLF
HSMM
NHEK
HMEC
Super Enhancers


C.
HCE SE
713
79
92
257
179
371
255
NHEK SE
HMEC SE
D.
HCE SE in other epithelial cells
H3K4me1
TE
SE
1.2
1.0
0.8
0.6
0.4
0.2
0.0
HCE SE in +NHEK
HCE SE in HMEC


E.
Cornea Specific TE
positive regulation of apoptosis
regulation of cell motion
response to hypoxia
ectoderm development
regulation of phosphorylation
0.0
0.5
1.0
1.5
2.0
2.5
-log(p-value)
F.
Cornea Specific SE
exocrine system development
regulation of kinase activity
neuron development
regulation of cell proliferation
regulation of cell migration
actin cytoskeleton organization
sensory organ development
0
1
2
3
4
5
-log(p-value)


G.
RXRA
EHF
H.
EHF
AP1
KLF

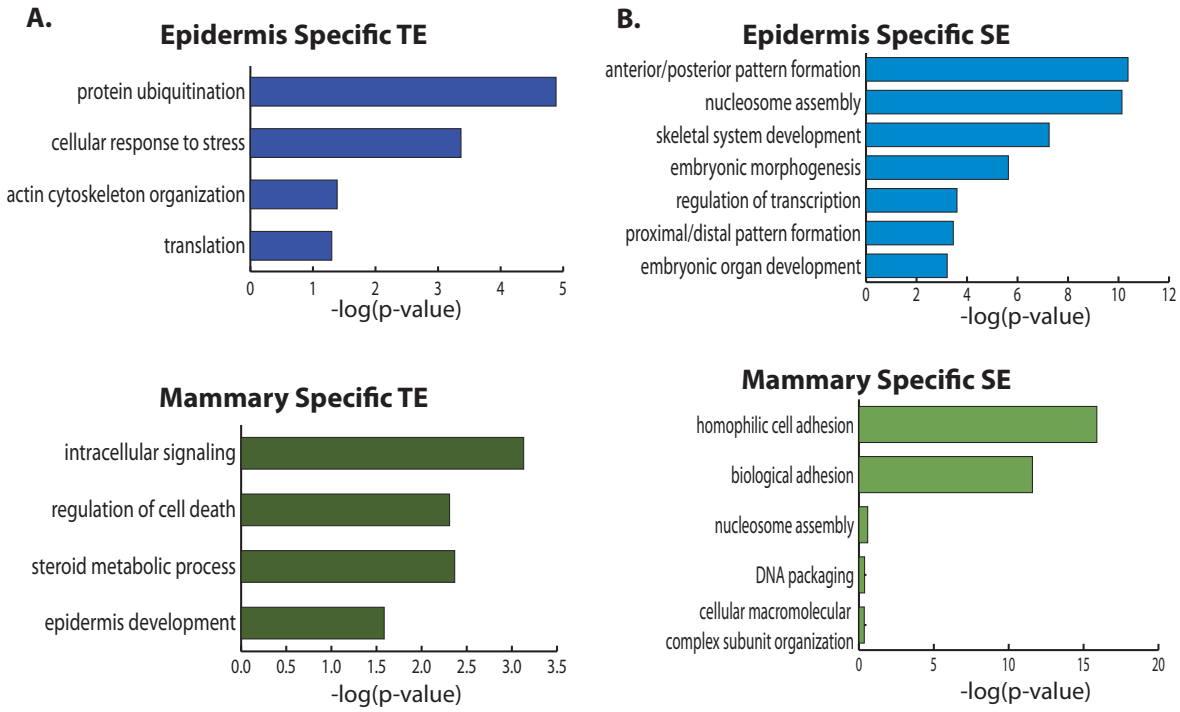


Figure 3-4. Unique SE and TE demonstrate important properties of each cell type. A) Enriched gene ontology for epidermis-specific TE. B) Enriched gene ontology for epidermis-specific SE. C) Enriched gene ontology for mammary epithelial-specific TE. D) Enriched gene ontology for mammary epithelial-specific SE.

Figure 3-5. KLF7 and KLF4 have opposing expression patterns and regulate corneal epithelial genes antagonistically. A) KLF4 expression in the cornea across the mouse lifespan: embryonic day (E) 14.5 to 2 years (yr). B). KLF7 expression in the cornea across the mouse lifespan. B) Enriched gene ontology for genes upregulated (blue bars) or downregulated (red bars) by knockdown of KLF7 (siKLF7). C). Enriched gene ontology categories for genes upregulated (red bars) or downregulated (blue bars) by knockdown of KLF4 (siKLF4). D) Overlap between genes downregulated by siKLF7 and genes upregulated by siKLF4. E) Overlap between genes upregulated by siKLF7 and genes downregulated by siKLF4.

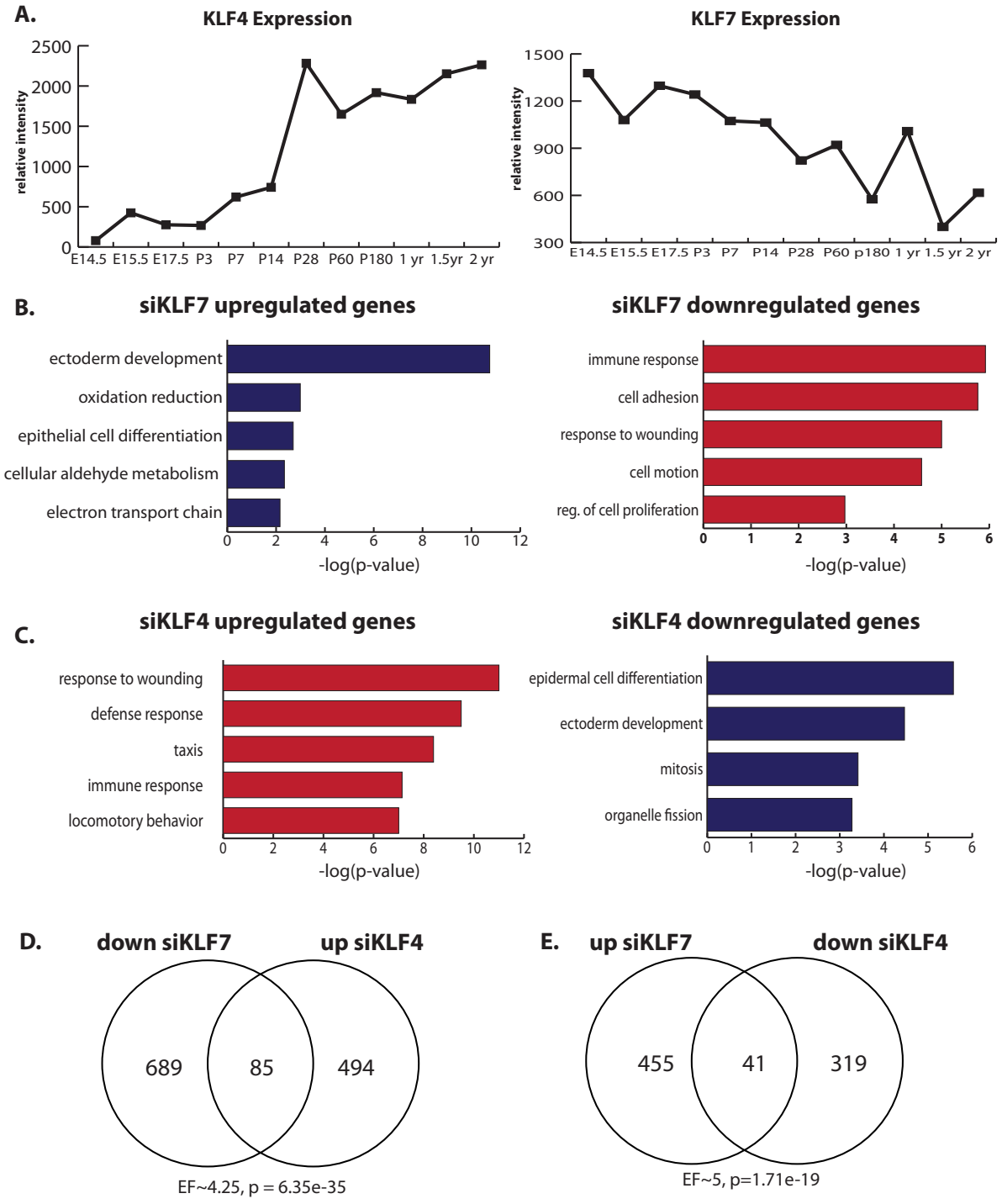
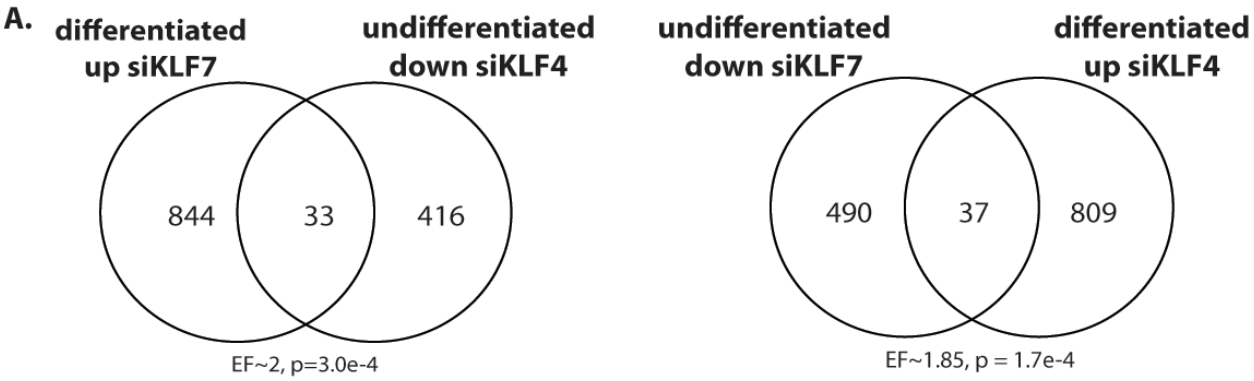
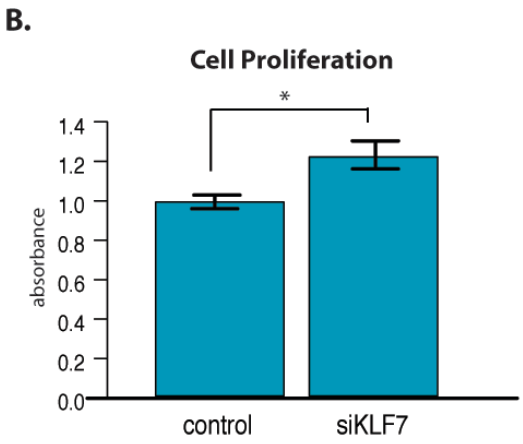
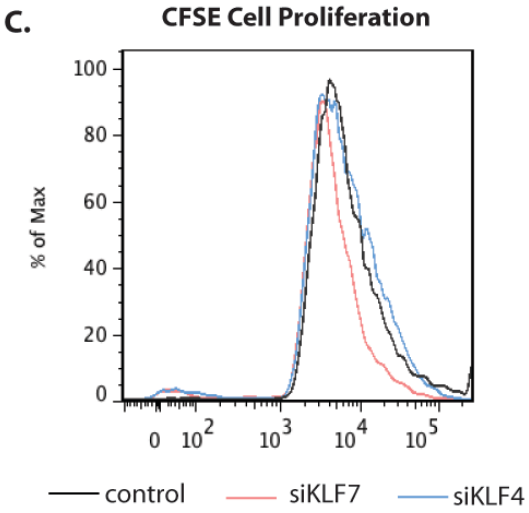
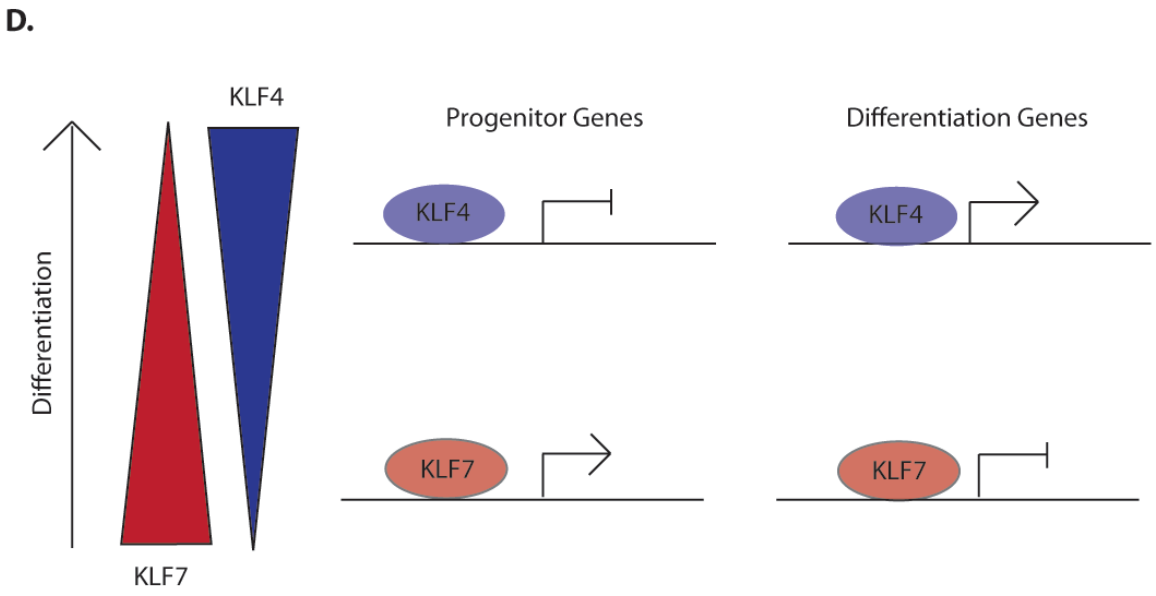

A.
KLF4 Expression
relative intensity
2500
2000
1500
1000
500
0
E14.5 E15.5 E17.5 P3 P7 P14 P28 P60 P180 1 yr 1.5yr 2 yr
KLF7 Expression
relative intensity
1500
1200
900
600
300
E14.5 E15.5 E17.5 P3 P7 P14 P28 P60 p180 1 yr 1.5 yr 2 yr
B.
siKLF7 upregulated genes
ectoderm development
oxidation reduction
epithelial cell differentiation
cellular aldehyde metabolism
electron transport chain
0 2 4 6 8 10 12
-log(p-value)
siKLF7 downregulated genes
immune response
cell adhesion
response to wounding
cell motion
reg. of cell proliferation
0 1 2 3 4 5 6
-log(p-value)
C.
siKLF4 upregulated genes
response to wounding
defense response
taxis
immune response
locomotory behavior
0 2 4 6 8 10 12
-log(p-value)
siKLF4 downregulated genes
epidermal cell differentiation
ectoderm development
mitosis
organelle fission
0 1 2 3 4 5 6
-log(p-value)
D.
down siKLF7
up siKLF4
689
85
494
EF~4.25, p = 6.35e-35
E.
up siKLF7
down siKLF4
455
41
319
EF~5, p=1.71e-19

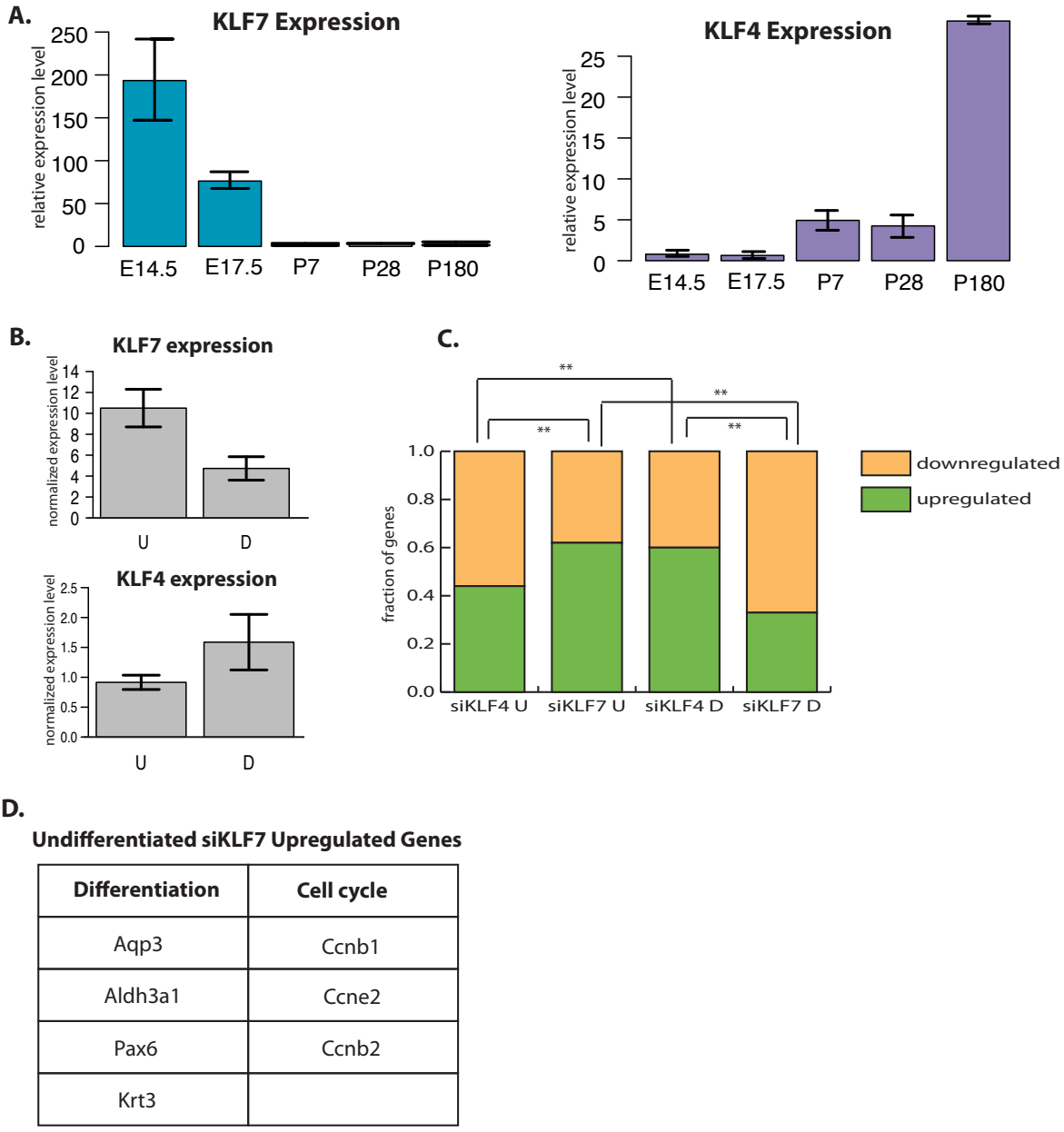
Figure 3-6. KLF7 and KLF4 also show an antagonistic regulatory relationship across differentiation, and KLF7 is a negative regulator of proliferation. A). Overlap between genes upregulated by siKLF7 in differentiating HCE (HCE-D) and genes downregulated by siKLF4 in undifferentiated HCE (HCE-U), overlap between genes downregulated by siKLF7 in HCE-U and upregulated by siKLF4 in HCE-D. B) MTT cell proliferation assays in HCE in siRNA control and siKLF7 treated cells, n=6. C) CFSE staining of siRNA control, siKLF4, and siKLF7 treated cells, n=1. D) Model for KLF7- and KLF4-mediated regulation of corneal epithelial differentiation.
* $p < .05$, ** $p < .01$.


A.
differentiated
up siKLF7
undifferentiated
down siKLF4
844
33
416
EF~2, p=3.0e-4
undifferentiated
down siKLF7
differentiated
up siKLF4
490
37
809
EF~1.85, p = 1.7e-4


B.
Cell Proliferation
*
absorbance
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
control
siKLF7


C.
CFSE Cell Proliferation
% of Max
100
80
60
40
20
0
0
10^2
10^3
10^4
10^5
control
siKLF7
siKLF4


D.
KLF4
Differentiation
KLF7
Progenitor Genes
KLF4
KLF7
Differentiation Genes
KLF4
KLF7



D.

Undifferentiated siKLF7 Upregulated Genes

Differentiation	Cell cycle
Aqp3	Ccnb1
Aldh3a1	Ccne2
Pax6	Ccnb2
Krt3	

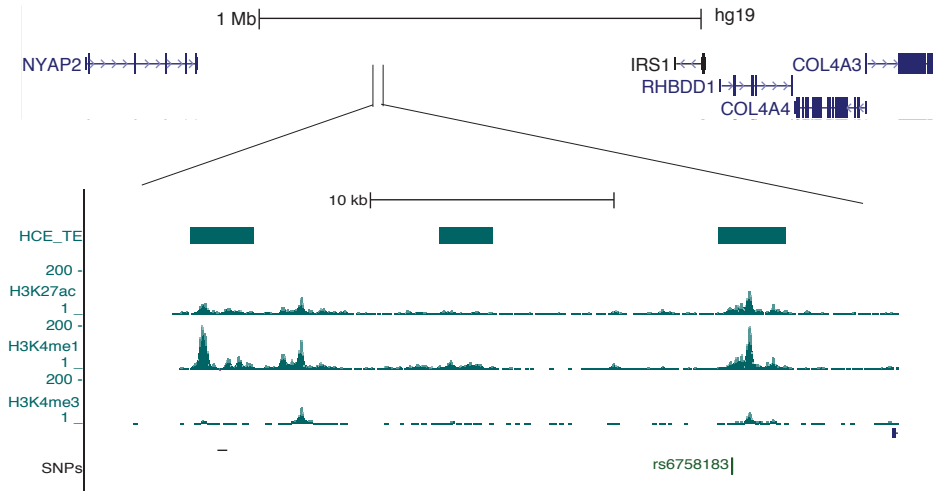
Figure 3-7. KLF7 and KLF4 antagonistically regulate HCE differentiation. A) qPCR validation of KLF7 and KLF4 expression in mouse whole cornea at selected time points across mouse lifespan, n=3 for E14.5, E17.5, P28. n=5 for P7, n=6 for P180. B) KLF7 and KLF4 expression in 48 hour differentiated HCE. C) Proportion of upregulated and downregulated genes by siKLF4 or siKLF7 in undifferentiated HCE (U) or differentiated HCE (D). D) Cell cycle and corneal epithelial differentiation genes upregulated by siKLF7.
* $p < .05$, ** $p < .01$.

Figure 3-8. HCE enhancers overlap SNPs linked to corneal disease. A). Table of GWAS SNP information. B) Genomic view of rs6758183 with HCE enhancer data. C) Luciferase reporter assays for WT and mut (disease-associated allele) SNPs, n=6. D) Analysis of motif disruption by SNP rs6758183, scores calculated by TRAP: differential score= mut score – WT score. E) EHF ChIP in HCE at rs6758183.
* $p < .05$, ** $p < .01$.

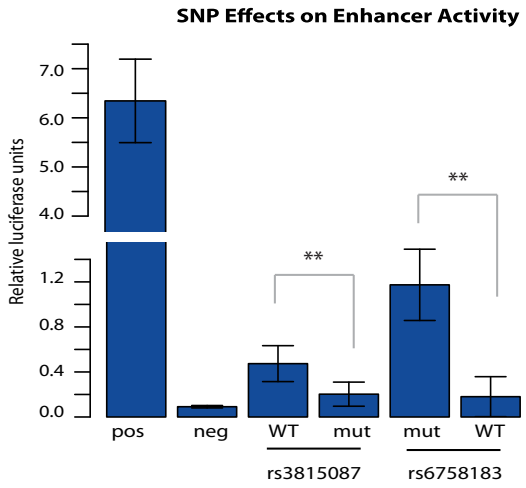
A.

SNP	GWAS Disease	Odds ratio	p-value	Nearest genes
rs6758183	Corneal Astigmatism	1.39	3x10^-6	*Irs1, Nyap2*
rs3815087	Steven's Johnson Syndrome	1.53	3x10^-7	*Psors1c1*

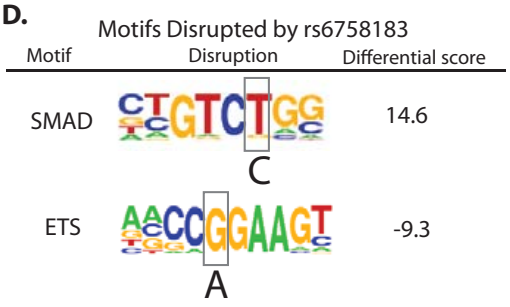
B.



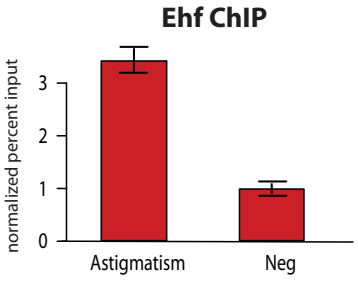
C.



D.



E.



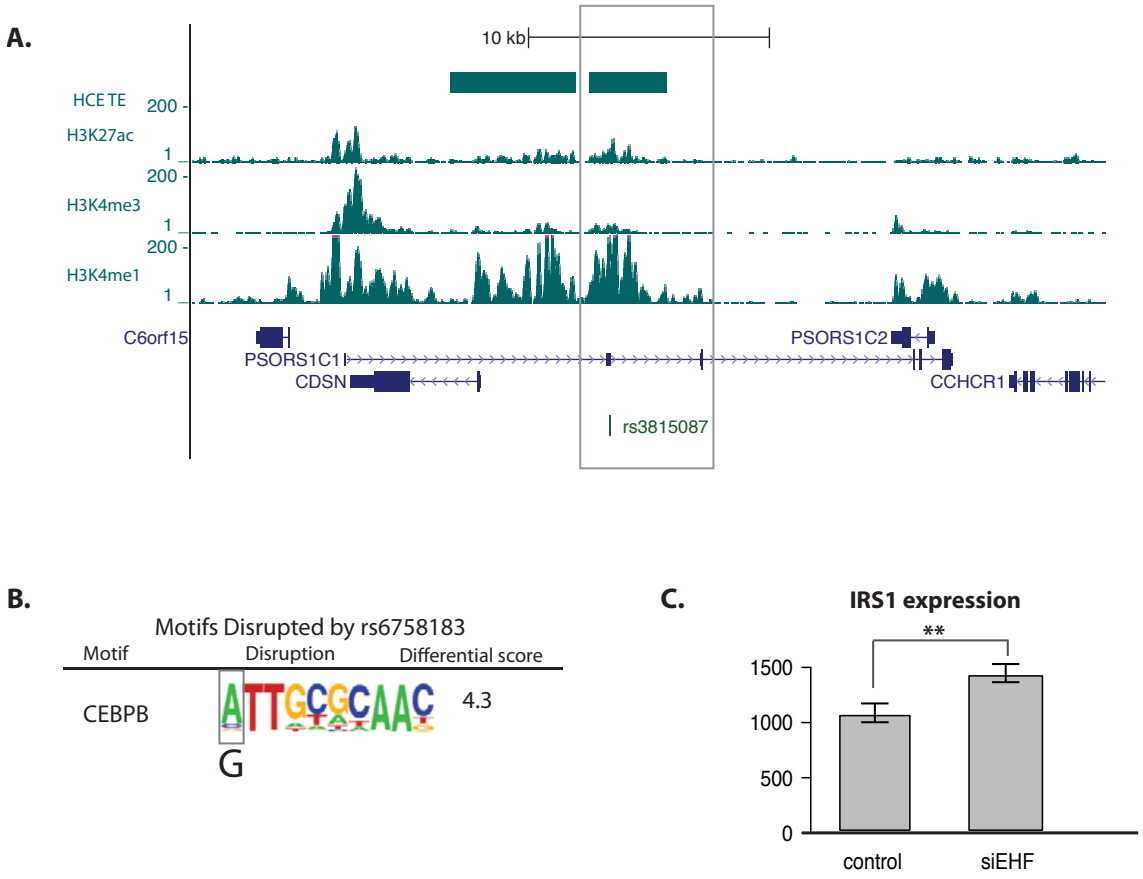


Figure 3-9. HCE enhancers overlap SNPs linked to corneal disease. A) Genomic view of rs3815087 with genes, HCE enhancers, and chromatin marks. B) Motif-disruption analysis for rs3815087, scores calculated by TRAP: differential score= mut score – WT score. C) IRS1 expression in siRNA control and siEHF treated HCE cells, n=2.
* $p < .05$, ** $p < .01$.

Chapter 4. GRHL3 interacts with super enhancers and the neuronal repressor REST to regulate keratinocyte differentiation and migration

Introduction

Stem cells are unique in their pluripotency and self-renewing potential, characteristics that require unique chromatin organization. As these cells divide and differentiate, they undergo dramatic shifts in gene expression, as well as large-scale re-organization of chromatin. The process of differentiation is tightly controlled, and such complex coordination of gene expression requires many layers of regulation. One such mechanism of regulation occurs through distal genomic regions called enhancers, which are believed to act as concentrating sites for transcription factors and other regulators, forming loops to contact promoters and enhance transcription. Active enhancers have been shown to have high levels of H3K4me1 and H3K27ac modified nucleosomes, with low levels of H3K4me3, while poised developmental enhancers have been shown to have H3K27me3 in place of H3K27ac [38]. Additionally, in each cell type, a select few enhancers have been found to be unusually long (more than 12.5kb) and show high intensity of transcription factor binding and chromatin enhancer signature marks [56]. These regions, termed super enhancers, are frequently tissue-specific and are thought to regulate the unique gene expression programs in each tissue. A great deal of work has been done to document the chromatin and transcriptional changes that occur as stem cells differentiate; however, whether similar mechanisms mediate the different functional states of a single cell lineage is not known.

The epidermal keratinocyte is an excellent example of a cell that undergoes such transitions. In the epidermis, progenitor cells reside in the basal layer; as they divide and move upward in the tissue, they progressively differentiate, losing proliferative potential, and acquiring the gene expression program required to form the epidermal barrier. After injury to the

epidermis, keratinocytes must activate a cell migration gene expression program and move into the wound, before differentiating to restore the barrier. Both migration and differentiation require large-scale changes in gene expression, however, the extent of chromatin changes between these states have not yet been characterized.

Grainyhead-like epithelial transactivator (GET1/GRHL3) is a transcription factor expressed in numerous epithelial tissues, and has been identified as a key regulator of epidermal terminal differentiation and barrier formation. *Grhl3* knockout mice die at birth, with numerous epidermal abnormalities including altered expression of adhesion proteins and changes in lipid composition, resulting in increased barrier permeability [26]. Intriguingly, GRHL3 also plays a role migration during of wound healing, an epidermal function distinct from barrier formation, which requires activation of coordinated movement of keratinocytes and suppression of the differentiation program. In this process, GRHL3 is expressed in keratinocytes at the leading edge of migrating cells in both embryonic and adult wounds [27], where it was shown to regulate genes involved in actin polymerization, possibly by affecting planar cell polarity pathways [28].

The role of GRHL3 in promoting both epidermal differentiation and cell migration provides an interesting opportunity to study the regulation of transitions between different keratinocyte cell states through the actions of a single transcription factor. We also sought to place GRHL3-mediated epidermal regulation in the context of the different chromatin states within a single cell type, defining an interaction between GRHL3 and chromatin in the process; characterization of super enhancers revealed an unexpectedly strong enrichment of GRHL3 binding events at these domains, pointing to the key role of this factor at the regulatory regions governing transitions between the different keratinocyte cell states. Further analysis of GRHL3 ChIP-Seq data revealed mechanistic differences between modes of GRHL3-mediated gene

regulation in differentiating and migrating keratinocytes, and identified a role for the transcriptional repressor REST in epidermal keratinocyte migration.

Results:

The changing landscape of active enhancers points to key regulatory mechanisms for the transition of keratinocytes between different cellular states.

To study in more detail the differences between keratinocytes as they transition between the different cellular states, we performed ChIP-Seq to identify active regulatory regions based on histone modification signatures in migrating and differentiating keratinocytes, and compared to data from proliferating keratinocytes from the ENCODE project [101]. Active typical enhancers were identified based on the presence of H3K4me1 and H3K27ac, and absence of high levels of H3K4me3. We identified approximately 20,000 typical enhancers in proliferating, 31,000 in differentiating, and 11,000 in migrating keratinocytes (Figure 5-1A). Due to their importance in cell type specific functions, we also identified super enhancers (SEs) using the H3K27ac mark. We identified 783, 761, and 363 SEs in proliferating, differentiating, and migrating keratinocytes, respectively (Figure 5-1A). 2144 typical enhancers (TEs) overlapped between all three conditions, while the majority of enhancers were active under only one or two conditions, pointing to differences in transcriptional regulation required in each situation (Figure 5-1B). Only 89 super enhancers were found in all 3 cell states, again pointing to the unique gene expression programs active under each cellular condition (Figure 5-1B).

Genes near typical enhancers active uniquely in one cell type are enriched for categories important to all epidermal cells, including keratinization and fatty acid and phospholipid synthesis (Figure 5-2A). In contrast, genes linked to SEs in each cell type are enriched for cell-

type specific functions (Figure 5-2B). In proliferating keratinocytes, where cells adhere to the basal lamina, super enhancers are located near genes important for hemidesmosome assembly. Super enhancers in proliferating keratinocytes are also located near important epidermal identity genes, including p63. SEs in differentiating keratinocytes are located at important differentiation regulators, including GRHL3, and ZNF750 (Figure 5-2C)**.** SEs in migrating cells are located at genes involved in actin filament regulation, and include FLNA, among others (Figure 5-2D).

To understand the scale of chromatin changes between different cellular states, we compared enhancers (both TEs and SEs) in migrating and differentiating keratinocytes to the chromatin state at these same regions in the proliferating cells from which they are derived. Approximately one third of active enhancers in differentiating or migrating cells are also active enhancers in proliferating keratinocytes (Figure 5-1C). H3K4me1 marks another third alone in proliferation, suggesting these regions, while not active, are already marked for future activation. A very small fraction of enhancers derive from the poised state in proliferating cells, and less than 10 percent do not show any histone marks associated with enhancers in the proliferating state (Figure 5-1C). These results suggest that the chromatin landscape is already well demarcated in proliferating keratinocytes, and that the poised enhancer chromatin state may be more important during development and lineage specification than within a single cell type transitioning between different states.

To examine how SEs change in the transition between different cellular states compared to the transition from embryonic stem cell (ESC) to keratinocyte progenitor cell (NHEK-P). Between ESC and NHEK-P, only 8 percent of SEs persist, 16 percent are converted to TEs, 43 percent lose active enhancer marks and only retain H3K4me1, and 32 percent of ESC SEs lose all enhancer chromatin marks in NHEK-P. In contrast, all SEs in NHEK-P retain some chromatin

marks in the transition to differentiating (NHEK-D) or migrating (NHEK-M) cells; the majority are converted to TEs in each case (Figure 5-1D). This suggests that the regulatory regions for all keratinocyte states are already defined in NHEK-P, and that changes during transitions between the different cell states primarily involve alterations in the intensity and activity of these regulatory regions, rather than the creation of new regulatory regions.

SNPs that are linked to disease often fall outside of coding regions of the genome; it is thought that these SNPs occur in key regulatory regions including enhancers, and that the polymorphism can disrupt transcription factor binding, leading to alterations in enhancer activity and subsequent gene expression. To determine if the identified enhancers in keratinocytes contained SNPs linked to disease, we overlapped them with GWAS SNP coordinate data. Three keratinocyte enhancers overlap SNPs linked to psoriasis, an inflammatory skin disease characterized by both epidermal and immune dysfunction; one of these enhancers is active in only proliferating cells, while the other two are active only in migrating cells (Table 1). To test if these SNPs could have an effect on enhancer function, we cloned both alleles of SNP rs4649203 into luciferase reporter vectors, and tested for ability to enhance luciferase expression. This region significantly enhances transcription, and the psoriasis-associated variant caused a decrease in enhancer activity, highlighting the importance of these distal regulatory regions to keratinocyte function (Figure 5-1E).

Closer examination of the genomic locations of super enhancers in differentiating and proliferating keratinocytes revealed that many such regions are positioned at the edges of the epidermal differentiation complex on chromosome 1 (Figure 5-3A), and at the edges of the keratin complexes on chromosomes 12 and 17 (Figure 5-4), which may be concentrating and loading sites for transcription factors that activate gene expression in these domains, and may

also be indicative of higher order chromatin structure at these loci. Intriguingly, these flanking SEs are not found in migrating keratinocytes, suggesting a reorganization in chromatin structure may occur during this transition. Comparing our enhancer data to a previously published microarray time course of NHEK differentiation, we found a strong correlation between genes upregulated during mid and late differentiation and the location of differentiation SEs (Figure 5-3B).

We next examined the motifs enriched in SEs under the different conditions; AP-1 motifs are significantly enriched in SEs from all three classes of keratinocytes, consistent with a general activating role for this factor in epithelial tissues (Figure 5-3C). KLF and AP-2 motifs were also enriched in proliferating and differentiating keratinocytes (Figure 5-3C). A motif similar to FOXD1 was more highly enriched in differentiating keratinocytes compared to the other cell states, whereas the MYC motif was found in more migrating cells, and a half-site for NF1 was more highly enriched in proliferating cells (Figure 5-3C).

To understand how GRHL3-mediated gene regulation integrates with the chromatin landscape in each situation, we combined newly generated GRHL3 ChIP-Seq data in migrating keratinocytes with previously published data for GRHL3 in NHEK-D [52]. We compared GRHL3 binding across different chromatin features. GRHL3 was more highly enriched in SEs compared to TEs for each cell condition. Additionally, GRHL3 was highly enriched in regions marked by H3K4me3, particularly in migrating keratinocytes (Figure 5-3D).

A number of key transcriptional regulators of keratinocyte cell states have been identified and characterized [106, 107]; we used published ChIP-Seq data for these factors to identify their rate of overlapping binding with super enhancers in the three keratinocyte states. Intriguingly, while factors like KLF4, ZNF750, and MAFB bound to a significant fraction of super enhancers,

GRHL3 bound to the largest proportion of super enhancers of any factor tested (Figure 5-3E). On average, GRHL3 bound 3 times to each of its target super enhancers, and approximately one third of SEs had greater than 5 GRHL3 binding events, further underscoring the important role of GRHL3 in keratinocyte gene regulation. There was much less overlap of GRHL3 peaks (in migrating and differentiating keratinocytes) with proliferation super enhancers (Figure 5-3E).

Grhl3 is required for epidermal keratinocyte migration through regulation of a set of genes distinct from terminal differentiation.

Further analysis of the GRHL3 migration ChIP-Seq data revealed that about 30 percent of binding events were found in proximal promoters of annotated genes (Figure 5-5A), a much larger fraction than in differentiation [52]. To identify direct targets of GRHL3 in migrating keratinocytes, we combined the ChIP-Seq peak data with microarrays for siRNA knockdown of GRHL3 in migrating keratinocytes. About half of the genes affected by the siRNA knockdown of Grhl3 in migration conditions had a Grhl3 peak in their promoter (Figure 5-5B), suggesting they are direct targets of Grhl3 during keratinocyte migration. These direct targets are enriched in gene ontology categories including cell adhesion, epidermis development, and cell motion (Figure 5-5C).

Our data points to a key role for GRHL3 in modulating the transitions between different keratinocyte states; to study this in more detail, we compared GRHL3 binding in the different cellular conditions. To identify statistically significant differentially bound GRHL3 sites between migration and differentiation we used the R program DiffBind [108]. GRHL3 binds to 1800 unique genomic regions in differentiating cells, and 2700 unique regions in migrating cells (Figure 5-5E).

Comparing the direct targets of Grhl3 in the two conditions of epidermal differentiation and migration, we found an interesting difference. Whereas the majority of direct targets of Grhl3 in differentiating cells are downregulated when Grhl3 is knocked down, indicating Grhl3 binds and activates these genes, the majority of direct targets of Grhl3 in migration are upregulated when Grhl3 is knocked down, indicating Grhl3 binds and represses these genes (Figure 5-5F). Based on this observation, it appears that Grhl3 primarily acts as a transcriptional activator during epidermal differentiation and as a transcriptional repressor during keratinocyte migration. Interestingly, the majority of unique migration peaks are found in the proximal, 3kb promoter, while the majority of unique peaks in differentiation are found outside the promoter (Figure 5-5G).

Further analysis of the genes repressed by Grhl3 in migration revealed that a number of these genes have been previously shown to act as inhibitors of migration in various different systems. One example of such a gene is SMAP1 (Figure 5-5D), a small GTPase-activating protein that regulates E-cadherin endocytosis during cell migration. Overexpression of SMAP1 causes a reduction in epithelial migration [109], suggesting it functions to inhibit cell migration, and that GRHL3-mediated downregulation of such genes promotes migration.

Based on the differences in Grhl3 binding and potential mechanisms of regulation, we next asked if these differences were related to differential co-factor association under the two conditions. We performed *de novo* motif searches on the DNA sequences from the unique migration and differentiation ChIP-Seq peaks. In both sets of unique peaks, we found enrichment of motifs with significant similarity to those bound by STAT and ETS factors, which have well established functions in epithelial tissues (Figure 5-5H). The differentiation peaks also had enrichment for a motif that is significantly similar to the KLF4 motif. This is consistent with the

known role of KLF4 in epidermal differentiation [107], and suggests GRHL3 and KLF4 could be acting together to regulate shared targets during differentiation.

In the unique migration Grhl3 ChIP-Seq peaks, a motif matching the transcriptional repressor REST was significantly enriched (Figure 5-5H). This is intriguing, as a role for REST in epidermal keratinocyte migration has not been conclusively identified. The enrichment of the REST motif fits with the predicted role of GRHL3 in repressing gene expression during migration, as REST commonly functions as a repressive factor, recruiting HDACs to turn off gene expression. Further support for this idea comes from the finding that Semaphorin 3a, which plays a role in inhibiting keratinocyte migration, is a direct target of REST in keratinocytes [110].

To test the effect of REST on keratinocyte migration we performed scratch assays after siRNA knockdown of REST. Cells depleted for REST closed the scratch at a slower rate than scramble control cells, indicating REST is required for keratinocyte migration (Figure 5-7A, B). We also tested the effect of depleting both REST and GRHL3 and found no additional decrease in rate of migration compared to either siRNA alone, suggesting REST and GRHL3 act on a similar set of targets to regulate cell migration in epidermal keratinocytes (Figure 5-7A, B).

To explore the role of REST in migrating keratinocytes, we knocked down REST with siRNA. We found a significant overlap of genes affected by REST knockdown and genes affected by GRHL3 knockdown (Figure 5-7C); the majority of these overlapping genes are upregulated by knockdown of REST or GRHL3 (Figure 5-7D), further supporting the repressive co-regulatory role for these two factors suggested by the motif analysis of GRHL3 peaks. Many of the genes fall into both cell migration and neuron projection categories, suggesting genes that

have been characterized as crucial components of regulation of neuronal migration during development are also important for cell projection and movement during keratinocyte migration.

Discussion:

In this study, we have identified regulatory regions in epidermal keratinocytes under different physiological states, providing insights into the mechanisms that govern gene expression in these different conditions, and the transitions between different states. We have defined roles for both SE and TE in these processes, and characterized the relationships between regulatory regions in different cellular states. We also identified regulatory regions with links to psoriasis and demonstrated their enhancer ability. Our work places GRHL3 regulation of the transitions between physiological states in the context of the chromatin landscape, where we found an unexpected enrichment of GRHL3 binding in migrating and differentiating SEs. Further analysis of GRHL3 regulation of differentiation and migration identified different mechanisms of regulation in these two physiological cell states, and identified REST as a co-factor of GRHL3 in migrating keratinocytes.

Through analysis of enhancers in migrating and differentiating keratinocytes, we found that, in contrast to the transition between embryonic stem cell and keratinocyte, cells transitioning inside the keratinocyte lineage do so within the confines of a pre-defined regulatory landscape; differences in gene regulation are likely due to changing the activity of these regulatory regions, rather than creating new regions. In comparing SEs to TEs, we have confirmed that, as in other tissues, NHEK SEs tend to show more tissue-specific function than TEs. We have also identified a potential role for SEs at the boundaries of gene clusters important for epidermal function, including the EDC and keratin clusters. These data suggest that

SEs could act as transcription factor loading and enrichment sites for coordinated regulation of genes occurring in these clusters.

Our data shows that GWAS SNPs linked to psoriasis risk that are found in enhancers affect the regulatory potential of these regions, and points to mechanisms of disease progression. While psoriasis had been thought to be primarily an immune disorder, more recent work implicates the epidermis in disease development [53]. Our data support these findings, and suggest that alterations in the expression levels of certain epidermally-expressed genes, due to nucleotide differences in distal regulatory regions, can contribute to disease.

The unexpected finding that GRHL3 binding is highly enriched in SEs of migrating and differentiating NHEK affirms the importance of GRHL3 in regulation of these keratinocyte states. Indeed, many of the genes linked to SEs with the highest number of GRHL3 binding events have been implicated in psoriasis or other diseases of the epidermis, suggesting that GRHL3 is very important to the proper functioning of the epidermis.

A single transcription factor performing unique regulatory functions under two conditions is not uncommon in biological systems, and our data adds to the understanding of the mechanisms behind such regulation. We find that GRHL3 regulates genes through very distinct mechanisms in each case; binding to distal regulatory regions and activating genes in differentiation, and binding to promoters and repressing migration genes. Our data suggest that GRHL3 may be able to accomplish these different forms of regulation through differential association with co-factors under the two conditions. Association with KLF4 in differentiation and REST in migration may determine the mechanism of GRHL3 action. Such a system might influence the specificity of other transcription factors with dual roles.

The association of GRHL3 and REST is a novel finding; REST was originally defined as an inhibitor of neuronal genes in non-neuronal tissues, but additional roles for this factor in tissue-specific regulatory processes are being uncovered. While targets of REST have been shown to have a role in keratinocyte migration, our study is the first to identify a direct role for REST in regulation of this process. Many of the shared targets of REST and GRHL3, the majority of which are repressed by the two factors, are inhibitors of cell migration and of axon projection in neurons. Consistent with our finding for GRHL3 and REST in repressing inhibitors of migration including genes regulating cell projection, GRHL3 has been shown to act in leading edge cells during migration and *Grhl3*-/- mice have reduced filopodia projections. Our results suggest GRHL3 and REST co-regulate a group of genes involved in neuronal and keratinocyte axon projection during cell migration and development.

SNP	GWAS Disease	Odds ratio	p-value	Nearest genes
rs4649203	Psoriasis	1.13	7x10^-8	IL28RA, GRHL3
rs1265181	Psoriasis	22.62	--	HLA-G, Pou5F1, HCG27
rs280519	Psoriasis	1.13	4x10^-9	TYK2

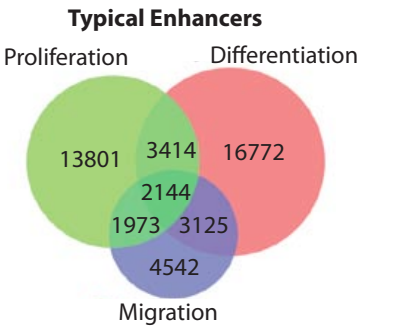
Table 1. SNPs linked to psoriasis that overlap enhancers in keratinocytes.

Figure 4-1. Identification of Typical (TE) and Super Enhancers (SE) in differentiating (NHEK-D) and migrating (NHEK-M) epidermal keratinocytes. A). Quantification and comparison of regulatory domains in proliferating, differentiating, and migrating keratinocytes. B. Comparison of the chromatin state in proliferating cells at TE locations in migrating or differentiating NHEK. C) Overlap of TE between the three cell conditions, overlap of SE between the three cell conditions, overlap of SE between the three cell conditions. D) Comparison of changes in SE and TE as cells differentiate from embryonic stem cells (ESC) to keratinocyte progenitors (P), with changes that occur in the transition between keratinocyte cell states (P to M), (P to D). E) Genome view of SNP rs4649203, luciferase reporter assays with WT(G) and disease-linked allele (A), n=6.
* $p < .05$, ** $p < .01$.

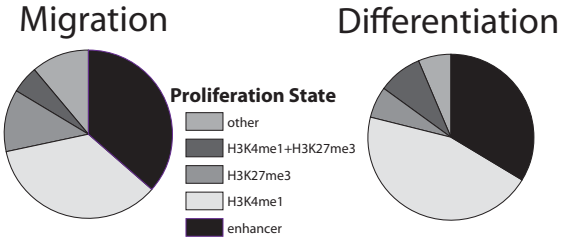
A. **Chromatin Domains**

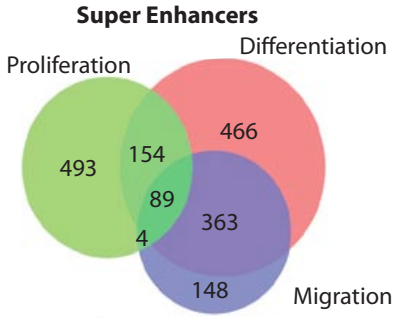
	Proliferation	Differentiation	Migration
Active Transcription (H3K4me3)	13873	14253	14221
Typical Enhancer	20084	30933	10599
Super Enhancer	783	761	363

B.



C.

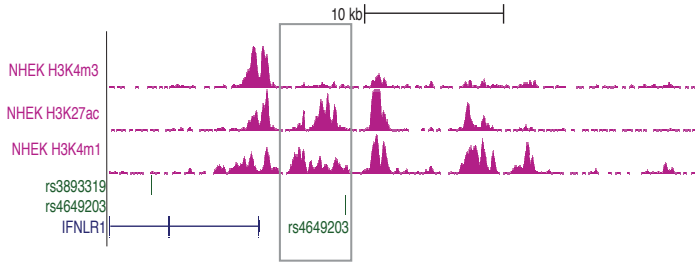




D.

	ESC -->P	P -->D	P -->M
Super Enhancer → Super Enhancer	.079	.299	.106
Super Enhancer → Typical Enhancer	.165	.604	.493
Super Enhancer → H3K4me1	.434	.092	.400
Super Enhancer → Unmarked Region	.322	0.000	0.000

E.



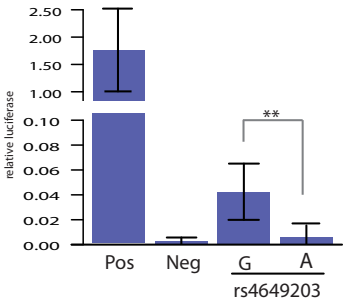
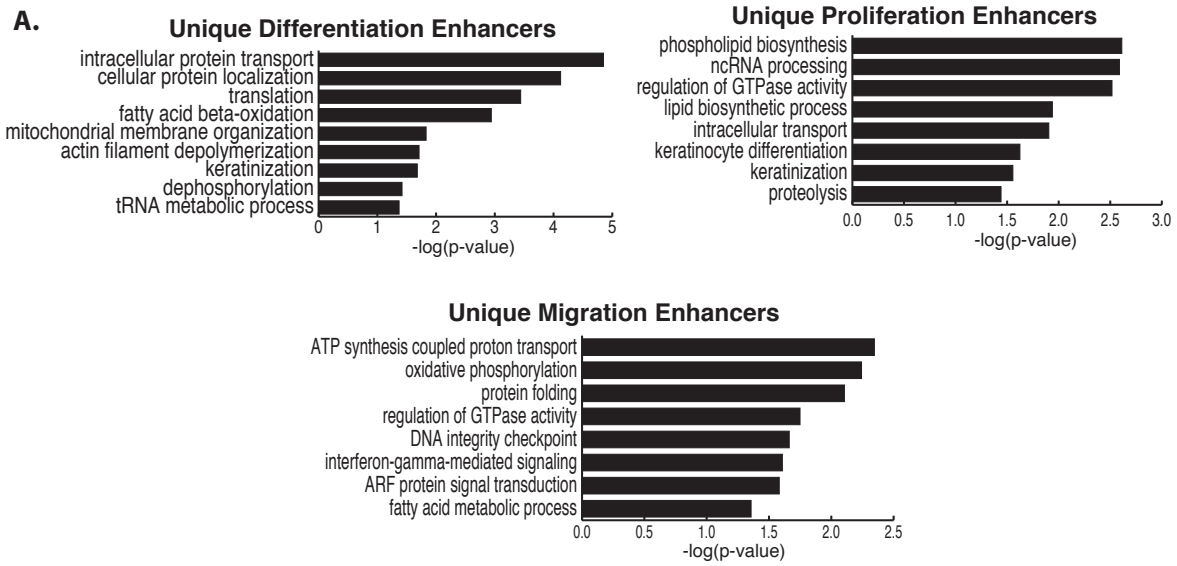


Figure 4-2. SE are found near genes with more cell-type specific functions than TE. A) Enriched gene ontology terms for genes near TE unique to differentiating, proliferating, or migrating cells. B) Enriched gene ontology terms for genes near SE in proliferating NHEK, SE near epidermal fate regulator P63. C) Enriched gene ontology terms for genes near SE in differentiating NHEK, SE near epidermal fate regulator GRHL3. D) Enriched gene ontology terms for genes near SE in migrating NHEK, SE near epidermal fate regulator FLNA.


A.
Unique Differentiation Enhancers
intracellular protein transport
cellular protein localization
translation
fatty acid beta-oxidation
mitochondrial membrane organization
actin filament depolymerization
keratinization
dephosphorylation
tRNA metabolic process
0 1 2 3 4 5
-log(p-value)
Unique Proliferation Enhancers
phospholipid biosynthesis
ncRNA processing
regulation of GTPase activity
lipid biosynthetic process
intracellular transport
keratinocyte differentiation
keratinization
proteolysis
0.0 0.5 1.0 1.5 2.0 2.5 3.0
-log(p-value)
Unique Migration Enhancers
ATP synthesis coupled proton transport
oxidative phosphorylation
protein folding
regulation of GTPase activity
DNA integrity checkpoint
interferon-gamma-mediated signaling
ARF protein signal transduction
fatty acid metabolic process
0.0 0.5 1.0 1.5 2.0 2.5
-log(p-value)

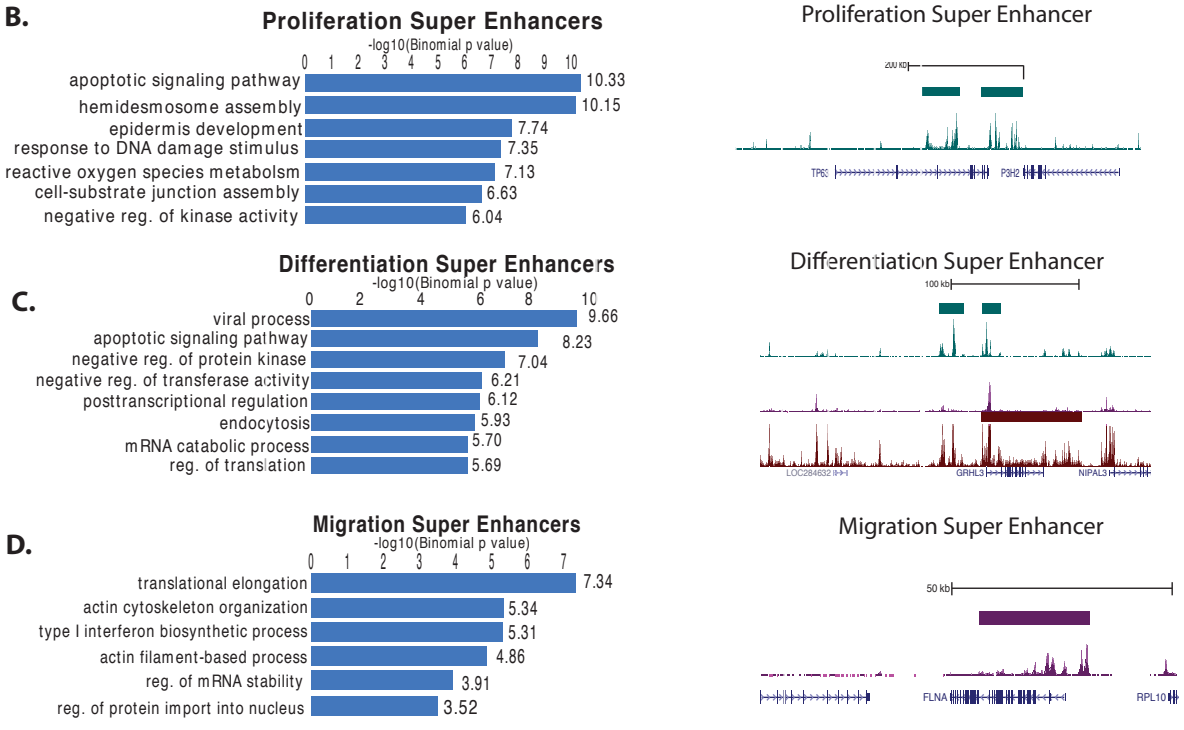
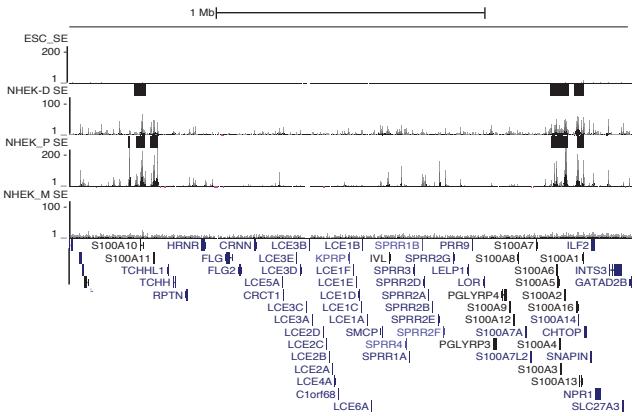

B.
Proliferation Super Enhancers
-log10(Binomial p value)
0 1 2 3 4 5 6 7 8 9 10
apoptotic signaling pathway 10.33
hemidesmosome assembly 10.15
epidermis development 7.74
response to DNA damage stimulus 7.35
reactive oxygen species metabolsm 7.13
cell-substrate junction assembly 6.63
negative reg. of kinase activity 6.04
Proliferation Super Enhancer
200 kb
TP63
P3H2
C.
Differentiation Super Enhancers
-log10(Binomial p value)
0 2 4 6 8 10
viral process 9.66
apoptotic signaling pathway 8.23
negative reg. of protein kinase 7.04
negative reg. of transferase activity 6.21
posttranscriptional regulation 6.12
endocytosis 5.93
mRNA catabolic process 5.70
reg. of translation 5.69
Differentiation Super Enhancer
100 kb
LOC284632
GRHL3
NIPAL3
D.
Migration Super Enhancers
-log10(Binomial p value)
0 1 2 3 4 5 6 7
translational elongation 7.34
actin cytoskeleton organization 5.34
type I interferon biosynthetic process 5.31
actin filament-based process 4.86
reg. of mRNA stability 3.91
reg. of protein import into nucleus 3.52
Migration Super Enhancer
50 kb
FLNA
RPL10

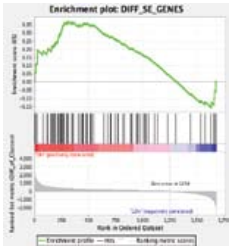
Figure 4-3. SEs are found near differentiation genes; GRHL3 binds to super enhancers during differentiation and migration. A) Genomic scale view of the Epidermal Differentiation Complex (EDC) of genes on chromosome 1 flanked by SE. B) Differentiation SEs correlate strongly with mid (24 hour) and late (96 hour) differentiation genes. C) Motif analysis on SEs in each condition. D) Distribution of GRHL3 peaks in chromatin domains. E) Fraction of SEs bound by epidermal transcription factors in proliferating (NHEK-P), differentiating (NHEK-D), and migrating (NHEK-M) cells.

A.



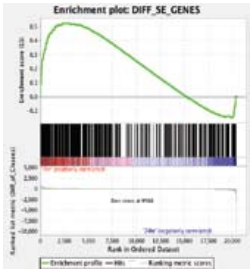
B.

Early differentiation genes



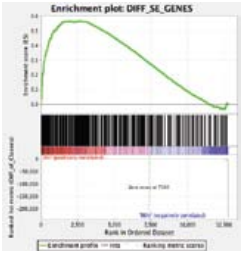
p = .09

Mid differentiation genes



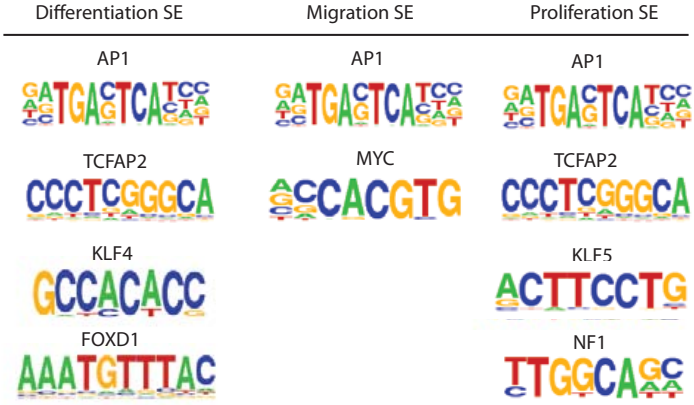
p <.001

Late differentiation genes



p <.001

C.



D.

	GRHL3-Diff	GRHL3-Mig
Super Enhancer	5281	1830
Typical Enhancer	2171	937
H3K4me3	5063	11714
H3K4me1-only	2562	230
Total GRHL3 Peaks	25340	18674

E.

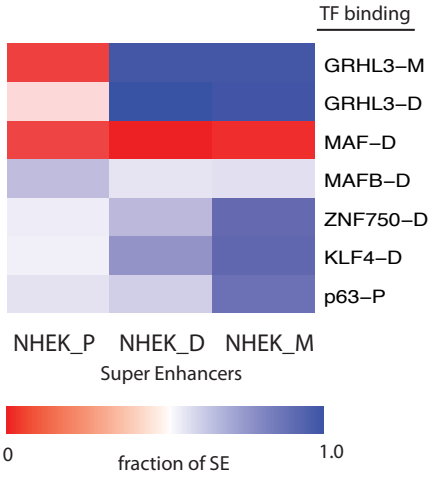


Figure 4-4. NHEK TEs overlap SNPs linked to psoriasis; SEs are located at edges of gene clusters. A) Genomic view of NHEK TEs overlapping SNPs linked to psoriasis. B-C) Genomic views of SEs at edges of *Krt* clusters.

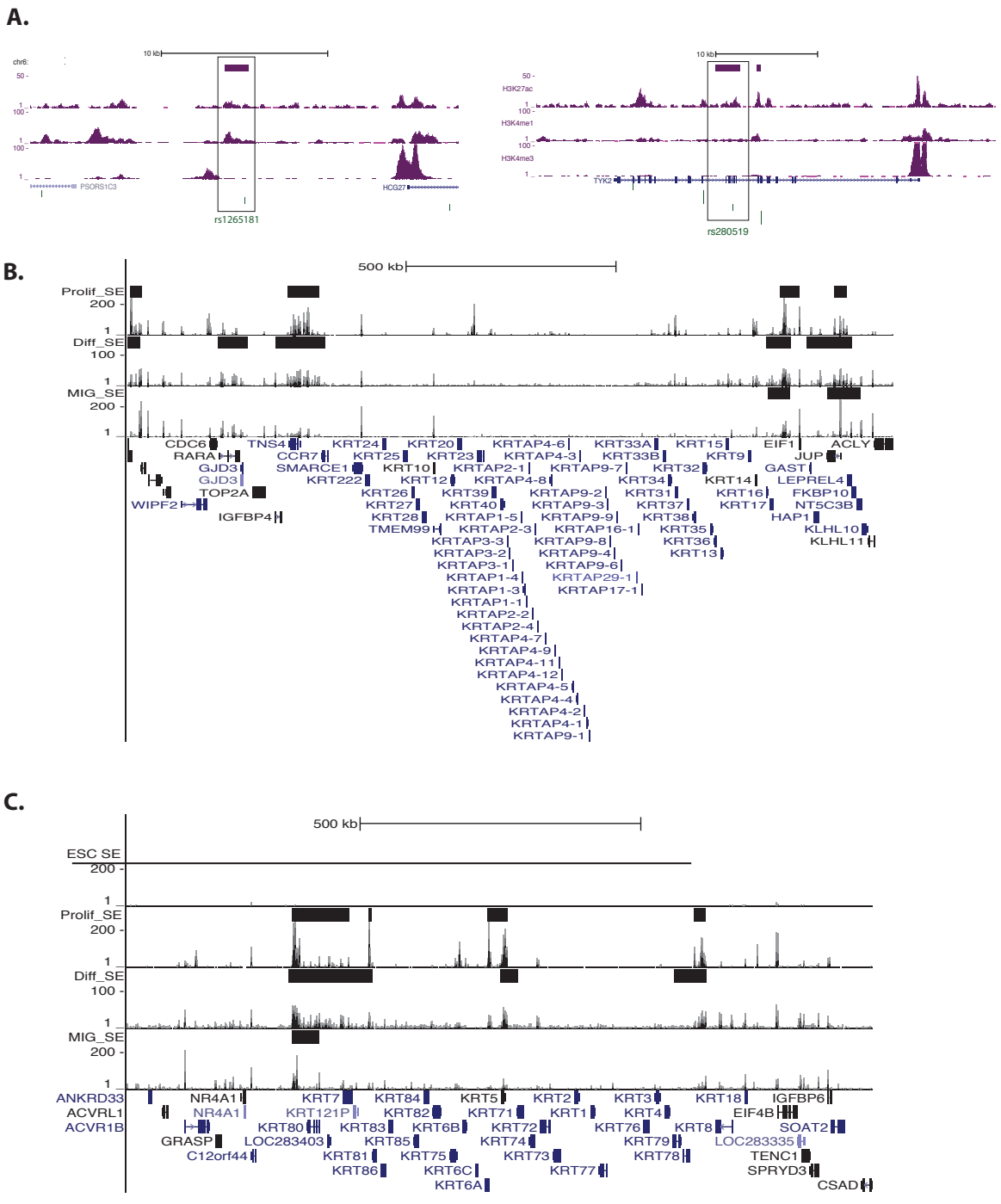
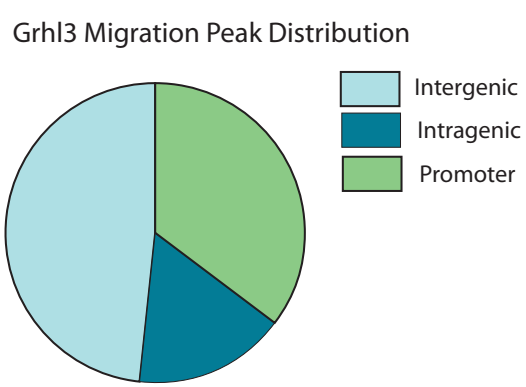
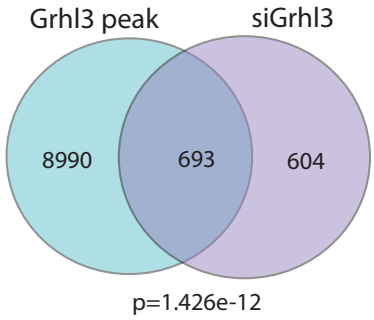

A.
rs1265181
rs280519
B.
500 kb
Prolif_SE
Diff_SE
MIG_SE
C.
500 kb
ESC SE
Prolif_SE
Diff_SE
MIG_SE

Figure 4-5. GRHL3 regulates genes through unique mechanisms in migrating and differentiating keratinocytes. A) Distribution of GRHL3 peaks in NHEK-M across genomic features. B) Overlap of genes with GRHL3 peaks and genes affected by siRNA knockdown of GRHL3 (siGRHL3). C) Enriched gene ontology categories for overlapping genes in (B). D) Genomic view of GRHL3 peak at *Smap1* promoter. E) Identification of significantly differentially bound GRHL3 sites in migration and differentiation. Proportion of GRHL3 bound genes upregulated or downregulated by siGRHL3 in differentiation and migration. G) Percentage of GRHL3-unique peaks in promoter in differentiation and migration. H) Enriched motifs in unique migration or differentiation GRHL3 peaks.

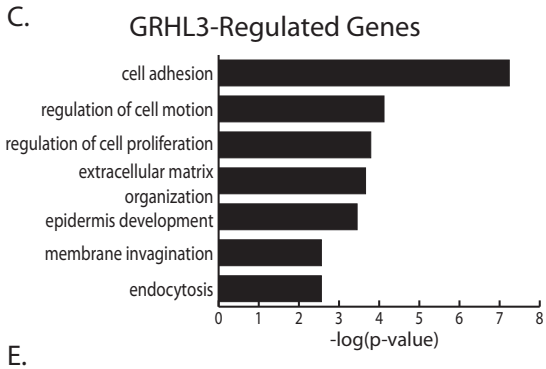
A.



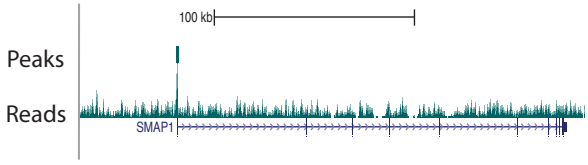
B.



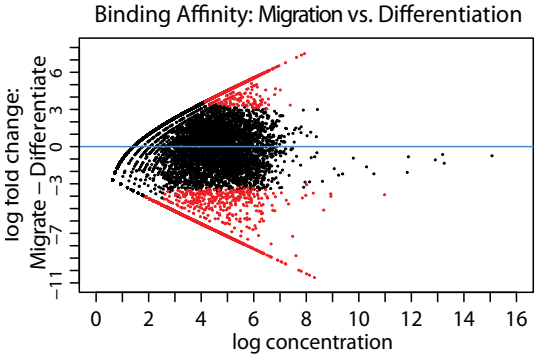
C.



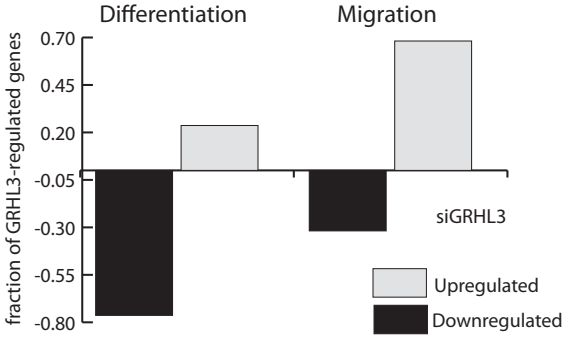
D.



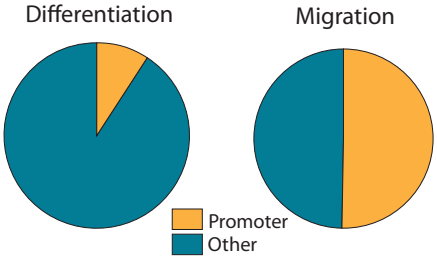
E.



F.



G.



H. **Top Motifs Enriched in Unique Grhl3 Peaks**

Migration only	Differentiation only
STAT1/3	STAT1/3
ETS	ETS
REST	KLF4

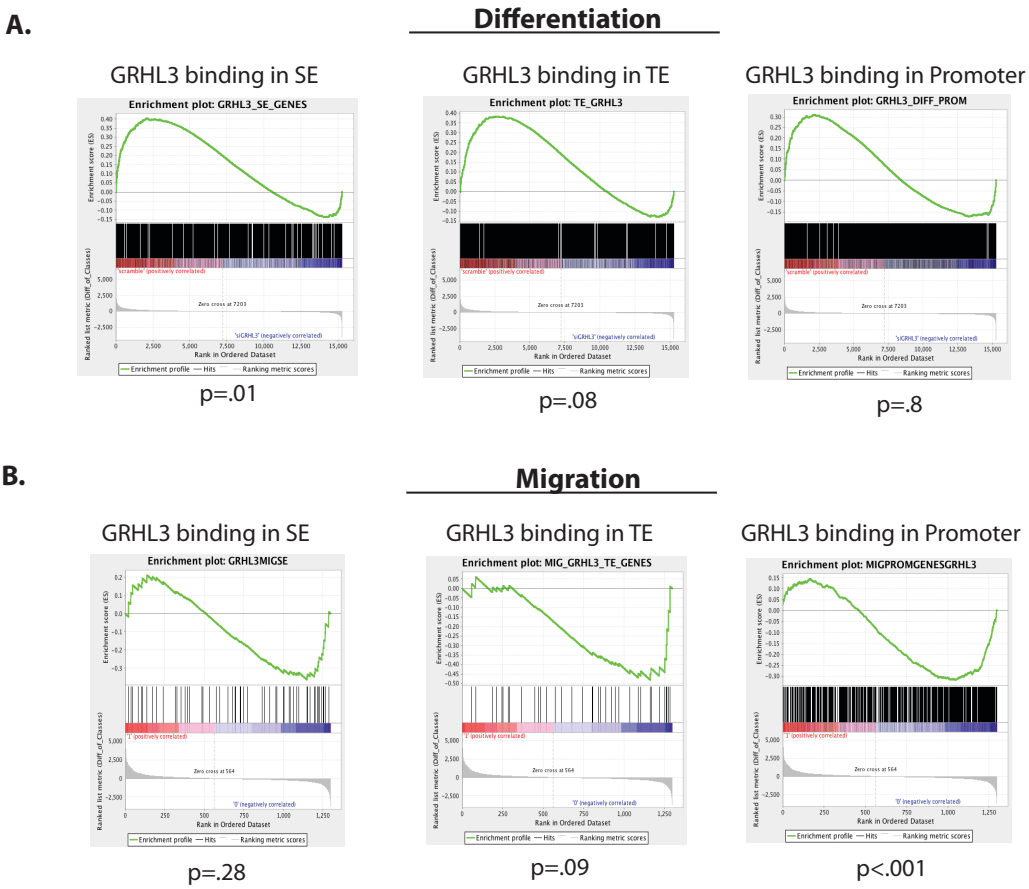
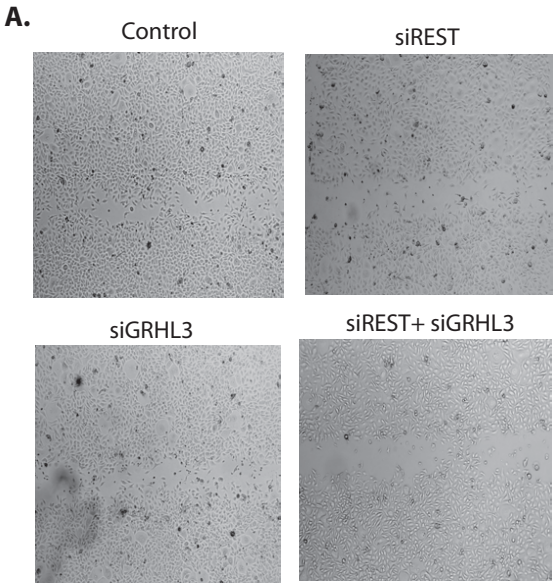
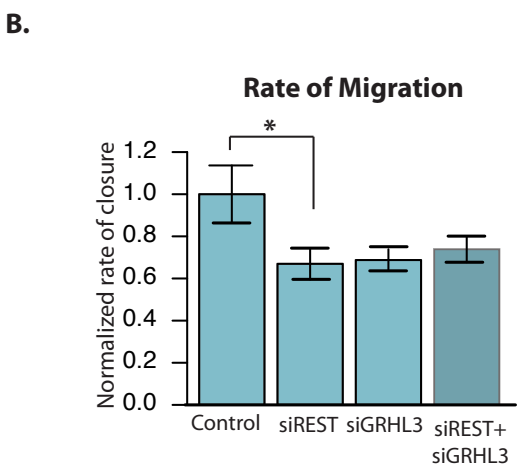
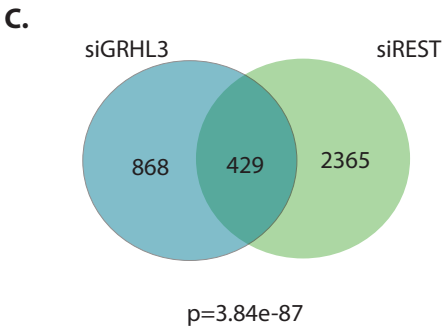


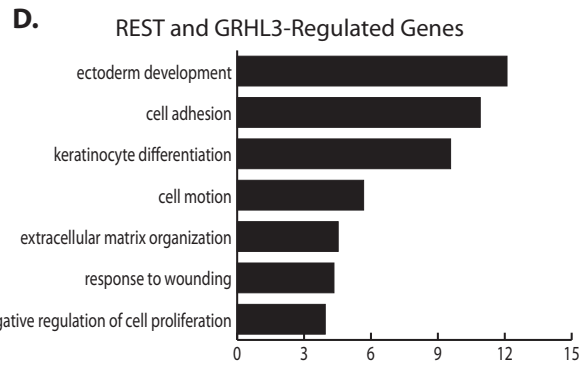
Figure 4-6. GRHL3 regulates differentiation genes through SEs, and migration genes through promoters. A) GSEA analysis correlating GRHL3 binding in SE, TE, or promoters with genes affected by siGRHL3 in differentiating NHEK. B) GSEA analysis correlating GRHL3 binding in SEs, TEs, or promoters with genes affected by siGRHL3 in migrating NHEK.

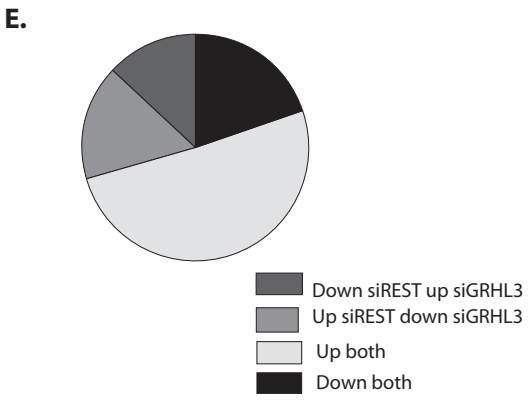
Figure 4-7. GRHL3 and REST regulate shared targets during keratinocyte migration. A) Images of migration assays in siControl, siREST, siGRHL3, and combined siREST and siGRHL3 treated cells. B) Quantification of rate of migration in siControl, siREST, siGRHL3, and combined siREST+siGRHL3, n=12 control, siREST, n=10 siGRHL3, siGRHL3+siREST. C) Overlap of genes affected ($p<.05$) by siGRHL3 and siREST in NHEK migration ($p<.05$). D) Enriched gene ontology terms for overlapping genes in C. E) Percentage of genes upregulated in both siGRHL3 and siREST (up both) treated cells, downregulated in both siGRHL3 and siREST (down both), upregulated in siREST and downregulated in siGRHL3, or downregulated in siREST and upregulated in siGRHL3 treated cells. F) Enriched gene ontology terms for genes upregulated in both siREST and siGRHL3 treated cells.

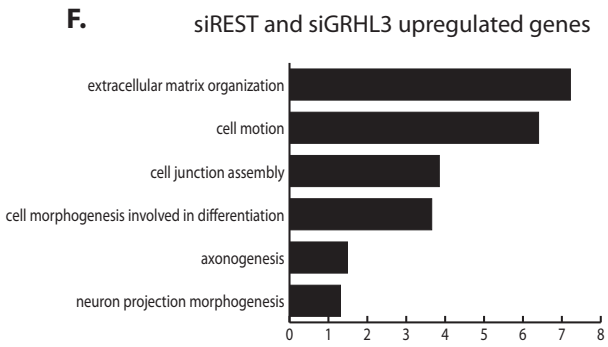

A.
Control
siREST
siGRHL3
siREST+ siGRHL3


B.
Rate of Migration
*
Normalized rate of closure
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Control
siREST
siGRHL3
siREST+ siGRHL3


C.
siGRHL3
siREST
868
429
2365
p=3.84e-87


D.
REST and GRHL3-Regulated Genes
ectoderm development
cell adhesion
keratinocyte differentiation
cell motion
extracellular matrix organization
response to wounding
negative regulation of cell proliferation
0
3
6
9
12
15


E.
Down siREST up siGRHL3
Up siREST down siGRHL3
Up both
Down both


F.
siREST and siGRHL3 upregulated genes
extracellular matrix organization
cell motion
cell junction assembly
cell morphogenesis involved in differentiation
axonogenesis
neuron projection morphogenesis
0
1
2
3
4
5
6
7
8

Chapter 5. Methods

Cell culture:

Normal Human Corneal Epithelial Cells (HCEC) were purchased from LifeLine Technologies and grown according to the manufacturer's instructions in OcuLife medium (LifeLine Tech) supplemented with OcuLife growth factors (LifeLine Tech). Telomerase- immortalized hTCEpi cells were grown in KBM medium (Lonza) supplemented with SingleQuots (Lonza).

Normal Human Epidermal Keratinocytes (NHEK) were purchased from LifeLine Technologies and grown according to the manufacturer's instructions in DermaLife medium (LifeLine Tech) supplemented with DermaLife growth factors (LifeLine Tech). Immortalized (HaCaT) Keratinocytes were grown in Epilife medium (Life Technologies) supplemented with EDGS (Life Technologies). Cells were induced to differentiate by addition of high calcium medium (1.8mM Ca2+). To induce migration, cells were grown to a confluent monolayer, and the scratched with a pipette tip.

Culture of primary mouse corneal epithelial cells:

After sacrifice, mice were sprayed with 70% EtOH and a drop of Betadine applied to each eye for 15 seconds. Eyes were flushed with PBS + Pen/Strep (100u/100ug per ml penicillin/streptomycin) and enucleated. Whole eye globes were washed in PBS + Pen/Strep 3x10 minutes and whole corneas dissected, removing all non-corneal tissues. Corneal epithelium was isolated by digestion in EMEM + dispase (20x dispase, CellNTec) with 50 mM sorbitol and Pen/Strep for two hours at 37°C, vortexing the whole eye globes gently every 15 minutes to aid separation. Corneal epithelial sheets were then peeled off the eye globes. The corneal epithelial

sheets were trypsinized for 15 min at 37C to obtain a single cell suspension and plated 100,000 cells/ well in a 6 well dish, using cells from one eye for each well. Prior to plating, culture dishes were coated with a 1:15 dilution of PureCol, incubated for two hours at 37C, and rinsed twice with PBS + Pen/Strep. Cells were grown in CnT50 medium (CellNTec). To verify the identity of the cultured primary corneal epithelial cells after two weeks in culture, we performed QPCR on a panel of limbal and basal cell markers, corneal differentiation markers, and epidermal differentiation markers (Supplemental Figure 1B). The isolated epithelial cells have high expression of progenitor cell markers Abcg2 and Krt19, and low expression of corneal epithelial differentiation markers like Krt12 at day 0 of isolation. After two weeks in culture, the expression of these progenitor markers still remains high compared to differentiation markers. Additionally, loricirn, a marker of epidermal differentiation is low at all timepoints, indicating these cells are not transitioning to an epidermal cell fate, as has been noted in other studies. These results suggest the cells remain progenitor-like after two weeks in culture.

Chromatin Immunoprecipitation Assays:

ChIP assays were performed as previously described [54] with the following changes: 24ug of sonicated chromatin was used for each IP, magnetic Dynabeads (Invitrogen) were used for immunoprecipitation and, for ChIP- qPCR analysis, enrichment was calculated over IgG and normalized to an intergenic negative control region. The following antibodies were used: IgG (Sigma), H3K4me1 (AbCam), H3K27ac (Millipore), H3K4me3 (Milipore), GRHL3 (Andersen Lab).

ChIP-Seq:

Sequencing libraries were generated using the NEB Next reagents and Illumina adaptors and oligos, according to the Illumina protocol for ChIP-Seq library prep, with some modification. Following the protocol by Schmidt et.al., after adaptor ligation, PCR amplification was performed prior to size selection of the library [111]. Clustering and 50-cycle single end sequencing were performed on the Illumina Hi-Seq 2000 Genome Analyzer.

Bioinformatics Analysis:

ChIP-Seq reads were aligned using Bowtie [112], and only uniquely aligning reads were retained (-m 1). Histone modification peaks were called using SICER [113] with a 200bp gap for H3K4me3 and H3K27ac, and a 400bp gap from H3K4me1. Enhancers were identified through intersection of H3K4me1 peaks with H3K27ac peaks, and subtraction of H3K4me3 peaks using BEDTools [114]. GRHL3 peaks were called using MACS [115]. Input samples were used as the control for both methods of peak calling. Galaxy software was used for further analysis identifying overlaping and unique peaks [116-118]. DAVID and GREAT were used to identify enriched gene ontology terms for the nearest genes for each super enhancer or typical enhancer (dtermined by BEDTools nearestBed function) [85, 119]. Enriched motifs were identified from the top 1000 peaks *de novo* usng MEME, and pekas were scanned for known motifs using FIMO [120]. Super enhancers were called using ROSE, with default settings [55, 121], including exlusion of the 2.5kb TSS regions. To identify enriched motifs within the broad super enhancer domains, nucleosome free regions for SE and TE were identified using Homer's –nfr option [122]. Enriched motifs were identified within these nfr regions using Homer- 5 iterations of peak calling were performed per sample and only motifs found enriched in multiple runs were

reported [120]. Galaxy [116, 117] software was used for further analysis of overlaping and unique regions. Gene Enrichment Analysis was performed using GSEA [123, 124].

Quantitative Real Time PCR:

For mRNA expression analysis, cDNA was prepared using iScript cDNA kit and RT-PCR was performed using SsoFast for Probes and SsoFast EvaGreen (Biorad Laboratories) master mixes in CFX384 Real-Time PCR Detection System (Biorad Laboratories). GAPDH or RPLPO were used as endogenous controls. Primer and probe sequences available upon request.

RNA extraction from cell culture:

Cells were collected and lysed in Trizol, followed by Chloroform extraction. RNA was extracted from the aqueous phase as previously described. RNA concentration and quality were quantified on a NanoDrop.

Isolation of RNA from tissues for microarray:

Eye globes were removed from *K14-DN-Clim* transgenic mice and wild type littermates immediately after sacrifice. Whole corneas were dissected, removing all non-corneal tissues but retaining the peripheral cornea, or limbal region. Total RNA from whole cornea was isolated with Trizol (Life Technologies) and further purified with the Qiagen RNEasy Micro Kit. Corneal RNA sample purity was validated by QPCR expression of tissue specific markers and absence of markers for adjacent tissues.

P3 samples were prepared for microarray with the NuGEN Ovation RNA Amplification System V2 and NuGEN FL-Ovation cDNA Biotin Module V2 (NuGEN Technologies, San Carlos CA). Gene expression was assessed with Affymetrix Mouse Gene 1.0 ST Arrays in triplicate.

Luciferase Assays:

Approximately 1kb regions surrounding each SNP were cloned into PGL3 reporter vectors. Site directed mutagenesis was performed following the manufacturer's specifications (Agilent Quick Change II Site Directed Mutagenesis kit). Successful mutagenesis was verified by sequencing. Cells were transfected with each plasmid and a renilla control, and collected after 72 hours. Luciferase was quantified using the Promega Luciferase Assay System.

MTT Assays:

Cells were plated to a density of 10,000 cells per well in 96 well plates, and transfected the following day. Proliferation was assayed after 72 hours using the Cell Titer proliferation assay (Promega). 20 uL MTT assay reagent (Promega) was added to each well and incubated in the dark for 2 hours. Absorbance was read at 490nm. Wells were transfected with the following vectors: empty vector (pcDNA3.1), DN-CLIM, h19 (Dharmacon 3449920), and siKLF7 (Dharmacon L-019540-00-0005).

Microarray Analysis:

Gene expression analysis for HCE cells was performed with biological duplicates, as previously described [26] except Affymetrix Human Gene 1.0 ST arrays (26,869 probe sets) were used and washed according to manufacturer's recommendations (Affymetrix, Santa Clara, CA. Gene Ontology analysis was performed using DAVID [125]. Statistical significance of overlaps was calculated using the Fisher's exact test. The RNA collected from the corneal epithelial cells was hybridized to Affymetrix Human Gene 1.0ST arrays in duplicate. Plier analysis was performed,

and the data were then filtered for expression levels. Probes with raw expression values below 200 were considered not expressed for subsequent analysis. CyberT was used to compare the cornea gene expression to both undifferentiated and differentiated epidermal keratinocyte gene expression data previously generated in the lab.

Colony-forming assays:

Primary corneal epithelial cells were isolated as described above and plated 10,000 cells/well in a 12 well dish. Medium was changed every 3 days. After 2 weeks, cells were fixed with 10% neutral buffered formalin, washed 3 x 5 min with PBS, and stained with Giemsa (Sigma) for 15 min. Wells were then rinsed with dH_2O, dried, and photographed under a dissection microscope.

BrdU labeling of *K14-DN-Clim* mice:

K14-DN-Clim mice and wild type littermates were injected intraperitoneally with 10mg/ml BrdU solution (BD Biosciences) to a concentration of 50 ug/g of body weight two hours prior to sacrifice. Whole eye globes were dissected, fixed in neutral buffered formalin at 4C O/N, then processed for paraffin embedding. Eyes were sectioned at 8 um and stained with rat antiBrdU (Abcam) and biotinylated antiRat IgG (Vector Laboratories). Labeling was detected with the Vectastain ABC Elite Kit (Vector) and DAB/DAB+ Chromogen Solution (Dako).

Transfection:

Primary human corneal epithelial cells were transfected with Lipofectamine LTX with Plus reagent (Life Technologies), according to manufactorer's protocol. Medium was changed after 16 hours, and cells were collected 72 hours after the addition of transfection reagents.

siRNA Transfection:

Lipofectamine RNAi Max (Life Technologies) in OptiMEM medium was used for siRNA transfections: 30nM pooled siRNA was used for knockdown each of GRHL3 (Dharmacon L-014017-02), REST (Qiagen hs_REST_5: S104153765, hs_REST_1: s100701407), siKLF7 (Dharmacon L-019540-00-0005) and siKLF4 (Qiagen and scramble control (Dharmacon D-001810-10-05). Experiments were performed 72 hours after transfection.

Animal model and procedures:

All procedures were reviewed and approved by the University of California Irvine Institutional Animal Care and Use Committee (IACUC animal protocol number 2001-2239).

Migration Assays:

72 hours after transfection, cells were incubated for 2 hours with 3.5ug mitomyosin C. Scrapes were made with a pipette tip and medium was changed. Images were taken immediately after scratching (0 hour), 12, 14, 16, 20, 24, and 36 hours after scratching. The area of the scratches was measured by Image J.

Cloning:

Approximately 1kb of enhancer regions containing the identified GWAS SNPs were cloned into PGL3 luciferase reporter vectors (Promega). Site directed mutagenesis at the SNP locus was performed using the Quick Change II Kit (Agilent Technologies) and confirmed by sequencing. Each reporter construct (containing WT or disease associated variant) was transfected into

HaCaT immortalized keratinocytes using Lipofectamine LTX and Plus reagent (Life Technologies). After 48 hours, cells were lysed and luciferase levels measured using a luminometer and the Promega luciferase assay system (Promega).

Chapter 6. Summary and Conclusions

Cornea epithelium and epidermis share many functional and regulatory similarities: they both serve important barrier functions, protecting the organism and underlying tissues from the external environment, they also both undergo constant cell turnover, requiring a precise balance of proliferation and differentiation to maintain tissue integrity. To accomplish this balancing act, they make use of similar mechanisms, and integrate information from pre-existing regulatory domains with the expression and actions of lineage-specific transcriptional regulators. While many regulatory regions and transcription factors are important across a wide range of epithelial cell types, the specificity in each tissue arises from the actions of a select group of unique regulatory domains and transcription factors.

Super enhancers make a significant contribution to the cell type specificity of gene expression, and are an important part of regulation for both cornea epithelium and epidermis, enhancing the expression of key genes. We found that many SE-linked genes are themselves transcriptional regulators, suggesting SEs may promote cell-specific functions through activation of lineage-specific transcription factors. Indeed, *Grhl3,* a key regulator of epidermal differentiation and migration, contains a super enhancer within its gene body. In addition, GRHL3 binding is also highly enriched in epidermal SEs, demonstrating that transcription factors with cell-specific functions can be regulated by SEs and also bind to and act upon SEs to regulate gene expression. We have also identified super enhancers positioned at the edges of gene clusters, suggesting a function in recruiting and loading transcription factors and the transcriptional machinery necessary for coordinated regulation of these genes.

The importance of distal regulatory regions is further underlined by the observation that many SNPs linked to increased risk of disease fall outside of protein-coding regions of the

genome. It has been hypothesized that these SNPs overlap and interfere with important regulatory domains, and that they exert their effect on disease risk through alterations in levels of gene expression, rather than through changes to the amino acid sequence of a protein. Our data in both cornea epithelium and epidermis support this hypothesis; we find that SNPs linked to disease in each tissue alter the enhancing ability of the region in which they fall. Furthermore, we find evidence for disruption of transcription factor binding motifs by these SNPs, suggesting a mechanism for the observed alterations in regulatory potential.

The transcriptional regulator GRHL3 had previously been shown to regulate both epidermal terminal differentiation and cell migration. We now provide mechanistic insight into how it carries out these unique functions within the same tissue under different conditions. GRHL3 appears to act very distinctly under the two conditions; it is associated with gene activation of structural and barrier genes in differentiating keratinocytes, and with repression of migration inhibitors in migrating cells. It acts mainly through enhancers and super enhancers in differentiation, and through promoter binding in migration. We find evidence that these different mechanisms of GRHL3-mediated regulation correlate with the unique associations of GRHL3 with co-factors under the different keratinocyte cell conditions. This brought to light a previously undescribed and uncharacterized interaction between GRHL3 and the transcriptional repressor REST. We find that they both target and repress a shared set of migration inhibitor genes that also play a role in neuronal axon projection, suggesting the role of REST as an inhibitor of neuronal genes is conserved in this context. The interaction between GRHL3 and REST may be a more general epithelial mechanism of gene regulation, particularly under migration conditions, as GRHL3 is expressed in a wide range of epithelial tissues including bladder, intestine, and trachea.

Roles for several KLF family members have been defined in epithelial tissues, however the high epithelial expression of so many members of this large gene family indicate that the fully story of KLF function has not yet been defined. We have identified a novel role for KLF7 in maintaining corneal epithelial progenitor cells in an undifferentiated state and preventing uncontrolled proliferation. KLF7 acts in an opposing manner to KLF4, a factor known to promote differentiation; the two appear to regulate the same genes in opposite directions. We also find a contrasting expression patterns of these two KLFs: KLF7 is highly expressed in progenitor cells and decreases during corneal epithelial differentiation, whereas KLF4 expression increases during differentiation, suggesting the two may compete for binding to shared targets, exerting opposing effects on transcription at these loci, and the more abundantly expressed KLF determines the readout at a given locus. KLF7 and KLF4 have similar expression patterns in epidermis compared to cornea epithelium; KLF7 is most highly expressed in progenitor cells, and KLF4 in differentiated cells. From this, it is possible to speculate that a similar mechanism is at work in the epidermis and other epithelial tissues, where KLF7 promotes progenitor cell maintenance, and KLF4 promotes the terminal differentiation program.

Chromatin structure is also a key aspect of transcriptional regulation, and factors like CLIMs/LDBs, which can interact with transcription factors at two different loci through homodimerization, are thought to aid in the formation of chromatin loops between distal regulatory regions and promoters. Our work has highlighted the importance of CLIMs in corneal epithelium development and maintenance, where we have found a role for these factors in regulating proliferation dynamics, acting in part through the noncoding RNA h19. This factor is an important growth inhibitor that helps to balance proliferation during embryonic development and in progenitor cells. The structure of the h19 locus and its regulatory domains is very

important for proper expression; the downregulation of h19 by DN-CLIM suggests that CLIMs are important mediators of the structure at this locus. Additionally, there is evidence for sex-specific expression of this gene, as h19 is much more highly expressed in the female cornea compared to the male. Consistent with this, we find estrogen receptor motifs co-localize with CLIM binding to the h19 gene. Our results suggest that CLIMs regulate the chromatin structure required for proper expression of h19, which acts to ensure the balanced proliferation necessary for correct development and maintenance of the cornea epithelium.

Transcriptional regulation is a dynamic and highly complex process. It involves contributions from more general regulatory domains and transcription factors, as well as input from unique regulatory regions and transcriptional regulators in each cell type. Reciprocal interactions, between transcription factors and chromatin state are also important to generate cell-type specific gene expression patterns. Our studies define the global regulatory landscape in epidermal and corneal epithelial cells, and examine the specific roles of key transcription factors within the chromatin context, providing mechanistic insight into transcriptional regulation in these tissues.

REFERENCES:

1. Fuchs E, Raghavan S. Getting under the skin of epidermal morphogenesis. Nat Rev Genet. 3. England2002. p. 199-209.
2. Blanpain C, Fuchs E. Epidermal homeostasis: a balancing act of stem cells in the skin. Nat Rev Mol Cell Biol. 10. England2009. p. 207-17.
3. Courville CB, Smolek MK, Klyce SD. Contribution of the ocular surface to visual optics. Exp Eye Res. 2004;78(3):417-25. Epub 2004/04/27. PubMed PMID: 15106921.
4. Kaye GI, Pappas GD. Studies on the cornea. I. The fine structure of the rabbit cornea and the uptake and transport of colloidal particles by the cornea in vivo. J Cell Biol. 1962;12:457-79. Epub 1962/03/01. PubMed PMID: 14454675; PubMed Central PMCID: PMCPmc2106044.
5. Cotsarelis G, Cheng SZ, Dong G, Sun TT, Lavker RM. Existence of slow-cycling limbal epithelial basal cells that can be preferentially stimulated to proliferate: implications on epithelial stem cells. Cell. 1989;57(2):201-9. Epub 1989/04/21. PubMed PMID: 2702690.
6. Zhao J, Mo V, Nagasaki T. Distribution of label-retaining cells in the limbal epithelium of a mouse eye. J Histochem Cytochem. 2009;57(2):177-85. Epub 2008/11/13. doi: 10.1369/jhc.2008.952390. PubMed PMID: 19001638; PubMed Central PMCID: PMCPmc2628319.
7. Stephens DN, Klein RH, Salmans ML, Gordon W, Ho H, Andersen B. The Ets transcription factor EHF as a regulator of cornea epithelial cell identity. J Biol Chem. 2013;288(48):34304-24. Epub 2013/10/22. doi: 10.1074/jbc.M113.504399. PubMed PMID: 24142692; PubMed Central PMCID: PMCPmc3843046.
8. Swamynathan SK, Davis J, Piatigorsky J. Identification of candidate Klf4 target genes reveals the molecular basis of the diverse regulatory roles of Klf4 in the mouse cornea. Invest Ophthalmol Vis Sci. 2008;49(8):3360-70. Epub 2008/05/13. doi: 10.1167/iovs.08-1811. PubMed PMID: 18469187; PubMed Central PMCID: PMCPmc2774783.
9. Kenchegowda D, Harvey SA, Swamynathan S, Lathrop KL, Swamynathan SK. Critical role of Klf5 in regulating gene expression during post-eyelid opening maturation of mouse corneas. PLoS One. 2012;7(9):e44771. Epub 2012/10/02. doi: 10.1371/journal.pone.0044771. PubMed PMID: 23024760; PubMed Central PMCID: PMCPmc3443110.
10. Xu X, Mannik J, Kudryavtseva E, Lin KK, Flanagan LA, Spencer J, et al. Co-factors of LIM domains (Clims/Ldb/Nli) regulate corneal homeostasis and maintenance of hair follicle stem cells. Dev Biol. 2007;312(2):484-500. Epub 2007/11/10. doi: 10.1016/j.ydbio.2007.09.052. PubMed PMID: 17991461; PubMed Central PMCID: PMCPmc2494569.
11. Sivak JM, West-Mays JA, Yee A, Williams T, Fini ME. Transcription Factors Pax6 and AP-2alpha Interact To Coordinate Corneal Epithelial Repair by Controlling Expression of Matrix Metalloproteinase Gelatinase B. Mol Cell Biol. 2004;24(1):245-57. Epub 2003/12/16. PubMed PMID: 14673159; PubMed Central PMCID: PMCPmc303332.
12. Zhang X, Tseng H. Basonuclin-null mutation impairs homeostasis and wound repair in mouse corneal epithelium. PLoS One. 2007;2(10):e1087. Epub 2007/11/01. doi: 10.1371/journal.pone.0001087. PubMed PMID: 17971852; PubMed Central PMCID: PMCPmc2034529.

13. Birger Y, Davis J, Furusawa T, Rand E, Piatigorsky J, Bustin M. A role for chromosomal protein HMGN1 in corneal maturation. Differentiation. 2006;74(1):19-29. Epub 2006/02/10. doi: 10.1111/j.1432-0436.2006.00054.x. PubMed PMID: 16466397.
14. Nakamura T, Ohtsuka T, Sekiyama E, Cooper LJ, Kokubu H, Fullwood NJ, et al. Hes1 regulates corneal development and the function of corneal epithelial stem/progenitor cells. Stem Cells. 2008;26(5):1265-74. Epub 2008/02/23. doi: 10.1634/stemcells.2007-1067. PubMed PMID: 18292208.
15. Nakatsu MN, Ding Z, Ng MY, Truong TT, Yu F, Deng SX. Wnt/beta-catenin signaling regulates proliferation of human cornea epithelial stem/progenitor cells. Invest Ophthalmol Vis Sci. 2011;52(7):4734-41. Epub 2011/03/02. doi: 10.1167/iovs.10-6486. PubMed PMID: 21357396; PubMed Central PMCID: PMCPmc3175950.
16. Vereb Z, Albert R, Poliska S, Olstad OK, Akhtar S, Moe MC, et al. Comparison of upstream regulators in human ex vivo cultured cornea limbal epithelial stem cells and differentiated corneal epithelial cells. BMC Genomics. 2013;14:900. Epub 2013/12/19. doi: 10.1186/1471-2164-14-900. PubMed PMID: 24344983; PubMed Central PMCID: PMCPmc3880589.
17. Wilson PA, Hemmati-Brivanlou A. Induction of epidermis and inhibition of neural fate by Bmp-4. Nature. 1995;376(6538):331-3. Epub 1995/07/27. doi: 10.1038/376331a0. PubMed PMID: 7630398.
18. Byrne C, Tainsky M, Fuchs E. Programming gene expression in developing epidermis. Development. 1994;120(9):2369-83. Epub 1994/09/01. PubMed PMID: 7525178.
19. Pei YF, Rhodin JA. The prenatal development of the mouse eye. Anat Rec. 1970;168(1):105-25. Epub 1970/09/01. doi: 10.1002/ar.1091680109. PubMed PMID: 5469558.
20. Ramaesh T, Collinson JM, Ramaesh K, Kaufman MH, West JD, Dhillon B. Corneal abnormalities in Pax6+/- small eye mice mimic human aniridia-related keratopathy. Invest Ophthalmol Vis Sci. 2003;44(5):1871-8. Epub 2003/04/26. PubMed PMID: 12714618.
21. Hay ED. Development of the vertebrate cornea. Int Rev Cytol. 1980;63:263-322. Epub 1979/01/01. PubMed PMID: 395131.
22. Tanifuji-Terai N, Terai K, Hayashi Y, Chikama T, Kao WW. Expression of keratin 12 and maturation of corneal epithelium during development and postnatal growth. Invest Ophthalmol Vis Sci. 2006;47(2):545-51. Epub 2006/01/25. doi: 10.1167/iovs.05-1182. PubMed PMID: 16431949.
23. Pei YF, Rhodin JA. Electron microscopic study of the development of the mouse corneal epithelium. Invest Ophthalmol. 1971;10(11):811-25. Epub 1971/11/01. PubMed PMID: 5116630.
24. Liu CY, Zhu G, Westerhausen-Larson A, Converse R, Kao CW, Sun TT, et al. Cornea-specific expression of K12 keratin during mouse development. Curr Eye Res. 1993;12(11):963-74. Epub 1993/11/01. PubMed PMID: 7508359.
25. Yu H, Luscombe NM, Qian J, Gerstein M. Genomic analysis of gene expression relationships in transcriptional regulatory networks. Trends Genet. 2003;19(8):422-7. Epub 2003/08/07. doi: 10.1016/s0168-9525(03)00175-6. PubMed PMID: 12902159.
26. Yu Z, Lin KK, Bhandari A, Spencer JA, Xu X, Wang N, et al. The Grainyhead-like epithelial transactivator Get-1/Grhl3 regulates epidermal terminal differentiation and interacts functionally with LMO4. Dev Biol. 299. United States2006. p. 122-36.

27. Yu Z, Bhandari A, Mannik J, Pham T, Xu X, Andersen B. Grainyhead-like factor Get1/Grhl3 regulates formation of the epidermal leading edge during eyelid closure. Dev Biol. 319. United States2008. p. 56-67.
28. Caddy J, Wilanowski T, Darido C, Dworkin S, Ting SB, Zhao Q, et al. Epidermal wound repair is regulated by the planar cell polarity signaling pathway. Dev Cell. 19. United States: 2010 Elsevier Inc; 2010. p. 138-47.
29. Swamynathan SK. Kruppel-like factors: three fingers in control. Hum Genomics. 2010;4(4):263-70. Epub 2010/06/01. PubMed PMID: 20511139; PubMed Central PMCID: PMCPmc2975451.
30. Zhang P, Andrianakos R, Yang Y, Liu C, Lu W. Kruppel-like factor 4 (Klf4) prevents embryonic stem (ES) cell differentiation by regulating Nanog gene expression. J Biol Chem. 2010;285(12):9180-9. Epub 2010/01/15. doi: 10.1074/jbc.M109.077958. PubMed PMID: 20071344; PubMed Central PMCID: PMCPmc2838337.
31. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126(4):663-76. Epub 2006/08/15. doi: 10.1016/j.cell.2006.07.024. PubMed PMID: 16904174.
32. McConnell BB, Ghaleb AM, Nandan MO, Yang VW. The diverse functions of Kruppel-like factors 4 and 5 in epithelial biology and pathobiology. Bioessays. 2007;29(6):549-57. Epub 2007/05/18. doi: 10.1002/bies.20581. PubMed PMID: 17508399; PubMed Central PMCID: PMCPmc2211634.
33. Jaubert J, Cheng J, Segre JA. Ectopic expression of kruppel like factor 4 (Klf4) accelerates formation of the epidermal permeability barrier. Development. 2003;130(12):2767-77. Epub 2003/05/09. PubMed PMID: 12736219.
34. Zhang Z, Wang H, Sun Y, Li H, Wang N. Klf7 modulates the differentiation and proliferation of chicken preadipocyte. Acta Biochim Biophys Sin (Shanghai). 2013;45(4):280-8. Epub 2013/02/27. doi: 10.1093/abbs/gmt010. PubMed PMID: 23439665.
35. Laub F, Lei L, Sumiyoshi H, Kajimura D, Dragomir C, Smaldone S, et al. Transcription factor KLF7 is important for neuronal morphogenesis in selected regions of the nervous system. Mol Cell Biol. 2005;25(13):5699-711. Epub 2005/06/21. doi: 10.1128/mcb.25.13.5699-5711.2005. PubMed PMID: 15964824; PubMed Central PMCID: PMCPmc1157008.
36. Vermeulen M, Mulder KW, Denissov S, Pijnappel WW, van Schaik FM, Varier RA, et al. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. Cell. 131. United States2007. p. 58-69.
37. Margueron R, Reinberg D. The Polycomb complex PRC2 and its mark in life. Nature. 469. England2011. p. 343-9.
38. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. Nature. 470. England2011. p. 279-83.
39. Biterge B, Schneider R. Histone variants: key players of chromatin. Cell Tissue Res. 2014;356(3):457-66. Epub 2014/05/02. doi: 10.1007/s00441-014-1862-4. PubMed PMID: 24781148.
40. Bai L, Morozov AV. Gene regulation by nucleosome positioning. Trends Genet. 26. England: 2010 Elsevier Ltd; 2010. p. 476-83.

41. He HH, Meyer CA, Shin H, Bailey ST, Wei G, Wang Q, et al. Nucleosome dynamics define transcriptional enhancers. Nat Genet. 42. United States2010. p. 343-7.
42. Pekowska A, Benoukraf T, Zacarias-Cabeza J, Belhocine M, Koch F, Holota H, et al. H3K4 tri-methylation provides an epigenetic signature of active enhancers. Embo j. 2011;30(20):4198-210. Epub 2011/08/19. doi: 10.1038/emboj.2011.295. PubMed PMID: 21847099; PubMed Central PMCID: PMCPmc3199384.
43. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet. 2007;39(3):311-8. Epub 2007/02/06. doi: 10.1038/ng1966. PubMed PMID: 17277777.
44. Hubert A, Henderson JM, Ross KG, Cowles MW, Torres J, Zayas RM. Epigenetic regulation of planarian stem cells by the SET1/MLL family of histone methyltransferases. Epigenetics. 2013;8(1):79-91. Epub 2012/12/14. doi: 10.4161/epi.23211. PubMed PMID: 23235145; PubMed Central PMCID: PMCPmc3549883.
45. Justin N, De Marco V, Aasland R, Gamblin SJ. Reading, writing and editing methylated lysines on histone tails: new insights from recent structural studies. Curr Opin Struct Biol. 2010;20(6):730-8. Epub 2010/10/20. doi: 10.1016/j.sbi.2010.09.012. PubMed PMID: 20956082.
46. Ernst P, Vakoc CR. WRAD: enabler of the SET1-family of H3K4 methyltransferases. Brief Funct Genomics. 2012;11(3):217-26. Epub 2012/06/02. doi: 10.1093/bfgp/els017. PubMed PMID: 22652693; PubMed Central PMCID: PMCPmc3388306.
47. Hu D, Gao X, Morgan MA, Herz HM, Smith ER, Shilatifard A. The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. Mol Cell Biol. 2013;33(23):4745-54. Epub 2013/10/02. doi: 10.1128/mcb.01181-13. PubMed PMID: 24081332; PubMed Central PMCID: PMCPmc3838007.
48. Ezhkova E, Pasolli HA, Parker JS, Stokes N, Su IH, Hannon G, et al. Ezh2 orchestrates gene expression for the stepwise differentiation of tissue-specific stem cells. Cell. 136. United States2009. p. 1122-35.
49. Mardaryev AN, Gdula MR, Yarker JL, Emelianov VU, Poterlowicz K, Sharov AA, et al. p63 and Brg1 control developmentally regulated higher-order chromatin remodelling at the epidermal differentiation complex locus in epidermal progenitor cells. Development. 141. England2014. p. 101-11.
50. Fessing MY, Mardaryev AN, Gdula MR, Sharov AA, Sharova TY, Rapisarda V, et al. p63 regulates Satb1 to control tissue-specific chromatin remodeling during development of the epidermis. J Cell Biol. 194. United States2011. p. 825-39.
51. Bao X, Tang J, Lopez-Pajares V, Tao S, Qu K, Crabtree GR, et al. ACTL6a enforces the epidermal progenitor state by suppressing SWI/SNF-dependent induction of KLF4. Cell Stem Cell. 2013;12(2):193-203. Epub 2013/02/12. doi: 10.1016/j.stem.2012.12.014. PubMed PMID: 23395444; PubMed Central PMCID: PMCPmc3661004.
52. Hopkin AS, Gordon W, Klein RH, Espitia F, Daily K, Zeller M, et al. GRHL3/GET1 and Trithorax Group Members Collaborate to Activate the Epidermal Progenitor Differentiation Program. PLoS Genet. 8. United States2012. p. e1002829.
53. Gordon WM, Zeller MD, Klein RH, Swindell WR, Ho H, Espetia F, et al. A GRHL3-regulated repair pathway suppresses immune-mediated epidermal hyperplasia. J Clin

Invest. 2014;124(12):5205-18. Epub 2014/10/28. doi: 10.1172/jci77138. PubMed PMID: 25347468; PubMed Central PMCID: PMCPmc4348962.
54. Yu Z, Mannik J, Soto A, Lin KK, Andersen B. The epidermal differentiation-associated Grainyhead gene Get1/Grhl3 also regulates urothelial differentiation. Embo j. 2009;28(13):1890-903. Epub 2009/06/06. doi: 10.1038/emboj.2009.142. PubMed PMID: 19494835; PubMed Central PMCID: PMCPmc2711180.
55. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, et al. Super-enhancers in the control of cell identity and disease. Cell. 2013;155(4):934-47. Epub 2013/10/15. doi: 10.1016/j.cell.2013.09.053. PubMed PMID: 24119843; PubMed Central PMCID: PMCPmc3841062.
56. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, et al. Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. Cell. 2013;153(2):307-19. doi: 10.1016/j.cell.2013.03.035. PubMed PMID: PMC3653129.
57. Krivtsov AV, Armstrong SA. MLL translocations, histone modifications and leukaemia stem-cell development. Nat Rev Cancer. 2007;7(11):823-33. Epub 2007/10/25. doi: 10.1038/nrc2253. PubMed PMID: 17957188.
58. Chang MJ, Wu H, Achille NJ, Reisenauer MR, Chou CW, Zeleznik-Le NJ, et al. Histone H3 lysine 79 methyltransferase Dot1 is required for immortalization by MLL oncogenes. Cancer Res. 2010;70(24):10234-42. Epub 2010/12/17. doi: 10.1158/0008-5472.can-10-3294. PubMed PMID: 21159644; PubMed Central PMCID: PMCPmc3040779.
59. Saigo K, Yoshida K, Ikeda R, Sakamoto Y, Murakami Y, Urashima T, et al. Integration of hepatitis B virus DNA into the myeloid/lymphoid or mixed-lineage leukemia (MLL4) gene and rearrangements of MLL4 in human hepatocellular carcinoma. Hum Mutat. 2008;29(5):703-8. Epub 2008/03/06. doi: 10.1002/humu.20701. PubMed PMID: 18320596.
60. Pickering CR, Zhou JH, Lee JJ, Drummond JA, Peng SA, Saade RE, et al. Mutational landscape of aggressive cutaneous squamous cell carcinoma. Clin Cancer Res. 2014;20(24):6582-92. Epub 2014/10/12. doi: 10.1158/1078-0432.ccr-14-1768. PubMed PMID: 25303977; PubMed Central PMCID: PMCPmc4367811.
61. Martin D, Abba MC, Molinolo AA, Vitale-Cross L, Wang Z, Zaida M, et al. The head and neck cancer cell oncogenome: a platform for the development of precision molecular therapies. Oncotarget. 2014;5(19):8906-23. Epub 2014/10/03. PubMed PMID: 25275298; PubMed Central PMCID: PMCPmc4253406.
62. Miyake N, Koshimizu E, Okamoto N, Mizuno S, Ogata T, Nagai T, et al. MLL2 and KDM6A mutations in patients with Kabuki syndrome. Am J Med Genet A. 2013;161a(9):2234-43. Epub 2013/08/06. doi: 10.1002/ajmg.a.36072. PubMed PMID: 23913813.
63. Jones WD, Dafou D, McEntagart M, Woollard WJ, Elmslie FV, Holder-Espinasse M, et al. De novo mutations in MLL cause Wiedemann-Steiner syndrome. Am J Hum Genet. 2012;91(2):358-64. Epub 2012/07/17. doi: 10.1016/j.ajhg.2012.06.008. PubMed PMID: 22795537; PubMed Central PMCID: PMCPmc3415539.
64. Yang J, Corces VG. Insulators, long-range interactions, and genome function. Curr Opin Genet Dev. 22. England: 2012 Elsevier Ltd; 2012. p. 86-92.
65. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature. 473. England: 2011 Macmillan Publishers Limited. All rights reserved; 2011. p. 43-9.

66. Wang D, Garcia-Bassets I, Benner C, Li W, Su X, Zhou Y, et al. Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. Nature. 2011;474(7351):390-4. Epub 2011/05/17. doi: 10.1038/nature10006. PubMed PMID: 21572438; PubMed Central PMCID: PMCPmc3117022.
67. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, et al. Widespread transcription at neuronal activity-regulated enhancers. Nature. 2010;465(7295):182-7. Epub 2010/04/16. doi: 10.1038/nature09033. PubMed PMID: 20393465; PubMed Central PMCID: PMCPmc3020079.
68. Zhang J, Poh HM, Peh SQ, Sia YY, Li G, Mulawadi FH, et al. ChIA-PET Analysis of Transcriptional Chromatin Interactions. Methods: 2012. Published by Elsevier Inc.; 2012.
69. Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, Wang P, et al. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell. 148. United States: 2012 Elsevier Inc; 2012. p. 84-98.
70. Jurata LW, Kenny DA, Gill GN. Nuclear LIM interactor, a rhombotin and LIM homeodomain interacting protein, is expressed early in neuronal development. Proc Natl Acad Sci U S A. 1996;93(21):11693-8. Epub 1996/10/15. PubMed PMID: 8876198; PubMed Central PMCID: PMCPmc38120.
71. Agulnick AD, Taira M, Breen JJ, Tanaka T, Dawid IB, Westphal H. Interactions of the LIM-domain-binding factor Ldb1 with LIM homeodomain proteins. Nature. 1996;384(6606):270-2. Epub 1996/11/21. doi: 10.1038/384270a0. PubMed PMID: 8918878.
72. Bach I, Carriere C, Ostendorff HP, Andersen B, Rosenfeld MG. A family of LIM domain-associated cofactors confer transcriptional synergism between LIM and Otx homeodomain proteins. Genes Dev. 1997;11(11):1370-80. Epub 1997/06/01. PubMed PMID: 9192866.
73. Storbeck CJ, Wagner S, O'Reilly P, McKay M, Parks RJ, Westphal H, et al. The Ldb1 and Ldb2 transcriptional cofactors interact with the Ste20-like kinase SLK and regulate cell migration. Mol Biol Cell. 2009;20(19):4174-82. Epub 2009/08/14. doi: 10.1091/mbc.E08-07-0707. PubMed PMID: 19675209; PubMed Central PMCID: PMCPmc2754931.
74. Li L, Jothi R, Cui K, Lee JY, Cohen T, Gorivodsky M, et al. Nuclear adaptor Ldb1 regulates a transcriptional program essential for the maintenance of hematopoietic stem cells. Nat Immunol. 2011;12(2):129-36. Epub 2010/12/28. doi: 10.1038/ni.1978. PubMed PMID: 21186366.
75. Dey-Guha I, Mukhopadhyay M, Phillips M, Westphal H. Role of ldb1 in adult intestinal homeostasis. Int J Biol Sci. 2009;5(7):686-94. Epub 2009/11/18. PubMed PMID: 19918297; PubMed Central PMCID: PMCPmc2777272.
76. Salmans M, Yu Z, Watanabe K, Cam E, Sun P, Smyth P, et al. The Co-Factor of LIM domains (CLIM/LDB/NLI) maintains basal mammary epithelial stem cells and promotes breast tumorigenesis. PLOS Genetics; 2014.
77. Soler E, Andrieu-Soler C, de Boer E, Bryne JC, Thongjuea S, Stadhouders R, et al. The genome-wide dynamics of the binding of Ldb1 complexes during erythroid differentiation. Genes Dev. 2010;24(3):277-89. Epub 2010/02/04. doi: 10.1101/gad.551810. PubMed PMID: 20123907; PubMed Central PMCID: PMCPmc2811829.
78. Mukhopadhyay M, Teufel A, Yamashita T, Agulnick AD, Chen L, Downs KM, et al. Functional ablation of the mouse Ldb1 gene results in severe patterning defects during

gastrulation. Development. 2003;130(3):495-505. Epub 2002/12/20. PubMed PMID: 12490556.
79. Deng W, Lee J, Wang H, Miller J, Reik A, Gregory PD, et al. Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. Cell. 2012;149(6):1233-44. Epub 2012/06/12. doi: 10.1016/j.cell.2012.03.051. PubMed PMID: 22682246; PubMed Central PMCID: PMCPmc3372860.
80. Kaffer CR, Grinberg A, Pfeifer K. Regulatory Mechanisms at the Mouse Igf2/H19 Locus. Molecular and Cellular Biology. 2001;21(23):8189-96. doi: 10.1128/MCB.21.23.8189-8196.2001. PubMed PMID: PMC99983.
81. Reinius B, Kanduri C. Elevated expression of H19 and Igf2 in the female mouse eye. PLoS One. 2013;8(2):e56611. Epub 2013/02/26. doi: 10.1371/journal.pone.0056611. PubMed PMID: 23437185; PubMed Central PMCID: PMCPmc3577879.
82. Coster DJ, Aggarwal RK, Williams KA. Surgical management of ocular surface disorders using conjunctival and stem cell allografts. Br J Ophthalmol. 1995;79(11):977-82. Epub 1995/11/01. PubMed PMID: 8534667; PubMed Central PMCID: PMCPmc505310.
83. Secker GA, Daniels JT. Corneal epithelial stem cells: deficiency and regulation. Stem Cell Rev. 2008;4(3):159-68. Epub 2008/07/16. doi: 10.1007/s12015-008-9029-x. PubMed PMID: 18622724.
84. Chen JJ, Tseng SC. Corneal epithelial wound healing in partial limbal deficiency. Invest Ophthalmol Vis Sci. 1990;31(7):1301-14. Epub 1990/07/01. PubMed PMID: 1694836.
85. Sherman BT, Huang da W, Tan Q, Guo Y, Bour S, Liu D, et al. DAVID Knowledgebase: a gene-centered database integrating heterogeneous gene annotation resources to facilitate high-throughput gene functional analysis. BMC Bioinformatics. 2007;8:426. Epub 2007/11/06. doi: 10.1186/1471-2105-8-426. PubMed PMID: 17980028; PubMed Central PMCID: PMCPmc2186358.
86. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009;37(1):1-13. Epub 2008/11/27. doi: 10.1093/nar/gkn923. PubMed PMID: 19033363; PubMed Central PMCID: PMCPmc2615629.
87. Zhou M, Li XM, Lavker RM. Transcriptional profiling of enriched populations of stem cells versus transient amplifying cells. A comparison of limbal and corneal epithelial basal cells. J Biol Chem. 2006;281(28):19600-9. Epub 2006/05/06. doi: 10.1074/jbc.M600777200. PubMed PMID: 16675456.
88. Folgueras AR, Guo X, Pasolli HA, Stokes N, Polak L, Zheng D, et al. Architectural niche organization by LHX2 is linked to hair follicle stem cell function. Cell Stem Cell. 2013;13(3):314-27. Epub 2013/09/10. doi: 10.1016/j.stem.2013.06.018. PubMed PMID: 24012369.
89. Povinelli BJ, Nemeth MJ. Wnt5a Regulates Hematopoietic Stem Cell Proliferation and Repopulation Through the Ryk Receptor. Stem Cells. 2013. Epub 2013/08/14. doi: 10.1002/stem.1513. PubMed PMID: 23939973.
90. Li J, Ying J, Fan Y, Wu L, Ying Y, Chan AT, et al. WNT5A antagonizes WNT/beta-catenin signaling and is frequently silenced by promoter CpG methylation in esophageal squamous cell carcinoma. Cancer Biol Ther. 2010;10(6):617-24. Epub 2010/07/07. PubMed PMID: 20603606.

91. Singh P, Lee D-H, Szabó PE. More than insulator: multiple roles of CTCF at the H19-Igf2 imprinted domain. Frontiers in Genetics. 2012;3:214. doi: 10.3389/fgene.2012.00214. PubMed PMID: PMC3471092.
92. Andersson Gronlund M, Dahlgren J, Aring E, Kraemer M, Hellstrom A. Ophthalmological findings in children and adolescents with Silver-Russell syndrome. Br J Ophthalmol. 2011;95(5):637-41. Epub 2010/09/02. doi: 10.1136/bjo.2010.184457. PubMed PMID: 20805133.
93. Azzi S, Abi Habib W, Netchine I. Beckwith-Wiedemann and Russell-Silver Syndromes: from new molecular insights to the comprehension of imprinting regulation. Curr Opin Endocrinol Diabetes Obes. 2014;21(1):30-8. Epub 2013/12/11. doi: 10.1097/med.0000000000000037. PubMed PMID: 24322424.
94. Johnsen SA, Gungor C, Prenzel T, Riethdorf S, Riethdorf L, Taniguchi-Ishigaki N, et al. Regulation of estrogen-dependent transcription by the LIM cofactors CLIM and RLIM in breast cancer. Cancer Res. 2009;69(1):128-36. Epub 2009/01/02. doi: 10.1158/0008-5472.can-08-1630. PubMed PMID: 19117995; PubMed Central PMCID: PMCPmc2713826.
95. Wang SB, Hu KM, Seamon KJ, Mani V, Chen Y, Gronert K. Estrogen negatively regulates epithelial wound healing and protective lipid mediator circuits in the cornea. Faseb j. 2012;26(4):1506-16. Epub 2011/12/22. doi: 10.1096/fj.11-198036. PubMed PMID: 22186873; PubMed Central PMCID: PMCPmc3316908.
96. Rhee H, Polak L, Fuchs E. Lhx2 maintains stem cell character in hair follicles. Science. 2006;312(5782):1946-9. Epub 2006/07/01. doi: 10.1126/science.1128004. PubMed PMID: 16809539; PubMed Central PMCID: PMCPmc2405918.
97. Mardaryev AN, Meier N, Poterlowicz K, Sharov AA, Sharova TY, Ahmed MI, et al. Lhx2 differentially regulates Sox9, Tcf4 and Lgr5 in hair follicle stem cells to promote epidermal regeneration after injury. Development. 2011;138(22):4843-52. Epub 2011/10/27. doi: 10.1242/dev.070284. PubMed PMID: 22028024.
98. Zieske JD. Corneal development associated with eyelid opening. Int J Dev Biol. 2004;48(8-9):903-11. Epub 2004/11/24. doi: 10.1387/ijdb.041860jz. PubMed PMID: 15558481.
99. Yoon JJ, Ismail S, Sherwin T. Limbal stem cells: Central concepts of corneal epithelial homeostasis. World J Stem Cells. 2014;6(4):391-403. Epub 2014/09/27. doi: 10.4252/wjsc.v6.i4.391. PubMed PMID: 25258661; PubMed Central PMCID: PMCPmc4172668.
100. Kleinjan DA, Seawright A, Schedl A, Quinlan RA, Danes S, van Heyningen V. Aniridia-associated translocations, DNase hypersensitivity, sequence comparison and transgenic analysis redefine the functional domain of PAX6. Hum Mol Genet. 2001;10(19):2049-59. Epub 2001/10/09. PubMed PMID: 11590122.
101. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57-74. Epub 2012/09/08. doi: 10.1038/nature11247. PubMed PMID: 22955616; PubMed Central PMCID: PMCPmc3439153.
102. Swamynathan SK, Katz JP, Kaestner KH, Ashery-Padan R, Crawford MA, Piatigorsky J. Conditional deletion of the mouse Klf4 gene results in corneal epithelial fragility, stromal edema, and loss of conjunctival goblet cells. Mol Cell Biol. 2007;27(1):182-94. Epub 2006/10/25. doi: 10.1128/mcb.00846-06. PubMed PMID: 17060454; PubMed Central PMCID: PMCPmc1800665.

103. Swamynathan S, Buela KA, Kinchington P, Lathrop KL, Misawa H, Hendricks RL, et al. Klf4 regulates the expression of Slurp1, which functions as an immunomodulatory peptide in the mouse cornea. Invest Ophthalmol Vis Sci. 2012;53(13):8433-46. Epub 2012/11/10. doi: 10.1167/iovs.12-10759. PubMed PMID: 23139280.
104. Genin E, Schumacher M, Roujeau JC, Naldi L, Liss Y, Kazma R, et al. Genome-wide association study of Stevens-Johnson Syndrome and Toxic Epidermal Necrolysis in Europe. Orphanet J Rare Dis. 2011;6:52. Epub 2011/08/02. doi: 10.1186/1750-1172-6-52. PubMed PMID: 21801394; PubMed Central PMCID: PMCPmc3173287.
105. Fan Q, Zhou X, Khor CC, Cheng CY, Goh LK, Sim X, et al. Genome-wide meta-analysis of five Asian cohorts identifies PDGFRA as a susceptibility locus for corneal astigmatism. PLoS Genet. 2011;7(12):e1002402. Epub 2011/12/07. doi: 10.1371/journal.pgen.1002402. PubMed PMID: 22144915; PubMed Central PMCID: PMCPmc3228826.
106. Lopez-Pajares V, 321.083px dd-c-wsl, 444.583px t, 20px f-s, sans-serif f-f, scaleX(1.002) t, et al. A LncRNA-MAF:MAFB Transcription Factor Network Regulates Epidermal Differentiation. Developmental Cell; 2015.
107. Boxer LD, Barajas B, Tao S, Zhang J, Khavari PA. ZNF750 interacts with KLF4 and RCOR1, KDM1A, and CTBP1/2 chromatin regulators to repress epidermal progenitor genes and induce differentiation genes. Genes Dev. 28. United States: 2014 Boxer et al.; Published by Cold Spring Harbor Laboratory Press.; 2014. p. 2013-26.
108. Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature. 2012;481(7381):389-93. Epub 2012/01/06. doi: 10.1038/nature10730. PubMed PMID: 22217937; PubMed Central PMCID: PMCPmc3272464.
109. Kon S, Tanabe K, Watanabe T, Sabe H, Satake M. Clathrin dependent endocytosis of E-cadherin is regulated by the Arf6GAP isoform SMAP1. Exp Cell Res. 2008;314(7):1415-28. Epub 2008/03/12. doi: 10.1016/j.yexcr.2007.11.006. PubMed PMID: 18331728.
110. Kurschat P, Bielenberg D, Rossignol-Tallandier M, Stahl A, Klagsbrun M. Neuron restrictive silencer factor NRSF/REST is a transcriptional repressor of neuropilin-1 and diminishes the ability of semaphorin 3A to inhibit keratinocyte migration. J Biol Chem. 2006;281(5):2721-9. Epub 2005/12/07. doi: 10.1074/jbc.M507860200. PubMed PMID: 16330548.
111. Schmidt D, Wilson MD, Spyrou C, Brown GD, Hadfield J, Odom DT. ChIP-seq: using high-throughput sequencing to discover protein-DNA interactions. Methods. 2009;48(3):240-8. Epub 2009/03/12. doi: 10.1016/j.ymeth.2009.03.001. PubMed PMID: 19275939.
112. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10(3):R25. Epub 2009/03/06. doi: 10.1186/gb-2009-10-3-r25. PubMed PMID: 19261174; PubMed Central PMCID: PMCPmc2690996.
113. Xu S, Grullon S, Ge K, Peng W. Spatial clustering for identification of ChIP-enriched regions (SICER) to map regions of histone methylation patterns in embryonic stem cells. Methods Mol Biol. 2014;1150:97-111. Epub 2014/04/20. doi: 10.1007/978-1-4939-0512-6_5. PubMed PMID: 24743992; PubMed Central PMCID: PMCPmc4152844.
114. Quinlan AR. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. Curr Protoc Bioinformatics. 2014;47:11.2.1-.2.34. Epub 2014/09/10. doi:

10.1002/0471250953.bi1112s47. PubMed PMID: 25199790; PubMed Central PMCID: PMCPmc4213956.
115. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9(9):R137. Epub 2008/09/19. doi: 10.1186/gb-2008-9-9-r137. PubMed PMID: 18798982; PubMed Central PMCID: PMCPmc2592715.
116. Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, et al. Galaxy: a web-based genome analysis tool for experimentalists. Curr Protoc Mol Biol. 2010;Chapter 19:Unit 19.0.1-21. Epub 2010/01/14. doi: 10.1002/0471142727.mb1910s89. PubMed PMID: 20069535.
117. Goecks J, Nekrutenko A, Taylor J. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol. 2010;11(8):R86. Epub 2010/08/27. doi: 10.1186/gb-2010-11-8-r86. PubMed PMID: 20738864; PubMed Central PMCID: PMCPmc2945788.
118. Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, et al. Galaxy: a platform for interactive large-scale genome analysis. Genome Res. 2005;15(10):1451-5. Epub 2005/09/20. doi: 10.1101/gr.4086505. PubMed PMID: 16169926; PubMed Central PMCID: PMCPmc1240089.
119. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol. 2010;28(5):495-501. Epub 2010/05/04. doi: 10.1038/nbt.1630. PubMed PMID: 20436461.
120. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009;37(Web Server issue):W202-8. Epub 2009/05/22. doi: 10.1093/nar/gkp335. PubMed PMID: 19458158; PubMed Central PMCID: PMCPmc2703892.
121. Loven J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell. 2013;153(2):320-34. Epub 2013/04/16. doi: 10.1016/j.cell.2013.03.036. PubMed PMID: 23582323; PubMed Central PMCID: PMCPmc3760967.
122. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010;38(4):576-89. Epub 2010/06/02. doi: 10.1016/j.molcel.2010.05.004. PubMed PMID: 20513432; PubMed Central PMCID: PMCPmc2898526.
123. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-50. Epub 2005/10/04. doi: 10.1073/pnas.0506580102. PubMed PMID: 16199517; PubMed Central PMCID: PMCPmc1239896.
124. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet. 2003;34(3):267-73. Epub 2003/06/17. doi: 10.1038/ng1180. PubMed PMID: 12808457.
125. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44-57. Epub 2009/01/10. doi: 10.1038/nprot.2008.211. PubMed PMID: 19131956.