Functional characterization of SPRR5 in neoplasia and epidermal homeostasis

D I S S E R T A T I O N

zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)
der Fakultät für Biologie und Vorklinische Medizin
der Universität Regensburg

vorgelegt von Christian Ziegler

aus Bad Mergentheim

im Jahr 2018

Das Promotionsgesuch wurde eingereicht am:

03. Juli 2018

Die Arbeit wurde angeleitet von:

PD Dr. Markus Kretz

Unterschrift:

Table of Contents

1 Abstract

Constituting an effective barrier against environmental challenges such as pathogen invasion, UV-radiation or prevention of extensive water loss, the epidermis as the outermost layer of the human skin has to be constantly regenerated in order to maintain its protective function. During this highly elaborate and delicate process of epidermal rejuvenation, progenitor keratinocytes conduct a terminal differentiation program which is accompanied by fluctuating expression of differentiation proteins and ultimately results in tightly agglutinated, dead and flattened cells that are embedded in an extracellular lipid layer.

In light of these profound alterations in gene expression, it is not surprising that epidermal homeostasis is controlled by an extensive network of signaling pathways, transcription factors and more recently also the involvement of long non-coding RNAs (lncRNAs) for this vital process has been uncovered. One of these regulatory lncRNAs might be the 326 nt spanning transcript SPRR5_326, which was initially identified from full transcriptome sequencing as a lncRNA that is induced over the course of keratinocyte differentiation and at the same time suppressed in squamous cell carcinoma samples. According to this particular expression pattern, a dual role of SPRR5_326 for both these processes seemed promising and the impact of sustained SPRR5_326 expression on skin carcinogenesis was tested by an *in vitro* invasion assay and an *in vivo* tumor formation assay with inconclusive results. Surprisingly, a subsequent in-depth isoform characterization for the novel SPRR5 locus revealed that the assumed SPRR5_326 transcript was not final but part of a longer (726 nt) SPRR5 transcript that is controlled by the transcription factor p63 and seems to be primarily localized in the cytoplasm of differentiated keratinocytes.

In contrast to the results from the conducted tumor formation assays, an indispensable functional impact of SPRR5 for keratinocyte differentiation could be established for *in vitro* differentiated keratinocytes as well as in regenerated organotypic epidermal tissue. Additionally, a global RNA-sequencing experiment ultimately proved the necessity of SPRR5 for a proper terminal differentiation program and hinted towards an epigenetic mechanism. Therefore, an Assay for Transposase Accessible Chromatin (ATAC-Seq) as well as a Chromatin Immunoprecipitation (ChIP-Seq) experiment for the histone marks H3K27me3, H3K4me1 and H3K27ac was conducted in SPRR5 deficient and control keratinocytes, however the results from these experiments could not support this initial assumption.

Interestingly, the SPRR5 transcript also encloses an open reading frame for a putative SPRR5 protein, which shares a high degree of similarity to the human small proline-rich protein (SPRR) family. On the other hand, a subsequent phylogenetic analysis uncovered the separate evolution of SPRR5 apart from the other SPRRs and thus suggests a different mode of action for SPRR5. Based on these observations, the actual presence of the predicted SPRR5 protein was investigated and only minor protein amounts could be detected with a highly targeted and sensitive mass spectrometry approach, raising the question whether SPRR5 functions as a protein, a lncRNA or a combination of both. Eventually this enigma should be solved by rescue experiments, though unfortunately the siRNA-mediated SPRR5 depletion lead to inconclusive results and also the generation of a homozygous SPRR5 knockout cell line was unsuccessful, precluding a final statement about the functional SPRR5 molecule in keratinocytes.

In conclusion an in-depth characterization of the human SPRR5 locus, including the detection as well as the functional testing of minor SPRR5 protein amounts has been conducted and the hypothesized necessity of SPRR5 for epidermal homeostasis could be ultimately confirmed during this work.

List of Abbreviations

AA	amino acid
A/A	antibiotic/antimycotic
AEBSF	4-(2-aminoethyl) benzenesulfonyl fluoride
Amp	Ampicillin
ANCR	anti-differentiation ncRNA
APS	ammonium persulfate
ATAC	Assay for Transposase Accessible Chromatin
BANCR	BRAF-regulated lncRNA
BCA	bicinchoninic acid
BCC	basal cell carcinoma
BCS	bovine calf serum
bp	base pair
Cas9	CRISPR associated protein 9
cDNA	complementary DNA
Cdk4	cyclin-dependent kinase 4
CDS	coding sequence
ChEA	ChIP Enrichment Analysis
ChIP	Chromatin Immunoprecipitation
chr	Chromosome
CMV	Cytomegalovirus
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CSF	codon substitution frequency
Ctrl	Control
Cy5	Cyanine5
d, D	Day
Da	Dalton (unified atomic mass unit)
DAG	Diacylglycerol
DAPI	4',6-diamidino-2-phenylindole
DB	dialysis buffer
diff	differentiated
DMSO	dimethyl sulfoxide
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
DNMT1	DNA (cytosine-5) methyltransferase 1
dNTP	deoxynucleotide triphosphate
DPBS	Dulbecco's phosphate-buffered saline
DTT	Dithiothreitol
dTTP	deoxythymidine triphosphate
et al.	et alia
EDC	epidermal differentiation complex
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
ERCC	External RNA Controls Consortium
EZH	Enhancer of zeste homolog
f	femto
FBS	fetal bovine serum
FC	fold change
Fendrr	FOXF1 adjacent non-coding developmental regulatory RNA
FISH	fluorescence *in situ* hybridization
FOXC1	forkhead box C1
FRIP	fraction of reads in called peak regions
g	Gram
GADD	growth arrest and DNA damage-inducible protein
GEO	Gene Expression Omnibus
GFP	green fluorescent protein
GO	gene ontology
goi	gene of interest
gRNA	guideRNA
h	hour
HaCaT	human adult skin keratinocytes propagated under low Ca^{2+} conditions and elevated temperature
HDAC	histone deacetylase
HDR	homology-directed repair
HEK	human embryonic kidney
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HF	high fidelity
hg19 / hg38	human genome build 19/38
HKGS	human keratinocyte growth supplement
HOTAIR	HOX transcript antisense RNA
hPGK	human phosphoglycerate kinase
IF	Immunofluorescence
IGV	Integrative Genomics Viewer
IP_3	inositol trisphosphate
JMJD3	Jumonji domain containing protein 3
k	kilo
KC	Keratinocyte
KD	knockdown
KGM	keratinocyte growth medium

KO	knockout
l	liter
LB	lysogeny broth
LCE	late cornified envelope
lnc-mg	myogenesis-associated lncRNA
lncRNA	long non-coding RNA
LTR	long terminal repeat
m	milli
M	molar
µ	micro
MAPK	mitogen-activated protein kinase
miRNA	microRNA
min	minute
mio	million
MOPS	3-morpholinopropane-1-sulfonic acid
mRNA	messenger RNA
MUNC	MyoD upstream no-coding
mut	mutated
n	nano
n	number (of samples)
NA	not available
n.d.	not determined
NHEJ	non-homologous end joining
NLS	nuclear localization signal
nt	nucleotides
OD	optical density
ORF	open reading frame
p	pico
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCA	principal component analysis
PCR	polymerase chain reaction
pH	power of hydrogen
PICSAR	p38 inhibited cutaneous squamous cell carcinoma associated lincRNA
PIP2	phosphatidylinositol bisphosphate
PNK	polynucleotide kinase
PRC	polycomb repressive complex
PRINS	psoriasis susceptibility-related RNA gene induced by stress
RACE	rapid amplification of cDNA ends
rcf	relative centrifugal force
RIPA	radio immunoprecipitation assay
rpm	rounds per minute
RNA	ribonucleic acid
RT	room temperature
RT-qPCR	reverse transcription-quantitative polymerase chain reaction
s	second
SCC	squamous cell carcinoma
SDS	sodium dodecyl sulfate
SENCR	smooth muscle and endothelial cell-enriched migration/differentiation-associated lncRNA
Seq	sequencing
SFM	serum-free medium
si	short interfering
SLIC	sequence- and ligation-independent cloning
SMRT-2	SCC misregulated transcript 2
SPRR	small proline-rich protein
SRA1	steroid receptor RNA activator 1
SRAP	steroid receptor RNA activator protein
TAE	Tris acetate EDTA buffer
TBS	Tris-buffered saline
TE	Tris EDTA buffer
TEMED	tetramethylethylenediamine
TGS	tris glycine SDS buffer
TINCR	terminal differentiation induced ncRNA
tracrRNA	trans-activating CRISPR RNA
Tris	tris(hydroxymethyl)aminomethane
U	Unit
UCSC	University of California at Santa Cruz
USA	United States of America
UTX	Ubiquitously transcribed tetratricopeptide repeat, X chromosome
UV	ultraviolet
V	Volt
v.	Version
v/v	volume by volume
VSV-G	G glycoprotein of the vesicular stomatitis virus
WB	wash buffer
WB	western blot
w/v	weight by volume
XIST	X-inactive-specific transcript
YFP	yellow fluorescent protein

2 Introduction

2.1 The human skin

With a surface area of almost two square meters and a portion of 15- 20% of the total body weight, the human skin is not only the largest but probably also the most important organ of the human body since it defends the body against a plethora of hazardous environmental challenges on a daily basis[1,2]. However, the protective function of the skin is not limited to interception of external influences such as microbial invasion, ultraviolet radiation, mechanical or chemical assaults but also comprises regulation of the body temperature or prevention of extensive water loss[3,4]. The secret in accomplishing all these duties lies in the sophisticated composition of the human skin, which can be divided into three layers (Figure 1). The innermost hypodermis consists mainly of fat tissue and is essential for thermal insulation, whereas the dermis as middle layer retains water, provides stability and includes sensory receptors, sweat glands, hair follicles and blood vessels which are crucial for oxygen and nutrient supply of epidermal cells[5–7]. Furthermore, the connective tissue of the dermis provides a specialized anchor platform for epidermal cells, called basement membrane, which separates the dermis from the overlying epidermis[5,7].

2.1.1 Composition and regeneration of the human epidermis

Although keratinocytes are the dominant cell type within the epidermis (90-95% of cells), interspersed neuroendocrine Merkel cells, specialized cells of the immune system as well as pigment-producing melanocytes are also present in the human epidermis and are crucial for a fully functional skin barrier[7]. Upon UV-B radiation for example, melanocytes produce melanin which has a photoprotective effect against tissue damages from reoccurring exposure to ultraviolet radiation[7,8].

As the outermost layer of the human skin, the epidermis is the first and most effective barrier against external influences and is composed of four different strata, which can be distinguished by the presence of keratinocytes in distinct differentiation states (see Figure 1)[6,7].

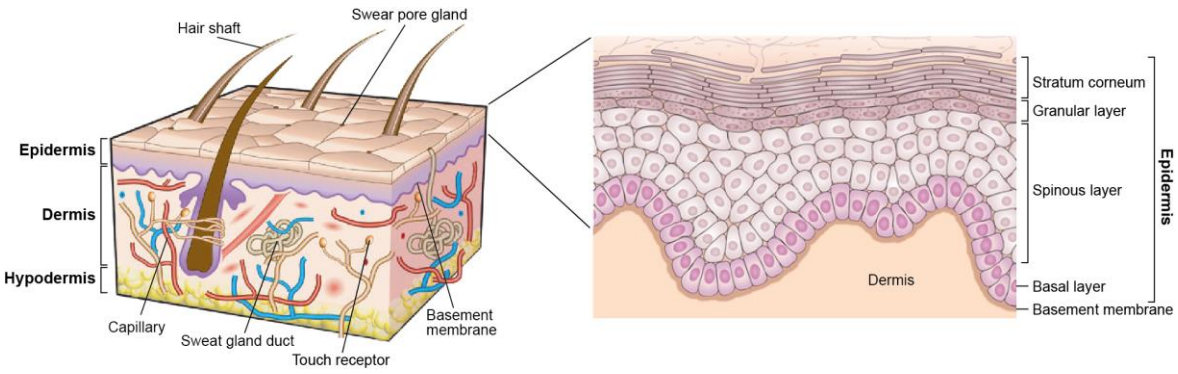


Figure 1: Cross section of the human skin and schematic overview of the epidermal layers
Schematic overview of the human skin, showing the three layers, hypodermis, dermis and epidermis as well as the therein included appendages (left; modified after[9]). Furthermore, a magnification of the epidermis, comprising the basal layer, spinous layer, granular layer and the stratum corneum is shown on the right (modified after[10]).

The deepest of the four epidermal strata is the basal layer, which mainly contains undifferentiated keratinocytes that upon cell division undergo a terminal differentiation program as they migrate through the spinous and the granular layer until they eventually reach the stratum corneum, where they undergo apoptosis and form an impenetrable protective barrier[6,11]. Due to its exposed position within the human body, this barrier is constantly weakened by environmental influences or ruptured by mechanical stress and thus has to be continuously renewed by succeeding keratinocytes. In humans, this highly elaborate process, balancing the reservoir of progenitor cells and terminally differentiating keratinocytes takes 28 days and is also known as epidermal homeostasis[12].

The process of epidermal homeostasis starts in the basal layer with undifferentiated keratinocytes, which are characterized by the expression of keratin 5 and keratin 14. In the classic model for epidermal regeneration those keratinocytes are claimed to be stem cells which give rise to short-lived progenitors known as transit-amplifying cells (TA cells)[10,11]. Upon several rounds of cell division, these TA cells would eventually withdraw from the cell cycle and conduct a terminal differentiation program as the cells migrate to the apical site of the epidermis[13,14]. Recent studies however contradict this model and claim that one uniform population of committed progenitor keratinocytes undergoes symmetric and asymmetric cell divisions with a fixed probability[15]. Whereas the cells from the symmetric cell division maintain the pool of committed progenitor cells, the progeny from asymmetric cell division is committed to terminal differentiation[10,14]. Activation of this differentiation program is initially triggered by increasing calcium concentrations in combination with signals from cell-cell contacts and becomes first apparent in the spinous layers, where the

existing keratins 5 and 14 are replaced by newly synthesized keratin 1 and keratin 10 filaments as well as involucrin[16–18]. By the time the differentiating keratinocytes reach the granular layer they express structural proteins such as loricrin, trichohyalin, small proline-rich proteins (SPRRs) and late cornified envelope proteins (LCEs) as well as profilaggrin, a highly phosphorylated oligomer of filaggrin repeats, which is stored together with keratins and loricrin in electron-dense keratohyalin granules[7,11,19–21]. Furthermore, lamellar bodies, a second type of lipid-filled granules containing mainly glycosphingolipids, free sterols and phospholipids can also be detected in granular cells[22].

During the transition from the granular layer to the stratum corneum keratinocytes undergo drastic changes and even apoptosis to eventually form a protective skin barrier. For that matter, increasing intracellular calcium levels trigger the release of profilaggrin from the keratohyalin granules, which ultimately becomes dephosphorylated and cleaved into filaggrin monomers that are subsequently bundled with the present keratin into macrofibrils, resulting in gradual flattening of the cells[11,23]. Concomitantly, transglutaminases 1, 3 and 5 introduce isopeptide bonds between these macrofibrils and other structural proteins like members of the S100 protein family, loricrin, LCEs, SPRRs and involucrin forming a rigid protein shell termed "cornified envelope"[11,20,24,25]. Finally, the lamellar bodies fuse with the plasma membrane and secrete the enclosed lipids into the extracellular space, generating a lipid lamella which is crucial for maintaining the epidermal water barrier[24].

As a result of this process, the cornified envelopes of terminally differentiated keratinocytes are tightly connected via modified desmosomes and embedded in extracellular lipid lamellae, providing the indispensable barrier against external challenges and loss of essential body fluids[20,26].

2.1.2 Regulation of epidermal homeostasis

Since generation of a functional epidermis is accompanied by profound alterations in gene expression, multiple regulatory mechanisms are required to orchestrate those changes.

By far the most important stimulus for keratinocyte differentiation arises from increasing levels of extra- and intracellular calcium concentrations, which is not only crucial because several differentiation proteins are stabilized by Ca^{2+}-ions or require Ca^{2+} for their proper function but also multiple calcium-dependent signalling pathways are involved over the course of keratinocyte differentiation[16,17,27,28]. In addition to that, cell-cell contacts which initiate a signaling cascade via the E-cadherin–catenin complex or the activated form of vitamin D_3 ($1,25(OH)_2D_3$) are also fundamental regulators of keratinocyte

differentiation[4,29,30]. Amongst other effects, all these stimuli lead to increased activity of phospholipase C which hydrolyzes phosphatidylinositol bisphosphate (PIP_2) to the second messengers inositol trisphosphate (IP_3) and diacylglycerol (DAG)[27,31]. Next, IP_3 releases calcium from intracellular storage granules and might even stimulate the influx of extracellular calcium which further expedites the differentiation program[27]. DAG on the other hand activates protein kinase C, which triggers intracellular signaling pathways including the mitogen-activated protein kinase (MAPK) pathway and ultimately results in the activation of transcription factors and thus induction of gene expression[27,32].

Apart from the described exemplary mechanism, several other signaling pathways are crucial for keratinocyte differentiation (including Notch, TGF-β and IKK/NF-κB), however almost all signaling pathways lead to the activation of specific transcription factors that ultimately induce the expression of differentiation genes[33,34]. Over the years, many transcription factors (for example members of the c-jun/c-fos family, GRHL3, Klf4, MAF:MAFB, OVOL1/2 or ZNF750) have been identified as fundamental mediators of keratinocyte differentiation but p63 was revealed as the master regulator of keratinocyte differentiation humans and mice[33,35–39]. Besides its direct effect on keratinocyte differentiation, p63 also controls the proliferative capacity of keratinocytes in the basal layer of the epidermis, which contributes to its stimulatory effect on keratinocyte differentiation and impressively accentuates its pivotal role for keratinocyte propagation and maturation[40,41]. Moreover, loss of p63 in keratinocytes results in severe defects in keratinocyte cell adhesion, basement membrane formation as well as epidermal stratification[40,42,43]. All these effects can be either reasoned by a direct p63-mediated induction of differentiation gene expression or the p63-dependent regulation of several chromatin modifiers, resulting in an epigenetic control of keratinocyte differentiation[43]. Furthermore, p63 is able to recruit the histone deacetylases 1 and 2 (HDAC1/2) in order to omit the expression of cell cycle arrest genes or alters the accessibility of epidermal enhancers in combination with the BAF chromatin remodeling complex[44,45].

Finally, keratinocyte differentiation is also regulated by several epigenetic mechanisms including chromatin remodeling or DNA methylation by DNMT1 or GADD45A/B for example[46,47]. Additionally, several histone modifications have been identified as regulators of keratinocyte differentiation. The repressive histone mark H3K27me3 for example is established in a conjoined action of the polycomb repressive complex (PRC) and its methyltransferase subunits EZH1 and EZH2, whereas target gene repression is restored by the demethylases JMJD3 and UTX[46,48–50].

A fourth layer of epigenetic regulation originates from large scale nuclear remodeling during which the size of the nucleus decreases and the epidermal differentiation complex (a genomic region on the human chromosome 1 encoding many differentiation genes) is relocated to the nuclear periphery and becomes associated with active nucleoli for induction of differentiation gene expression[51].

In conclusion, a conjoined effort of a plethora of intracellular signaling pathways, transcription factors as well as epigenetic mechanisms is mandatory to orchestrate the profound changes in the gene expression pattern during epidermal homeostasis.

2.1.3 The human epidermal differentiation complex and the SPRR protein family

Spanning 1.9 Mb on the human chromosome 1 and encoding roughly sixty proteins essential for proper keratinocyte differentiation and epidermal barrier formation, the epidermal differentiation complex (EDC) is indispensable for epidermal homeostasis[52,53]. Interestingly, gene distribution within the EDC is not random but genes encoding for proteins with similar functional and structural properties are clustered accordingly into distinct genomic locations (Figure 2). Flanked by the family of S100 proteins which are characterized by two calcium-binding EF-hands are the S100-fused type proteins containing profilaggrin, trichohyalin or repetin as well as the clusters for the late cornified envelope (LCE) and small proline-rich proteins (SPRR)[54,55]. In combination with the interspersed genes for loricrin and involucrin, the LCE and SPRR cluster account for the majority of proteins present in the cornified envelope of terminally differentiated keratinocytes[53].

As the result of a series of intra- and intergenic gene duplications from a single gene progenitor, today the human SPRR protein family comprises 10 protein-coding genes as well as one pseudogene (SPRR2C) and is massively induced during the course of keratinocyte differentiation by cyclic-AMP, retinoids, phorbols, interferon-γ, UV-radiation as well as the conjoined action of several transcription factors[56–58].

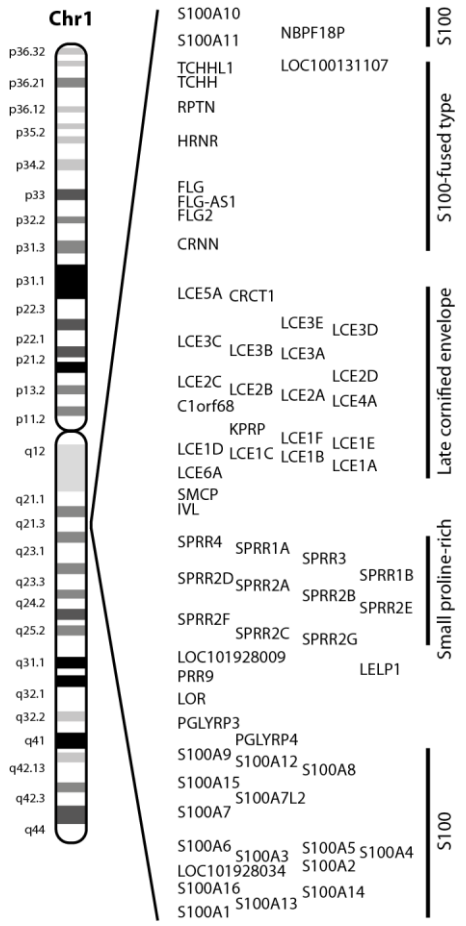


Figure 2: Overview of the human epidermal differentiation complex and its encoded genes
The human chromosome 1 is shown on the left and the respective cytogenetic bands are indicated. Furthermore, the area of the epidermal differentiation complex with its encoded genes (according to the RefSeq release 2016) is magnified and the main gene clusters are indicated on the right.

Given the evolution from a common ancestor, SPRR proteins as well as their respective genes share a high degree of similarity. However, based on their sequence conservation two major subfamilies for the human SPRR genes can be identified: the SPRR2 subfamily (SPRR2A-G) and a second group comprising SPRR1A, SPRR1B, SPRR3 as well as the most recently identified SPRR4[57,59]. Moreover, close examination of the SPRR amino-acid sequences reveals a characteristic composition, as their N- and C-termini are rich in glutamine and lysine residues that are used for protein-protein crosslinking and thus ensures the mechanical resistance of the epidermis. In contrast to this, their central domain consists of varying proline-rich repeats, allowing the formation of a flexible and non-organized structure which is essential for epidermal elasticity[11,57]. It is this combination of features that render the "spring-like" 6-18 kDa SPRR proteins excellent crosslinkers for other cornified envelope proteins like loricrin, LCEs, filaggrin or involucrin and it is therefore not surprising that this is also the best described function for SPRR proteins[11,57,59]. Apart from that, SPRR proteins are also able to detoxify reactive oxygen species (ROS) by building intra- and

intermolecular disulfide bonds between cysteine residues, which was first reported as an essential function of SPRRs in injured skin tissue[60]. Further studies eventually uncovered the SPRR2A mediated deacetylation of the p53, providing a mechanistic link between SPRR proteins and cell migration during wound healing[57,61]. Later, the antioxidant effect of small proline-rich proteins was also identified in non-wounded tissue and SPRRs were inferred as part of an effective antioxidant defense mechanism within the epidermis[62]. Interestingly, the recently identified SPRR4 exhibits the highest antioxidant potential and is induced upon UV-radiation, suggesting a specialized role of SPRR4 during the adaptive UV-response in epidermal tissue[59,62]. Furthermore, compared to the other human SPPR proteins, SPRR4 has the highest DNA binding capacity, which is directly controlled by the concentration of reactive oxygen species but unfortunately exact target sites as well as the functional significance of this DNA binding *in vivo* remain elusive[63].

Aberrant expression of SPRR proteins is also described for various forms of cancer, including breast, colorectal, esophageal as well as several types of skin cancer. Depending on the analyzed cancer type, the corresponding expression levels were found to be either up- or downregulated, leading to an inconclusive picture of their role during cancer development and progression. Additionally, several skin diseases are marked by altered expression of SPRR proteins, highlighting their importance and particular role for skin development in general but surprisingly no defective epidermis development or aberrant epidermal differentiation program in SPRR knockout mice has been reported so far[57,64,65].

2.1.4 Defective epidermal homeostasis results in skin cancer progression

Given the complex composition and the highly elaborate regulatory networks involved in epidermal homeostasis, one can easily visualize that already incremental changes within this process lead to cancer development. Moreover, the three major forms of skin cancer occur in the epidermis and arise either from keratinocytes like the basal cell carcinoma (BCC) and the squamous cell carcinoma (SCC) or from pigment producing melanocytes in the case of melanoma[66,67]. Common risk factors for skin cancer include alcohol consumption, fair skin pigmentation, chronic cutaneous infection and immune suppression but by far the most prevalent factor is the continuous exposure to UV-light and accompanying reoccurring sun burns[66–68].

Although less common than the keratinocyte derived BCC and SCC, melanoma accounts for a higher number of deaths and in general correlates with a poor prognosis. However, if recognized and treated prior to metastasis, melanoma is almost always curable by surgical

excision, which is the reason for focusing clinical efforts mainly on early melanoma detection and removal[66,69,70].

Basal cell carcinoma on the other hand, is the most common type of skin cancer and is frequently found in sun-exposed areas of skin, especially heaped in areas on and around the nose[67]. As the name already indicates, BCC arises from nonkeratinizing progenitor keratinocytes located in the basal layer of the epidermis and is characterized by slow growing and rarely metastasizing cells. Nevertheless, BCC patients frequently suffer from major tissue damage resulting in serious mutilating deformations and loss of vital structures[66,67,71]. From a mechanistic point of view, BCCs often include a characteristic mutation in the patched 1 tumor suppressor gene (PTCH1) but also other activating mutations in the sonic hedgehog signaling pathway have been reported, which might offer a promising therapeutic opportunity for BCC besides surgical excision[66,72]. Interestingly, signaling via the sonic hedgehog pathway in BCC patients also activates the Ras/Raf signaling pathway by increasing the expression of the platelet-derived growth factor receptor α (PDGFRα)[73].

Activation the Ras/Raf pathway is also commonly detected in cutaneous squamous cell carcinoma (SCC), which is only the second most common form of skin cancer in the USA but at the same time the one with the highest metastatic potential and the highest mortality after formation of aggressive metastasis[66,74,75]. Apart from greatly increased incidences of SCC in patients receiving immunosuppressive therapy, which is presumably caused by concomitant human papilloma virus infection and a number of heritable conditions favoring SCC development, UV-radiation-mediated mutagenesis is thought to be the major environmental factor promoting this type of cancer[66]. Thus, it is not surprising that UV-induced p53 loss of function mutations are the key event in cutaneous SCC carcinogenesis[75]. Moreover, mutations in the CDKN2A locus are frequently observed in SCC samples and result in an unleashed cell cycle control which is at least partly caused by loss of p16 mediated inhibition of cyclin-dependent kinase 4 (Cdk4) activity. Additionally, increased expression of Cdk4 is often associated with SCC development, which is especially interesting since Cdk4 overexpression is sufficient to induce SCC neoplasms in human epidermis when it is co-expressed with oncogenic forms of Ras proteins[66,76]. Consistent with this, release of the cell cycle control by a blockade in the NF-κB signaling pathway and simultaneous expression of oncogenic Ras can also transform human epidermis into a highly aggressive invasive neoplasia which presents itself indistinguishable from SCC[77,78]. In light of this central role of Ras proteins in SCC samples one can easily envision that expression of constitutively active Ras variants might be sufficient to cause SCC

development. Surprisingly however, expression of oncogenic Ras proteins alone drives keratinocytes into senescence and only the combination of oncogenic Ras and an unleashed cell cycle, for example by blocking the NF-κB signaling pathway or overexpressing Cdk4 (see above), result in SCC generation[77,79,80]. Despite this fundamental involvement of constitutively active Ras proteins in SCC generation, it is astonishing that amplification of Ras genes or activating mutations have only been reported for the minority of SCC samples, hence the observed increased levels of active GTP-bound Ras in the majority of human SCC tissues are presumably caused by the faulty activation of several upstream factors like receptor tyrosine kinases for example[66,75,81]. Histologically and biochemically, the keratinocyte derived SCC cells can be distinguished from healthy keratinocytes by their altered proliferative capacity as well as their reduced expression of keratinocyte differentiation proteins. In fact, suppression of gene expression for several key differentiation genes is a hallmark of SCC development and was frequently reported over the past years[82–84].

Apart from the above described proteins, several miRNAs have been reported as important stage specific biomarkers in the diagnosis of SCC and others were found to be gradually increased or decreased, implicating a crucial gene regulatory function for miRNAs during SCC progression[85,86]. Ultimately, even members of the novel class of long non-coding RNAs, which are essential regulators of fundamental cellular processes and tissue development have been implicated in SCC carcinogenesis (for details see 2.2.4)[87–91].

2.2 Long non-coding RNAs

2.2.1 Identification and classification of long non-coding RNAs

- DNA is transcribed into RNA, which is eventually translated into a functional protein. - For years, this sentence has been the irrevocable central dogma of molecular biology. With fundamental cellular functions ranging from energy metabolism to structural components, over signal transduction to being key regulators of gene expression, proteins were attributed great scientific attention while the RNA was contemplated as the inevitable intermediary required for protein production. However, this picture changed dramatically as high-throughput sequencing data revealed that more than two-thirds of the human genome are actively transcribed into RNA but only <2% actually encodes for proteins[92,93]. Although a role of ncRNAs for several basic cellular functions like translation (transfer RNAs,

ribosomal RNAs), RNA editing (small nucleolar RNAs) or splicing (small nuclear RNAs) has been known for quite a long time and the more recently identified shorter ncRNAs (20-30 nt in length) including microRNAs, endogenous short-interfering RNAs or piwi-associated RNAs as crucial regulators of gene expression were also identified, another tremendous piece of the ncRNA puzzle seemed to be missing[94–97]. Ultimately, this gap was closed with the identification of the novel class of long non-coding RNAs (lncRNAs), which has gained increasing scientific interest over the past few years and we are only beginning to appreciate their exigency and significant role for cellular processes as well as tissue homeostasis[98–100].

Per definition lncRNAs are over 200 nucleotides in length, can be spliced, capped and/or polyadenylated and are localized either in the nucleus or the cytoplasm of the cell[87,92,101]. Furthermore, lncRNAs in general lack a protein-coding potential, although recent reports about lncRNAs associating with ribosomes and translation of small functional peptides from lncRNA templates challenge this rather strict classification[102–104]. One of the first reports about those bifunctional RNAs was SRA1 and its protein counterpart steroid receptor RNA activator protein (SRAP), which are both co-regulators of steroid receptors as well as MyoD, a transcription factor important for skeletal myogenesis[105–107]. More recently, the peptide DWORF was identified in the lncRNA LOC100507537 and many additional putative peptides are predicted to arise from lncRNAs. In contrast to this, other global studies indicate that the majority of lncRNAs is not translated and thus indeed non-coding[102,103,108–110].
Scattered all over the genome, lncRNAs can either emerge as divergent transcripts from a neighboring gene promotor or from completely autonomous genomic loci (intergenic). Additionally, they can represent the overlapping antisense transcript of a protein-coding gene or originate from an intronic region. Depending on the overlap and the orientation to the host gene, this rather general classification can be further refined into several more defined subclasses[111,112]. Apart from being included into other transcriptional units, lncRNAs themselves can host protein-coding genes or other ncRNAs like the recently identified circular RNAs (circRNAs), tRNAs, miRNAs and especially snoRNAs[98,113–117]. This high degree of complexity in lncRNA loci is even further increased by the presence of an average of 2.3 to 3.9 different isoforms per gene locus, resulting from a combination of alternative splicing, alternative polyadenylation and employing alternative transcription start sites, which highlights the requirement for a careful lncRNA locus characterization[98,100,118,119].

Interestingly, many lncRNAs feature a strict cell-type and also differentiation state specific expression pattern which already implies a central function during different cellular processes. And indeed, lncRNAs have been proven essential regulators for fundamental processes such as cellular differentiation, epigenetic imprinting, cell cycle control, apoptosis, X-chromosome inactivation, promoter-specific gene regulation and nuclear import[88,120–127]. Furthermore, aberrant lncRNA signatures are a hallmark of several severe diseases including numerous types of cancer, which impressively highlights their tremendous significance on almost every aspect of the life cycle[87,128].

2.2.2 Molecular mechanisms of lncRNA function

Although we are aware of several thousand different lncRNAs in humans today and new lncRNAs are identified on a regular basis, only a small number of lncRNAs has been mechanistically characterized in detail[129]. What we know today is that in general lncRNAs can either act as guides or decoys for other RNAs and proteins, exert a signaling function, or act as a scaffold during the assembly of larger ribonucleoprotein complexes[130]. By employing solely one of the above described archetypes or a combination of those, lncRNAs are able to fulfill their multifaceted functions which are mostly reasoned by a regulatory function on gene expression at both, the transcriptional and the posttranscriptional level. Furthermore, depending on the localization of the lncRNA gene in respect to the position of transcriptional control, *cis*- (the lncRNA regulates gene expression in the vicinity of its own locus) or *trans-acting* (spatial separation of lncRNA expression and transcriptional control) lncRNAs can be distinguished[130,131].

One of the best studied examples for transcriptional gene regulation in *cis* is the lncRNA X-inactive-specific transcript (XIST), which is crucial for mammalian dosage compensation. To this end, XIST is exclusively expressed from one of the two X-chromosomes in females, subsequently coats this chromosome and triggers a series of events that lead to conformational reorganization, chromatin modifications and ultimately transcriptional silencing of this X-chromosome[132,133]. In contrast to this, dosage compensation in flies is achieved by transcriptional activation of the single X-chromosome in males. During this process, the functionally redundant lncRNAs roX1 and roX2 target the male-specific lethal complex to the X-chromosome, where hypertranscription is achieved by increasing H4K16 acetylation[134,135]. Another well-characterized lncRNA is the 2.2 kb HOX transcript antisense RNA (HOTAIR), which was the first identified lncRNA featuring a *trans*-acting mechanism. HOTAIR is expressed from the HOXC locus but represses transcription of the distant HOXD

locus by scaffolding and guiding the polycomb repressive complex 2 (PRC2) as well as the LSD1/CoREST/REST complex to the HOXD cluster, resulting in H3K27 trimethylation and H3K4 demethylation[120,122]. Besides the above-mentioned histone and chromatin modifications, lncRNAs were also found to directly control Pol II activity, associate with transcription factors and exert coregulatory functions during transcriptional regulation of gene expression[136].

Moreover, lncRNAs are also able to control gene expression on a post-transcriptional level by regulating RNA splicing, RNA editing, mRNA stability, mRNA translation efficiency as well as miRNA-mediated mRNA destabilization[136,137]. The antisense lncRNA BACE1-AS for instance, is upregulated in Alzheimer's disease and increases the stability of its antisense transcript BACE mRNA, by masking its binding sites for miR-485-5p[138]. Apart from competing for miRNA binding sites in other transcripts, lncRNAs - in this context also known as miRNA sponges or competing endogenous RNAs (ceRNAs) - can efficiently sequester miRNAs themselves like Linc-MD1 for example, which activates muscle-specific gene expression in human and murine myoblasts via sponging miR-133 and miR-135[139].

2.2.3 LncRNAs control tissue homeostasis and organ development

Given the largely cell-type and also differentiation state specific expression pattern of lncRNAs, it is not surprising that lncRNAs are crucial regulators of organ and tissue development and homeostasis[87,140]. Besides the already mentioned functions of the bifunctional SRA1 and linc-MD1 during muscle development, skeletal muscle differentiation and regeneration is also promoted by the lncRNAs H19, SENCR, MUNC or lnc-mg. Furthermore, the small peptide DWORF, which is encoded on the lncRNA LOC100507537, enhances muscle contractility by stimulating SERCA mediated Ca^{2+}-uptake into the sarcoplasmic reticulum[110,140]. Along this line, even the homeostasis of the most important muscle – the heart – is orchestrated by several lncRNAs. One of the first discovered examples for this was Braveheart, which activates a core cardiovascular gene network and possibly mediates epigenetic regulation of cardiac commitment by interacting with components of the PRC2 complex[141]. Additionally, proper murine heart development is ensured by the recruitment of the histone modifying complexes PRC2 and TrxG/MLL through the lncRNAs Fendrr and Carmen[142,143].

Moreover, several lncRNAs are exclusively expressed in the brain and the nervous system and ensure proper brain development, synapse formation and function, stress responses and age-associated changes. Evf2 for example recruits transcription factors in *cis* and *trans*

during murine ventral forebrain development, while Pinky regulates neurogenesis in the embryonic and postnatal brain in combination with the splicing factor PTBP1[124,144]. Interestingly, a recent single-cell analysis of the lncRNA signature from human neocortex cells at different developmental stages revealed numerous cell type-specific lncRNAs, suggesting a highly complex regulatory lncRNA network that is essential for brain development[145].

Apart from the indispensable roles during muscle, heart and brain development, lncRNAs are also well-established regulators of literally every organ, including bones, lung, liver, fat as well as intestinal tissue and even the homeostasis of the largest human organ – the skin – is tightly controlled by several lncRNAs[140].

2.2.4 Roles of long non-coding RNAs in epidermal tissue homeostasis and skin diseases

The 3.7 kb terminal differentiation-induced ncRNA (TINCR) was one of the first lncRNAs featuring a crucial function for maintenance of epidermal homeostasis. During this highly complex process, TINCR regulates differentiation gene expression via a post-transcriptional mechanism in which it forms short double stranded RNA duplexes with several differentiation protein mRNAs. Specific interactions in this instance are mediated by short 25 nt TINCR-box motifs, which become subsequently bound by the protein Staufen1, ultimately resulting in mRNA stabilization and therefore sustained expression of differentiation genes[88]. Whereas TINCR is essential for keratinocyte differentiation, the anti-differentiation ncRNA (ANCR) has an opposing effect on keratinocyte differentiation (Figure 3). ANCR is widely expressed in progenitor keratinocytes and was found to be significantly downregulated upon induction of the terminal differentiation program. Furthermore, loss of ANCR in undifferentiated keratinocytes induced ectopic expression of differentiation markers, ultimately proving that ANCR is essential for maintaining the undifferentiated cell state within the epidermis[146]. Interestingly, a recent study was able to show that TINCR and ANCR both control the transcription factors MAF and MAFB and thus exert their regulatory function in epidermal homeostasis at least partially via the same pathway. While TINCR potentiates the differentiation state through enhancing MAF/MAFB mRNA stability, ANCR suppresses MAF/MAFB expression in association with the PRC2 component EZH2[36].

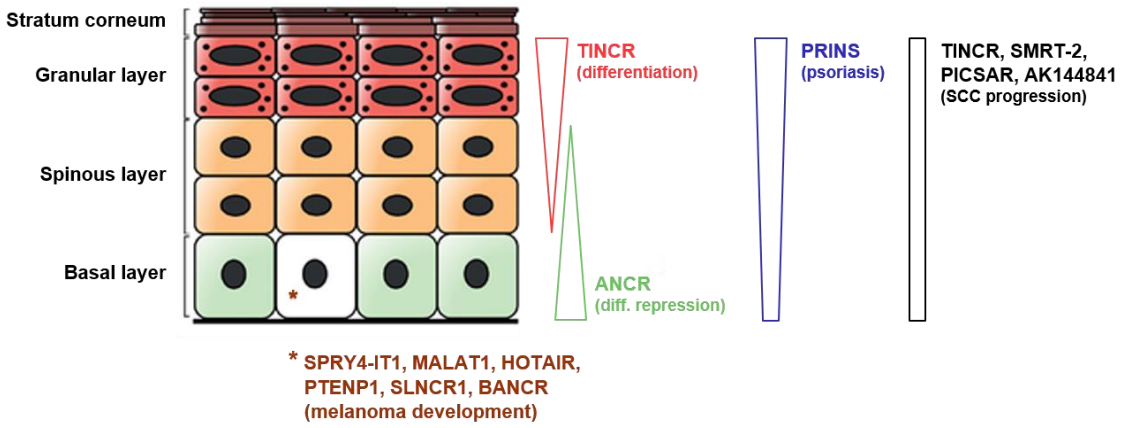


Figure 3: Expression pattern of several mammalian lncRNAs in the skin
Schematic overview of the epidermis, depicting keratinocytes in varying differentiation states and their occurrence in the corresponding four epidermal strata (left). Furthermore, the expression pattern of the two antagonizing lncRNAs TINCR (red) and ANCR (green), controlling keratinocyte differentiation are shown next to the psoriasis affecting lncRNA PRINS (blue) as well as lncRNAs featuring altered expression in SCC (black) or melanoma (brown) (modified after[147]). diff. = differentiation

However, not only proper keratinocyte differentiation is maintained by lncRNAs but also generation and progression of several skin diseases and cancer types are accompanied by aberrant lncRNA expression (Figure 3). TINCR as well as the recently identified lncRNA SMRT-2 (SCC misregulated transcript 2) for example are not only crucial factors for normal keratinocyte differentiation but both exhibit reduced expression in cutaneous squamous cell carcinoma (SCC) samples[88,89]. Given the fact that keratinocyte derived SCC represents the second most common cancer in the United States and has the highest metastatic potential amongst all forms of skin cancer, it is surprising that besides the above-mentioned examples and the recently identified PICSAR as well as the murine AK144841, only a handful of lncRNAs have been functionally characterized in SCC[89,90]. Also in melanoma, the most dangerous form of skin cancer originating from melanocytes, several lncRNAs like SPRY4-IT1, MALAT1, HOTAIR, PTENP1 or SLNCR1 feature an aberrant expression pattern, indicating a lncRNA involvement in melanoma progression[69,147,148]. Another well-known example for a melanoma lncRNA is the 693 bp BRAF-regulated lncRNA (BANCR), which controls the migratory capacity of melanoma cells and hence directly correlates with their metastatic potential[149]. Further studies addressing the mechanistic details of BANCR-mediated melanoma progression were able to unravel that BANCR controls the MAPK pathway activity and leads to upregulation of Notch2 via sponging of miR-204[150,151]. Moreover, diverging lncRNA signatures have been found in several other skin diseases. The primate-specific lncRNA PRINS (psoriasis susceptibility-related RNA gene induced by

stress) for example affects psoriasis susceptibility as well as the innate immune response of the skin and generally functions in the keratinocyte stress response[152,153].

In summary, lncRNAs have been proven to be crucial regulators of epidermal homeostasis and their aberrant expression is tightly linked to several skin diseases, including multiple forms of skin cancer. Thus, lncRNAs are not only valuable markers for skin diseases or keratinocyte differentiation but expanding our knowledge about the underlying modes of action might also offer novel angles for the development of therapeutic approaches for skin diseases.

2.3 Detection of SPRR5_326 and preliminary results

Initially identified from full transcriptome sequencing of differentiated versus undifferentiated keratinocytes, SPRR5_326 was discovered as a novel long non-coding RNA which is encoded within the epidermal differentiation complex, drastically induced during keratinocyte differentiation and suppressed in SCC samples as compared to site-matched tissue control samples from the same donors[88,89,146]. Since no annotation for this lncRNA was present at the time of identification, RACE analyses were conducted and overlaid with the results from the performed RNA-Seq experiments, leading to the prediction of a 326 nt long transcript lacking an open reading frame, which was hence termed SPRR5_326. Furthermore, given the facts that lack of differentiation is a hallmark of SCC cells, SPRR5_326 is induced during keratinocyte differentiation, repressed in SCC samples and is encoded within the human epidermal differentiation complex where many essential differentiation genes are clustered, SPRR5_326 seemed to control both epidermal homeostasis as well as SCC progression.

3 Objective

Given the proposed dual function of SPRR5_326 during regular keratinocyte differentiation as well as its dysregulation in human squamous cell carcinoma, the aim of this PhD project was to shed more light on the exact role of SPRR5_326 throughout those processes. To this end, first the impact of sustained SPRR5_326 expression on cancer progression should be assessed. Furthermore, this novel and previously uncharacterized lncRNA gene locus will be characterized in more detail with respect to its upstream regulating transcription factors, all of its expressed isoforms as well as their expression pattern during keratinocyte differentiation. Surprisingly, results from these studies indicated that the previously anticipated SPRR5_326 transcript is not the dominant transcript but instead a longer isoform spanning 762 nt, from now on termed SPRR5, is massively induced in differentiating keratinocytes and also comprises a predicted open reading frame. Due to these findings, the actual presence of the predicted SPRR5 protein had to be tested and its functional relevance was probed in rescue experiments in order to decipher whether SPRR5 functions as a lncRNA, a protein or a combination of both.

Finally, the impact of SPRR5 deficiency on epidermal homeostasis should be evaluated and insights from these studies should be used to postulate and test a hypothesized molecular mechanism of SPRR5 during the course of keratinocyte differentiation.

4 Results

4.1 Linking SPRR5 and neoplasia

The fact that SPRR5_326 is downregulated in human squamous cell carcinoma samples compared to their site matched healthy tissue controls already implies an important function for SPRR5_326 in either cancer onset or development and lead to the hypothesis that sustained SPRR5_326 expression might prevent neoplasia or tumor progression. In order to test this hypothesis, two experiments were performed in collaboration with the Stanford University School of Medicine in the laboratory of Prof. Dr. Paul Khavari. First, the impact of sustained SPRR5_326 expression was studied in an invasive three-dimensional organotypic neoplasia tissue model, which recapitulates natural features of tumor progression and second the influence of SPRR5_326 on tumor development was assessed by an *in vivo* tumor formation assay[154].

Since for both experiments adequate and long-lasting SPRR5_326 overexpression was essential, this was ensured in a preliminary experiment during which the observed overexpression lasted over the entire time period of more than two weeks and SPRR5_326 expression was increased by 27- to 77-fold as compared to the lacZ control (Supplementary Figure 1). Next, the effect of this overexpression was analyzed in an invasive organotypic neoplasia tissue model. To this end, primary keratinocytes were transformed into "tumorigenic keratinocytes" by forced overexpression of oncogenic human HRasG12V and simultaneous overexpression of Cdk4 in order to omit the reported onset of cellular senescence caused by the single overexpression of oncogenic Ras variants[76,79,80]. Next, these tumorigenic keratinocytes, with either lacZ or SPRR5_326 overexpression, were seeded onto a dermal matrix with embedded fibroblasts and the resulting tissue was harvested six or eight days after seeding when invasion of keratinocytes into the dermal matrix was analyzed by immunofluorescence analyses[76,154]. Figure 4A clearly shows that the transformation into tumorigenic keratinocytes was successful as lacZ and SPRR5_326 overexpressing cells exhibited identical and elevated expression of oncogenic Ras as well as Cdk4 as compared to wild type cells. Furthermore, these tumorigenic keratinocytes were also invasive, since they crossed the degraded basement membrane and could be detected within the dermal matrix which is a distinct feature of tumorigenic keratinocytes (an exemplary picture is given in Figure 4C). With roughly 35-fold higher levels as compared to the lacZ control, also the overexpression of SPRR5_326 was satisfactory (Figure 4B). Thus, the invasion depth of keratinocytes into the dermal matrix was analyzed as an indicator

for the neoplastic potential, since this process never occurs in healthy epidermal tissue but represents a hallmark feature of skin cancer progression[154–156].

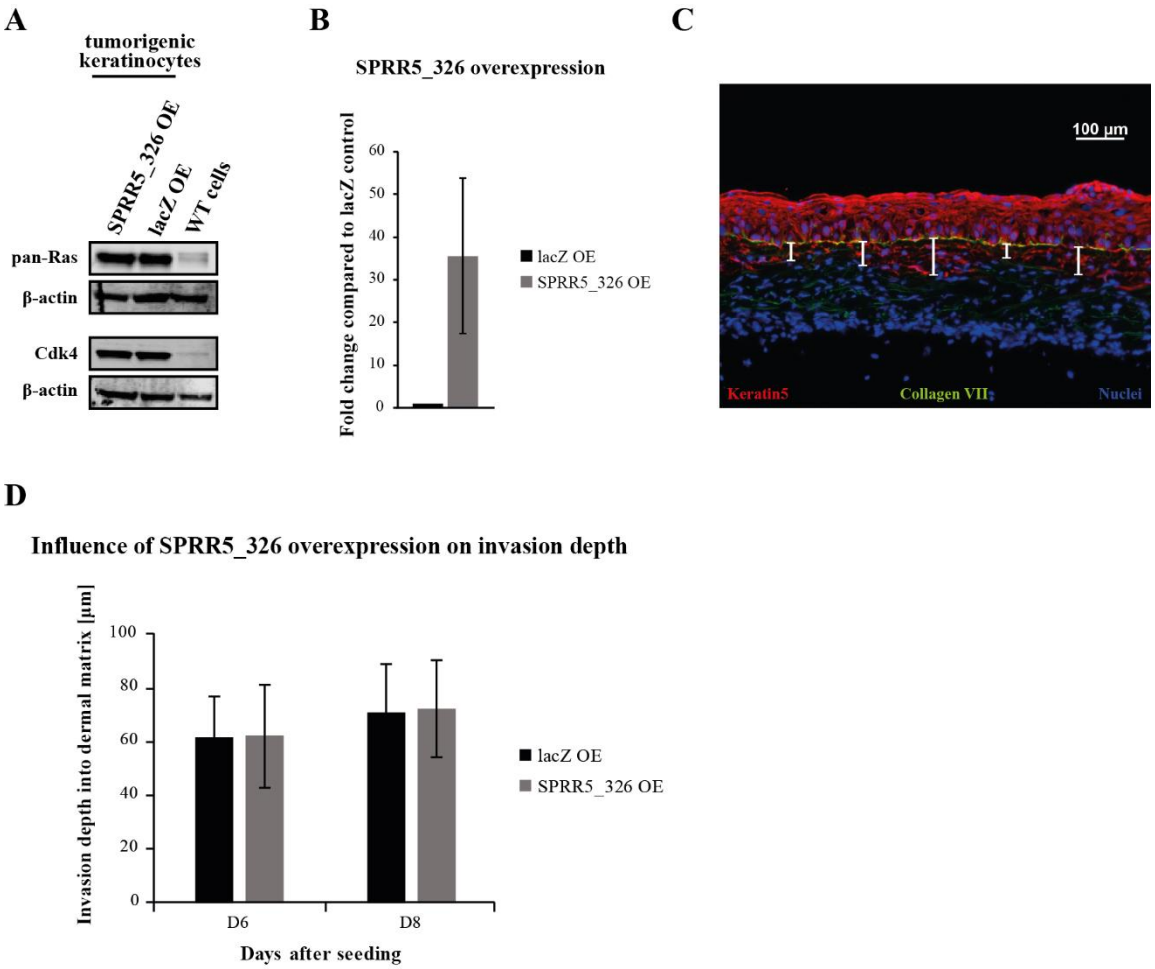


Figure 4: Generation and analysis of invasive neoplastic tissue
Western blot confirms that tumorigenic keratinocytes showed similar and adequate overexpression for HRasG12V and Cdk4 as compared to wild type (WT) cells (A). Furthermore, the obtained overexpression of SPRR5_326 in tumorigenic cells utilized for generation of invasive neoplastic tissue was analyzed by RT-qPCR (n=2), normalized to L32 and compared to expression in lacZ overexpressing cells (B). (C) shows an exemplary picture from the performed evaluation of measuring the invasion depth of tumorigenic keratinocytes into the dermis at five different sites per section (shown in white). Keratin5 (marker for keratinocytes) is shown in red, nuclei are shown in blue and collagen VII (marker for the basement membrane) is depicted in green. Mean and standard deviation of all measured invasion depths for all biological replicates per group (n=3) and timepoint were calculated and plotted for each timepoint. OE = overexpression

Regarding the results from this experiment (Figure 4D), it is striking that no tumor suppressive effect of SPRR5_326 overexpression could be detected, since the invasion depth for SPRR5_326 overexpressing keratinocytes was almost identical to the invasion depth in the lacZ control at both assessed timepoints. Taking the thickness of the overlying epidermis into account and analyzing the ratio of invasion depth to epidermis thickness did also not show a clear difference between lacZ and SPRR5_326 overexpression (data not shown).

In a second experiment, the influence of SPRR5_326 overexpression on tumor formation and growth should be addressed *in vivo*. Therefor the same tumorigenic keratinocytes as described above, with additional overexpression of a luciferase-YFP fusion protein, were injected into the flank of immunodeficient mice and tumor growth was monitored over time by *in vivo* luciferase measurements and volumetric tumor measurements.

Since all eight mice showed growing nodules after tumor cell injection, the tumor formation in general was successful, however at the end of the experiment only three tumors in the SPRR5_326 overexpression group were detectable (see Supplementary Figure 2), which is the reason why only the three biggest tumors from each group were used for data evaluation. Examination of the obtained tumor volumes as well as estimates on the tumor mass (Figure 5) showed the slight tendency that forced SPRR5_326 overexpression lead to slower tumor growth *in vivo* compared to the lacZ control. Nevertheless, it should be mentioned that the standard deviation of the mean tumor sizes was very high and that tumor sizes varied dramatically, which can also be seen by the size of the explanted tumors and the *in vivo* luciferase signal measurements (see Supplementary Figure 2A+C). Moreover, the evaluation of the mean luciferase signal for each group lead to no conclusive result, as at first SPRR5_326 overexpressing tumors showed slower tumor progression as compared to the lacZ control but on day 47 this trend was inverted, which is why luciferase measurements were only recorded until day 64 (see Supplementary Figure 2B).

In summary, no indication for a tumor suppressive effect of SPRR5_326 can be deduced from the performed experiments and subsequent studies with more replicates are required, which could not be done in Regensburg due to the missing infrastructure and permissions.

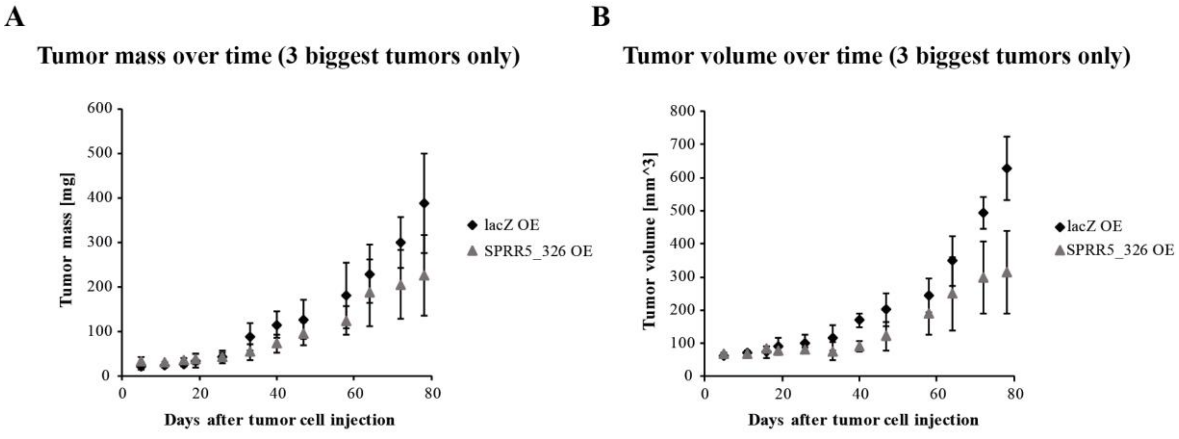


Figure 5: Tumor growth over time
Tumor size was monitored by caliper measurements and the corresponding tumor mass (A) as well as the tumor volume (B) were calculated for the three biggest tumors for each group (lacZ or SPRR5_326 OE) and mean and standard deviation were plotted against the time after injection. OE = overexpression

4.2 Transcript characterization for SPRR5

4.2.1 Subcellular localization of SPRR5

Unravelling the cellular localization of a given lncRNA might already give insights into its possible modes of action, which is why a subcellular fractionation approach followed by RT-qPCR analysis was chosen to reveal the localization of SPRR5 in differentiated keratinocytes (Figure 6).

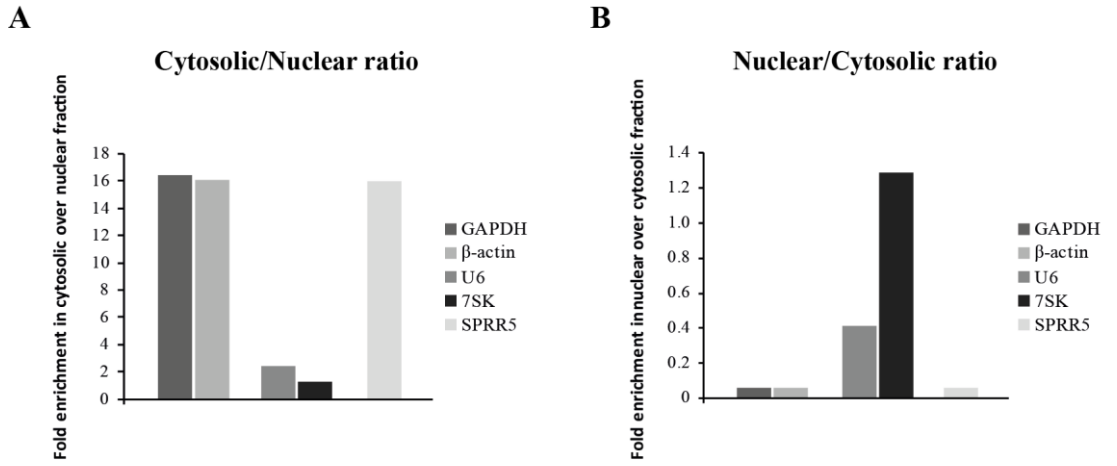


Figure 6: Subcellular localization of SPRR5
Analysis of nuclear and cytoplasmic RNA fractions of day 6 differentiated keratinocytes by RT-qPCR. The localization of cytoplasmic markers (GAPDH and β-actin), nuclear markers (U6 and 7SK) as well as SPRR5 was assessed and the enrichment in each fraction relative to the other fraction was calculated and plotted.

Regarding the enrichment of the assessed marker transcripts clearly indicates a decent cytoplasmic fraction since the cytosolic markers GAPDH as well as β-actin are roughly 16-fold enriched in the cytosolic over the nuclear fraction, whereas the nuclear markers U6 and 7SK are not enriched in the cytosol (see Figure 6A). The inverse analysis (enrichment in nuclear over cytoplasmic fraction) however showed only a slight enrichment of the nuclear marker 7SK in the nuclear fraction and no clear enrichment of U6 (Figure 6B). Given the facts, that almost all subcellular fractionation approaches are not able to completely separate the cytosolic and nuclear fraction and that the primary function of differentiated keratinocytes is the formation of a tightly sealed cornified envelope which impedes with their separation prior to the cellular fractionation, no perfect separation into a nuclear and cytosolic fraction was anticipated in this experiment and the depicted results are the best results after numerous protocol optimization iterations. Furthermore, the expected pattern for the cytosolic and nuclear marker transcripts could be observed in general and thus the localization of SPRR5 could be analyzed in this experiment.

Figure 6A shows that SPRR5 is enriched in the cytosolic fraction (when compared to the nuclear fraction) with similar enrichment values as the cytoplasmic markers GAPDH and β-actin. Hence, SPRR5 is presumably mainly localized in the cytoplasm of differentiated keratinocytes, which matches the results from previous fractionation experiments (data not shown).

Complementary to this subcellular fractionation approach, RNA fluorescence *in situ* hybridization (RNA-FISH) was performed by Bianca Förstl as a second method to study the localization of SPRR5. In principal, this method was applicable to differentiated keratinocytes since the utilized positive controls repeatedly exhibited the expected subcellular localization and the negative control was characterized by a complete lack of signal. Despite various attempts with different labelling reagents, amplification strategies and numerous variations of the applied protocol however, no conclusive results for SPRR5 could be obtained by RNA-FISH. Thus, it appears that RNA-FISH for this particular transcript is especially challenging and no final statement for the SPRR5 localization can be made since due to the technical difficulties of the subcellular fractionation approach, results from this analysis should be interpreted as a first indication rather than a final proof of the SPRR5 localization.

4.2.2 p63 controls SPRR5 expression

The initial transcriptomic screen of undifferentiated versus differentiated keratinocytes (2.3) indicated a strong induction of SPRR5 during terminal differentiation, raising the question about the accountable factor for SPRR5 induction. Thus, upstream transcription factors for SPRR5 were predicted with the ARCHS[4] website, resulting in p63 as the most likely regulator of SPRR5 expression in keratinocytes (Figure 7A)[157]. Furthermore, analysis of GEO deposited p63 ChIP-Seq peaks from Kouwenhoven et al. and Bao et al. indicated p63 binding sites in the vicinity of the SPRR5 transcript (Figure 7C)[44,158–160]. Eventually, p63 could be confirmed as a regulator of SPRR5 expression by pan-p63 knockdown in keratinocytes, resulting in decreased SPRR5 amounts on D3 of differentiation (Figure 7B).

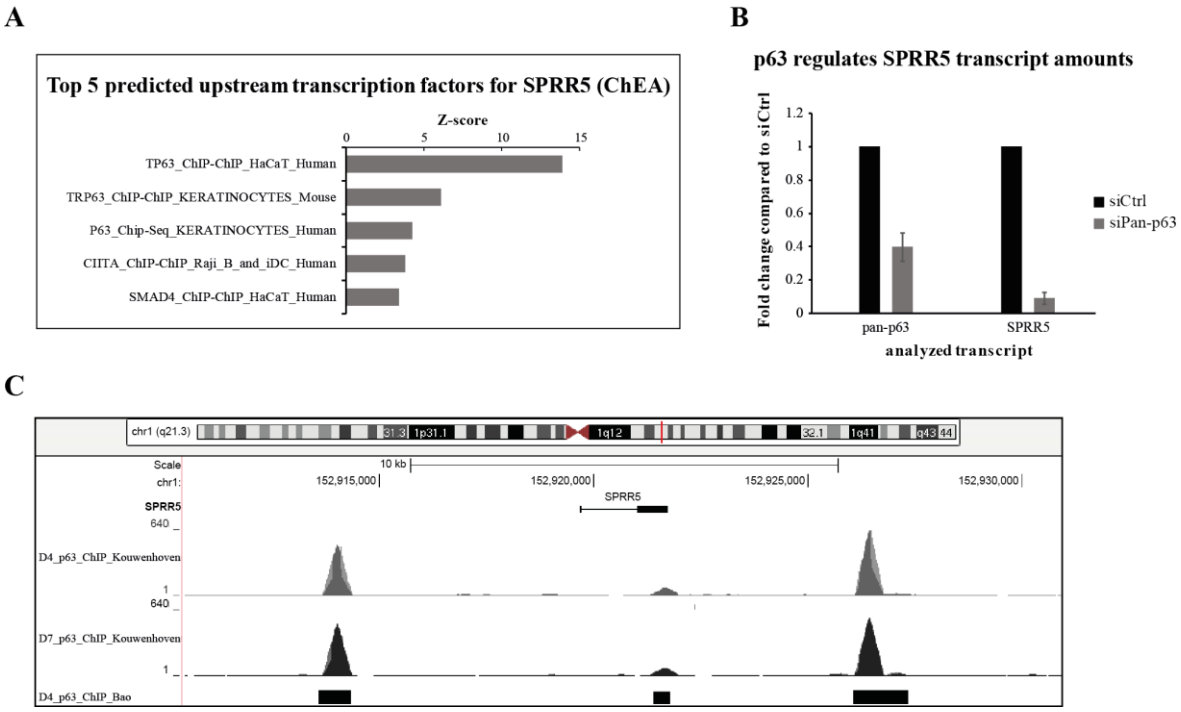


Figure 7: p63 controls SPRR5 expression
Based on the ChEA algorithm, ARCHS[4] predicted p63 as most likely upstream transcription factor for SPRR5 (A), which was further supported by the presence of p63 ChIP-Seq peaks in the vicinity of SPRR5 (C). Knockdown of pan-p63 in keratinocytes (n=4) and assessing SPRR5 levels on day 3 of differentiation eventually confirmed p63 as a regulator of SPRR5 expression.

4.2.3 Recent annotations are not valid for the SPRR5 locus

At the beginning of this project, no annotation was present for the SPRR5 gene locus and the assumed SPRR5_326 transcript was based on the distribution of RNA sequencing reads and confirmed by RACE analyses. However, during progression of the project first a second lncRNA, LINC01527 which is encoded from the opposite strand as SPRR5_326 was annotated for this gene locus. Later, the potential protein-coding SPRR5 transcript appeared in the Ensembl genome browser and in the recently published FANTOM5 dataset a similar but unspliced transcript with a different 5'end was annotated (ENCT00000012178)[161,162]. Furthermore, a third, longer and partially SPRR5 overlapping transcript (MICT00000021193) was included into the FANTOM5 dataset, necessitating an exact characterization of the transcripts arising from the genomic SPRR5 locus (an overview of all annotations is given in Figure 8).

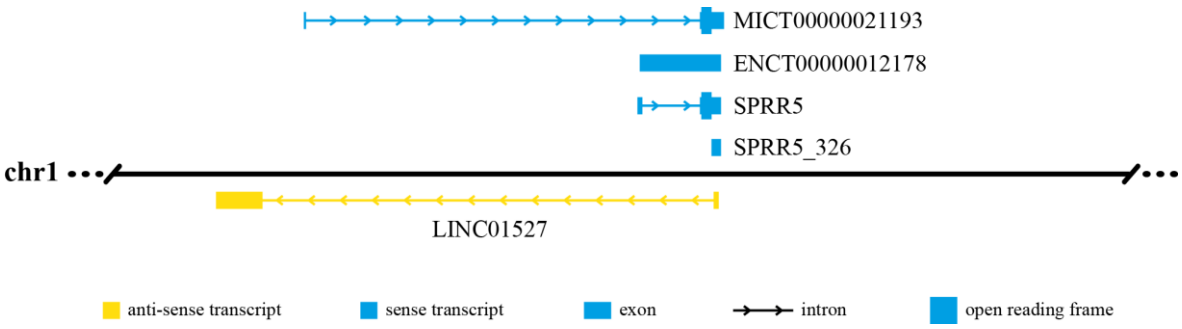


Figure 8: Overview of the recently annotated and self-identified transcripts from the SPRR5 locus
The assumed transcript SPRR5_326 as well as recent annotations from the Ensembl genome browser and the FANTOM5 dataset are included, highlighting the imperative of careful isoform characterization.

Regarding the question whether LINC01527 is expressed in keratinocytes, mapped sequencing reads from full transcriptome sequencing of organotypic epidermal tissue (4.4.3) were analyzed first, showing that they do not match the LINC01527 transcript (Figure 9 lower panel). To omit the possibility of missing indications of LINC01527 expression due to an unsuitable timepoint of differentiation or the assessment of solely one sample, the same analysis was repeated for an already published RNA-Seq dataset of undifferentiated, day 3 and day 6 differentiated keratinocytes from Kretz et al. leading to the same result (data not shown)[88]. Furthermore, LINC01527 could not be detected by northern blot analyses with different DNA oligo probes or by RT-qPCR measurements using three different primer sets, one of which was already published by Gao et al. for LINC01527 detection (data not shown)[163]. In summary, LINC01527 expression can be neglected for keratinocytes.

Deciphering which sense transcript arises from the SPRR5 locus commenced again with analysis of RNA-Seq read distribution from day 4 differentiated keratinocytes and the differentiation time course from Kretz et al. (Figure 9 lower panel, data from Kretz et al. not shown). This data shows that RNA-Seq reads cover solely the second exon of the annotated SPRR5 transcript, rendering ENCT00000012178 and MICT00000021193 very unlikely isoforms in keratinocytes. On the other hand, the SPRR5_326 transcript was supported by RNA-Seq reads but at least a second, longer isoform had to be present in order to explain the complete coverage of the second exon of the Ensembl annotated SPRR5. For the purpose of addressing this hypothesis, northern blot analyses with different probes were performed (Figure 9 top panel), which consistently showed one band at roughly 900 nt for all utilized probes, proving the presence of one dominant isoform in keratinocytes that is longer than the previously anticipated SPRR5_326 transcript. Furthermore, the presence of two exons as proposed by the Ensembl annotation of SPRR5 but with a shorter first exon was indicated

by northern blot analysis, as the probe Out_Exon1_SPRR5 did not show a signal at all, whereas the Exon1_SPRR5 probe was able to detect this first exon of SPRR5. For this reason, the actual 5' end of the verified SPRR5 transcript was inferred from the FANTOM5 annotated ENCT00000012178, for which a strong CAGE signal from keratinocytes was present and the predicted length of the first exon as well as the complete SPRR5 transcript, including its assumed exon-intron structure, could be verified by a novel RACE analyses with the Expand Long Template PCR system from Roche (work by Bianca Förstl).

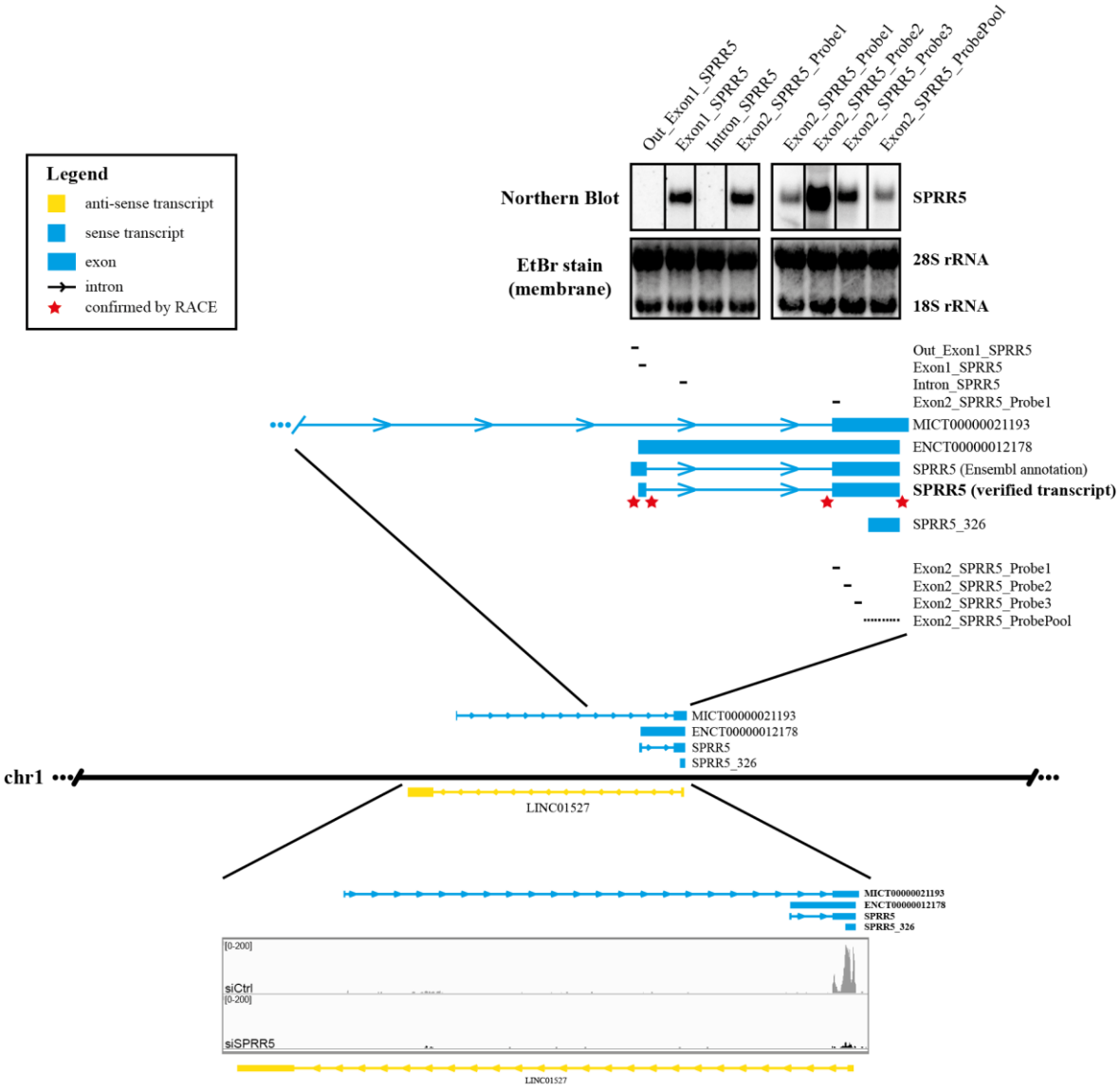


Figure 9: Isoform detection for the SPRR5 locus
Overview of the genomic localization of possible transcripts from the SPRR5 gene locus as well as the verified SPRR5 isoform. The lower panel shows aligned and scaled sequencing reads from full transcriptome sequencing of day 4 organotypic epidermal tissue with and without SPRR5 knockdown and the localization of possible transcripts. The top panel illustrates the location of utilized antisense DNA oligo probes with respect to the possible SPRR5 sense transcripts as well as the obtained northern blot signals for the depicted probes next to the ethidium bromide staining of the respective blots as a loading control. Summarizing the results, in contrast to the Ensembl annotation the verified SPRR5 transcript is shorter on the 5' end, which could also be confirmed by RACE analysis (red stars).

In summary, one dominant SPRR5 transcript could be identified in keratinocytes (the complete sequence and genomic coordinates are given in the appendix 9.2.2 and Table 37), which contains a predicted open reading frame and features some characteristics of the Ensembl SPRR5 and some of the FANTOM5 ENCT00000012178 annotation. With the verified SPRR5 transcript structure in hand, the challenge whether the included open reading frame gives rise to a functional protein or if SPRR5 acts as a lncRNA remained and will be addressed in the next chapter.

4.3 SPRR5 – protein or lncRNA?

4.3.1 SPRR5 is presumably protein-coding but evolved differently than other human SPRRs

Considering the observed discrepancies between the hitherto annotations for the SPRR5 gene locus and the eventually verified SPRR5 transcript (4.2.3), the possible translation from the predicted SPRR5 open reading frame was also examined in more detail in order to elucidate whether this hypothesized open reading frame is indeed translated or solely a misleading annotation. As a first step in this analysis, bioinformatic predictions about the coding potential of the SPRR5 transcript were performed. Both, iSeeRNA and the Coding Potential Calculator, predict that the SPRR5 transcript is indeed protein-coding (Figure 10A+B)[164,165]. Furthermore, PhyloCSF showed a positive score matching the predicted SPRR5 open reading frame (Figure 10C, PhyloCSF+2 track)[166]. Since a positive PhyloCSF score is an indicator for high frequencies of silent or synonymous codon substitutions and low frequencies of missense or nonsense substitutions, it helps to distinguish protein-coding from non-coding transcripts and in this case PhyloCSF predicts that SPRR5 represents a protein-coding transcript.

A

Transcript ID	Prediction	Noncoding Score
SPRR5	coding	0.0128

B

ID	C/NC	CODING POTENTIAL SCORE
SPRR5	coding (weak)	0.294612

C

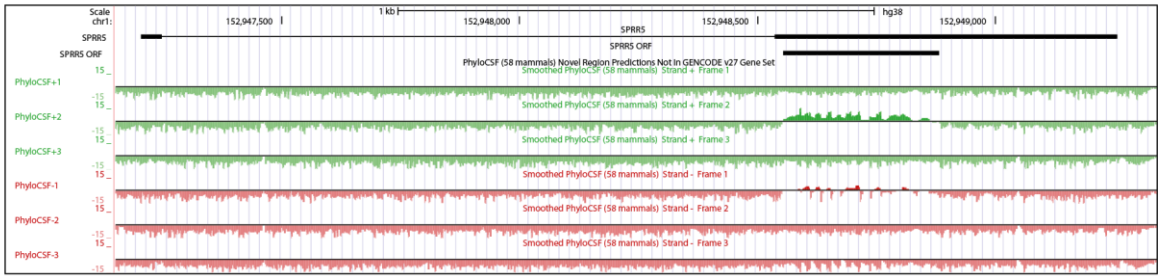


Figure 10: Bioinformatic approach to unravel the coding potential of SPRR5
The tools iSeeRNA (A) and Coding Potential Calculator (B) predict that SPRR5 is a protein-coding transcript, which is supported by inspection of PhyloCSF tracks in UCSC (C), where a positive score in the PhyloCSF+2 track matches the predicted SPRR5 open reading frame (ORF), indicating a strong conservation and thus coding potential for the putative SPRR5 ORF.

Due to the fact that the SPRR5 transcript was only recently annotated, very limited data about the transcript itself as well as the potential protein is available and therefore literature research for both molecules was not very productive in general. However, one study by Sendoel et al. addressing the translational landscape during tumor development included ribosome profiling data from *in vivo* mouse skin and since SPRR5 has a mouse homolog, this dataset could be utilized for addressing the question whether the SPRR5 transcript is indeed translated *in vivo* (Figure 11)[167].

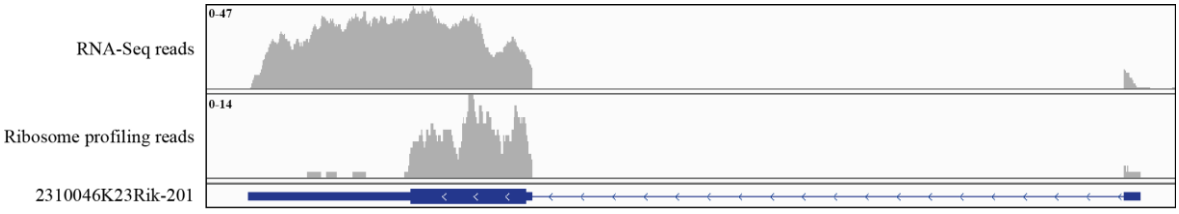


Figure 11: Mouse epidermal ribosome profiling data indicates translation of SPRR5
The mouse homolog of SPRR5 (2310046K23Rik-201) is shown in blue with its intron (thin line with arrows), both exons (blue bars) and the open reading frame (indicated by a thicker blue bar with white arrows inside). Additionally, the coverage of RNA-Seq reads (top) and ribosome profiling reads (middle) from mouse epidermis, extracted from a study by Sendoel et al. is depicted, indicating translation from the open reading frame of 2310046K23Rik-201.

While the whole 2310046K23Rik-201 transcript (mouse homolog of the human SPRR5 transcript) is covered by RNA-Seq reads (Figure 11 top), the ribosome profiling reads (Figure 11 middle) mainly cover the open reading frame, indicating translation of 2310046K23Rik-201 in mouse epidermis. In combination with the results from the bioinformatic coding potential predictions, these findings suggest the translation of a human SPRR5 protein in keratinocytes and raised the question about similarities between SPRR5 and other members of the adjacently encoded human SPRR protein family.

The human SPRR protein family counts 11 genes giving rise to 10 SPRR proteins (SPRR2C is a pseudogene) and all of these proteins share a very characteristic composition, featuring a glutamine- and lysine-rich N- and C-terminus which is used for protein-protein crosslinking, and a central domain consisting of varying proline-rich repeats[57]. Since the predicted SPRR5 protein would have a similar composition like the other SPRRs and its gene locus is directly adjacent to the human SPRR gene cluster, the annotation as a novel small proline-rich protein seems reasonable, although the existence of the SPRR5 protein was never shown. Apart from those similarities, a striking functional difference for SPRR5 and the other SPRRs exists, as the best described function of epidermal SPRR proteins is

contributing to the generation of a cornified envelope, by being crosslinked to other differentiation proteins like loricrin, involucrin or keratins for example[168]. Although some recent publications claim a protective effect of SPRR proteins against reactive oxygen species and an involvement in wound healing and cell migration, no effect on keratinocyte differentiation or an abnormal epidermis development in SPRR knockout mice has been reported so far[60,64,65]. Given the induced expression of the SPRR5 transcript during keratinocyte differentiation and the observed keratinocyte differentiation deficiency upon SPRR5 depletion (4.4), SPRR5 seems to be strikingly different from other human SPRR proteins.

In order to address this assumption, a phylogenetic analysis of the coding sequences from all human SPRR proteins, including the newly identified SPRR5 was conducted, unravelling that SPRR5 is drastically different from the other members of the human SPRR family, as SPRR5 clusters separately (Figure 12, bioinformatic analysis by Prof. Dr. Rainer Merkl). Thus, SPRR5 presumably originated from gene duplication of an SPRR ancestor but evolved in a different direction.

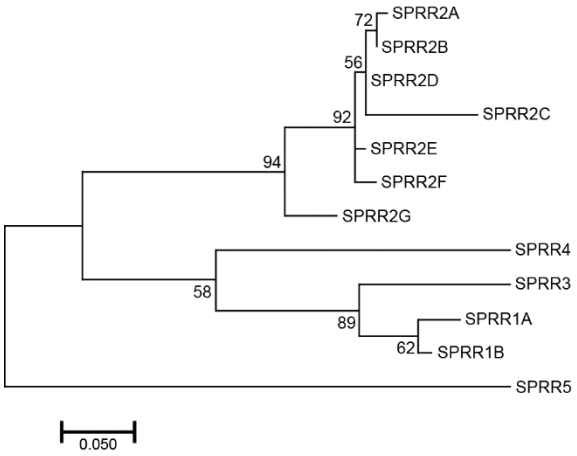


Figure 12: Phylogenetic tree of human SPRR-coding sequences
Molecular phylogenetic analysis by maximum likelihood method of coding sequences for all known human SPRRs indicating a common ancestor of SPRR5 and the other SPRRs but a separate evolution of SPRR5.

Taken together, bioinformatic analysis indicate that SPRR5 might be a protein-coding transcript, which evolved differently from other human SPRRs and additionally its murine homolog (2310046K23Rik-201) seems to be translated in mouse epidermis which might be an indicator that the predicted SPRR5 protein is indeed expressed in the human skin. Nevertheless, recent publications revealed striking differences between gene expression in murine and human[169]. One example for this is the transcription factor forkhead box C1 (FOXC1) which is expressed in the murine hair follicle but absent from the surrounding

epidermal tissue[170]. In contrast to this, FOXC1 is not only expressed in the human epidermis but exerts an important function as a vital regulator of the keratinocyte differentiation program[171]. Thus, the existence of the predicted SPRR5 protein in human keratinocytes as well as its potential functional relevance on epidermal homeostasis had to be investigated, since the SPRR5 transcript itself could still function as a lncRNA rather than serving as a blueprint for the predicted SPRR5 protein.

4.3.2 Small amounts of SPRR5 protein are detectable in keratinocytes

Predictions, annotations and phylogenetic analysis are a first indication but no proof for the actual presence of a SPRR5 protein in keratinocytes. So, in order to address this topic, two polyclonal antibodies against two peptides from the predicted SPRR5 protein were generated in cooperation with Eurogentec. Unfortunately, neither of them was able to detect the endogenous SPRR5 protein and both antibodies showed only weak binding to overexpressed and tagged SPRR5 protein variants (data from Katrin Hartinger; data not shown).

Since the generated antibodies were not suitable to detect endogenous SPRR5 protein, mass spectrometry was chosen as an alternative approach to identify the predicted protein within differentiated keratinocytes. Therefor protein lysates from keratinocytes on day 5 of differentiation were separated on an SDS-PAGE because a high expression of the SPRR5 transcript at this timepoint (Figure 19) should maximize the chances for the detection of the hypothesized SPRR5 protein. Following the protein separation step, samples from sections 1-14 (corresponding to a size range of roughly 5 to 20 kDa) were analyzed by mass spectrometry, as the predicted SPRR5 protein with its molecular weight of 11.9 kDa should be included in this part of the gel (Figure 13). However, once again no evidence for the SPRR5 protein in keratinocytes was found as not a single peptide from the predicted SPRR5 protein could be found within the analyzed sections. To exclude the possibility that the employed mass spectrometry approach was unsuitable for the detection of SPRR proteins in general, the obtained dataset was revaluated and several SPRR proteins like SPRR1A (9.9 kDa), SPRR2G (8.2 kDa) or SPRR3 (18.2 kDa) for example were amongst the measured proteins, proving that SPRR proteins with similar sizes as SPRR5 were detectable in this experiment.

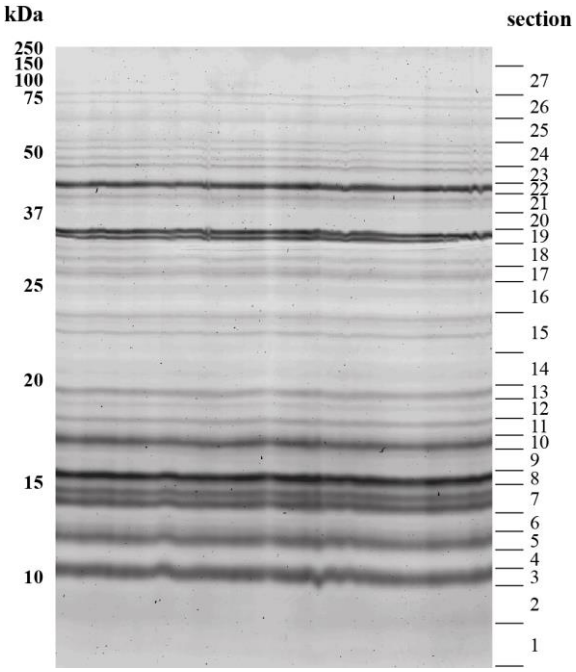


Figure 13: Schematic overview of the obtained mass spectrometry gel
Picture of the coomassie stained SDS-PAGE after separation of protein lysate from day 5 differentiated keratinocytes. A rough size estimation can be obtained in comparison to the utilized protein standard (indicated on the left) and the cutting sites as well section identifiers for each gel section are given on the right. kDa = kilo Dalton

As a final experiment, single reaction monitoring (SRM) measurements were performed because they feature an increased sensitivity during the analysis due to the prior calibration of the mass spectrometer with corresponding heavy labelled peptides. For this purpose, four labelled SPRR5 peptides were ordered and tested (sequence and location are given in 9.2.3) but only the N-terminal peptide was suitable for quantification as the other peptides were not detectable during preliminary mass spectrometry measurements. In order to omit the possibility of missing any SPRR5 derived peptide due to the limitation to sections 1-14, this time samples from the whole gel (molecular weight between 5 and 250 kDa, sections 1-27) were analyzed and now the SPRR5 protein could be detected in sections 13 and 14. Surprisingly, the protein standard indicates a molecular weight of roughly 20 kDa for these sections, which is considerably higher than the predicted 11.9 kDa for the SPRR5 protein. However, this size discrepancy was also observed for tagged SPRR5-variants during antibody testing (data not shown) and thus presumably originates from an aberrant protein migration behaviour.

Despite an increased sensitivity, single reaction monitoring also offers the possibility of protein quantification since the number of spiked-in peptides is known. Thus, subsequent quantification with the utilized spiked-in heavy peptide was performed and resulted in roughly 3,000 SPRR5 protein molecules per cell, which represents a rather low expression level for a human protein[172]. In summary, only small amounts of SPRR5 protein could be detected in human keratinocytes and only with a highly specific and targeted mass

spectrometry approach, which leaves the question whether those small amounts of protein exert the molecular function of SPRR5 during the course of keratinocyte differentiation or are solely the product of pervasive translation and the actual functionality arises from the SPRR5 transcript itself.

4.3.3 Rescue experiments for SPRR5

In order to determine whether SPRR5 exerts its cellular function as a protein or a lncRNA, rescue experiments in SPRR5 depleted keratinocytes were performed because SPRR5 deficiency results in differentiation defects (see 4.4) and thus introducing the functional SPRR5 molecule should rescue this phenotype. To this end, overexpression of either the wild type SPRR5 transcript (SPRR5_WT), which could function as a lncRNA and could be translated into a protein, or the same RNA with a destroyed start codon to obtain only the potential lncRNA function (SPRR5_Mut) were utilized. Furthermore, functional relevance of the SPRR5 protein was tested by introducing only the SPRR5-coding sequence (SPRR5_ORF) and the identical RNA sequence with a mutated start codon (SPRR5_MutORF) was used to exclude the possibility of a functional relevance of the RNA sequence alone (an overview of the overexpressed transcripts is given in Figure 14).

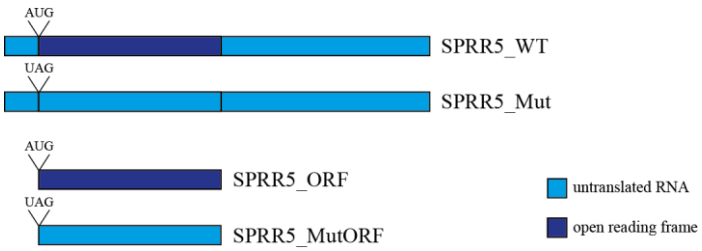


Figure 14: Overview of overexpressed SPRR5 transcripts
The SPRR5 wild type mRNA is depicted at the top, untranslated regions are shown in light blue, whereas the open reading frame for the SPRR5 protein, including its start codon (AUG), is indicated in dark blue. Overexpressed transcripts were either wild-type SPRR5 (top), wild-type RNA with a mutated start codon so that no functional protein could be made (SPPR5_Mut; 2nd from the top), the protein-coding open reading frame without any flanking RNA elements (SPRR5_ORF; 3rd from the top) or the open reading frame with no start codon (SPRR5_MutORF; bottom). WT = wild type, Mut = mutated, ORF = open reading frame

As a reference for the effect of SPRR5 depletion, lacZ (for the full-length transcripts) or a luciferase-YFP protein (for the short transcripts) were overexpressed in SPRR5 depleted keratinocytes, since these transcripts should not affect keratinocyte differentiation. Comparing the expression of several differentiation markers in keratinocyte differentiation cultures from SPRR5 depleted (siSPRR5) and control keratinocytes (siCtrl) however showed no clear differentiation defect upon SPRR5 depletion neither for the lacZ nor the luciferase-YFP control (Figure 15A-C). This rather unexpected result (one would expect a

lack of differentiation after SPRR5 depletion) was presumably due to the high cellular stress levels induced by the siRNA treatment in combination with the lentiviral transduction (see also 5.3.2) and made a conclusion about a potential rescue effect of either of the employed overexpressed SPRR5 constructs from this dataset impossible (Figure 15A-C).

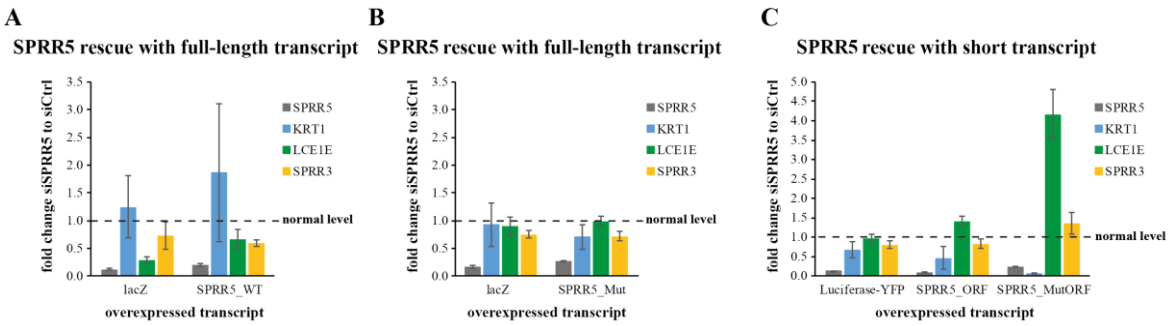


Figure 15: SPRR5 knockdown and rescue in calcium-induced keratinocyte differentiation
Expression of differentiation genes was quantified on day 3 of calcium-induced keratinocyte differentiation and a fold change between SPRR5 depleted (siSPRR5) and control (siCtrl) cells was calculated and plotted for keratinocytes with simultaneous overexpression of the indicated transcripts. (A) The possible rescue effect of the wild-type SPRR5 transcript (SPRR5_WT) is tested, whereas (B) depicts overexpression of the same RNA with a mutated start codon (SPRR5_Mut); both shown next to values obtained for the lacZ control (n=2-5). (C) Only the protein-coding region of the SPRR5 transcript is tested (SPRR5_ORF) and compared to the same RNA with a mutated start codon (SPRR5_MutORF) as well as the luciferase-YFP control (n=2-3).

Due to the inconclusive results with *in vitro* differentiated keratinocytes (Figure 15), the rescue experiment was repeated in organotypic epidermal tissue cultures since these represent a more natural environment for keratinocyte differentiation and hence might exhibit the anticipated differentiation defect upon SPRR5 depletion even in this experimental rescue setting. Surprisingly, once again no reduced differentiation rate between SPRR5 deficient and control keratinocytes with luciferase-YFP overexpression could be observed (Figure 16B). In fact, it appeared that SPRR5 depletion in this setup lead to increased expression of differentiation genes and these elevated transcript levels were not drastically altered by overexpression of SPRR5_ORF or SPRR5_MutORF, although introducing the latter seemed to slightly intensify this effect (Figure 16B).

SPRR5 depletion in keratinocytes overexpressing lacZ eventually resulted in the expected differentiation defect and thus a potential rescue of the full-length SPRR5 transcripts could be analyzed (Figure 16A). Whereas overexpression of the lncRNA version SPRR5_Mut did not rescue the effect of SPRR5 depletion, introducing the SPRR5 wild type transcript (SPRR5_WT) reversed the differentiation defect for SPRR5 depleted keratinocytes and even lead to an increased expression of differentiation genes (Figure 16A). Therefore, it seemed that the SPRR5 protein is the functional molecule in keratinocytes.

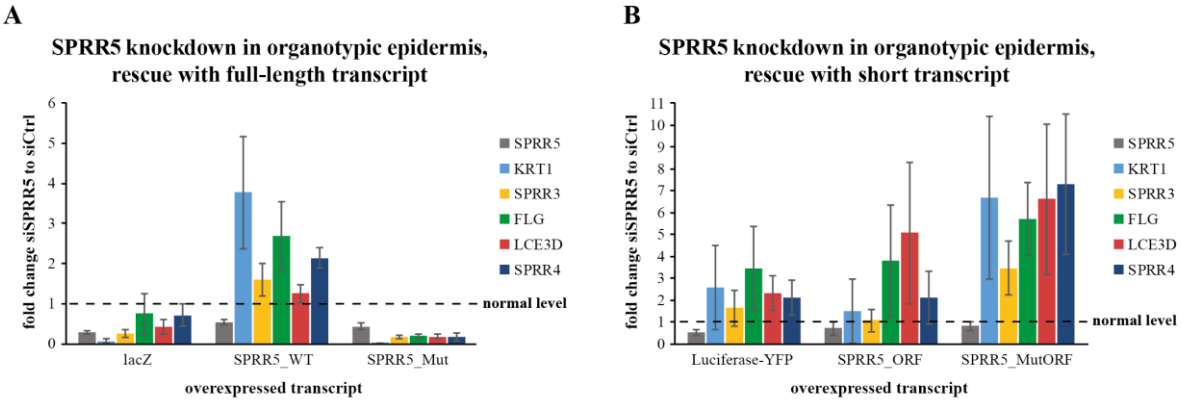


Figure 16: SPRR5 rescue experiment in organotypic epidermal tissue on day 3 of differentiation
Differentiation marker levels in regenerated organotypic epidermal tissue (D3) were assessed by RT-qPCR for SPRR5 depleted (siSPRR5) and control (siCtrl) cells with simultaneous overexpression of the indicated transcripts and the fold change between siSPRR5 and siCtrl was calculated and plotted for each condition. (A) The potential rescue effect of wild-type SPRR5 transcript (SPRR5_WT) and same RNA with a mutated start codon (SPRR5_Mut) was analyzed and compared to the lacZ control (n=2). (B) The SPRR5 protein-coding region with an intact (SPRR5_ORF) or a destroyed start codon (SPRR5_MutORF) as well as the luciferase-YFP control were tested for their potential to rescue SPRR5 deficiency (n=3).

To exclude the possibility, that the observed SPRR5_WT-mediated rescue was simply due to an incomparable differentiation program in this experimental setting, differentiation marker mRNA levels for the knockdown control samples ("siCtrl") with simultaneous overexpression of lacZ, SPRR5_WT or SPRR5_Mut were compared to the lacZ overexpression control (Figure 17).

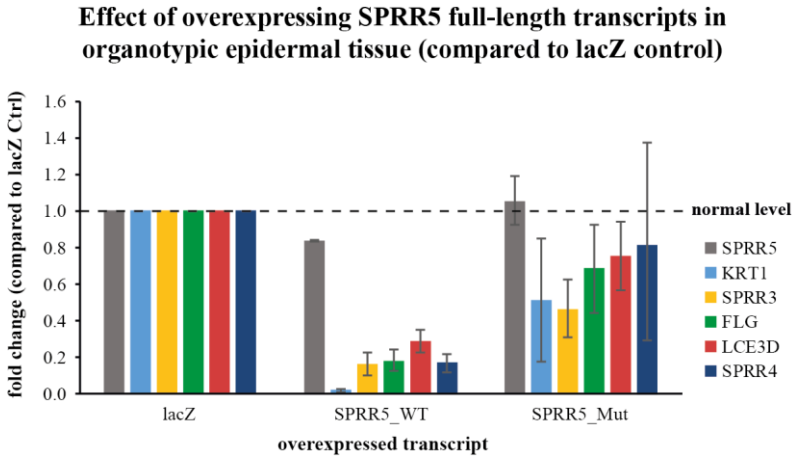


Figure 17: Effect of overexpressing full-length SPRR5 transcripts in organotypic epidermis
Expression of differentiation genes in day 3 organotypic epidermal tissue obtained from knockdown controls (cells treated with a scrambled siRNA control) was measured by RT-qPCR and the expression levels after overexpression of either lacZ, the wild-type SPRR5 transcript (SPRR5_WT), or the same RNA with a destroyed start codon (SPRR5_Mut) were evaluated in relation to the expression in cells overexpressing lacZ (n=2). Ctrl = control

This analysis clearly showed, that overexpression of SPRR5_Mut had only a negligible effect on epidermal regeneration, whereas SPRR5_WT overexpressing cells exhibited a dramatically reduced terminal differentiation state (Figure 17), indicating that the attenuated differentiation rate in SPRR5_WT overexpressing cells might be the reason for the observed rescue effect of SPRR5_WT. In fact, the reduced differentiation in SPRR5_WT overexpressing cells could not be further decreased by SPRR5 depletion and hence the difference between SPRR5 depleted keratinocytes and their respective knockdown controls was abolished, which misleadingly appeared like a genuine rescue effect during the previous analysis (Figure 16). Moreover, the decreased differentiation rate for the SPRR5_WT overexpressing cells once more highlights the fluctuating differentiation programs caused be simultaneous siRNA treatment and lentiviral transduction of primary keratinocytes. Additional support for this assumption comes from evaluation of the previous rescue experiments (Figure 15 and Figure 16), where no differentiation defects could be observed after SPRR5 depletion which is in sharp contrast to the general necessity of SPRR5 for proper keratinocyte differentiation (4.4).

In summary, no conclusion about the identity of the functional SPRR5 molecule can be drawn from the conducted rescue experiments due to the fluctuating differentiation program in this particular experimental setting.

4.3.4 Generation of SPRR5 knockout cell lines

Due to the inconclusive results from the conducted knockdown-based rescue experiments (4.3.3), another approach for SPRR5 depletion was inevitable and SPRR5 knockout cells were chosen as a substitute. However, primary keratinocytes were unsuitable for this approach since they possess only a very limited number of cell doublings *in vitro*[173,174]. Thus, the immortalized HaCaT keratinocyte cell line which is described to differentiate properly in an epidermal environment was utilized for the generation of SPRR5 knockout cell lines. As a first step towards the generation of a SPPR5 knockout cell line, the ability of HaCaT cells to differentiate in our calcium-induced keratinocyte differentiation cultures as well as in regenerated organotypic epidermal tissue cultures was assessed[175]. This pilot experiment clearly indicated that HaCaT cells are a suitable model as they express SPRR5 and conduct a similar differentiation program like primary keratinocytes, although with a slightly reduced expression of differentiation markers and absence of loricrin protein expression (Supplementary Figure 3).

In theory, knockout cell line generation works as described in the following: After transfection of a plasmid encoding chimeras of the desired gRNAs and the required trans-activating CRISPR RNA (tracrRNA), the gRNAs should target the Cas9 endonuclease (also encoded on the plasmid) to the intended genomic cutting sites via base complementarity between the gRNA and the DNA[176,177]. Next, Cas9 cleaves the DNA specifically at the targeted sites and the resulting DNA double-strand breaks are repaired either via non-homologous end joining (NHEJ) or the homology-directed repair (HDR) mechanism[176]. In this case however only the NHEJ of two cutting sites flanking the SPRR5 locus would generate the desired genomic SPRR5 deletion, as HDR with potentially remaining wild type alleles would reconstitute the initial state with no genomic alteration. Since the introduction of double-strand breaks is a crucial step in this process, two guide RNAs on each side of the intended genomic deletion were used simultaneously to increase the chances of a successful SPRR5 knockout. Furthermore, it should be mentioned that at the time of SPRR5 knockout cell line generation, the first exon of SPRR5 was not known and hence only the large second exon, including the open reading frame, was targeted by the deletion approach (Figure 18).

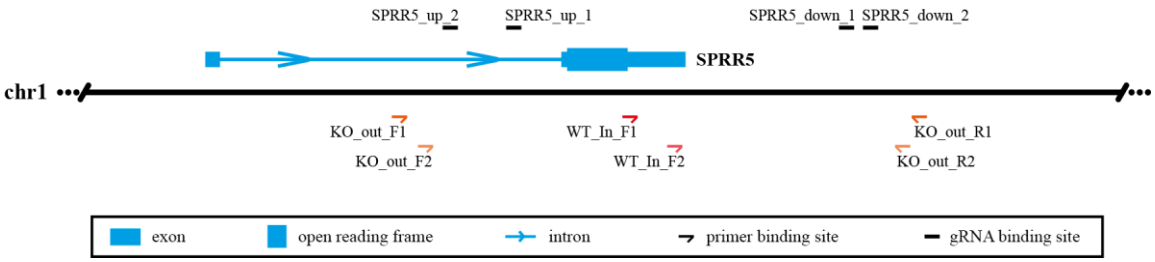


Figure 18: Localization of gRNAs and screening primers at the SPRR5 locus
The genomic SPRR5 locus and its transcript (light blue) is shown. Additionally, the localization of the utilized gRNAs for SPRR5 knockout (black bars), as well as the position of the employed screening primers for the detection of knockout alleles (orange arrows) and the forward primers for the detection of wild-type alleles (red arrows) are depicted. up = upstream, down = downstream, KO = knockout, WT = wild type

The results from the PCR-based screening of potential SPRR5 knockout cell lines, obtained by transient transfection of a plasmid encoding all four gRNAs and the Cas9 endonuclease and subsequent generation of a clonal cell line, are summarized in Table 1. Considering that during the first round of transfection only 3.5% of screened clones were heterozygous for genomic SPRR5 deletions and no homozygous knockout cell line was obtained, it became evident, that the genomic editing was not very efficient. For this reason, three heterozygous cell lines from the first experiment, were subjected to another cycle of transfection and

single-cell seeding in order to eventually get a homozygous SPRR5 knockout cell line but despite screening of 402 clones, no homozygous SPRR5 knockout clone was found (Table 1).

Table 1: Results from SPRR5 knockout cell line screening (transient approach)

	1st round	2nd round
Heterozygous SPRR5 knockout	24 (3.5%)	n.d.
Homozygous SPRR5 knockout	-	-
Total number of generated cell lines	692	402

Since the transient transfection approach did not yield any homozygous SPRR5 knockout cell lines, a lentiviral delivery system was chosen for stable genomic integration of the four gRNAs as well as the Cas9 endonuclease into wild type HaCaT cells which would ensure prolonged and constant expression of Cas9 as well as the gRNAs and thus might increase the probability for genomic deletions. And indeed, after only one cycle of infection, selection and clonal expansion, the portion of heterozygous clones could be raised from 3.5% to 13.4% but once more no homozygous SPRR5 knockout cell line could be identified (Table 2). Therefore, six heterozygous clones from the first lentiviral infection cycle were chosen and the presence of any remaining SPRR5 wild type alleles was assessed after repeating the lentiviral-based knockout procedure. Surprisingly, even this rather drastic treatment did not result in one single homozygous SPRR5 knockout cell line (Table 2) which is the reason for ceasing this subproject.

Table 2: Results from SPRR5 knockout cell line screening (lentiviral approach)

	1st round	2nd round
Heterozygous SPRR5 knockout	41 (13.4%)	n.d.
Homozygous SPRR5 knockout	-	-
Total number of generated cell lines	306	435

In summary, the identity of the functional SPRR5 molecule in differentiated keratinocytes could not be ultimately deciphered. Although bioinformatic predictions indicate that SPRR5 does not act as a non-coding transcript, only minor amounts of the SPRR5 protein could be detected in differentiated keratinocytes and phylogenetic analyses show that SPRR5 differs drastically from other human SPRR proteins. Furthermore, the conducted rescue experiments as well as the generation of an SPRR5 knockout cell line were inconclusive, leaving the question whether SPRR5 exerts its function as a protein or lncRNA unsolved.

4.4 Epidermal homeostasis is controlled by SPRR5

4.4.1 SPRR5 is required but not sufficient for keratinocyte differentiation

Despite the ambiguity about the functional molecule, the role of the SPRR5 transcript for keratinocyte differentiation could be assessed, since the level of both molecules (lncRNA and protein) depends on the available amount of SPRR5 RNA.

As a first step towards this functional characterization, SPRR5 transcript amounts were monitored over the course of keratinocyte differentiation, demonstrating a strong induction of SPRR5 during terminal differentiation (Figure 19A+B), which already implied a functional importance of SPRR5 for keratinocyte differentiation. In order to test this hypothesis, siRNA-mediated knockdown of SPRR5 was performed in keratinocyte differentiation cultures and the expression of several key differentiation markers was assessed by RT-qPCR as well as western blot analysis (Figure 20).

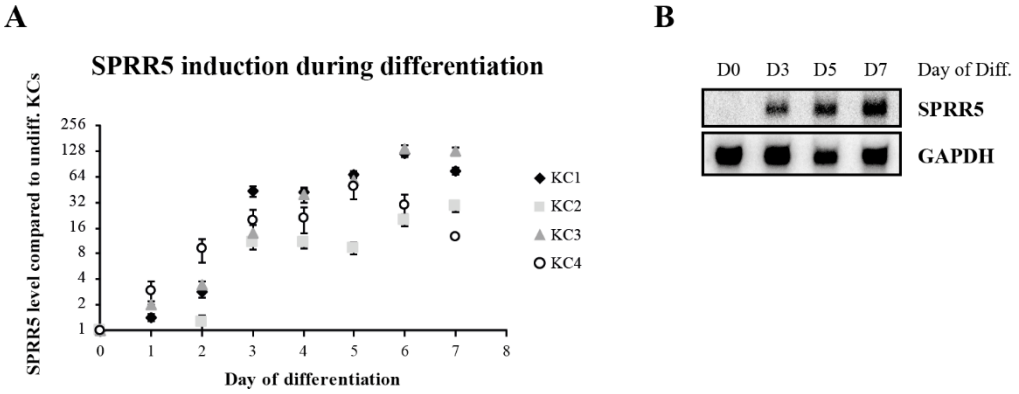


Figure 19: SPRR5 induction during keratinocyte differentiation
SPRR5 transcript levels were analyzed by RT-qPCR and compared to the transcript levels in undifferentiated keratinocytes in four different primary keratinocyte isolates (KC1-4; n=2), revealing a strong SPRR5 induction during keratinocyte differentiation (A). Furthermore, induction of SPRR5 was confirmed by northern blot analysis, using GAPDH as a loading control (B). KC = keratinocyte, undiff. = undifferentiated, Diff. = differentiation, GAPDH = glyceraldehyde 3-phosphate dehydrogenase

For all three timepoints of differentiation, an efficient knockdown of SPRR5 transcript levels was obtained (<5% SPRR5 remaining), which resulted in a reduced expression of the evaluated differentiation marker mRNAs small proline-rich protein 3 (SPRR3), keratin 1 (KRT1), as well as filaggrin (FLG) (Figure 20A+B). Furthermore, also the protein levels of keratin 1 (KRT1) and loricrin (LOR) were diminished upon SPRR5 depletion (Figure 20C), proving that SPRR5 is essential for proper terminal keratinocyte differentiation.

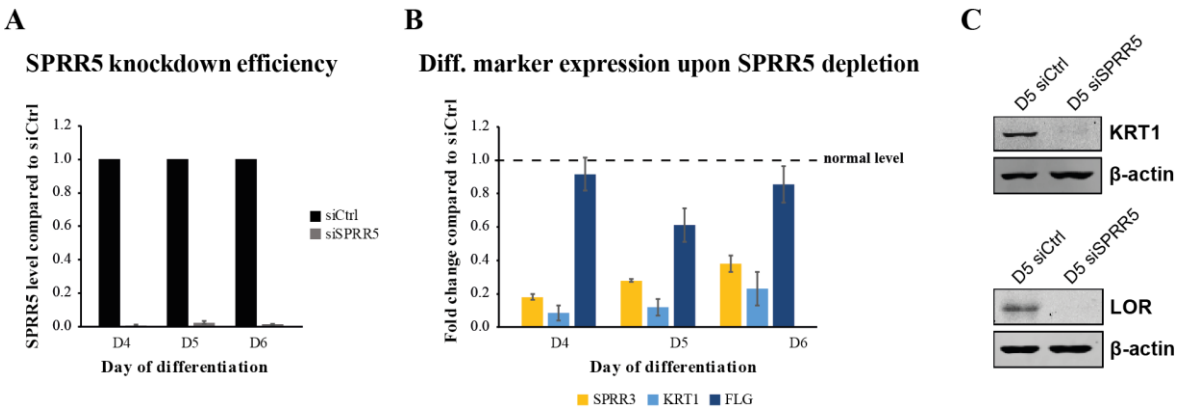


Figure 20: SPRR5 depletion leads to differentiation defects
Depletion of SPRR5 by siRNAs (A) leads to a decreased expression of differentiation marker mRNAs (B) as obtained by RT-qPCR (n=3-4). Western blot analysis reveals a reduction of differentiation protein levels upon SPRR5 knockdown (C). Ctrl = control, Diff. = differentiation

Whether SPRR5 alone is sufficient to ectopically induce the keratinocyte differentiation program was examined by overexpression of SPRR5 followed by RT-qPCR analysis of differentiation marker mRNA levels. Compared to the lacZ control, a moderate overexpression of roughly threefold higher SPRR5 levels was achieved (Figure 21A), however no elevated transcript levels for the small proline-rich protein 3 (SPRR3), keratin 1 (KRT1) or filaggrin (FLG) could be detected and the mRNA amounts of the late cornified envelope proteins 1E and 3A (LCE1E and LCE3A) were only slightly increased on day 3 and day 4 of differentiation (Figure 21B). Thus, the SPRR5 transcript is essential but not sufficient for proper keratinocyte differentiation.

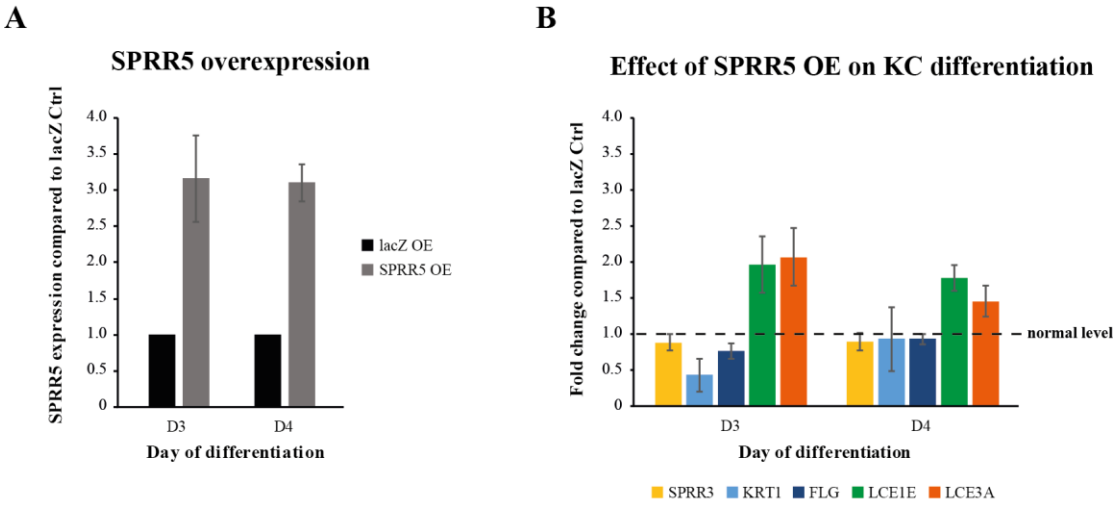


Figure 21: Effect of SPRR5 overexpression on keratinocyte differentiation
SPRR5 (or lacZ as a control) was overexpressed in keratinocytes and differentiation was monitored on day 3 and day 4 by evaluating overexpression levels of SPRR5 (A) and the resulting effect on mRNA amounts for key differentiation markers by RT-qPCR (B) (n=2-3). OE = overexpression, KC = keratinocyte, Ctrl = control

4.4.2 Epidermal tissue homeostasis necessitates SPRR5

Since SPRR5 regulates keratinocyte differentiation *in vitro* (4.4.1), it was interesting to test whether SPRR5 also promotes keratinocyte differentiation in a mature tissue environment and hence SPRR5 knockdown was conducted in epidermal tissue cultures.

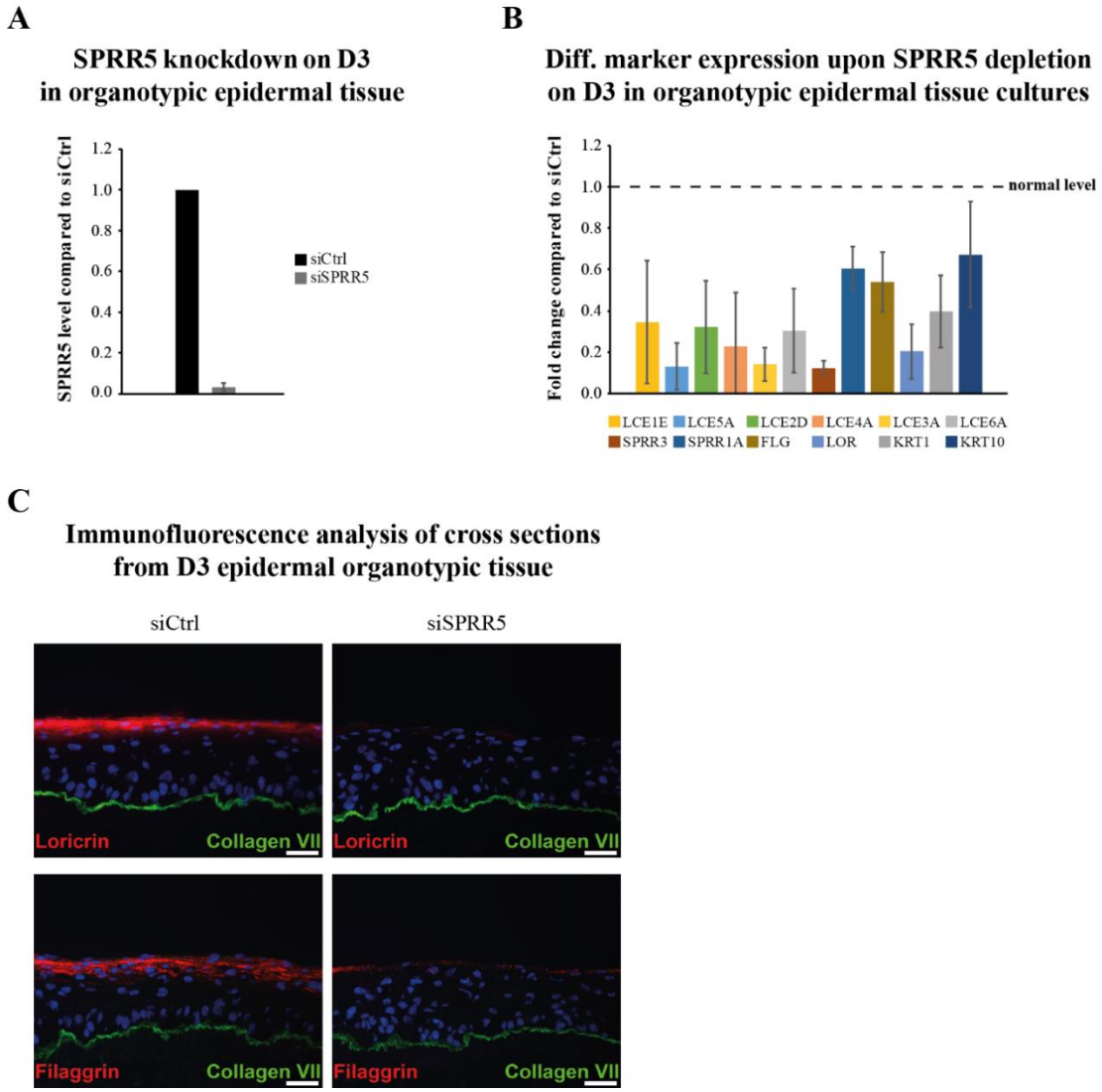


Figure 22: SPRR5 controls epidermal tissue homeostasis on day 3 of differentiation
Efficient depletion of SPRR5 in organotypic epidermal tissue cultures (A) results in diminished expression of differentiation marker mRNAs (B) as obtained by RT-qPCR analysis (n=4). Exemplary pictures of immunofluorescence analysis of cross sections from the corresponding regenerated epidermal tissue (C) reveals reduced amounts of the differentiation proteins loricrin and filaggrin (red) in SPRR5 depleted tissue. Collagen VII (green) is depicted for orientation purposes as it separates the epidermis from the beneath lying dermis, nuclei are shown in blue and the scale bar (white) indicates 50 µm. Diff. = differentiation, Ctrl = control

On day 3 of differentiation, SPRR5 transcript levels were efficiently diminished down to 5% (as compared to the control) (Figure 22A), resulting in an extensive loss of differentiation marker expression as illustrated by the reduced expression of several members of the late cornified envelope proteins (LCEs), small proline-rich proteins (SPRRs), keratins (KRTs)

as well as filaggrin (FLG) and loricrin (LOR) (Figure 22B). Moreover, cross sections from the obtained epidermal tissue also revealed decreased levels for the differentiation proteins filaggrin and loricrin, whereas no stratification defect or aberrant epidermis thickness was observed (Figure 22C). Additionally, analysis of SPRR5 depleted organotypic epidermal tissue on day 4 of differentiation lead to similar results (Supplementary Figure 4), proving that SPRR5 is not only required for *in vitro* differentiation of keratinocytes but also essential for maintaining epidermal tissue homeostasis.

4.4.3 SPRR5 regulates keratinocyte differentiation on a global level

In order to obtain a comprehensive picture of the global impact of SPRR5 knockdown on epidermis development, a full transcriptome sequencing approach of SPRR5 depleted and control tissue was performed. To this end, three to five biological replicates of regenerated organotypic epidermis with and without SPRR5 depletion were harvested on day 3 or on day 4 of differentiation and the isolated RNA was subjected to a poly-A enrichment followed by library preparation and next generation sequencing. With 24 to 33 million mapped reads per sample, a sufficient sequencing depth was received and the subsequent principal component analysis proved comparability between the transcriptomes of the biological replicates for each timepoint and nice clustering in a control and SPRR5 depletion group (Supplementary Figure 5). Thus, differential gene expression analysis was done for each timepoint (work by Dr. Nicholas Strieder) using a customized DeSeq2 script and the obtained results were successfully verified by RT-qPCR for selected transcripts, confirming the validity of the obtained RNA-Seq analysis (Supplementary Figure 6).

First, the efficient knockdown of SPRR5 with 3% and 9% remaining transcript amounts on day 3 or day 4 respectively (Figure 23B) should be emphasized which ultimately resulted in numerous differentially expressed genes ($-0.5 > \log_2(\text{fold change}) > 0.5$ and an adjusted p-value < 0.05; lists of differential expressed genes are included in the appendix). Whereas on day 3 of differentiation a total of 249 genes was differentially expressed and 209 genes were downregulated and 40 genes upregulated upon SPRR5 depletion (Figure 23C), a similar pattern could be observed for the day 4 samples as with 323 downregulated and 56 upregulated genes, the majority of the 379 total regulated genes showed decreased expression upon loss of SPRR5 (Figure 23D). In combination with the results from the previous studies that indicated a reduced differentiation pattern upon SPRR5 depletion (4.4.1 and 4.4.2), it seemed reasonable to proceed only with the downregulated genes and test whether differentiation genes are enriched within this subset.

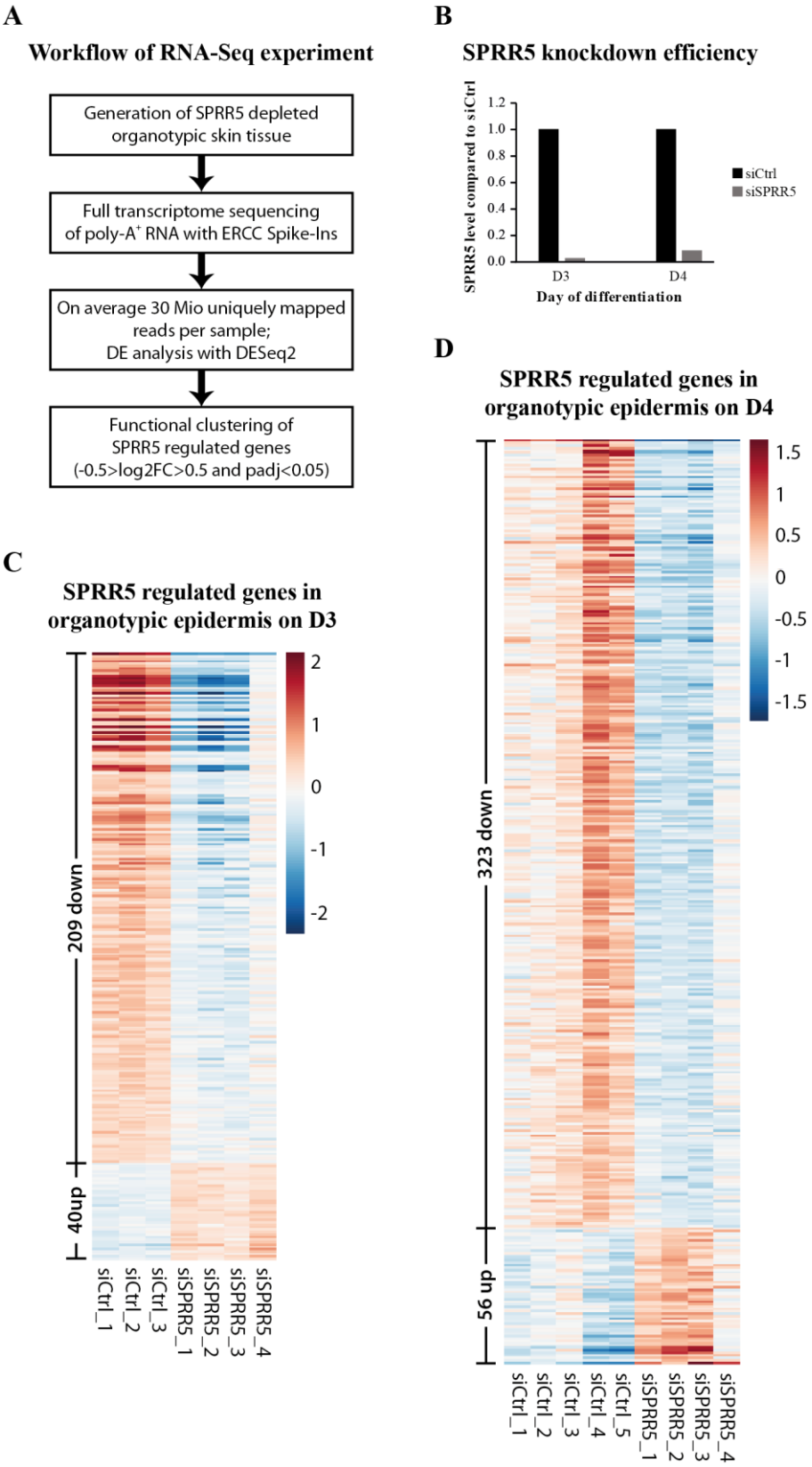


Figure 23: SPRR5 depletion severely alters the transcriptome in epidermal tissue
Workflow of the performed RNA-Seq experiment (A) and the obtained knockdown efficiency of SPRR5 for each timepoint of differentiation as obtained from the RNA-Seq analysis (B). Heatmaps of genes with a significantly altered expression (adjusted p-value < 0.05 and -0.5 > log2 fold change > 0.5) for day 3 (C) and day 4 of differentiation (D) reveal that SPRR5 depletion changes the expression of a plethora of genes. log2FC = $\log_2$(fold change), padj = adjusted p-value, Ctrl = control, up = upregulated, down = downregulated

And indeed, functional annotation clustering of the genes with decreased expression upon SPRR5 depletion showed highly significant enrichment for gene ontology terms like "cornified envelope", "keratinocyte differentiation" or "epidermis development" (Figure 24), ultimately proving that SPRR5 is indispensable for proper keratinocyte differentiation in an epidermal tissue environment. In contrast to this, a functional annotation clustering of genes that are upregulated upon SPPR5 knockdown did not disclose any significant enrichment for a gene ontology term or reveal any conclusive link between the subset of upregulated genes (data not shown).

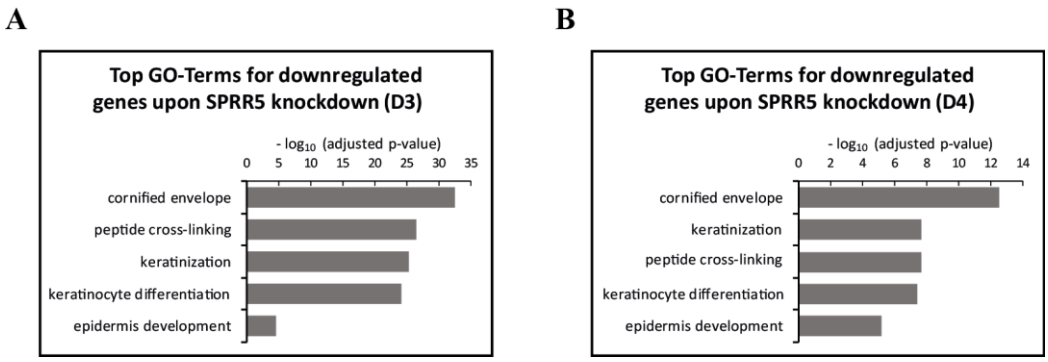


Figure 24: GO-Term analysis of genes with decreased expression upon SPRR5 depletion
GO-Term enrichment analysis was done for significantly downregulated genes after SPRR5 knockdown with the DAVID tool and the negative decadic logarithm of the 5 most enriched GO-terms of the classes "biological process" and "cellular compartment" is plotted for day 3 (A) and day 4 (B) of differentiation.
GO = gene ontology

Interestingly, analysis of the genomic distribution of SPRR5 regulated genes revealed only one significant hit, namely chromosome 1 cytogenetic band 21, which is especially striking as this represents the genomic localization of the human epidermal differentiation complex (EDC) where many keratinocyte differentiation genes as well as SPRR5 itself are encoded. Thus, a more detailed analysis of the SPRR5 regulated genes within the EDC was performed, revealing the regulation of almost the complete late cornified envelope protein (LCE) gene cluster on day 3 of differentiation and the small proline-rich protein cluster (SPRR) on day 4, so precisely those gene clusters flanking the genomic SPRR5 locus (Figure 25A). Furthermore, it seemed like SPRR5 mainly regulated a specific subset of genes for each timepoint of differentiation individually and only a smaller set of concordantly regulated genes. This initial assumption could be confirmed as only 87 genes (16%) are regulated by SPRR5 at both timepoints, whereas SPRR5 controls 162 (30%) genes exclusively on day 3 and 292 genes (54%) only on day 4 of differentiation (Figure 25B). In conclusion, SPRR5 is not only essential for keratinocyte differentiation but controls different gene clusters during the course of differentiation indicating a possible epigenetic mechanism of SPRR5.

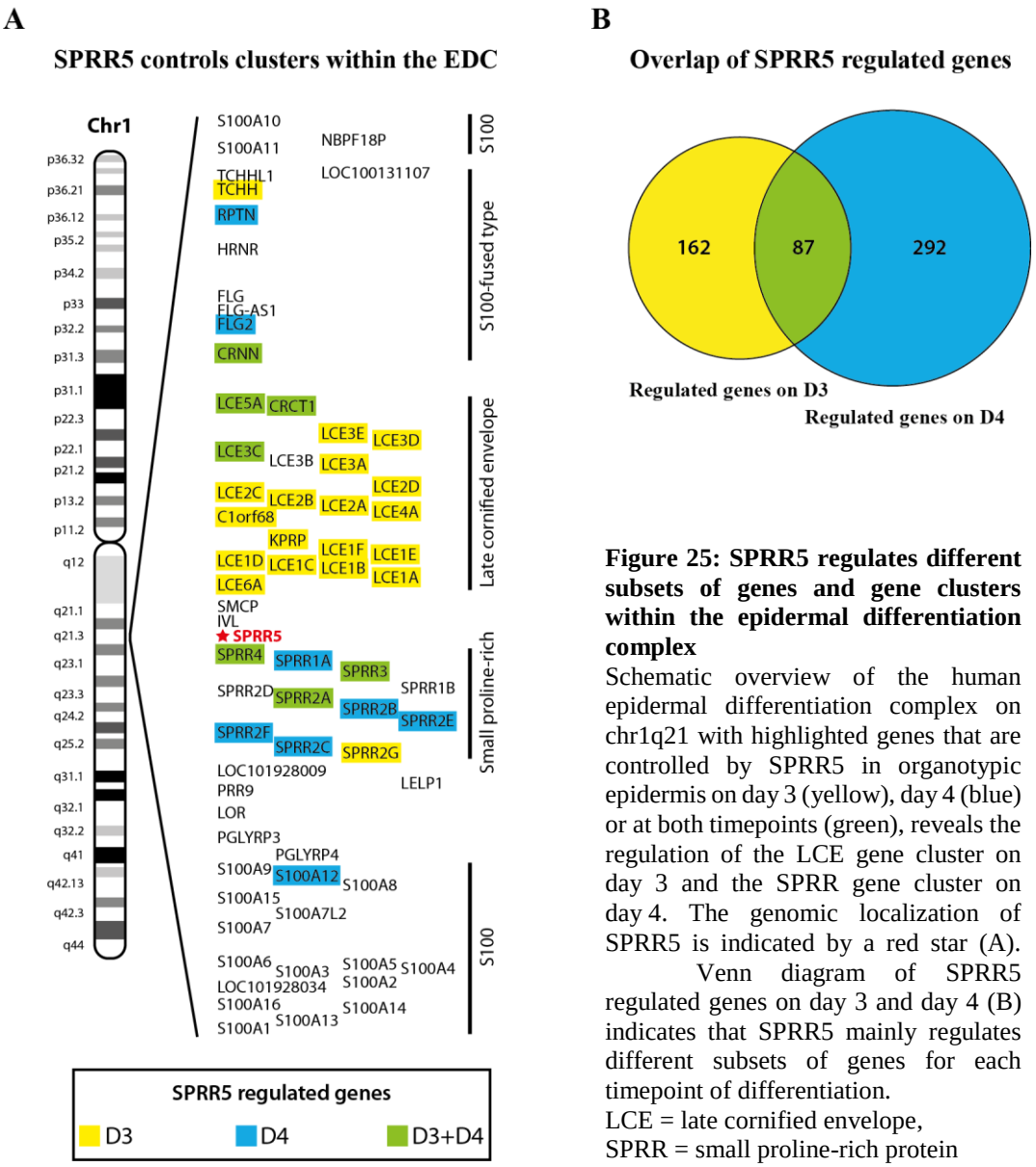


Figure 25: SPRR5 regulates different subsets of genes and gene clusters within the epidermal differentiation complex
Schematic overview of the human epidermal differentiation complex on chr1q21 with highlighted genes that are controlled by SPRR5 in organotypic epidermis on day 3 (yellow), day 4 (blue) or at both timepoints (green), reveals the regulation of the LCE gene cluster on day 3 and the SPRR gene cluster on day 4. The genomic localization of SPRR5 is indicated by a red star (A). Venn diagram of SPRR5 regulated genes on day 3 and day 4 (B) indicates that SPRR5 mainly regulates different subsets of genes for each timepoint of differentiation.
LCE = late cornified envelope,
SPRR = small proline-rich protein

4.4.4 Investigating the potential epigenetic mechanism of SPRR5

Given the observed regulation of adjacent gene clusters by SPRR5 (4.4.3) and the numerous reports about epigenetic control of keratinocyte differentiation[46,49,178–180], an epigenetic mechanism for SPRR5 seemed promising. In order to address this hypothesis, the effect of SPRR5 depletion on the accessibility of genomic regions was probed with ATAC-Seq in SPRR5 depleted and regular keratinocytes for day 4 of differentiation as well as in undifferentiated cells (in collaboration with Katrin Hartinger and the AG Rehli).

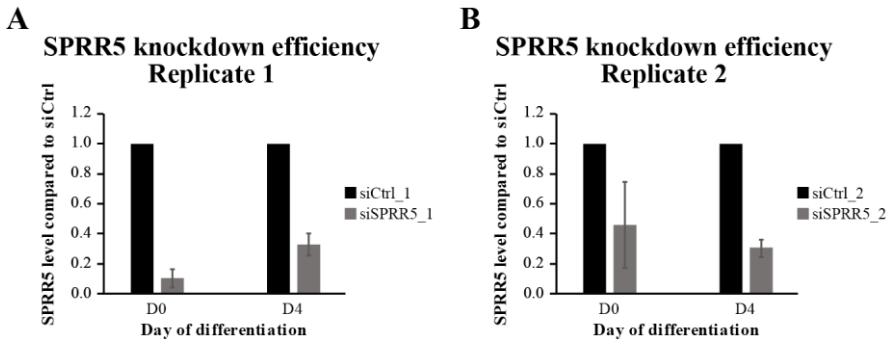


Figure 26: Knockdown efficiency for ATAC-Seq replicates
SPRR5 knockdown efficiencies in undifferentiated (D0) and differentiated (D4) keratinocytes as obtained by RT-qPCR for the biological ATAC-Seq replicates 1 (A) and 2 (B).

As a first step in this analysis successful depletion of SPRR5 was validated for both biological replicates (Figure 26), followed by a quality control of the obtained sequencing data, revealing an overall satisfying data quality with only few mitochondrial reads. Nevertheless, it should be mentioned that with 6.6 to 73.8 million reads per sample, the total number of sequencing reads per sample varied drastically (Table 3).

Table 3: Quality control of ATAC-Seq data (FRIP = fraction of reads in called peak regions)

Sample	Total reads	Mapped reads	Mitochondrial reads	Unique reads	FRIP
D0_siCtrl_1	29.2 mio	94.2%	0.71%	18.1	34.5%
D0_siCtrl_2	6.9 mio	95.9%	3.38%	5.1	40.2%
D0_siSPRR5_1	16.5 mio	88.5%	0.67%	11.8	37.8%
D0_siSPRR5_2	73.8 mio	96.2%	2.27%	31.5	41.9%
D4_siCtrl_1	12.2 mio	93.9%	0.67%	8.5	29.7%
D4_siCtrl_2	6.6 mio	95.9%	0.56%	5.1	22.3%
D4_siSPRR5_1	33.7 mio	95.1%	0.67%	20.3	28.2%
D4_siSPRR5_2	46.4 mio	96.4%	0.67%	29.5	22.7%

Additionally, current standards for ATAC-Seq data from the ENCODE consortium were applied and with two biological replicates, over 80% mapped reads, >20% fraction of reads in called peak regions (FRIP) and over 100,000 called peaks for each timepoint also met in general[181,182]. However, the ENCODE required sequencing depth of over 25 million unique reads was not reached for almost all samples (Table 3) and further criteria were not investigated because replicate reproducibility for example was analyzed via principal component analysis (PCA). Since in these PCA plots the second biological control replicate (siCtrl_2) did cluster far away from all the other samples and since siCtrl_2 was the sample with the lowest amount of sequencing reads for each timepoint, this sample was excluded during further analysis (Figure 27 A+B). Rerunning the PCA analysis without siCtrl_2 showed no drastic differences in terms of chromatin accessibility for undifferentiated

samples (D0) but a nice clustering for day 4 of differentiation, matching the observation that SPRR5 is preferably crucial for gene expression in differentiated keratinocytes (Figure 27 C+D). Since in undifferentiated cells no noteworthy changes in chromatin accessibility due to SPRR5 depletion were detectable (data not shown), only the results from day 4 of differentiation are shown here.

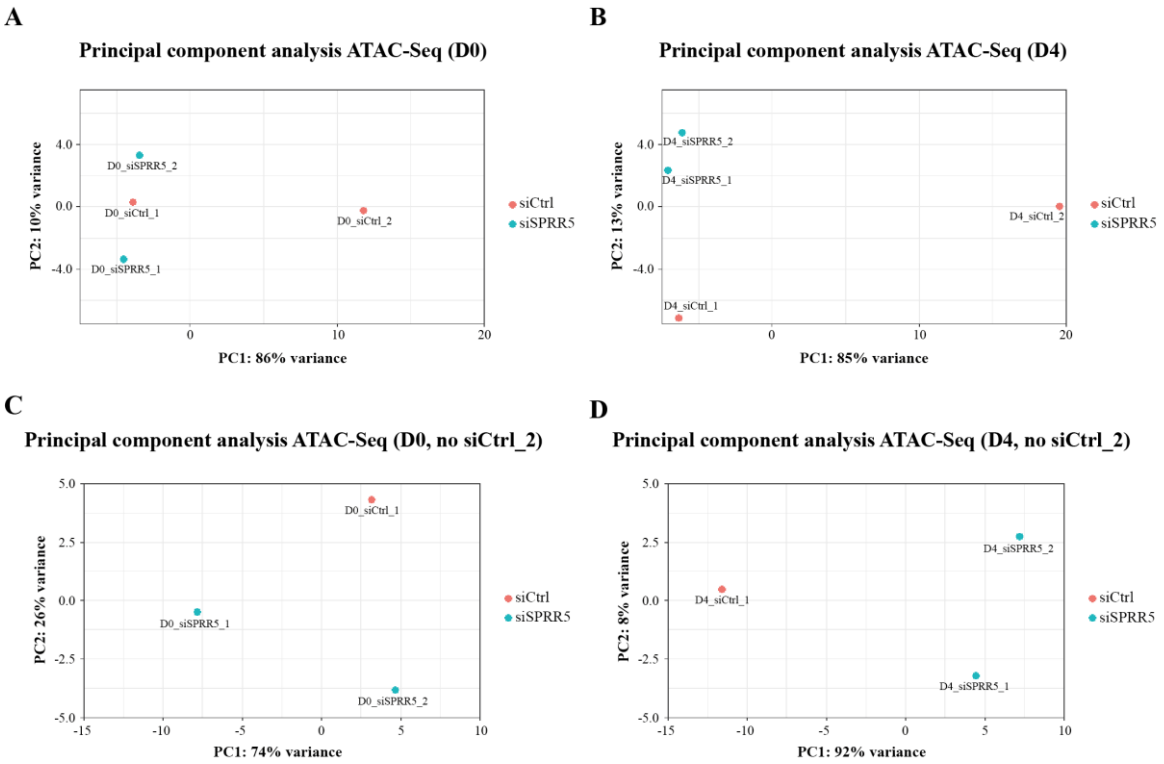


Figure 27: Principal component analysis of ATAC-Seq samples
Principal component analysis of ATAC-Seq samples on D0 (A) and D4 (B) showed a drastic difference for the control replicate 2 (siCtrl_2), which is why this was excluded for subsequent analysis. Repeating the PCA analysis without siCtrl_2 for D0 (C), indicates that all three samples are similar, whereas for D4 (D) the SPRR5 depleted samples cluster separately. Ctrl = control, PC = principal component

Regarding the changes in genomic accessibility upon SPRR5 depletion in differentiated keratinocytes (Figure 28, left), no clear difference between opened and closed regions on a genome wide level could be detected, although the vast downregulation of differentiation gene expression identified by RNA-Seq (4.4.3) would suggest more closed regions upon SPRR5 depletion. To omit the possibility that SPRR5 does control accessibility of the adjacent LCE and SPRR gene clusters but fails to regulate chromatin accessibility on a global level, the analysis was restricted to the epidermal differentiation complex resulting in four opened and four closed regions (Figure 28, right).

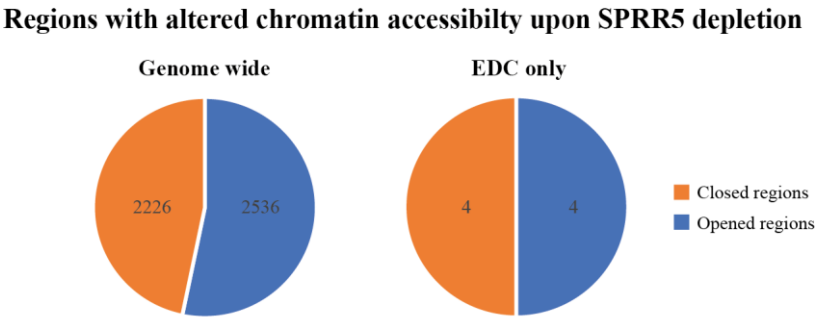


Figure 28: SPRR5 controls genomic accessibility
Regions exhibiting altered accessibility in SPRR5 deficient keratinocytes for the whole genome (left) or solely the epidermal differentiation complex (EDC; right) on day 4 of keratinocyte differentiation are shown. SiCtrl_2 was excluded for this analysis and opened regions (blue) were defined by a $\log_2$(fold change) higher than 0.5, whereas closed regions (orange) had to meet a $\log_2$(fold change) lower than -0.5.

Given the formulated hypothesis that SPRR5 could control gene expression of the LCE and the SPRR gene clusters, SPRR5 depletion should result in an enrichment of closed chromatin in these areas. However, only one of the four closed regions localized to the SPRR gene cluster whereas the remaining regions were found roughly 470 kbp upstream and 220 kbp downstream of the LCE or the SPRR cluster. In combination with the finding that 50% of the opened regions were present within the genomic LCE and SPRR region, no evidence for an epigenetic regulation of chromatin accessibility for those gene clusters by SPRR5 could be observed.

Apart from opening or closing parts of the genome, SPRR5 might affect the distribution of histone modifications and exert its regulatory function via this epigenetic mechanism. Thus, the occurrences of the repressive histone mark H3K27me3 as well as the activating marks H3K4me1 and H3K27ac were compared between regular and SPRR5 depleted keratinocytes with ChIP-Seq (in collaboration with Katrin Hartinger and the AG Rehli)[183,184].

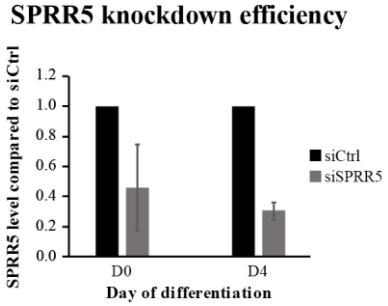


Figure 29: SPRR5 knockdown efficiency for ChIP-Seq
The SPRR5 knockdown efficiency in undifferentiated (D0) and differentiated (D4) keratinocytes is shown for the cells utilized for ChIP-Seq sample generation.
Ctrl = control

Similar to the ATAC-Seq analysis, successful SPRR5 knockdown validation (Figure 29) was followed by a general quality control of the obtained sequencing data (Table 4). With total numbers of reads ranging from 13.9 to 38.4 million and 69 to 77% unique reads (after removal of duplicate reads and low-quality mapping reads), the current ENCODE standards

for ChIP-Seq demanding 20 to 45 million reads, over 90% unique reads and biological duplicates were not completely matched[181,185]. Nevertheless, the obtained data quality was high enough to get a first impression about a possible involvement of SPRR5 in modifying the histone code in keratinocytes.

Table 4: Quality control of ChIP-Seq samples (FRIP = fraction of reads in called peak regions)

Sample	Total reads	Mapped reads	Unique reads	FRIP
D0_siCtrl_Input	31.7 mio	97.7%	75.0%	n.d.
D0_siCtrl_H3K4me1	24.0 mio	98.2%	77.0%	21.9%
D0_siCtrl_H3K27me3	17.6 mio	98.2%	76.7%	27.3%
D0_siCtrl_H3K27ac	25.0 mio	97.4%	69.0%	31.0%
D0_siSPRR5_Input	38.4 mio	97.4%	72.6%	n.d.
D0_siSPRR5_H3K4me1	30.0 mio	97.9%	74.2%	19.7%
D0_siSPRR5_H3K27me3	25.2 mio	98.2%	74.8%	29.5%
D0_siSPRR5_H3K27ac	22.4 mio	98.0%	74.9%	29.9%
D4_siCtrl_Input	26.2 mio	97.5%	74.6%	n.d.
D4_siCtrl_H3K4me1	15.9 mio	97.8%	76.7%	8.2%
D4_siCtrl_H3K27me3	13.9 mio	97.6%	77.0%	17.5%
D4_siCtrl_H3K27ac	17.7 mio	97.6%	75.7%	29.6%
D4_siSPRR5_Input	28.5 mio	98.1%	76.2%	n.d.
D4_siSPRR5_H3K4me1	20.4 mio	98.6%	76.0%	9.1%
D4_siSPRR5_H3K27me3	15.1 mio	98.2%	75.0%	13.8%
D4_siSPRR5_H3K27ac	19.3 mio	98.4%	75.0%	33.3%

Moreover, with about 20 to 30%, the fraction of reads in called peak regions (FRIP) was satisfactory for almost all samples, however a clearly lower FRIP was obtained for H3K4me1 and H3K27me3 on day 4 of differentiation. For this reason and as an additional second quality control, the obtained ChIP-Seq data for control samples on day 4 of differentiation were compared to already analyzed and published ChIP-Seq data from Bao et al. (an exemplary view is given in Supplementary Figure 7)[44]. Since both datasets nicely correlated for the called peaks, the overall distribution of obtained reads as well as for the observed signal-to-noise ratio, a subsequent analysis of genome wide alterations in histone mark deposition upon SPRR5 depletion was performed with DeSeq2.

Table 5: Results from ChIP-Seq analysis with DeSeq2
(only regions that met $-0.5>\log_2$(fold change) or $\log_2$(fold change)>0.5 are shown, siSPRR5 compared to siCtrl)

Sample	Total altered regions	Increased signal	Decreased signal
D0_H3K4me1	361	228	113
D0_H3K27me3	1,049	247	802
D0_H3K27ac	494	147	347
D4_H3K4me1	11,657	4,381	7,276
D4_H3K27me3	9,787	7,354	2,433
D4_H3K27ac	5,484	3,535	1,949

The results of this global analysis (Table 5) revealed that SPRR5 affects the histone landscape mainly in differentiated cells because in undifferentiated keratinocytes the number of total altered regions is drastically lower than at day 4 of differentiation. Furthermore, in differentiated cells the majority of altered regions for the active chromatin mark H3K4me1 exhibited a decreased signal whereas the repressive mark H3K27me3 predominantly was enriched upon SPRR5 depletion, matching the observed differentiation deficiency upon SPRR5 depletion. Surprisingly, the other active chromatin mark H3K27ac showed also more enriched regions in SPRR5 deficient cells as compared to the control sample.

In order to test whether SPRR5 directly regulates the histone marks at the LCE and SPRR gene cluster, the analysis was restricted to the epidermal differentiation complex (EDC) and only the results for differentiated keratinocytes are shown here, since the previous analysis for undifferentiated cells revealed only minor changes in their histone marks (Table 5).

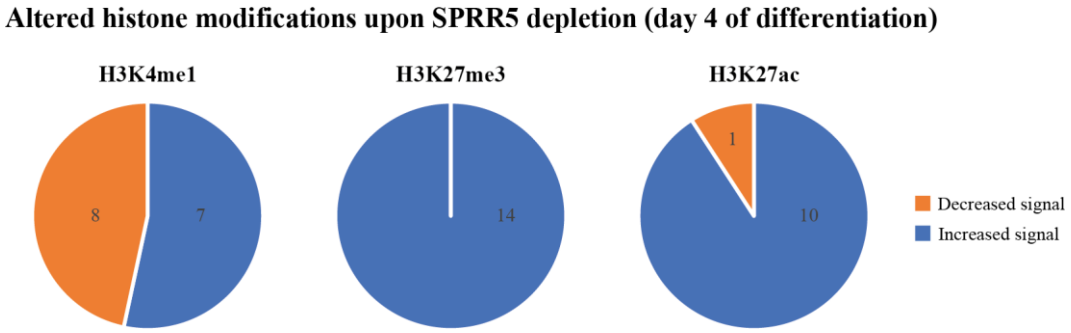


Figure 30: Altered histone modifications upon SPRR5 depletion within the EDC
Histone modifications exhibiting changes upon SPRR5 depletion on day 4 of keratinocyte differentiation and are included within the epidermal differentiation complex (EDC) are depicted. A decreased signal (orange) is defined by a $\log_2$(fold change) lower than -0.5, whereas an increased signal (blue) had to meet a $\log_2$(fold change) above 0.5. Altered signals for the histone modification H3K4me1 are shown on the left, next to H3K37me3 (middle) and H3K27ac on the right.

This close examination of the altered histone marks within the EDC revealed no clear difference between increased and decreased deposition of H3K4me1 but a strong increase in H3K27me3 and H3K27ac sites in SPRR5 deficient cells (Figure 30). Similar to the results from the global analysis, this increase in H3K27 acetylation in SPRR5 deficient keratinocytes would rather contradict the positive effect of SPRR5 on keratinocyte differentiation but since almost all altered peaks were not within the LCE or the SPRR gene cluster, this histone mark could be neglected for studying the possible epigenetic regulation of SPRR5 for those two gene clusters. Interestingly, SPRR5 depletion led to exclusively

more deposition of the repressive histone mark H3K27me3 within the EDC (Figure 30 middle), suggesting an involvement of SPRR5 for prevention this modification during keratinocyte differentiation. However, an analysis of the exact location of the identified increased H3K27me3 signals revealed only two sites within the LCE cluster, none for the SPRR cluster and no overlap with the recently identified enhancer regions of the EDC, rendering an epigenetic control of those regions by SPRR5 rather unlikely[54,186,187].

Given the inconclusive results from the analysis above and the observed rather small signal changes for histone mark signals, the analysis was repeated with another analysis pipeline for ChIP-Seq data called HOMER[188]. Surprisingly, this second analysis lead to strikingly different and partially completely contrary results in the distribution of histone marks upon SPRR5 depletion (results not shown). Thus, no evidence for an epigenetic regulation of SPRR5 during keratinocyte differentiation could be obtained from the analyzed ATAC-Seq and ChIP-Seq data sets and the molecular mechanism of SPRR5 remains elusive.

5 Discussion and Outlook

5.1 Linking SPRR5 and neoplasia

Considering the vast amount of lncRNAs featuring an aberrant expression pattern in cancer and in particular in skin cancer progression (examples given in 2.1.4), we are only beginning to realize their significant role for carcinogenesis. Interestingly, a recent study from Piipponen et al. unraveled that the lncRNA PICSAR promotes growth of cutaneous squamous cell carcinoma (SCC) by regulating ERK1/2 activity[90]. Additionally, the identification of seven annotated but uncharacterized lncRNAs that are differentially expressed in SCC impressively highlights that lncRNAs are crucial regulators of SCC progression and that our current knowledge about their numbers and molecular modes of action is very limited[89]. Furthermore, a dual role during proper keratinocyte differentiation on the one hand and in SCC generation on the other hand seems to be characteristic for several skin specific lncRNAs such as TINCR and SMRT-2, which have been reported to be involved during both processes and are induced during the course of keratinocyte differentiation and suppressed in SCC samples[88,89]. SPRR5_326 exhibits the exact same expression pattern and since lack of differentiation is a hallmark of skin cancer, sustained SPRR5_326 expression might prevent neoplasia. In order to test this hypothesis, an invasion assay for tumorigenic keratinocytes in neoplastic epidermal tissue cultures and an *in vivo* tumor formation assay were conducted as valid metrics for tumorigenesis (4.1)[154–156]. Surprisingly, no effect of forced SPRR5_326 expression could be observed for the invasion depth of tumorigenic keratinocytes at both assessed timepoints (Figure 4D). One possible explanation for this result could be insufficient SPRR5_326 overexpression levels. Although an initial upregulation of roughly 35-fold more SPRR5_326 was obtained, overexpression levels might have been too low to see an effect on invasion depth. Furthermore, no data is available for the SPRR5_326 overexpression at the time point of tissue harvest but long-lasting overexpression was detected during preliminary experiments, which would suggest reasonable SPRR5_326 overexpression levels over the entire experiment. In addition to that, the timepoint of tissue harvest might have been not optimal and choosing an earlier timepoint, grasping commencing invasion might reveal the hypothesized tumor protective effect of SPRR5_326. Nevertheless, one of the most likely explanations for the undetectable tumor suppressive effect of SPRR5_326, is presumably that SPRR5_326 corresponds solely to the 3’ half of the entire SPRR5 transcript, which was not known at that

time. Thus, the experiment has to be repeated with the full length SPRR5 transcript and with more timepoints in order to draw a final conclusion about its link to carcinogenesis.

The *in vivo* tumor formation assay as second method to study the effect of SPRR5_326 on tumor development showed only a slight reduction in tumor mass and volume for SPRR5_326 overexpressing tumors but inconclusive luciferase signals and drastically varying tumor sizes after explantation, which might have been caused by commencing tissue necrosis (Figure 5 and Supplementary Figure 2). As discussed above for the invasion assay, the obtained overexpression might have been to weak and possibly not high enough over the entire duration of this experiment. Furthermore, not the whole SPRR5 transcript was employed which might explain the partially mixed results. Strikingly, the formed tumors in this experiment did grow drastically slower than anticipated for tumors overexpressing oncogenic Ras and Cdk4. Whereas in the initial publication from Lazarov et al. the tumors reached volumes >500 mm^3 28 days after injection, it took over 70 days for the control tumors to reach that size during this experiment[76]. Hence, the observed slow tumor growth might also be an indication for aberrant tumor development in this particular setting, which in combination with overexpression of a partial SPRR5 version, might have masked the tumor suppressive effect of SPRR5. To finally answer whether SPRR5 controls tumor development, both experiments (generation of invasive neoplastic tissue and the *in vivo* tumor formation assay) need to be repeated with a stronger overexpression of the full length SPRR5 and more biological replicates, which was not possible in Regensburg due the lacking infrastructure and permission for animal experiments as well as for the missing requirements to overexpress oncogenic Ras variants. Unfortunately, a repetition at the Stanford University was also not an option due to a strict time limitation of the short research visit to four weeks.

To shed further light on the possible link between SPRR5 and carcinogenesis, the commonly used colony formation assay, scratch assay or a transwell invasion assay might be applied in the future[189,190]. Interestingly, preliminary results from a transwell migration assay in two SCC cell lines already indicated a reduced invasive capacity of SCC cells upon overexpression of the full-length SPRR5 transcript, encouraging the supposed tumor suppressive effect of SPRR5 (data not shown, work from Kunal Das Mahapatra, in collaboration with the Pivarcsi lab, Karolinska Institute, Sweden).

5.2 Transcript characterization for SPRR5

Unravelling the cellular localization of a lncRNA is not only part of a comprehensive transcript characterization but might also shed first light on the molecular mechanism of a given lncRNA. To this end, subcellular fractionation in combination with RT-qPCR was employed during this project, hinting towards a cytoplasmic localization of SPRR5 (4.2.1). Regarding the results from this analysis however, only preliminary conclusions should be drawn, since no satisfying enrichment of the nuclear markers was obtained from differentiated keratinocytes. Given the formation of a tightly connected cornified envelope during the course of keratinocyte differentiation, this was presumably due to insufficient separation of differentiated keratinocytes in combination with inevitable cell damage prior to fractionation. On the other hand, employing undifferentiated keratinocytes would facilitate the cell separation prior to the fractionation but would have the drawback of low SPRR5 abundance and additionally a switch in the subcellular localization of SPRR5 during the course of keratinocyte differentiation could not be excluded.

In general, it appears that subcellular fractionation is a common obstacle in the keratinocyte field, since other reports about lncRNA localization in skin utilized single molecule fish analysis for example rather than a biochemical fractionation approach[88]. Unfortunately, so far RNA-FISH for SPRR5 lead to inconclusive and partly contradictory results, despite the proven applicability of this method for differentiated keratinocytes (as in our hands other transcripts like Klf4 could be reproducibly detected in the expected cellular compartment and the employed negative controls showed no signal at all). Even numerous RNA-FISH protocol optimization steps and different signal amplification strategies were unable to finally reveal the subcellular localization of SPRR5 (work by Bianca Förstl). Therefore, the more promising single molecule FISH approach should be employed in future experiments but since this approach demands a high resolution microscope in combination with a cooled CCD-camera it could not be used in this study[191]. Additionally, both conventional RNA-FISH as well as the conducted subcellular fractionation should be further adapted in order to eventually unravel the subcellular localization of SPRR5 in differentiated keratinocytes.

SPRR5 regulates epidermal homeostasis but what controls SPRR5 expression? By answering this question, it might be possible to integrate SPRR5 into the complex regulatory network of keratinocyte differentiation and furthermore roughly infer the unacquainted promotor region for SPRR5. Hence a bioinformatic prediction of upstream transcription factors was performed, leading to p63 as the most likely regulator of SPRR5 expression

(Figure 7), which was further supported by the presence of p63 ChIP-Seq peaks in the vicinity of the human SPRR5 locus and ultimately confirmed via p63 knockdown in primary keratinocytes[44,158]. In light of the already well-established role of p63 as a master regulator of murine and human epidermal homeostasis, the observed regulation of SPRR5 by p63 once more supported the hypothesized involvement of SPRR5 during keratinocyte differentiation[39,40].

Interestingly, the area of the upstream p63 ChIP peak also exhibited a polymerase II as well as a ChIP-Seq signal for the transcriptional co-activator p300, indicating that this area might represent the promotor for SPRR5[44,192]. Furthermore, the assumed SPRR5 promotor region was investigated for additional ChIP signals of known epidermal transcription factors revealing overlapping ChIP peaks for ZNF750 and KLF4 but not MAF or MAFB which suggests a specialized role of the p63-ZNF750-KLF4 axis for SPRR5 regulation (data not shown)[35,36]. Since KLF4 is controlled by ZNF750, which in turn is a downstream target of p63, and loss of either p63, ZNF750 or KLF4 in keratinocytes results in disrupted epidermal differentiation, the occurrence of ChIP signals for all three transcription factors at the hypothesized SPRR5 promotor accentuates its inevitable necessity for epidermal homeostasis because its expression seems to be directly regulated by all three members of the p63-ZNF750-KLF4 axis individually[35,40,193,194].

With the advance in sequencing technologies over the recent years, numerous novel lncRNA transcripts have been identified and it became evident that several lncRNA isoforms can arise from one gene locus through a combination of alternative splicing, polyadenylation or promotor usage[100,119,129,195]. With SPRR5 being a novel lncRNA locus whose transcript has only been provisionally annotated as a 326 nt long lncRNA and the recent miscellaneous annotations for this gene locus in the latest Ensembl and FANTOM5 release, an exact isoform characterization for SPRR5 became indispensable[161,162]. Surprisingly, this analysis was not only able to disprove all public annotations for this gene locus but revealed that the previously anticipated SPRR5_326 was indeed part of the dominantly expressed, longer SPRR5 transcript rather than representing a discrete alternative isoform (Figure 9). But why was the longer SPRR5 RNA not detected earlier? As no annotation for this gene locus was present at the time of identification, isoform detection commenced with inspection of the initially obtained RNA-Seq reads from differentiated versus undifferentiated keratinocytes, revealing one large pile-up of reads spanning roughly 60% of the second exon of SPPR5 (≈500 nt) and the exact same 3’ end as for SPRR5[88,146]. Thus, in order to elucidate the start and end point of the transcript with single nucleotide resolution, RACE analyses with a 5’cap

dependent and a 3' poly-A based strategy were conducted. Interestingly, results from these experiments confirmed the assumed 3'end and reproducibly detected a 5'start site corresponding to the 326 nt long SPRR5_326 (work by PD Dr. Markus Kretz, Dr. Sonja Hombach and Bianca Förstl). Due to the commonly used unstranded RNA-Seq approach, integration of RACE and RNA-Seq data lead to the presumed SPRR5_326 transcript and the prediction of a second antisense transcript which would explain the presence of RNA-Seq reads across the 5'end of SPRR5_326. This antisense transcript however was never detectable by northern blots and in combination with the results from the herein conducted stranded RNA-Seq experiment (4.4.3), the presence of an antisense transcript could be eventually declined.

At the same time, this recent RNA-Seq experiment raised again the assumption that SPRR5_326 might not be the exclusive transcript from the SPRR locus but a second, longer isoform had to be present in order to account for the observed RNA-Seq reads outside of the SPRR5_326 boundaries (Figure 9). For this reason, RACE studies were repeated with several protocol iterations and finally the use of a specialized polymerase revealed the presence of two exons and an upstream 5'end for SPRR5 explaining the observed RNA read distribution. So in retrospective, previous RACE analyses supposedly were hampered by a combination of the utilized DNA polymerase and a GC-rich region directly upstream of the 5'end of SPRR5_326, which has been described to impede with proper DNA amplification[196]. In conclusion, mainly one dominant capped and polyadenylated SPRR5 transcript spanning 762 nt is expressed in human keratinocytes, which could be ultimately confirmed by a series of northern blots with probes targeting different parts of SPRR5.

Recapitulating the obtained insights into the SPRR5 transcript characterization, the dominant isoform as well as the involvement of several epidermal transcription factors for SPRR5 could be revealed. However, the hypothesized direct regulation by KLF4 and ZNF750 need to be biochemically verified and also the possible promotor of SPRR5 has to be confirmed in future luciferase experiments. Moreover, it will be fascinating to uncover how SPRR5 expression is interlaced into other processes controlling epidermal homeostasis. Given the transcriptional control of TINCR and ANCR on the epidermal transcription factors MAF/MAFB and the recently identified induction of SPRR5 upon decreased levels of the progenitor specific lncRNA LINC00941, there seems to be a highly complex and elaborate network of lncRNAs and transcription factors regulating each other and in a conjoined effort controlling epidermal tissue homeostasis (Ziegler & Graf et al., under review)[36].

5.3 SPRR5 – protein or lncRNA?

Originally defined as transcripts longer than 200 nucleotides with no protein-coding capacity, recent studies revealed that several lncRNAs associate with ribosomes and do serve as ribosomal templates for the expression of small functional peptides[102–104]. Although the extent of this lncRNA translation is under current debate, a few elaborate reports could already confirm the existence of bifunctional lncRNAs acting as both RNA as well as through the encoded small peptide[102,103,107–110]. Moreover, based on large RNA-Seq studies numerous transcripts might have been misleadingly annotated as lncRNAs but their true function is serving as a mRNA for functional protein synthesis. Thus, it remains a challenging task in the lncRNA field to discriminate whether transcripts containing a predicted open reading frame, act as a coding or non-coding transcript or as a combination of both. Interestingly, this obstacle also occurred for the herein verified SPRR5 transcript which includes an annotated open reading frame for a putative SPRR5 protein consisting of 108 amino acids and a molecular weight of 11.9 kDa. This putative protein also reasons the name of the SPRR5 transcript, since it might represent a novel family member of the human small proline-rich proteins, whose gene cluster is directly adjacent to the genomic localization of SPRR5 and strikingly the predicted SPRR5 protein shares compositional similarities with other SPRR proteins as it contains a central proline-rich domain built from a repeating amino acid motif that is flanked by a N- and C-terminal domain, harboring several glutamine and lysine residues[197]. Nevertheless, the question whether the SPRR5 protein is indeed expressed in keratinocytes remained and was first addressed by applying several bioinformatic coding potential calculations (Figure 10). Concordantly, all employed methods predicted a coding nature of SPRR5 which was even further supported by the presence of translating ribosomes at the potential open reading frame of the mouse SPRR5 homolog in a ribosomal profiling study in murine keratinocytes[167]. Moreover, the observed cytoplasmic localization of the SPRR5 transcript would be in agreement with a potential translation of SPRR5.

As already indicated above, the putative SPRR5 protein might be a novel member of the human SPRR protein family. In order to shed some light on their evolutionary relationship a phylogenetic analysis of their coding sequences was performed, revealing that SPRR5 seems to have evolved differently than other SPRRs. This finding might already explain the significantly different phenotype observed after SPRR5 depletion. Whereas no epidermal defects for SPRR knockout mice have been reported so far, loss of SPRR5 clearly disrupts proper keratinocyte differentiation as shown above (4.4)[64,65].

Besides the acquisition of this novel gene function by SPRR5, its separate evolution apart from the other SPRR proteins might have also resulted in a different mode of action, in that SPRR5 exerts its molecular function as a lncRNA rather than as a functional SPRR5 protein. Interestingly, a recent report from Hezroni et al. supports this hypothesis, since it estimates that 5% of conserved mammalian lncRNAs arose from remnants of protein-coding genes[198].

5.3.1 Small amounts of endogenous SPRR5 protein are detectable

Of course, deciphering whether SPRR5 acts as a lncRNA or as a protein implies the actual presence of detectable SPRR5 protein amounts within keratinocytes. To this end, a global mass spectrometry driven approach was chosen, which was unable to detect a single peptide from the predicted SPRR5 protein (4.3.2). However, employing a highly sensitive and targeted approach using synthetic heavy labelled SPRR5 peptides for instrument calibration prior to the analysis of protein lysate from differentiated keratinocytes revealed the presence of roughly 3,000 SPRR5 protein molecules per cell. This rather low protein level might also explain the unsuccessful attempt to detect the endogenous SPRR5 protein in immunofluorescence analysis of epidermal tissue cross sections or in western blot analysis utilizing two polyclonal SPRR5 antibodies which have been proven to detect SPRR5 peptides in a preliminary dot blot experiments[172]. Furthermore, even the detection of overexpressed and tagged SPRR5 variants turned out to be a challenging endeavour, since only weak signals were obtained for both SPRR5 antibodies as well as antibodies recognizing the utilized protein tags which might be an evidence for an unstable SPRR5 protein (antibody testing was done in collaboration with Katrin Hartinger; data not shown)[199]. Of note, during all these studies, the SPRR5 protein was consistently and exclusively detected between 20 and 25 kDa, which is in contrast to the predicted molecular weight of 11.9 kDa but a slower migration pattern during SDS-PAGE is a known phenomenon for proline-rich proteins[200,201].

Considering the well-described function of other human SPRR proteins as flexible protein-crosslinkers during the formation of the cornified envelope, detection of at least some SPRR5 peptides in the higher molecular weight fractions was anticipated[11,57]. Although high-molecular weight protein complexes (drastically crosslinked aggregations of SPRR proteins and other differentiation proteins) might have been segregated during the performed centrifugation over a 100 kDa cut off spin column or the detection of potential SPRR5 peptides from these fractions might have been hampered by the aberrant peptide

mass after massive protein-protein crosslinking, at least sporadic SPRR5 peptides should have been observed. For instance, peptides from crosslinked SPRR5 proteins could arise from initial two-protein crosslinking intermediates, where SPRR5 would have been crosslinked to another differentiation protein like loricrin or involucrin for example. Peptides from these complexes should then have been measured in the targeted mass spectrometry approach within higher molecular weight fractions since the introduction of a protein crosslink solely in the area of the detectable SPRR5 peptide seems highly unlikely. Once more, these results indicate merely the presence of uncrosslinked SPRR5 protein in differentiated keratinocytes and additionally substantiate a potential novel function of SPRR5 apart from other human SPRR proteins.

In conclusion, only relatively small amounts of uncrosslinked human SPRR5 protein were detectable in human keratinocytes and solely with an extremely sensitive mass spectrometry approach, hinting towards either low protein stability or accidental translation of SPRR5. Moreover, this observation raised serious doubts whether those low amounts of SPRR5 protein are sufficient to account for the detected effect on keratinocyte differentiation. However, future experiments should address the subcellular localization of the SPRR5 protein in differentiated keratinocytes in order to omit the possibility of higher SPRR5 protein levels that might have been masked in this semi-quantitative approach by being crosslinked to other differentiation proteins at the cellular membrane.

5.3.2 Rescue experiments for SPRR5 lead to inconclusive results

Up to now, the gold standard in deciphering the functionality of a given molecule remains in rescuing its phenotypic effect by reintroducing the molecule of interest into a depletion background[202]. Thus, siRNA-mediated SPRR5 transcript depletion, resulting in defective keratinocyte differentiation (4.4.3), should be rescued by introducing either the SPRR5 protein or the SPRR5 lncRNA (4.3.3). Surprisingly, for *in vitro* differentiated keratinocytes no differentiation defect upon SPRR5 depletion was detectable, despite the reproducible lack of differentiation after siRNA-mediated SPRR5 depletion without simultaneous lentiviral-based overexpression (4.4.1). Furthermore, no disrupted but a rather intensified differentiation marker expression was observed between SPRR5 depleted and control cells when the rescue experiment was conducted in epidermal tissue cultures and the shorter SPRR5 open reading frame constructs or the unrelated luciferase-YFP fusion protein were overexpressed. Unfortunately, also the initially assumed rescue of the SPRR5 wild-type

transcript in organotypic epidermal tissue cultures turned out to be an artefact of a faulty differentiation program in the knockdown control for SPRR5_WT rather than the desired phenotypic rescue effect. Hence, no conclusion about the functional SPRR5 molecule could be drawn from the conducted siRNA-based rescue experiments. Considering the reported induction of keratinocyte differentiation upon an increase in cellular stress levels, the publications about stress responses triggered by lentiviral infection and electroporation of keratinocytes, the siRNA-mediated knockdown in combination with the simultaneous lentiviral transduction of overexpression constructs presumably prohibited proper keratinocyte differentiation and thus foredoomed the performed rescue experiments[203–205].

In order to reduce the cellular stress levels and eventually obtain a possible rescue effect by either the SPRR5 protein or the SPRR5 RNA, a CRISPR/Cas9 mediated SPRR5 depletion strategy was employed as a substitute for the electroporation based SPRR5 knockdown. Since primary keratinocytes were unsuitable for a SPRR5 knockout cell line generation due to their limited amount of cell passages, the commonly used HaCaT keratinocyte cell line was employed and successfully validated as a suitable model system for this experiment. However, neither the transient introduction of the Cas9 endonuclease and two guideRNAs flanking each site of the desired deletion, nor the stable genomic integration which prolonged the window of opportunity for genomic editing events resulted in a homozygous SPRR5 deletion, even though the latter approach drastically increased the rate of heterozygous deletions from 3.5% to 13.4%. Given the numerous reports about efficient genomic editing using the CRISPR/Cas9 system, it was puzzling to run into these kinds of trouble at first, however several parameters were identified that might have hindered the generation of a homozygous knockout cell line. First, the rather low portion of heterozygous clones should be mentioned. Although, in one study the introduction of a 1.3 kb genomic deletion was successful for 37.5% of the analyzed alleles, another study with different cells aiming at a 119 bp excision reported only 1.6% deletion efficacy, accentuating that there might be large differences in genome editing events depending on the cell line, the desired genomic deletion as well as the targeted locus[176,206,207]. Nevertheless, the herein observed deletion rates were rather at the lower end of the reported range which might be due to an inaccessible SPRR5 locus in undifferentiated keratinocytes. This hypothesis is substantiated by reports about the presence of a highly condensed chromatin structure of the genomic SPRR5 locus in undifferentiated keratinocytes and the lack of DNAse accessible sites upstream of the desired SPRR5 deletion[51,179,181,208–210]. In addition to that, also a

CRISPR/Cas9 mediated endogenous tagging approach for SPRR5 failed, despite an efficient delivery of exogenous DNA into HaCaT cells, which once more highlights the challenging character of genome editing at this particular gene locus in keratinocytes (work from Karina Berschneider, Bianca Förstl and Sonja Hombach).

Another important factor that should not be neglected is the reported aneuploidy of HaCaT cells. Strikingly, the presence of four copies of the targeted chromosome 1 did certainly not alleviate the intended homozygous SPRR5 knockout but might have even hampered this effort by serving as a template for the homology-directed repair pathway upon successful excision of SPRR5[175,176].

Since the applied knockout strategy did not result in the desired SPRR5 depletion so far, further rounds of transfection, selection and screening of already heterozygous knockout clones might ultimately result in a homozygous SPRR5 knockout cell line. In light of the reports about off-target modifications introduced by the CRISPR/Cas9 system and the accumulating number of random mutations over the course of clonal cell expansion and cultivation, it would be highly questionable whether such a cell line would still be comparable to the initially utilized HaCaT cells after several knockout cycles[211–214]. Alternatively, the HaCaT cells could be substituted with the immortalized N/Tert keratinocyte cell line which, in contrast to HaCaT cells, only contains two copies of the targeted chromosome 1 and hence might increase the chances of a homozygous SPRR5 knockout[215]. Moreover, the recently developed enhanced Cas9 should be combined with guideRNAs targeting the DNAse accessible regions surrounding the SPRR5 locus in order to combine fair chances for a knockout cell line with a reduced number of off-target mutations[216]. Eventually, the generation of a SPRR5 knockout cell line seems inevitable, as the rescue experiments with the siRNA-mediated depletion background failed and electroporation represents the only efficient delivery method for siRNAs into keratinocytes in our hands. Furthermore, simple overexpression of either the protein or the lncRNA moiety alone did not result in any ectopic or increased differentiation marker expression in keratinocytes and thus was also unable to solve the mystery about the functional SPRR5 molecule (data not shown and Figure 21).

In conclusion, the identity of the functional SPRR5 molecule in differentiated keratinocytes could not be ultimately deciphered due to the inconclusive results from the performed rescue experiments and the unsuccessful generation of SPRR5 knockout cell lines. Nevertheless, bioinformatic predictions indicate the translation of SPRR5 and minor amounts of soluble

SPRR5 protein could be detected in differentiated keratinocytes. The drastically different effect of SPRR5 in comparison to other known SPRR proteins, the diminutive intracellular protein levels and the conducted phylogenetic analysis in combination with a report about lncRNA evolution from formerly protein-coding genes on the other raised serious doubts whether the predicted protein is indeed functional or an unstable by-product of accidental protein synthesis[109,198,217]. Based on these ambiguous findings, three conceivable scenarios for the SPRR5 transcript remain: Either, SPRR5 solely serves as a mRNA and its translation results in a functional protein, or SPRR5 exerts its regulatory role during keratinocyte differentiation exclusively as a lncRNA and the detected protein represents the outcome of pervasive translation or a combination of both applies, meaning that SPRR5 acts as a lncRNA and as a functional protein. Interestingly, several of these bifunctional lncRNAs were recently characterized, demonstrating that the strict classification into coding and non-coding transcripts is not appropriate anymore and might also be applicable for SPRR5[104,106,110,218].

5.4 Epidermal homeostasis is controlled by SPRR5

Despite the ambiguity about the functional SPRR5 molecule, it could be shown that proper execution of the terminal keratinocyte differentiation program is strictly dependent on the presence of the SPRR5 transcript, as its depletion for *in vitro* differentiated keratinocytes as well as in organotypic epidermal tissue cultures resulted in reduced expression of several key differentiation markers on both mRNA and protein levels (4.4.1 and 4.4.2). Moreover, the global RNA-sequencing approach of SPRR5 deficient epidermal tissue versus control tissue revealed that SPRR5 controls gene expression for a plethora of genes and since the top GO-terms for genes with reduced expression upon SPRR5 depletion were associated with epidermal development and homeostasis, this experiment ultimately proved the exigency of SPRR5 for keratinocyte differentiation. Strikingly, SPRR5 apparently controls the gene expression of almost the whole SPRR and LCE gene cluster over the course of the differentiation program, suggesting a *cis*-regulatory role of SPRR5 on those neighboring gene clusters (Figure 25). According to a recent review, three distinct mechanisms for *cis-acting* lncRNAs can be distinguished with respect to their regulatory component. First, gene expression can be controlled by the lncRNA itself for example by guiding or scaffolding transcription factors to nearby gene loci. Second, the process of RNA transcription or splicing confers the regulatory effect or as a third possibility the presence

of regulatory DNA elements within the gene locus might be the mode of action for the expected *cis*-regulatory function[131]. Given the observed SPRR5 knockdown phenotype (4.4), the relevance of the SPRR5 transcript itself was unambiguously verified since siRNAs should mainly affect the steady-state levels of a given RNA transcript and thus, the first possibility seems the most promising. So, assuming that SPRR5 exerts its cellular function as a *cis-acting* lncRNA, again various modes of action are conceivable like the guidance of transcription factors, assembly of an active transcription machinery as well as altering the genomic accessibility or the histone marks at the regulated gene loci. Considering that analysis of the regulated genes from the RNA-Seq experiment did not hint towards a common transcription factor and numerous publications reported a crucial role of activating and repressing histone marks in epidermal differentiation, which are controlled by the demethylases JMJD3 and UTX, histone deacetylases 1 and 2, the PCR2 complex or the BAF chromatin remodeling complex (see also 2.1.2), an epigenetic mechanism for SPRR5 seemed promising and was tested with ATAC-Seq and ChIP-Seq experiments[44–46,48–50].

5.4.1 No evidence for an epigenetic mechanism of SPRR5

As discussed above, SPRR5 might act as a lncRNA and epigenetically control gene expression from the LCE and SPRR gene cluster in *cis* and other gene loci via a *trans-acting* mechanism but on the other hand epigenetic control of keratinocyte differentiation might also be accomplished by the SPRR5 protein which was an additional reason for moving forward into this direction. To this end, the genomic accessibility was probed with ATAC-Seq in SPRR5 deficient as well as in control keratinocytes at day 4 of differentiation as well as in progenitor cells. Surprisingly, loss of SPRR5 in differentiated keratinocytes resulted in no major differences in the number of opened or closed regions although the detected downregulation of many differentiation genes would suggest a generally more condensed chromatin structure in SPRR5 depleted cells. Furthermore, the observed dynamic range for the detected alterations was rather low as not a single region was altered by a factor of two or more. One reason for these findings certainly arises from the high sample variability within the replicates. Not only the knockdown efficiencies varied between the biological replicates but also gross differences in the sequencing depth as well as a high variance for the genomic accessibility became apparent (Figure 26, Figure 27, Table 3). Thus, it was also not surprising that none of the detected altered regions was statistically significant. On the other hand, the obtained sequencing data had an overall decent quality, so possibly existing fundamental alterations in the genomic accessibility upon SPRR5

depletion should have been detected even though with no statistical significance. Another important parameter might have been the sample preparation, in which the keratinocytes were detached and dead cells were removed via an Annexin V based depletion strategy. Although this step might have improved the overall data quality, exposing the cells to suboptimal growth conditions for such a long time (≈90 min) might have already induced changes into their chromatin accessibility which in turn might have masked SPRR5 dependent alterations. To this end, future ATAC experiments should be rather performed according to the described approach by Bao et al., where the transposase reaction takes place directly in the cultivation vessel and minimizes the window of opportunity for possibly occurring belated chromatin alterations[44]. Interestingly, in one of the few publications about ATAC-Seq in keratinocytes, Bao et al. found roughly 12% of open and accessible chromatin regions to be dependent on the BAF complex, which is an unexpectedly low rate given the well-established chromatin remodeling role of BAF and the also reported conjoined action with p63, which is the master regulator of keratinocyte differentiation[33,39,40,42,44]. Moreover, the reported nucleosomal shifting rates by an average of only 20 nucleotides substantiates the impression that in general rather incremental changes in the genomic accessibility of keratinocytes appear and that these might have been masked here due to the high sample variance originating from sample preparation and different sequencing depths[44]. In conclusion, SPRR5 does not seem to control the chromatin accessibility during keratinocyte differentiation, however future experiments with more replicates and an optimized sample preparation are necessary to substantiate this hypothesis.

Although no evidence for an epigenetic mechanism of SPRR5 could be deduced from the ATAC-Seq experiments, SPRR5 might regulate the deposition of histone marks, which was tested for the activating marks H3K4me1 and H3K27ac and the repressive mark H3K27me3 (4.4.4). As anticipated, SPRR5 depletion in progenitor keratinocytes did not lead to major changes in the histone landscape which can be reasoned by the low SPRR5 expression level in progenitor cells and their dormant differentiation program. In contrast to this, for day 4 of differentiation either a decreased deposition of the activating marks H3K4me1 or H3K27ac and/or an increased signal for the repressive mark H3K27me3 would be expected in SPRR5 deficient cells, since at this differentiation stage loss of SPRR5 results in an extensive downregulation of differentiation gene expression. Strikingly, this expectation was matched for H3K4me1 and H3K27me3 but not for the activating mark H3K27ac. Given the

hypothesized expression control of SPRR5 over the LCE and the SPRR cluster, in a next step only altered histone modifications within this area were assessed, which unfortunately revealed no dramatic alterations in their histone landscape. Additionally, no regulated ($-0.5 > \log_2(\text{fold change}) > 0.5$) histone signals were obtained within two recently characterized EDC enhancer regions and only sporadic and inconclusive areas were found within super-enhancer regions identified in the murine epidermal differentiation complex, excluding the possibility that SPRR5 alters histone modifications specifically in those enhancer regions and exerts its global gene regulation within the EDC via this mechanism[54,186,187]. Strikingly, reanalysis of the altered ATAC-Seq signals for these enhancer regions lead to similar results, also neglecting a SPRR5 involvement in regulating their accessibility. Thus, an epigenetic control of keratinocyte differentiation mediated by SPRR5 regulated deposition of the tested histone modifications seemed rather unlikely, which was further substantiated by reanalyzing the data set with the HOMER ChIP-Seq toolbox, leading to different and partly contradicting results[188]. Nevertheless, in order to draw a final conclusion on that, this experiment has to be repeated since no biological replicates were employed, which makes an estimation about the biological variance as well as the detection of significantly altered regions impossible. Interestingly, a recent report claimed that transcriptional control in keratinocytes is accompanied by modest, essentially quantitative changes in histone modifications rather than an all or nothing mechanism which highlights the indispensable requirement for biological replicates in order to detect those modest alterations with statistical significance[180]. Furthermore, it should also be considered that SPRR5 might alter the distribution of histone modifications during the course of differentiation but simply not the modifications that have been tested here. Although several differentiation-associated genes are known to be suppressed in progenitor cells by the repressive histone mark H3K27me3 and are activated by demethylation upon induction of the terminal differentiation program and a more recent report revealing broad differences for H3K4me1 and H3K27ac deposition between progenitor and differentiated keratinocytes, similar reports exist for the histone modifications H4K20me1 or H3K9ac for example[45,180,219,220]. In addition to that, SPRR5 might also affect the DNA methylation state, which has been found to be a crucial regulator of maintaining the keratinocyte progenitor state as loss of the DNA methyltransferase 1 (DNMT1), and thus a decreased DNA methylation rate at certain gene loci, leads to premature keratinocyte differentiation[178].

In summary, it appears that SPRR5 does not exert its molecular function on keratinocyte differentiation by altering the chromatin accessibility or the distribution of the histone modifications H3K4me1, H3K27ac or H3K27me3, however several other epigenetic roles of SPRR5 might apply as discussed above. Furthermore, the nuclear role of SPRR5 might also occur on the level of higher-order genome organization since it was shown that the coordinated induction of EDC genes over the course of keratinocyte differentiation depend on an active reorganization of the EDC locus[51]. Recent studies in keratinocytes for instance unraveled spatial interactions between an AP-1 bound enhancer element and several EDC genes and another group identified that the formation of long-range chromatin looping by the lncRNA CCAT1-L, which induces expression of the transcription factor MYC in colorectal cancer cells[186,221]. In conclusion, these findings suggest that also SPRR5 might establish spatial contacts between distinct genomic regions within the EDC which could be tested by a chromosome conformation capture analysis in SPRR5 deficient keratinocytes.

All of the above discussed possible mechanism for SPRR5 take place in nucleus but the preliminary proposed cytoplasmic localization of the SPRR5 transcript would rather contradict a direct nuclear lncRNA mechanism. Nevertheless, a cytoplasmic regulation of another lncRNA or protein mediator that could then exert the epigenetic control in the nucleus is conceivable. On the other hand, the cytosolic localization of the SPRR5 transcript is in agreement with the observed translation of SPRR5 and the resulting SPRR5 protein might be the effector of the hypothesized nuclear function. Along this line, first experiments about the localization of the SPRR5 protein in HEK cells however, revealed a largely cytosolic localization of the overexpressed SPRR5-GFP fusion-protein and bioinformatic predictions failed to detect a nuclear localization signal (NLS) for the SPRR5 protein (overexpression done by Katrin Hartinger)[222–225]. Interestingly, a recent report showed that overexpressed SPRR4, which also lacks a predicted NLS, is predominantly localized in the cytosol of HeLa cells but has the ability to bind double-stranded DNA which is even further increased at high levels of reactive oxygen species[63,222–225]. In combination with another report from the same group about consistent nuclear detection of overexpressed pEGFP-SPRR1B in migrating keratinocytes, the authors envisioned a nuclear role of human SPRR proteins, which might also be applicable for the SPPR5 protein[60,63].

5.4.2 Conceivable cytosolic mechanisms for SPRR5 and future directions

Apart from the already discussed possible nuclear functions of SPRR5 (5.4.1), also a cytoplasmic control of gene expression is imaginable and will be discussed here under the assumption that SPRR5 acts as a lncRNA. First, SPRR5 might exert a post-transcriptional regulation of gene expression similar to the well-elaborated role of the lncRNA TINCR, in which TINCR stabilizes several differentiation marker transcripts together with the protein Staufen 1[88]. But since no specific SPRR5-Staufen 1 association could be detected in a previous protein interaction study, a different mechanism seems to apply for SPRR5 (unpublished results from the Kretz lab)[88]. Furthermore, SPRR5 might intervene into cellular signaling pathways as shown for the NF-κB Interacting LncRNA (NKILA) which would be especially intriguing since blockade of the NF-κB signaling pathway is one factor that is able to induce neoplastic tumors from epidermal keratinocytes and a functional involvement of SPRR5 in SCC development is assumed[77,226]. Finally, SPRR5 might also fine-tune the levels of miRNAs, which in turn regulate the expression of keratinocyte differentiation markers. Interestingly, several crucial miRNAs have already been reported for epidermal homeostasis and the lncRNA H19 was shown to regulate keratinocyte differentiation by sponging miR-130b-3p[85,86,227–231]. In light of the plethora of possible mechanism under the assumption that SPRR5 represents a functional lncRNA, the detection of the SPRR5 protein drastically augments the number of conceivable cellular mechanisms, which will not be discussed here.

With the in-depth SPRR5 transcript characterization and the detailed functional data for the SPRR5 transcript, future endeavors for SPRR5 should be first focused on the elucidation whether SPRR5 controls epidermal tissue homeostasis as a lncRNA, a protein or a combination of both. To this end, rescue experiments in a keratinocyte-derived SPRR5 knockout cell line would be highly favorable but also the lentiviral introduction of short-hairpin RNAs targeting SPRR5 might be applicable as a different depletion strategy. Once the functional SPRR5 molecule is clear, unravelling its subcellular localization might shed first light onto the exact mode of action, which could then be complemented by targeted interaction studies for protein or RNA interactors and in case of a nuclear localization also by assaying possible DNA interactions. Furthermore, insights from these studies might interlace SPRR5 into the intriguing regulatory network of epidermal homeostasis and possibly also hint towards its involvement in skin cancer progression or the generation of skin diseases.

6 Material

6.1 Antibodies and beads

6.1.1 Antibodies

Table 6: Primary antibodies used during this thesis

Name	Source	Dilution (Fixative)	Application	Supplier
α-β-actin	mouse, monoclonal AC-15	1:5,000	WB	Sigma-Aldrich, A1978
α-β-actin	mouse, monoclonal AC-15	1:10,000	WB	Abcam, ab6276
α-Cdk4	rabbit, polyclonal	1:1,000	WB	Santa Cruz Biotechnology, sc-749
α-collagen VII	rabbit, polyclonal	1:400 (dependent on diff. marker)	IF staining	Merck Millipore, 234192
α-collagen VII	mouse, monoclonal LH7.2	1:800 (dependent on diff. marker)	IF staining	Merck Millipore, MAB1345
α-collagen VII	mouse, monoclonal clone32	1:1,000 (methanol/acetone 1:1)	IF staining (tissue: 7.2.6)	Merck Millipore, MAB2500
α-Filaggrin	mouse, monoclonal	1:50 (ethanol)	IF staining	Santa Cruz Biotechnology, sc-66192
α-H3K4me1	rabbit, polyclonal	12.5 ng/µl (final)	ChIP-Seq	Abcam, ab8895
α-H3K27ac	rabbit, polyclonal	12.5 ng/µl (final)	ChIP-Seq	Abcam, ab4729
α-H3K27me3	rabbit, polyclonal	12.5 ng/µl (final)	ChIP-Seq	Diagenode, C15410069
α-Loricrin	rabbit, polyclonal	1:800 (acetone)	IF staining	Covance, PRB-145P
α-Loricrin	rabbit, polyclonal	1: 1,000	WB	Covance, PRB-145P
α-Keratin1	rabbit, polyclonal	1:1,000	WB	Covance, PRB-149P
α-Keratin1	rabbit, polyclonal	1:2,000 (methanol)	IF staining	Covance, PRB-149P
α-Keratin5	rabbit, polyclonal	1:750 (methanol/acetone 1:1)	IF staining (tissue: 7.2.6)	Covance, PRB-160P
α-pan Ras	mouse, monoclonal C-4	1:1,000	WB	Santa Cruz Biotechnology, sc-166691

Table 7: Utilized secondary antibodies

Name	Source	Dilution	Application	Supplier
Alexa Fluor 488 goat α-rabbit IgG	goat, polyclonal	1:300	IF staining	Thermo Fisher Scientific, A-11008
Alexa Fluor 488 goat α-mouse IgG	goat, polyclonal	1:300	IF staining	Thermo Fisher Scientific, A-11001
Alexa Fluor 555 goat α-rabbit IgG	goat, polyclonal	1:300	IF staining	Thermo Fisher Scientific, A-21428
Alexa Fluor 555 goat α-mouse IgG	goat, polyclonal	1:300	IF staining	Thermo Fisher Scientific, A-21422
IRDye 680RD goat α-mouse IgG	goat, polyclonal	1:15,000	WB	LI-COR Biosciences, 926-68070
IRDye 800CW goat α-mouse IgG	goat, polyclonal	1:15,000	WB	LI-COR Biosciences, 926-32210
IRDye 800CW goat α-rabbit IgG	goat, polyclonal	1:15,000	WB	LI-COR Biosciences, 926-32211

6.1.2 Beads

Table 8: Overview of utilized beads

Name	Application	Supplier	Catalogue number
Agencourt AMPure XP beads	ATAC-Seq, ChIP-Seq	Beckman Coulter	A63880
nProtein A Sepharose 4 Fast Flow	ChIP-Seq	Merck	GE17-5280-04
Sepharose CL-4B beads	ChIP-Seq	Merck	CL4B200

6.2 Buffers and solutions

All buffers and solutions were prepared with deionized H_2O unless stated otherwise.

Table 9: Utilized buffers and solutions

Buffer/Solution	Composition	
10% APS	10% (w/v)	ammonium persulfate in water
2x RNA loading dye	20 mM	MOPS, pH 7.0
	5 mM	sodium acetate
	1 mM	EDTA
	5.9% (v/v)	formaldehyde
	45% (v/v)	formamide
	0.01% (w/v)	bromophenol blue
	5% (v/v)	glycerol
2x SSC	300 mM	sodium chloride
	30 mM	sodium citrate
		pH 7.0

20x SSC	3 M	sodium chloride
	0.3 M	sodium citrate
		pH 7.0
4x Resolving gel buffer	1.5 M	Tris
	0.4% (w/v)	SDS
		pH 8.8
4x Stacking gel buffer	0.5 M	Tris
	0.4% (w/v)	SDS
		pH 6.8
5x laemmli buffer	300 mM	Tris/HCl, pH 6.8
	10% (w/v)	SDS
	62.5% (v/v)	glycerol
	0.1% (w/v)	bromophenol blue
	10% (v/v)	β-mercaptoethanol, added fresh before use
ATAC resuspension buffer	10 mM	Tris/HCl, pH 7.4
	10 mM	NaCl
	3 mM	$MgCl_2$
Coomassie fixative	10% (v/v)	acetic acid
	50% (v/v)	methanol
ChIP-Seq elution buffer	1% (w/v)	SDS
	100 mM	sodium hydrogen carbonate
ChIP RIPA buffer	1% (v/v)	NP-40
	0.5% (w/v)	sodium deoxycholate
	0.1% (w/v)	SDS
	1 mM	EDTA
	add before use:	in PBS, pH 7.8
	1x	protease inhibitor (Roche)
	1 mM	AEBSF
	10 mM	sodium butyrate
ChIP swelling buffer	100 mM	Tris/HCl, pH 7.6
	10 mM	Potassium acetate
	15 mM	magnesium acetate
	1% (v/v)	NP-40
	add before use:	
	1x	protease inhibitor (Roche)
	1 mM	AEBSF
	10 mM	sodium butyrate
ChIP-Seq DB	20 mM	Tris/HCl, pH 7.4
	100 mM	sodium chloride
	2 mM	EDTA
	0.5% (v/v)	Triton X-100
	add before use:	
	1 mM	AEBSF
	10 mM	sodium butyrate
	1x	protease inhibitor (Roche)
ChIP-Seq WB I	20 mM	Tris/HCl, pH 7.4
	150 mM	sodium chloride
	0.1% (w/v)	SDS
	1% (v/v)	Triton X-100
	2 mM	EDTA
ChIP-Seq WB II	20 mM	Tris/HCl, pH 7.4
	500 mM	sodium chloride
	1% (v/v)	Triton X-100
	2 mM	EDTA
ChIP-Seq WB III	10 mM	Tris/HCl, pH 7.4

	250 mM	lithium chloride
	1% (v/v)	NP-40
	1% (w/v)	sodium deoxycholate
	1 mM	EDTA
DNA loading dye	0.25% (w/v)	bromophenol blue
	30% (v/v)	glycerol
Hoechst solution	4 µg/ml	Hoechst 33342
		dissolved in PBS
hybridization solution	750 mM	sodium chloride
	75 mM	sodium citrate
	20 mM	disodium phosphate, pH 7.2
	7% (w/v)	SDS
	0.02% (w/v)	albumin fraction V
	0.02% (w/v)	Ficoll400
	0.02% (w/v)	polyvinylpyrrolidon K30
LB medium	1% (w/v)	sodium chloride
	1% (w/v)	tryptone
	0.5% (w/v)	yeast extract
		pH 7.4, autoclaved before usage
LB-Amp	0.01% (w/v)	ampicillin
		added to LB medium after sterilization
LB-(Amp)agar	1.5% (w/v)	agar
	0.01% (w/v)	ampicillin
		dissolved in LB medium
NB wash I	750 mM	sodium chloride
	75 mM	sodium citrate
	1% (w/v)	SDS
		pH 7.0
NB wash II	150 mM	sodium chloride
	15 mM	sodium citrate
	1% (w/v)	SDS
		pH 7.0
PBS	140 mM	sodium chloride
	2.7 mM	potassium chloride
	10 mM	disodium phosphate
	1.8 mM	monopotassium phosphate
		pH 7.2
protein lysis buffer	25 mM	Tris/HCl, pH 7.5
	150 mM	sodium chloride
	5% (v/v)	glycerol
	2 mM	EDTA
	0.3% (w/v)	NP-40
	1 mM	dithiothreitol, added fresh
	1x	cOmplete, EDTA-free Protease Inhibitor Cocktail (Roche), added fresh
RIPA	50 mM	Tris/HCl, pH 7.5
	150 mM	sodium chloride
	0.1% (w/v)	SDS
	1% (w/v)	sodium deoxycholate
	1% (w/v)	NP-40
	1 mM	dithiothreitol, added fresh
	1x	cOmplete, EDTA-free Protease Inhibitor Cocktail (Roche), added fresh

RNA gel buffer	20 mM	MOPS
	5 mM	sodium acetate
	1 mM	EDTA
	1.9% (v/v)	formaldehyde
		pH 7.0
siRNA annealing buffer	60 mM	HEPES
	4 mM	magnesium acetate
	200 mM	potassium acetate
TAE	8 mM	Tris
	0.2 mM	EDTA
	4 mM	acetic acid
TBS-T	10 mM	Tris
	150 mM	sodium chloride
	0.02 % (w/v)	Tween-20
TE	10 mM	Tris
	1 mM	EDTA
		pH 8.0
TGS	25 mM	Tris
	192 mM	glycine
	0.1% (w/v)	SDS
		pH 8.3
TFBI	30 mM	potassium acetate
	50 mM	manganese chloride
	100 mM	rubidium chloride
	10 mM	calcium chloride
	15% (v/v)	glycerol
		pH 5.8, filtered sterile through 0.2 µm pore filter
TFBII	10 mM	MOPS sodium salt
	75 mM	calcium chloride
	10 mM	rubidium chloride
	15% (v/v)	glycerol
		pH 7.0, filtered sterile through 0.2 µm pore filter
YT-medium	0.8% (w/v)	tryptone
	0.5% (w/v)	yeast extract
	85 mM	sodium chloride
	20 mM	magnesium sulfate
	10mM	potassium chloride
		pH 7.5, autoclaved before usage
western blot transfer buffer	25 mM	Tris
	192 mM	glycine
	20% (v/v)	methanol
		pH 8.6

6.3 Chemicals, enzymes and peptides

Unless stated otherwise, chemicals were purchased from Applichem (Darmstadt, Germany), Bio-Rad (Hercules, USA), Carl Roth (Karlsruhe, Germany), Merck (Darmstadt, Germany), Roche (Basel, Switzerland), Sigma-Aldrich (St. Louis, USA), Thermo Fisher Scientific (Waltham, USA), VWR International (Radnor, USA) and Weckert Labortechnik (Kitzingen, Germany).

Radiochemicals were purchased from Hartmann Analytics (Braunschweig, Germany), restriction enzymes, enzymes for RNA and DNA modifications (ligases, polymerases etc.) and markers were purchased from Bio-Rad (Hercules, USA), Merck (Darmstadt, Germany), New England Biolabs (Ipswich, USA), Roche (Basel, Switzerland) and Thermo Fisher Scientific (Waltham, USA).

Heavy labeled peptides for quantitative mass spectrometry were obtained from JPT Peptide Technologies (Berlin, Germany) and contained a $^{13}C^{15}N$-label on the C-terminal arginine or lysine residue.

Table 10: Sequences of heavy labeled peptides

Peptide	Sequence (N to C terminus)	AA pos. in SPRR5
SPRR5_peptide1	CQEPCAPK	88-95
SPRR5_peptide2	QPCQPPPK	80-87
SPRR5_peptide3	CPPPQQCQTSK	96-106
SPRR5_peptide4	QCAPPQQCCPPPQQR	8-22

6.4 Commercial kits

Commercial kits were used according to their included manuals unless stated otherwise.

Table 11: List of commercial kits

Name	Supplier	Catalogue number
Agilent High Sensitivity DNA Kit	Agilent Technologies (Santa Clara, USA)	5067-4626
Agilent RNA 6000 Pico Kit	Agilent Technologies (Santa Clara, USA)	5067-1513
Annexin V MicroBead Kit	Miltenyi Biotec (Bergisch Gladbach, Germany)	130-090-201
Cytoplasmic and Nuclear RNA Purification Kit	Norgen Biotek (Thorold, Canada)	21000
DNase I, RNase-free	Thermo Fisher Scientific (Waltham, USA)	EN0521
ERCC RNA Spike In Mix	Thermo Fisher Scientific (Waltham, USA)	4456740
Expand Long Template PCR System	Sigma-Aldrich (St. Louis, USA)	11681834001
FirstChoice RLM-RACE Kit	Thermo Fisher Scientific (Waltham, USA)	AM1700
High Sensitivity D1000 ScreenTape and reagents	Agilent Technologies (Santa Clara, USA)	5067- 5584; 5067- 5585

Human Keratinocyte Nucleofector Kit	Lonza (Basel, Switzerland)	VVPD-1002
illustra MicroSpin G-25 Columns	GE Healthcare (Chalfont St Giles, Great Britain)	27532501
iScript cDNA Synthesis Kit	Bio-Rad (Hercules, USA)	170-8890
Lipofectamine 3000	Thermo Fisher Scientific (Waltham, USA)	L3000015
Monarch PCR & DNA Purification Kit	New England Biolabs (Ipswich, USA)	T1030S
NEBNext Multiplex Oligos for Illumina	New England Biolabs (Ipswich, USA)	E7335 (Set1) E7500 (Set 2)
NEBNext Ultra II DNA Library Prep Kit for Illumina	New England Biolabs (Ipswich, USA)	E7645
Nextera DNA Library Preparation Kit	Illumina (San Diego, USA)	FC-121-1030
Nextera Index Kit	Illumina (San Diego, USA)	FC-121-1011
NucleoBond Xtra Maxi EF	Macherey-Nagel (Dueren, Germany)	740424
NucleoBond Xtra Midi Kit	Macherey-Nagel (Dueren, Germany)	740410
NucleoSpin Gel and PCR Clean-up Kit	Macherey-Nagel (Dueren, Germany)	740609
NucleoSpin Plasmid (NoLid)	Macherey-Nagel (Dueren, Germany)	740499
pGEM-T Easy Vector Kit	Promega (Madison, USA)	A1360
Pierce BCA Protein Assay Kit	Thermo Fisher Scientific (Waltham, USA)	23225
PrecisionX Multiplex gRNA Cloning Kit	BioCat (Heidelberg, Germany)	CAS9-GRNA-KIT-SBI
QIAShredder	Qiagen (Hilden, Germany)	79654
QIAquick PCR Purification Kit	Qiagen (Hilden, Germany)	28106
Qubit dsDNA HS Assay Kit	Thermo Fisher Scientific (Waltham, USA)	Q32851
RNeasy Plus Mini Kit	Qiagen (Hilden, Germany)	74136
Roti-Quant	Carl Roth (Karlsruhe, Germany)	K015.1
SsoFast EvaGreen	Bio-Rad (Hercules, USA)	1725200
Takyon No ROX SYBR 2x MasterMix blue dTTP	Eurogentec (Luettich, Belgium)	UF-NSMT-B0701
TURBO DNA-free Kit	Thermo Fisher Scientific (Waltham, USA)	AM1907
TruSeq Stranded mRNA Library Prep	Illumina (San Diego, USA)	20020594

6.5 Consumables, membranes and screens

Consumables were, unless stated otherwise, purchased from Bio-Rad (Hercules, USA), Carl Roth (Karlsruhe, Germany), Eppendorf (Hamburg, Germany), Eurogentec (Luettich, Belgium), GE Healthcare (Chalfont St Giles, Great Britain), MP Biomedicals (Heidelberg, Germany), NeoLab (Heidelberg, Germany), Sarstedt (Nuembrecht, Germany), Thermo Fisher Scientific (Waltham, USA) or VWR International (Radnor, USA).

Table 12: List of membranes and screens

Name	Supplier	Catalogue number
Amersham Hybond-ECL	GE Healthcare (Chalfont St Giles, Great Britain)	RPN132D
Amersham Hybond-N+	GE Healthcare (Chalfont St Giles, Great Britain)	RPN203B
Storage Phosphor Screen GP	Kodak (Rochester, NY, USA)	1707843

6.6 Eukaryotic cell cultivation

Dermis for regenerated organotypic epidermal tissue cultures was prepared from frozen split-skin obtained from Biopredic International (Saint-Grégoire, France), Tissue Solutions (Glasgow, Great Britain) or was isolated from surgical specimens in the Khavari laboratory at the Stanford University School of Medicine (Stanford, USA). Utilized primary cells, cell lines, media components and their composition are listed below.

Table 13: Overview of primary eukaryotic cells and cell lines

Name	Details	Supplier
HaCaT cell line	Catalogue number 300493	Cell Lines Service (Eppelheim, Germany)
HEK293T cells	gift from the AG Meister	AG Meister (Regensburg, Germany)
Normal Human Epidermal Keratinocytes (NHEK), adult single donor	isolated from fresh surgical specimens at the University Hospital Regensburg	AG Kretz
Normal Human Epidermal Keratinocytes (NHEK), juvenile foreskin, pooled	Lot numbers: 1020401, 1040101 and 407Z001	PromoCell (Heidelberg, Germany)
Primary neonatal human keratinocytes	isolated from fresh surgical specimens at the Stanford University	Khavari laboratory (Stanford, USA)
Primary human fibroblasts	isolated from fresh surgical specimens at the Stanford University	Khavari laboratory (Stanford, USA)

Table 14: Reagents for eukaryotic cell cultures

Name	Supplier	Catalogue number
3,3',5'-Triiodo-L-thryonine	Sigma-Aldrich (St. Louis, USA)	T0281
Adenine hydrochloride hydrate	Sigma-Aldrich (St. Louis, USA)	A-9795
100x Antibiotic-Antimycotic	Thermo Fisher Scientific (Waltham, USA)	15240-096
Cholera toxin from Vibrio cholerae	Sigma-Aldrich (St. Louis, USA)	C8052
Corning Matrigel	Thermo Fisher Scientific (Waltham, USA)	11543550
DMSO	Carl Roth (Karlsruhe, Germany)	A994.1

Dulbecco's Modified Eagle Medium, high glucose, pyruvate (DMEM)	Thermo Fisher Scientific (Waltham, USA)	41966-029
Dulbecco's Phosphate-Buffered Saline, no calcium, no magnesium (DPBS)	Thermo Fisher Scientific (Waltham, USA)	14190-094
Epidermal Growth Factor human	Sigma-Aldrich (St. Louis, USA)	E9644
Fetal Bovine Serum	Thermo Fisher Scientific (Waltham, USA)	10270-106
Ham's F12	Lonza (Basel, Switzerland)	BE12-615F
Holo-Transferrin human	Sigma-Aldrich (St. Louis, USA)	T0665
Human Keratinocyte Growth Supplement	Thermo Fisher Scientific (Waltham, USA)	S-001-5
HyClone Bovine calf serum	Thermo Fisher Scientific (Waltham, USA)	SH3007303
HyClone Characterized Fetal Bovine serum	Thermo Fisher Scientific (Waltham, USA)	SH3007103
Hydrocortisone	Sigma-Aldrich (St. Louis, USA)	H0396
Insulin solution human	Sigma-Aldrich (St. Louis, USA)	I9278
Keratinocyte-SFM Serum Free Medium	Thermo Fisher Scientific (Waltham, USA)	17005-042
Medium 154	Thermo Fisher Scientific (Waltham, USA)	M-154-500
Opti-MEM Reduced Serum Medium	Thermo Fisher Scientific (Waltham, USA)	31985070
Penicillin-Streptomycin (10000 U/ml)	Thermo Fisher Scientific (Waltham, USA)	15140-122
Puromycin dihydrochloride	Carl Roth (Karlsruhe, Germany)	0240.3
Polybrene	Sigma-Aldrich (St. Louis, USA)	10768910G
Supplements for Keratinocyte-SFM	Thermo Fisher Scientific (Waltham, USA)	37000-015
Trypsin-EDTA 0.05%, phenol red	Thermo Fisher Scientific (Waltham, USA)	25000-054
Trypsin-EDTA 0.25%, phenol red	Thermo Fisher Scientific (Waltham, USA)	25200-056

Table 15: Components and composition of cell culture medium

Medium/Solution	Composition	
50:50 medium	500 ml	Keratinocyte SFM
	500 ml	Medium 154
	5 ml	Human Keratinocyte Growth Supplement
	10 ml	100x Antibiotic-Antimycotic
	1x	Supplements for Keratinocyte-SFM
Adenine stock solution	12 mg	Adenine hydrochloride hydrate
	6.75 ml	DMEM
		adjust pH to 7.5
Basic buffer	0.2 M	sodium chloride
	3 mM	sodium azid
	1 mM	EDTA
		adjust pH to 8.0

Cholera toxin solution	1 mg 1 ml 99 ml	Cholera toxin from *Vibrio cholera* Basic buffer DMEM
Conditioned 50:50 medium		50:50 medium which was used for cultivation of HaCaT cells for two days and subsequently filtered through a 0.2 µm pore filter
DMEM+BCS	500 ml 50 ml 5 ml	DMEM BCS Penicillin/Streptomycin
DMEM+FBS	500 ml 50 ml 5 ml	DMEM FBS (Thermo) Penicillin/Streptomycin
EGF stock solution	100 µg 10 ml	EGF ddH_2O
Hydrocortisone stock solution	5 mg 1 ml 24 ml	Hydrocortisone Ethanol DMEM
Insulin stock solution	5 mg 1 ml	Insulin solution human ddH_2O
KGM	330 ml 110 ml 5 ml 5 ml 50 ml 1 ml 0.5 ml 1 ml 0.5 ml 0.5 ml 0.5 ml	DMEM Ham's F12 Penicillin-Streptomycin 100x Antibiotic-Antimycotic FBS (HyClone) Adenine stock solution Cholera toxin solution Hydrocortisone stock solution T/T3 solution EGF stock solution Insulin stock solution
PBS+2x A/A	10 ml 500 ml	100x Antibiotic-Antimycotic DPBS
Polybrene solution	1 mg/ml	Polybrene in PBS
T/T3 solution	9.9 ml 100 µl	Transferrin stock solution Triiodo-L-thyronine stock solution
Transferrin stock solution	50 mg 10 ml	Holo-Transferrin human DPBS
Triiodo-L-thyronine stock solution	13.6 mg 100 ml	3,3',5'-Triiodo-L-thyronine ddH_2O

6.7 Instruments

General laboratory instruments and devices were purchased from Beckman Coulter (Brea, USA), Bio-Rad (Hercules, USA), Eppendorf (Hamburg, Germany), NeoLab (Heidelberg, Germany) and Thermo Fisher Scientific (Waltham, USA). Particular instruments are listed in the table below.

Table 16: List of instruments

Name	Supplier

2200 TapeStation System	Agilent Technologies (Santa Clara, USA)
Agilent 2100 Bioanalyzer	Agilent Technologies (Santa Clara, USA)
Amersham Ultrospec 3300 pro	GE Healthcare (Chalfont St Giles, Great Britain)
Centrifuge 5424 R	Eppendorf (Hamburg, Germany)
Centrifuge 5810	Eppendorf (Hamburg, Germany)
CO_2-Incubator HERAcell 240i	Thermo Fisher Scientific (Waltham, USA)
Cryostat Microm HM 500 OM	Thermo Fisher Scientific (Waltham, USA)
Electroporation Device Nucleofector II	Lonza (Basel, Switzerland)
FastPrep-24 Instrument	MP Biomedicals (Heidelberg, Germany)
Heraeus Megafuge 40R	Thermo Fisher Scientific (Waltham, USA)
Heraeus Multifuge 1S	Thermo Fisher Scientific (Waltham, USA)
HeraSafe KS	Thermo Fisher Scientific (Waltham, USA)
HiSeq 1000	Illumina (San Diego, USA)
Hybridization oven type T 5042	Heraeus (Hanau, Germany)
IKA MS3	Agilent Technologies (Santa Clara, USA)
Incubator Model B6200	Heraeus (Hanau, Germany)
Inverted microscope Axiovert 200 M	Carl Zeiss (Oberkochen, Germany)
Inverted microscope Diavert	Leitz (Wetzlar, Germany)
IVIS 100 In Vivo Imaging System	PerkinElmer (Waltham, USA)
Leica CM3050 S Cryostat	Leica Biosystems (Nussloch, Germany)
MaXis plus UHR-QTOF	Bruker (Billerica, USA)
MilliQ Q-Pod	Merck (Darmstadt, Germany)
Mx3000P	Agilent Technologies (Santa Clara, USA)
NanoDrop 1000	Thermo Fisher Scientific (Waltham, USA)
Nanophotometer Classic	Implen (Munich, Germany)
New Brunswick Innova 44 Shaker	Eppendorf (Hamburg, Germany)
Odyssey Imaging System	LI-COR Biosciences (Lincoln, USA)
PM1002 mobile anesthesia machine	Parkland Scientific (Coral Springs, USA)
PMI Personal Molecular Imager FX	Bio-Rad (Hercules, USA)
QTRAP 4500	SCIEX (Framingham, USA)
Qubit 2.0 Fluorometer	Thermo Fisher Scientific (Waltham, USA)
Real-Time PCR Cycler CFX96	Bio-Rad (Hercules, USA)
S220 Focused-ultrasonicator	Covaris (Woburn, USA)
Screen Eraser-K	Bio-Rad (Hercules, USA)
Shake 'n' Stack Hybridization Oven	Thermo Fisher Scientific (Waltham, USA)
Thermomixer comfort	Eppendorf (Hamburg, Germany)
Trans-Blot SD Semi-dry transfer cell	Bio-Rad (Hercules, USA)
Transilluminator Quantum ST4	PEQLAB (Erlangen, Germany)
UltiMate 3000 RSLCnano System with Acclaim PepMap100 C18 Nano-Trap column and Acclaim PepMap100 C18 column	Thermo Fisher Scientific (Waltham, USA)
UV Stratalinker 2400	Stratagene (La Jolla, USA)

6.8 Mouse strain

Immunodeficient mice (CBySmn.CB17-*Prkdc*^{scid}/J) were purchased from The Jackson Laboratory (Bar Harbor, USA).

6.9 Oligonucleotides

DNA oligonucleotides were ordered from Metabion (Martinsried, Germany) or Sigma-Aldrich (St. Louis, USA), siRNA Pools (mixture of 11-30 different siRNAs per target, exact sequences are available from siTools upon request) were designed and ordered from siTools (Munich, Germany) and single siRNAs were obtained from biomers.net (Ulm, Germany).

Table 17: Overview of utilized siRNAs

siRNA	Sequence (5' to 3')
Control_sense	GUAGAUUCAUAUUGUAAGG
Control_antisense	CCUUACAAUAUGAAUCUAC
pan_p63i_sense	CGACAGUCUUGUACAAUUU
pan_p63i_antisense	AAAUUGUACAAGACUGUCG
siCtrl	siPool
siSPRR5	siPool
siSPRR5_326	siPool

Table 18: List of sequencing primers

Name	Sequence (5' to 3')
5'Cas9_seq_R	CCGATGCTGTACTTCTTGTC
CMV_F	CGCAAATGGGCGGTAGGCGTG
KO_out_F2	CCAACTCTAAGAGAGGTAAGTATG
KO_out_R2	GTATGCAGTGTTTGCATAGACTGTC
lentiCRISPRv2_seq_F	GGACAGCAGAGATCCAGTTT
lentiCRISPRv2_seq_R	AGCCAATTCCCACTCCTTTC
pLARTA_F5	AGAATCGCAAAACCAGCAAG
PGK_F	TGTTCCGCATTCTGCAAGCC
pLARTA_Ins_F	CGAATCACCGACCTCTCTCC
pLARTA_R1	AAACCGTCTATCAGGGCGAT
pX459_F	TTACGGTTCCTGGCCTTTTG
pX459_R	TGTCTGCAGAATTGGCGCA
SP6	ATTTAGGTGACACTATAG
T7	TAATACGACTCACTATAGGG
WPRE-rev	CATAGCGTAAAAGGAGCAACA
WT_In_F2	CCCCTCAGCTCCTGTTTTCA

Table 19: List of primer sequences used for qRT-PCR

Name	Forward primer (5' to 3')	Reverse primer (5' to 3')
7SK	CCTGCTAGAACCTCCAAACAAG	GCCTCATTTGGATGTGTCTG
ALOX12B	AGACTGCAATTCCGGATCAC	TGTGGAATGCACTGGAGAAG
β-actin	GGACTTCGAGCAAGAGATGG	AGGAAGGAAGGCTGGAAGAG
CALB1	TGGCTTTGTCGGATGGAGGG	GGTTGCGGCCACCAACTCTA
C1orf68	TTCTGGCCCCCTCTCTGTTA	GGGACTGTACTAACTCTGGC
ELOVL3	TTCGAGGAGTATTGGGCAAC	GAAGATTGCAAGGCAGAAGG
FLG	AAAGAGCTGAAGGAACTTCTGG	AACCATATCTGGGTCATCTGG
GAPDH	GAAGAGAGAGACCCTCACTGCTG	ACTGTGAGGAGGGGAGATTCAGT
KRT1	TGAGCTGAATCGTGTGATCC	CCAGGTCATTCAGCTTGTTC

KRT10	GCAAATTGAGAGCCTGACTG	CAGTGGACACATTTCGAAGG
L32	AGGCATTGACAACAGGGTTC	GTTGCACATCAGCAGCACTT
LacZ	GTGCGGATTGAAAATGGTCT	GACCTGACCATGCAGAGGAT
LCE1A	GAAGCGGACTCTGCACCTAGAA	AGGAGACAGGTGAGGAGGAAATG
LCE1E	TGAAGTGGACCTTGACTTCCTC	CTCCAGGCAAGACTTCAAGC
LCE2A	TGGAGAAACTTGCAACCAGGA	CCTCACAAGGTGTGTCAGCC
LCE2D	GGACGTGTCTGTGCTTTTGC	CTTGGGAGGACATTTGGGAGG
LCE3A	TGTCTGCCTCCAGCTTCCT	AGTTGGAGCTCTGGCAACG
LCE3D	TCTTGATGCATGAGTTCCCAGA	TGGACATCAGACAGGAAGTGC
LCE4A	CCCCCTCCCAAGTGTCCTAT	GAGCCACAGCAGGAAGAGAT
LCE5A	CCCAGGTGCTGAAGATGTGT	ATGGAGTGAACATGGGCAGG
LCE6A	GTCCTGATCTCTCCTCTCGTCT	CAAGATTGCTGCTTCTGCTGT
Lenti_SPRR5	CCCTCAGCTCCTGTTTTCAA	GCTGCCTTGTAAGTCATTGG
LINC01527_1	GCCTCTCCTGCAAGTGTGA	TCCTCATTTATGACATTTTCAGTCTC
LINC01527_2	ACATGCCAGTGAAGTCTGTTCA	AGTCATGTGAGGCAGTTCCA
LINC01527_3	TCCCCCTACCTCTCATAGGC	AGATGAACAGACTTCACTGGCA
LOR	CTCTGTCTGCGGCTACTCTG	CACGAGGTCTGAGTGACCTG
Pan-p63	GACAGGAAGGCGGATGAAGATAG	TGTTTCTGAAGTAAGTGCTGGTGC
pLEX_SPRR5	CACGCTGTTTTGACCTCCATAG	CACTGCTTCTGCTTCTGCTGAG
Puro	CACCAGGGCAAGGGTCTG	GCTCGTAGAAGGGGAGGTTG
SPRR1A	CAGCCCATTCTGCTCCGTAT	GGCTGGCAAGGTTGTTTCAC
SPRR2A	ACACAGGGAGCTTCTTTCTCC	CCAGGACTTCCTTTGCTCAGT
SPRR2D	TCGTTCCACAGCTCCACTTG	CAGGCCACAGGTTAAGGAGA
SPRR3	CCTCGACCTTCTCTGCACAG	GGTTGTTTCACCTGCTGCTG
SPRR4	AGCCTCCAAGAGCAAACAGA	GCAGGAGGAGATGTGAGAGG
SPRR5	AGCAGCTGCAGTTTCCATCT	AAACAGGAGCTGAGGGGAAG
U6	CACATATACTAAAATTGGAACG	CTTCACGAATTTGCGTGTCATC

Table 20: List of primers used for molecular cloning and PCR

Name	Sequence (5’ to 3’)	Purpose
3’inner	TCTCAGCAGAAGCAGAAGCAGT	RACE verification of SPRR5
3’outer	CAAGTCCAGCCCAGAATGTCTC	RACE verification of SPRR5
5’inner	GAACCAGATCCTTGTCTGAGAT	RACE verification of SPRR5
5’outer	CTGCCAGAGGAATTCTGTTTTAATTG	RACE verification of SPRR5
AfeI_CMV_R	GATGATAGCGCTAGCTCTGCTTATATAG ACCT	Cloning CMV promotor and enhancer into lentiCRISPRv2
KO_out_F1	CTGCATGTGTGTGTCGGTAC	Screening for SPRR5 KO cells
KO_out_F2	CCAACTCTAAGAGAGGTAAGTATG	Screening for SPRR5 KO cells
KO_out_R1	TTACTGTCGGGTGTGTGTGT	Screening for SPRR5 KO cells
KO_out_R2	GTATGCAGTGTTTGCATAGACTGTC	Screening for SPRR5 KO cells
NheI_CMV_F	GATGATGCTAGCGACATTGATTATTGAC TAGT	Cloning CMV promotor and enhancer into lentiCRISPRv2
SPRR5_KO_1	AAAGGACGAAACACCGGATCCCTTCACA GATTTAAGGTTTTAGAGCTAGAAATAGC AAG	Generation of a gRNA cassette for SPRR5 KO with Multiplex gRNA Cloning Kit, Block1
SPRR5_KO_2	GCCAAAAAGACCTGCGCAGTGGATCCAA GGTGTCTCATAC	Generation of a gRNA cassette for SPRR5 KO with Multiplex gRNA Cloning Kit, Block1
SPRR5_KO_3	GCAGGTCTTTTTGGCGTTTTAGAGCTAG AAATAGCAAG	Generation of a gRNA cassette for SPRR5 KO with Multiplex gRNA Cloning Kit, Block2

SPRR5_KO_4	AGAGTGTGCTCTCTGAGTTCGGTGTTTCG TCCTTTCCAC	Generation of a gRNA cassette for SPRR5 KO with Multiplex gRNA Cloning Kit, Block2
SPRR5_KO_5	AGAGAGCACACTCTAGTTTTAGAGCTAG AAATAGCAAG	Generation of a gRNA cassette for SPRR5 KO with Multiplex gRNA Cloning Kit, Block3
SPRR5_KO_6	TTCTAGCTCTAAAACGGCAATTTATTCCA GATGCCGGATCCAAGGTGTCTCATAC	Generation of a gRNA cassette for SPRR5 KO with Multiplex gRNA Cloning Kit, Block3
WT_In_F1	CGTCCAAGCAGAAGTAAGCC	Screening for SPRR5 KO cells
WT_In_F2	CCCCTCAGCTCCTGTTTTCA	Screening for SPRR5 KO cells

Table 21: DNA probes for Northern Blot Analysis

Name	Sequence (5' to 3')
GAPDH #1	CATGGACTGTGGTCATGAGTCCTTCCACGATACCAAAGTT
GAPDH #2	GAATTTGCCATGGGTGGAATCATATTGGAACATGTAAACCATG TAGTTGAGG
Out_Exon1_SPRR5	GATGCTCTGAGCCTATCTACCTTTTATATACCTCTCATCCCTGC CTGAGG
Exon1_SPRR5	CAAGGTTCACCAGCGACTGGAGCAGATAGGTGTGGAGGGGTT T
Intron_SPRR5	CCCACACTACTCTGCTGGATAACTTGCTGGACCCAAGATAGGT TTGTCAC
Exon2_SPRR5_Probe1	CACACTGCTTCTGCTTCTGCTGAGACATTCTGGGCTGGACTTG CAACTGG
Exon2_SPRR5_Probe2	CAAGGCTGCTTGGTCTGTTGGGGTGGGGGGCAGTAC
Exon2_SPRR5_Probe3	GATCACATGCCCAGGGCTTACTTCTGCTTGGACGTCTGGCAC
Exon2_SPRR5_ProbePool_1	TGTGGAGTGTAAGATGCTGG
Exon2_SPRR5_ProbePool_2	AGAGTAGAGCTCGAGGAAGC
Exon2_SPRR5_ProbePool_3	CAGGGCAATGTAGATGCATG
Exon2_SPRR5_ProbePool_4	GCCAGAGGAATTCTGTTTTA
Exon2_SPRR5_ProbePool_5	TGCTGGCCCAGAAGGAACTT
Exon2_SPRR5_ProbePool_6	TGGTGAGCAGGGCTTTGCTT
Exon2_SPRR5_ProbePool_7	TGTGGCTGCATCCAGAAGCA
Exon2_SPRR5_ProbePool_8	CCTCTCCTGCAAGTGTGAAG
Exon2_SPRR5_ProbePool_9	ACTGCAGCTGCTAATTAAGA
Exon2_SPRR5_ProbePool_10	GATCCTTGTCTGAGATGGAA
Exon2_SPRR5_ProbePool_11	CATGCCGCTCATGTTTCCAG
Exon2_SPRR5_ProbePool_12	GGTAGGAGAAGATGCCTGTG
LINC01527_1	CACTTCACACTTGCAGGAGAGGCAGCACCAGG
LINC01527_2	GCTGGCGGCTTCCTCGAGCTCTACTCTTCTTA
LINC01527_3	GCAGCTGCAGTTTCCATCTCAGACAAGGATCTGG
LINC01527_4	CATCTTTCCTGACATGCATCTACATTGCCCTG
LINC01527_5	GGAAACATGAGCGGCATGCACAGGCATC

Table 22: GRNA sequences for SPRR5 KO cell generation

Name	Sequence (5' to 3')
SPRR5_down_1	ACTGCGCAGGTCTTTTTGGC
SPRR5_down_2	AACTCAGAGAGCACACTCTA
SPRR5_up_1	GATCCCTTCACAGATTTAAG
SPRR5_up_2	GGCATCTGGAATAAATTGCC

6.10 Plasmids

Table 23: List of plasmids

Plasmid	Properties	Origin
lentiCRISPRv2_complete	*Lentiviral transfer plasmid based on lentiCRISPRv2 from Addgene (#52961). The SPRR5 KO cassette has been inserted via SLIC cloning after BsmBI digest of the vector, encodes Cas9 from S. pyogenes*	*AG Kretz/Julia Junghans (supervised master student) (University of Regensburg)*
lentiCRISPRv2_complete_ CMVCas9	*Lentiviral transfer plasmid based on lentiCRISPRv2_complete, with Cas9 under the control of a CMV promotor*	*this work*
pCMV dR8.91	*Packaging plasmid for generation of lentiviral particles, encodes the viral gag, rev, tat and pol proteins under control of the CMV promotor*	*gift from the Khavari laboratory (Stanford University, USA)*
pGEM-T Easy	*vector for TA cloning of PCR products*	*Promega (Madison, USA)*
pLARTA_326	*Lentiviral transfer plasmid, contains the SPRR5_326 sequence under the control of the hPGK promotor, flanked by the LTR sequences for stable genomic integration*	*AG Kretz/Bianca Förstl (University of Regensburg)*
pLARTA_lacZ	*Lentiviral transfer plasmid, contains the lacZ sequence under the control of the hPGK promotor, flanked by the LTR sequences for stable genomic integration*	*AG Kretz/Johannes Graf (University of Regensburg)*
pLARTA_SPRR5	*Lentiviral transfer plasmid, contains the SPRR5 sequence under the control of the hPGK promotor, flanked by the LTR sequences for stable genomic integration*	*AG Kretz/Bianca Förstl (University of Regensburg)*
pLARTA_SPRR5_MutDB	*Lentiviral transfer plasmid, contains the SPRR5 sequence with mutated start codons under the control of the hPGK promotor, flanked by the LTR sequences for stable genomic integration*	*AG Kretz/Bianca Förstl (University of Regensburg)*
pLEX_Cdk4	*Based on the pLEX-MCS vector backbone (Thermo Scientific Open Biosystems, #OHS4735), lentiviral transfer plasmid for overexpression of human Cdk4*	*gift from the Khavari laboratory (Stanford University, USA)*
pLEX_HRasG12V	*Based on the pLEX-MCS vector backbone (Thermo Scientific Open Biosystems, #OHS4735), lentiviral transfer plasmid for overexpression of constitutively active human HRasG12V protein*	*gift from the Khavari laboratory (Stanford University, USA)*
pLEX_LY	*Based on the pLEX-MCS vector backbone (Thermo Scientific Open Biosystems, #OHS4735), lentiviral transfer plasmid for overexpression of a fusion protein of luciferase and YFP protein*	*gift from the Khavari laboratory (Stanford University, USA)*
pLEX_SPRR5_ORF	*Based on the pLEX-MCS vector backbone (Thermo Scientific Open Biosystems, #OHS4735), lentiviral transfer plasmid for overexpression of SPRR5 protein*	*AG Kretz/Bianca Förstl (University of Regensburg)*
pLEX_SPRR5_MutORF	*Based on the pLEX-MCS vector backbone (Thermo Scientific Open Biosystems, #OHS4735), lentiviral transfer plasmid for overexpression of the SPRR5 CDS with a mutated start codon*	*AG Kretz/Bianca Förstl (University of Regensburg)*

pUC-MDG	*Envelope plasmid for generation of lentiviral particles, encodes the viral VSV-G envelope protein under control of the CMV promotor*	*gift from the Khavari laboratory (Stanford University, USA)*
pX459_complete	*Plasmid for CRISPR/Cas9 mediated knockout of SPRR5, based on pX459 (Addgene, #62988) with the insertion of a SPRR5 KO cassette via SLIC cloning after BbsI digest of the vector backbone*	*AG Kretz/Julia Junghans (supervised master student) (University of Regensburg)*

6.11 Prokaryotic cells

Table 24: Overview of utilized *Escherichia coli* strains during this work

Strain	Genotype	Details
DH5α	F^- Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ^-	Propagation of plasmids
Stbl3	F^-*mcr*B *mrrhsd*S20(r_B^-, m_B^-) *rec*A13 *sup*E44 *ara*-14 *gal*K2 *lac*Y1 *pro*A2 *rps*L20(Str^R) *xyl*-5 λ^-*leumtl*-1	Propagation of lentiviral plasmids

6.12 Software

Table 25: Software used during this work

Name	Source of supply
2200 TapeStation Software	Agilent Technologies (Santa Clara, USA)
Agilent 2100 Expert Software B.02.08SI648	Agilent Technologies (Santa Clara, USA)
AxioVision 4.9.1.0	Carl Zeiss (Oberkochen, Germany)
Bio-Rad CFX Manager 3.1	Bio-Rad (Hercules, USA)
CASAVA1.8.2	https://biogist.wordpress.com/2012/10/23/casava-1-8-2-installation/
Galaxy server tools (indicated individually)	local installation after https://usegalaxy.org/
gRNA design tool	http://crispr.mit.edu/
HOMER (v4.9, 2-20-2017)	http://homer.ucsd.edu/homer/
Integrative Genomics Viewer 2.3.90 (IGV)	https://www.broadinstitute.org/igv/
IrfanView 4.42	http://www.irfanview.de/
Living Image 4.5.2	PerkinElmer (Waltham, USA)
MASCOT 2.5.1	Matrix Science (London, United Kingdom)
Microsoft Office	Microsoft (Redmond, USA)
MxPro QPCR Software	Agilent Technologies (Santa Clara, USA)
ND-1000 3.81	Thermo Fisher Scientific (Waltham, USA)
Odyssey 3.0.30	LI-COR Biosciences (Lincoln, USA)
ProteinScape4 3.1.3 461	Bruker (Billerica, USA)
Quantity One 4.6.9	Bio-Rad (Hercules, USA)
R version 3.3.1	https://www.r-project.org/
Rstudio version 1.0.136	https://www.rstudio.com/
R-packages (several, indicated individually)	Bioconductor
SnapGene Viewer 4.1.4	GSL Biotech LLC (Chicago, USA)
SonoLab Software 7.2	Covaris (Woburn, USA)
Skyline 4.1	University of Washington (McCoss Lab) https://skyline.ms/project/home/software/Skyline/begin.view
Zotero	https://www.zotero.org/

7 Methods

7.1 Bioinformatical data analysis

7.1.1 Analysis of full transcriptome sequencing data

Differential gene expression analysis for RNA sequencing (7.6.7) was done by Dr. Nicholas Strieder.

Preprocessing: Quality of sequencing data from the RNA-Seq libraries was examined using FASTQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and size, adapter and quality trimming was performed using Trimmomatic[232] (ver. 0.32, ILLUMINACLIP:TruSeq_s tranded_SE.fa:2:30:10 HEADCROP:0 TRAILING:26 LEADING:26 SLIDINGWINDOW:4:15 MINLEN:25). Trimmed reads were aligned to the Homo sapiens genome (ftp://ftp.ensembl.org/pub/release-85/fasta/homo_sapiens/dna/ Homo_sapiens.GRCh38.dna.primary_assembly.fa.gz) extended with the ERCC spike in sequence information[233] using STAR[234]. The mapped reads were then assessed on the gene level using featureCounts from the Rsubread R-library[235] based on the annotation information from Ensembl[161] (GRCh38.p7 release-85). Following alignment, further quality control was performed using QoRTs[236] and aligned reads were inspected with the UCSC genome browser[208] or the Integrative Genomics Viewer (IGV)[237,238].

Differential expression analysis: Count data at the gene level was analyzed with DESeq2[239] using ERCC spike ins for library size normalization and all comparisons were corrected for multiple testing using FDR[240]. Genes that met the indicated $\log_2$(fold change) restraint and had a false discovery rate < 0.05 ("padj") were considered significantly differential expressed.

Heatmap generation and functional annotation: Heatmaps for genes with altered expression (adj. p-value < 0.05 and -0.5 > log2FC > 0.5) were generated with pheatmap[235] after variance stabilization transformation of count data with DeSeq2. Functional annotation clustering was performed using the David 6.8 database[241,242] as well as the Enrichr tool[243,244] for the GO-Term classes "biological process" and "cellular compartment". Principal component analysis was done in R[245] with the plotPCA function of the DeSeq2 package after variance stabilizing transformation of read counts per gene.

7.1.2 Analysis of publicly available datasets for p63 regulation of SPRR5

For p63 regulation of SPRR5 in keratinocytes, published p63 ChIP-Seq datasets from Kouwenhoven et. al.[158] (GSE59827) and Bao et. al.[44] (GSE67382) were downloaded from the Gene Expression Omnibus (GEO) database[159,160] and uploaded into the UCSC genome browser[208] to visualize the peak location with respect to SPRR5. Furthermore, the prediction of upstream transcription factors for SPRR5 (using its former gene identifier RP1-20N18.10) was done with the ARCHS[4] tool from the Ma'ayan Lab[244].

7.1.3 Coding potential analysis for SPRR5

PhyloCSF[166] tracks were imported into the UCSC genome browser[208] by inserting "http://www.broadinstitute.org/compbio1/PhyloCSFtracks/trackHub/hub.txt" into the "my hubs" section and comparison of these tracks relative to the SPRR5 transcript[246]. Furthermore, the coding potential was assessed by running iSeeRNA[165] with a gtf file of the SPRR5 exons and by inspection of the spliced SPRR5 transcript with the Coding Potential Calculator (CPC)[164] by using their default settings.

7.1.4 Data analysis for ATAC-Seq

Data analysis for ATAC-Seq was done with a local installation of the Galaxy platform[247] and several Galaxy tools (indicated individually) with default settings unless indicated otherwise and the resulting output was inspected in IGV[237,238]. First, quality of obtained sequencing reads was monitored with FastQC (v. 0.70)[248] and reads were aligned to the human genome (hg19) using Bowtie2 (v. 2.3.3.1)[249,250]. Next, low quality mapping reads (MAPQ < 10) and duplicate reads were removed with BamTools (v. 2.4.1)[251] and rmDup (v. 2.0.1) of the SamTools[252] suite respectively. Statistics about the mapped reads were extracted with SamTools (IdxStats and Stats) and percentages of mapped reads were calculated in Excel. Peak calling for each ATAC sample was done individually with the callpeak function of MACS2 (v. 2.1.1.20160309.0, -q 0.05, --nomodel, --shift -100, --extsize 200)[253,254] without a control sample and the resulting bedgraph files were normalized to 10 million reads according to their corresponding read number after filtering and duplicate removal. A common peak set for each day of differentiation was generated by concatenating the individual peak files and merging overlapping and close peaks with the BedTools suite (v. 2.27.0.0)[255] by subsequently running "SortBed" and "MergeBed" (-d 100). The resulting peak file was then transformed into a gtf file, which was used to count the number of reads (after rmDup and quality filtering) within peaks using htseq-count (v. 0.6.1galaxy3,

stranded = no, mode = union, ID attribute = gene_id)[256] for each ATAC sample individually and Excel was used to calculate the fraction of reads in called peak regions (FRIP). Furthermore, obtained count tables were combined in Excel for each day of differentiation, imported into R[245], differential gene expression analysis was done with DeSeq2 (v. 1.14.1)[239] and PCA plots were generated with the plotPCA function of DeSeq2 after variance stabilizing transformation of read counts per peak location.

7.1.5 Data analysis for ChIP-Seq

In general, data analysis with Galaxy[247] and R[245] was performed as described for ATAC-Seq data (7.1.4) but with the following modifications. Peak calling for ChIP-Seq datasets was done against the respective input controls with MACS2[253,254] by "building the shifting model", --mfold 2,5000, --bw 200, --broad, --broad-cutoff=0.1 and "MergeBed" was run with -d 300. Furthermore, HOMER (v4.9, 2-20-2017)[188] was used for ChIP-Seq data analysis according to the HOMER documentation. Briefly, bedgraphs from filtered and deduplicated bamfiles were generated by "makeUCSCfile" (-res 10) and enriched peaks were identified with "getDifferentialPeaks" (-F 2 or -F 4), using the combined peak files from MACS2[253,254].

Additionally, already published ChIP-Seq datasets from Bao et al.[44] (GSE67382) were extracted from the GEO data repository[159,160] and processed as described above.

7.1.6 Mouse ribosome profiling data analysis

Reanalysis of mouse ribosome profiling data was done by Uwe Schwartz (University of Regensburg). Briefly, *in vivo* RNA-Seq and ribosome profiling reads from P4 mice (both replicate 0) from the Sendoel et al. paper[167] (GSE83332) were extracted, trimmed, deduplicated and aligned to the mouse genome (mm10) as described in the paper and the obtained alignment files were inspected with IGV[237,238].

7.1.7 Phylogenetic analysis of human SPRR coding sequences

The coding sequences for all human SPRR proteins were extracted from Ensembl[161], and the subsequent phylogenetic analysis was done by Prof. Dr. Rainer Merkl (University of Regensburg). Briefly, the evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model[257]. The tree with the highest log likelihood (-761.45) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial trees for the heuristic search were obtained

automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 5.2252)). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Codon positions included were $1^{st}+2^{nd}+3^{rd}$+Noncoding and all positions containing gaps and missing data were eliminated resulting in a total of 137 positions in the final dataset. Evolutionary analyses were conducted in MEGA7[258] and 1000 bootstrap iterations were performed.

7.2 Cell culture methods

All cell culture methods were performed in a biological safety cabinet under laminar air-flow to obtain sterile conditions. Lentiviral procedures were performed under biosafety level 2 precautions, whereas biosafety level 1 regulations applied to the remaining methods. Cell culture media, trypsin and DPBS (see Table 15) were pre-warmed to 37 °C prior to use and cultivation of cells was performed in a humidified incubator at 37°C and 5% carbon dioxide.

7.2.1 Cultivation of HEK293T cells and fibroblasts

HEK293T cells and fibroblasts were cultivated in DMEM+FBS (see Table 15) and passaged when they reached 80-90% confluency. Therefor the medium was aspirated, cells were washed once with DPBS and detached using 0.05% trypsin-EDTA. DMEM+FBS was used to quench the trypsin, cells were resuspended to obtain a single cell suspension and after centrifugation (200 rcf, RT, 5 min) cells were resuspended in an appropriate volume of DMEM+FBS and equally distributed onto fresh cell culture dishes.

7.2.2 Cultivation of keratinocytes

Primary keratinocytes and HaCaT cells were cultivated in 50:50 medium (see Table 15) until a maximal confluency of 80% to prevent premature differentiation. For passaging of keratinocytes, cells were washed once with DPBS and detached using 0.05% trypsin-EDTA (4 min, 37 °C). After tapping the plate to lift the cells and addition of at least four volumes of DMEM+BCS, cells were resuspended to obtain a single cell solution and pelleted by centrifugation (200 rcf, RT, 5 min). Cell pellets were resuspended in 50:50 medium and seeded onto appropriate cell culture dishes with at least 5% confluency.

Primary keratinocytes were not passaged more than 7 times.

7.2.3 Determination of cell numbers

Cells in suspension were added to a Neubauer Counting Chamber and the cells in four quadrats were counted. The number of cells per milliliter cell suspension was obtained by:

$$\text{Cells per ml} = \frac{Counted\ number\ of\ cells}{4\ (\#\ of counted\ quadrats)} \text{ x } 10{,}000$$

7.2.4 Electroporation of keratinocytes

Keratinocytes were nucleofected with siRNA pools or annealed siRNAs (see 7.6.1) utilizing the Human Keratinocyte Nucleofector Kit (Lonza). After detaching the cells with trypsin (see 7.2.1), the cell number was determined and six million keratinocytes were pelleted (200 rcf, RT, 5 min) and resuspended in 100 µl electroporation buffer. 95 µl of this suspension were mixed with 10 µl siRNA pool or annealed siRNAs (both 100 µM) and transferred to the electroporation cuvette and nucleofection was achieved with the Electroporation Device Nucleofector II (Lonza) and its preset program T-018 (Keratinocytes human, neonatal, high efficiency). Following nucleofection, the cells were transferred to 500 µl prewarmed (37°C) 50:50 medium and after recovery at 37°C for 30 minutes, seeded on a 15 cm dish and recovered for at least 24 hours at 37°C with 5% carbon dioxide.

7.2.5 Freezing and thawing of cells

For storage of cells, 1.5 million keratinocytes or 3 million HEK293T cells were resuspended in 1 ml freezing medium (cultivation medium for the cell type supplemented with 10% DMSO) and the temperature was slowly (1°C/min) lowered to -80°C. The next day, cryo-stocks were transferred to the vapor phase of liquid nitrogen for long term storage.

Cryopreserved cells were thawed in a water bath at 37°C, transferred to 10 ml cultivation medium (see Table 15) and DMSO was removed by aspirating the medium after centrifugation (200 rcf, RT, 5 min) and resuspending the cells in fresh cultivation medium followed by seeding the cells onto appropriate cell culture dishes.

7.2.6 Generation of invasive three-dimensional organotypic neoplastic tissue

Devitalized human dermis (see 7.2.11) was cut into 2.25 cm^2 pieces and each piece was spread with the basement membrane side down into a 12-well tissue culture dish. After air drying of the dermis 1 ml DMEM+FBS was added to each well.

In order to obtain dermis with embedded fibroblasts, fibroblasts were detached (see 7.2.1

and 100,000 fibroblasts in 500 µl DMEM+FBS were added to each well and centrifuged into the dermis (200 rcf, RT, 20 min). This step was repeated once with the opposite orientation of the 12-well plate during centrifugation to ensure even distribution of fibroblasts. After this step, fibroblasts were allowed to fully penetrate the dermis by cultivation at 37°C and 5% carbon dioxide and medium changes every second day.

Seven days after embedding of fibroblasts, the dermis was mechanically detached from the 12-well plate and placed with the basement membrane side upwards onto prepared inserts for organotypic epidermal cultures, the bottom was sealed with 90 µl Matrigel (Thermo Fisher Scientific) and 500,000 "tumorigenic keratinocytes" in 20 µl KGM were seeded onto the dermis after addition of KGM to the bottom of the prepared setup (for details see 7.2.7). "Tumorigenic keratinocytes", are primary keratinocytes that have been transduced with lentiviral particles to overexpress human HRasG12V and Cdk4 to unleash the cell cycle and either lacZ or SPRR5_326 overexpression (see 7.2.10).

After 6 or 8 days of cultivation with medium changes every second day, the tissue was harvested as described in 7.2.7, however only one third was used for RNA extraction and two thirds were embedded into Tissue-Tek (Weckert Labortechnik) for immunofluorescence analysis of invasion depth.

7.2.7 Generation of organotypic epidermal tissue

Setups for organotypic epidermal tissue cultures consisted of an insert with a squared (0.8 cm edge length) cavity that rested on glass beads that have been mounted in a 6 cm cell culture dish. Next, devitalized human dermis was cut, placed basement membrane side upwards over the cavity of the insert and 90 µl Matrigel (Thermo Fisher Scientific) were applied to the bottom side of the dermis to seal holes. Finally, KGM was added to the setups until it reached the bottom of the insert and thus nutrients and fluids could diffuse through the dermis.

Keratinocytes (untreated, nucleofected with siRNA or transduced with lentiviral particles) were detached, counted and pelleted (see 7.2.2 and 7.2.3) and 500,000 cells were resuspended in 20 µl KGM and equally distributed onto the dermis part covering the cavity of the insert. Stratification and differentiation was induced by raising the keratinocytes the air-liquid interface and KGM was exchanged every second day.

Organotypic epidermal tissue cultures were harvested by lifting the insert and carefully removing the Matrigel and excess devitalized human dermis. The remaining regenerated organotypic skin tissue was cut in the middle to give rise to two equal triangles and one half

was embedded for immunohistological analysis (7.3.1) into Tissue-Tek (Weckert Labortechnik), flash frozen on dry-ice and stored at -80°C, whereas the other half was used for RNA extraction (see 7.6.14).

7.2.8 Generation of SPRR5 knockout cell lines

The utilized plasmids were prepared by Julia Junghans (master student under my supervision). Briefly, guide RNAs for SPRR5 KO were designed with the gRNA Design tool from the Zhang Lab (http://crispr.mit.edu) and the two gRNAs upstream and two downstream of the SPRR5 coding sequence with the highest score were chosen and combined into a gRNA expression cassette using the PrecisionX Multiplex gRNA Cloning Kit (BioCat) and subsequently inserted into pX459 or lentiCRISPRv2 (both Addgene).

For the transient transfection-based approach, 150,000 HaCaT cells (WT or heterozygous KO cells) were seeded into a 6-well the day before transfection. The next day, the medium was changed to fresh 50:50 medium, the transfection mix was prepared according to Table 26 and after combining plasmid and lipofectamine solution and incubation at room temperature for 20 minutes, this mixture was added dropwise to the cells, followed by a medium change, six hours post transfection.

Table 26: Transfection mix for SPRR5 KO cell line generation

	6-well	Component
Plasmid dilution	125 µl	Opti-MEM
	2.5 µg	pX459_complete
	5 µl	P3000 enhancer reagent
Lipofectamine dilution	125 µl	Opti-MEM
	7,5 µl	Lipofectamine 3000 reagent

Selection of plasmid containing cells was achieved by changing the medium to 50:50 medium supplemented with 1 µg/ml puromycin 48 hours and 72 hours post transfection. After selection, medium was exchanged to conditioned 50:50 medium and cells were allowed to recover for one day.

For single cell seeding, cells were detached (7.2.2), counted (7.2.3) and diluted with conditioned 50:50 medium to a concentration of 0.5 cells per 100 µl and 100 µl were plated into a 96-well. Two weeks after seeding, 50 µl of a 1:1 mixture of fresh and conditioned 50:50 medium was added to each well and when a clonal cell would become more than 60% confluent this cell line was expanded by detaching the cells with trypsin (described in 7.2.2) and 60% of the cells for were used for further cultivation on a 48-well and 40% for screening purposes.

In order to screen potential KO cell lines, a cell lysate was prepared by pelleting the cells (5 min, RT, 13,000 rcf), resuspending the pellet in 40 µl 1x ThermoPol buffer (New England Biolabs) supplemented with 8 µg Proteinase K (Roche) and incubation at 65°C for one hour followed by 15 minutes at 95°C. This lysate was then used for the PCR-based screening approach (see 7.6.10) and heterozygous KO cell lines were further expanded while wild type cells were discarded.

The lentiviral approach was similar to the transfection-based approach, however the SPRR5 KO cassette was stably integrated into the genome of the HaCaT cells by lentiviral transduction (see 7.2.10 with the transfer plasmid: lentiCRISPRv2_complete_CMVCas9) rather than introducing it by transient transfection. Starting with the puromycin selection, cells from this approach were treated as described for the transient transfection approach.

7.2.9 Keratinocyte differentiation cultures

Differentiation of keratinocytes in monolayers was induced by seeding the keratinocytes after detachment and centrifugation (see 7.2.1) at full confluency and addition of 1.2 mM calcium chloride to the 50:50 medium. For differentiation cultures, the medium was renewed every day.

7.2.10 Lentivirus production and transduction of keratinocytes

Transfection of HEK293T cells was performed according to the table below and only endotoxin free plasmid preparations (see 7.6.10) were used for transfection in combination with Lipofectamine 3000 (Thermo Fisher Scientific). The lentiviral transfer vectors were either pLEX, pLARTA or lentiCRISPRv2 with varying inserts (see Table 23) and the employed amount for transfection was calculated depending on the size of the lentiviral transfer vector to obtain a molar ratio of 1:1:1 for all three plasmids.

Table 27: Transfection reaction mixtures for lentiviral particle generation

	10 cm dish	15 cm dish	Component
Plasmid dilution	1.5 ml	3.5 ml	Opti-MEM
	7.5 µg	17.5 µg	pCMV dR8.91
	3.72 µg	8.68 µg	pUC-MDG
	varied	varied	lentiviral transfer vector
	35 µl	81 µl	P3000 enhancer reagent
Lipofectamine dilution	1.5 ml	3.5 ml	Opti-MEM
	41 µl	95 µl	Lipofectamine 3000 reagent

Plasmid and lipofectamine dilution were mixed and incubated at room temperature for 25 minutes. After that, the transfection reaction was applied to the 10 cm (or 15 cm) dish and 21 million (50 million for 15 cm) HEK293T cells resuspended in 8 ml (or 18 ml for 15 cm) were added for reverse transfection. Six and 24 hours post transfection, the medium was changed to 12 ml (or 25 ml for 15 cm) DMEM+FBS and viral particles were harvested 48 hours post transfection by filtering the virus containing medium through a 0.45 µm pore-sized polyethersulfone membrane.

For lentiviral transduction 45,000 (or 250,000 per 10 cm dish) keratinocytes were seeded in a 6-well (or 10 cm dish) the day before infection. The next day, 2.5 ml (or 12 ml per 10 cm dish) transduction mix per 6-well was prepared, by diluting the viral particles in DMEM+FBS and addition of polybrene to a final concentration of 5 µg/ml. The optimal dilution was determined in preceding efficiency tests for each batch of lentiviral particles to ensure an adequate overexpression in combination with no cytotoxic effect. The finally employed dilutions can be obtained from Table 28.

Table 28: Overview of employed lentiviral dilutions for keratinocyte transduction

Lentiviral transfer vector	Dilution
lentiCRISPRv2_complete_CMVCas9	1:1.66
pLARTA_326/lacZ (for 7.2.6 and 7.4.2)	undiluted
pLARTA backbone with varying inserts (for all procedures except 7.2.6 and 7.4.2)	1:1.66
pLEX_Cdk4/HRasG12V/LY	1:2/1:3/1:2
pLEX backbone with varying inserts for Rescue experiments	1:6.66

Infection of keratinocytes was accomplished by exchanging the growth medium to the transduction mix and centrifugation at room temperature with 250 rcf for one hour. Afterwards, the transduction mix was aspirated, cells were washed twice with DPBS and recovered in 50:50 medium for at least 24 hours.

7.2.11 Preparation of human devitalized dermis

Human split-skin was washed twice in PBS+2x A/A and incubated in PBS+2x A/A at 37°C for 3-7 days until the epidermis could be mechanically detached from the dermis. Until further use, the dermis was stored in PBS+2x A/A at 4°C.

7.3 Histological analysis

7.3.1 Immunofluorescence analysis of cryosections from epidermal tissue

7 µm thick sections from embedded regenerated organotypic epidermal tissue (see 7.2.7) were prepared with the Cryostat Microm HM 500 OM (Thermo Fisher Scientific), transferred onto Polysine slides (Carl Roth) and dried for one hour at room temperature and stored at -20°C until staining.

Depending on the differentiation protein to be analyzed, skin sections were fixed in methanol, ethanol or acetone (see Table 6) for 10 min at -20°C and after gradually replacing the fixative with PBS, slides were blocked for 20 min at room temperature in PBS with 10% BCS (Thermo Fisher Scientific). Primary antibody solution in 1% BCS in PBS was applied and incubated on the sections for one hour at room temperature or overnight at 4°C in a humidified chamber (double staining of differentiation protein and collagen VII, dilutions are given in Table 6). Next, sections were washed with 1% BCS in PBS (three times, 5 min each) and fluorescently labeled secondary antibodies (see Table 7) diluted in 1% BCS in PBS were applied in a dark humidified chamber for one hour at room temperature, such as the differentiation protein was labeled with Alexa 555 and collagen VII with Alexa 488. Following this step, slides were briefly rinsed in PBS and nuclei were stained by incubating the sections with Hoechst solution for 5 min at room temperature. Slides were washed three times with PBS (5 min, RT), air dried, mounted using ProLong Gold Antifade Mountant (Thermo Fisher Scientific) and stored at 4°C.

Pictures from representative areas for each section were taken with the Inverted microscope Axiovert 200 M (Carl Zeiss) in combination with the AxioVision software (Carl Zeiss) by overlaying pictures from the DAPI, GFP and Cy5 channel taken at 20x magnification.

7.3.2 Immunofluorescence analysis of cryosections from neoplastic tissue

Invasive three-dimensional organotypic neoplastic tissue (7.2.6) was cut entirely into 10 µm thick sections with the Leica CM3050 S cryostat (Leica Biosystems) and every 7th section was processed as described above. Staining and analysis of sections was performed as described in 7.3.1 with the following modifications. Sections were fixed in a mixture of equal parts methanol and acetone, blocking was done with 10% goat serum (Thermo Fisher Scientific) and 0.1% Triton X-100 in PBS, antibodies (see Table 6 and Table 7) were diluted

in 1% goat serum (Thermo Fisher Scientific) and 0.1% Triton X-100 in PBS and washing steps were performed with 0.1% Triton X-100 in PBS. After the final washing step, Triton X-100 was removed by washing with PBS (twice, 5 min, RT) sections were air dried and instead of a Hoechst staining, nuclei were stained by using Duolink *In Situ* Mounting Medium with DAPI (Sigma-Aldrich). Pictures were obtained as described in 7.3.1 and all measurements were done with the AxioVision software (Carl Zeiss).

Invasion of keratinocytes into the dermal matrix is a hallmark of biologic malignancy in 90% of human cancers and thus can be used for evaluation of neoplasia onset[154]. For this reason, the invasion depth of keratinocytes (visualized by keratin 5 staining) into the dermal matrix was measured for each picture at five fixed and independent locations and the mean as well as the standard deviation were calculated and plotted across all three biological replicates per timepoint. Furthermore, the epidermis thickness above the invasion sites was also determined and the ratio of epidermis thickness to invasion depth was calculated and plotted as described above.

7.4 *In vivo* experiments

6-8 weeks old immunodeficient mice (CBySmn.CB17-*Prkdc*^{scid}/J, obtained from The Jackson Laboratory (Bar Harbor, USA)) were used and all experiments were performed according to an approved experimental protocol by the Stanford Panel on Laboratory Animal Care.

7.4.1 *In vivo* imaging of tumor size

Tumor size was monitored *in vivo* by measuring the luciferase signal originating from the luciferase overexpressing tumor cells. To this end, mice were anesthetized with isoflurane, back hair was clipped and 100 µl of 30 mg/ml D-luciferin were injected intraperitoneal. After 10 minutes, the luminescence was recorded with the IVIS 100 *In Vivo* Imaging System (PerkinElmer) and superimposed with the corresponding photographic picture. Exposure time, binning and distance to the camera were adjusted as required to omit saturation and kept constant between measuring lacZ and SPRR5_326 overexpressing tumors to allow cross-sample comparisons.

Data analysis was done with the Living Image software (PerkinElmer). Briefly, the luminescence signal was determined for each animal individually using the automatic detection function of the software and the values were extracted in “Photons” mode. The mean and standard deviation of the resulting total flux (photons/sec in each pixel summed

over the region of interest area) for each group was calculated per day and plotted against the days after injection. Additionally, some exemplary pictures from the measurements are shown.

7.4.2 *In vivo* tumor formation assay

Primary keratinocytes were consecutively transduced with lentiviral particles to overexpress a Luciferase-YFP fusion protein for imaging, HRasG12V and Cdk4 for transformation into "tumorigenic keratinocytes" and SPRR5_326 or lacZ to study the potential tumor suppressive effect of SPRR5_326 (for details see 7.2.10 and 6.10). After detachment and counting of the cells (see 7.2.2 and 7.2.3), cells were washed once with PBS (5 min, RT, 200 g) and 600,000 cells were resuspended in 100 µl PBS supplemented with 50 µl Matrigel (Thermo).

Mice were anesthetized with an isoflurane/oxygen mix (PM1002 mobile anesthesia machine, Parkland Scientific) and the back hair was removed. For i*n vivo* tumor generation, the prepared 600,000 cells were injected into the subcutaneous space on the flank of the mouse and for each condition (SPRR5_326 or lacZ overexpression) 4 biological replicates were prepared.

Tumor growth was monitored by caliper measurements and *in vivo* imaging (7.4.1) once or twice per week. For caliper measurements, length (L), height (H) width (W) for each tumor was determined and the tumor volume was calculated using the formula from Faustino-Rocha et al.[259].

$$\text{tumor volume} = \frac{W \; x \; W \; x \; L}{2}$$

Tomayko and Reynolds showed that the ellipsoid volume formula correlated best with the tumor mass and hence this formula was used to estimate the tumor mass[260].

$$\text{tumor mass} = \frac{\pi \; x \; L \; x \; W \; x \; H}{6}$$

85 days after injection of tumorigenic keratinocytes, the mice were euthanized and the tumors were extracted and photographed (done by Dr. Zurab Siprashvili, Stanford University).

7.5 Microbiological techniques

7.5.1 Cultivation of Escherichia coli

Bacteria were grown overnight at 37°C on LB-agar plates (with or without antibiotics). Liquid bacteria cultures were inoculated in an appropriate volume of LB-Amp medium and incubated at 37°C with 220 rpm overnight.

7.5.2 Preparation of chemically competent *Escherichia coli*

5 ml YT-medium were inoculated with a single colony of the selected *Escherichia coli* strain and grown overnight at 37°C and 220 rpm. The next day, 50 µl from this starter culture were mixed with 5 ml fresh YT-medium and incubated at 37°C and 220 rpm until an OD_{600} of 0.8 was reached, 100 ml YT-medium were added and the culture was grown to an OD_{600} of 0.5. Cells were spun down (2,000 rcf, 5 min, 4°C), resuspended in 20 ml ice-cold TFBI and incubated on ice for 10 minutes. After sedimentation of the cells (2,000 rcf, 5 min, 4°C), cells were resuspended in 4 ml ice-cold TFBII, divided in 100 µl aliquots, snap-frozen in liquid nitrogen and stored at -80°C.

7.5.3 Transformation of chemically competent *Escherichia coli*

100 µl of chemically competent cells were thawed on ice, mixed with either 50 ng of plasmid DNA (retransformation) or a heat inactivated ligation reaction and incubated for 30 min on ice. Following the heat shock required for plasmid uptake (42°C, 90 s), cells recovered 2 min on ice and after addition of 900 µl LB-medium (pre-warmed to 37°C), the cells were incubated for 1 h at 37°C and 600 rpm. Appropriate volumes of the transformation reaction were then spread on LB-Amp agar plates and incubated at 37 °C overnight.

7.6 Molecular biological methods

7.6.1 Annealing of siRNAs

For annealing of siRNAs, 1 nmol of sense and antisense single siRNAs (for sequences see Table 17) were mixed with 5 µl siRNA annealing buffer in a 10 µl reaction and heated to 95°C for 3 min followed by 1 hour at 37°C.

7.6.2 Assay for Transposase Accessible Chromatin (ATAC-Seq)

Sample preparation for ATAC-Seq was done by Katrin Hartinger (master student under my supervision) in collaboration with the AG Rehli (University Hospital Regensburg) based on the recently published Omni-ATAC protocol from Corces et al.[261]. Briefly, day 4 differentiated or undifferentiated keratinocytes with or without SPRR5 knockdown were harvested as described above (7.6.3 and 7.2.2). Dead cells were removed by magnetic bead depletion using the Annexin V MicroBead Kit (Miltenyi Biotec) according to the manufacturer's instructions and 50,000 living cells were used for the transposition reaction. To this end, cells were pelleted (500 rcf, 5 min, 4 °C), resuspended in 50 µl ice-cold ATAC resuspension buffer supplemented with 0.1% NP40, 0.1% Tween-20 and 0.01% digitonin and incubated for 3 min on ice. Next, cells were washed (500 rcf, 5 min, 4°C) with 1 ml of ice-cold ATAC resuspension buffer supplemented with 0.1% Tween-20, pelleted (500 rcf, 5 min, 4°C) and resuspended in 50 µl of transposition mixture (see Table 29, enzymes and buffers from the Nextera DNA Library Preparation Kit, Illumina). The transposition reaction took place at 37°C (30 min, 1,000 rpm).

Table 29: Preparation of the ATAC transposition mixture

Component	volume [µl]
2x ATAC transposition buffer (Nextera kit, Illumina)	25
Nextera Tn5 Transposase (Nextera kit, Illumina)	2.5
1% (w/v) digitonin	0.5
10% (v/v) Tween-20	0.5
PBS	16.5
H_2O	5

After transposition, the DNA was purified with the Monarch PCR & DNA Purification Kit (New England Biolabs) according to the manufacturer's instructions and recovered in 21 µl elution buffer. Indexes for next-generation sequencing were introduced by PCR. Therefor 10 µl of the transposed DNA were mixed with 10 µl of 5x Phusion HF buffer, 1 U Phusion DNA Polymerase (New England Biolabs), 2.5 µl of the specific i5 index primer, 2.5 µl of the specific i7 index primer (both from the Nextera Index Kit) and brought to a final concentration of 0.3 mM dNTPs and 1.3 M betaine in a 50 µl reaction. The utilized thermal cycling conditions are given in Table 30.

Table 30: Cycling conditions for amplification of ATAC-Seq samples

Step	Time	
72°C	5 min	
98°C	30 s	
98°C	10 s	11 x
63°C	30 s	
72°C	1 min	
4°C	∞	

Subsequent DNA purification was achieved with the Monarch PCR & DNA Purification Kit (New England Biolabs, first biological replicates) or AMPure XP beads (Beckman Coulter, second biological replicates) for size selection between 100 and 500 bp, according to the manufacturer´s instruction. For both purifications, DNA was eluted with 15 µl elution buffer from the Monarch PCR & DNA Purification Kit (New England Biolabs) and the obtained libraries were quantified on the Qubit 2.0 Fluorometer in combination with the Qubit dsDNA HS Assay Kit (both Thermo Fisher Scientific) according to the manufacturer´s instruction. Finally, DNA size distribution was analyzed on the TapeStation in combination with the High Sensitivity D1000 ScreenTape and its corresponding reagents (Agilent) and libraries were pooled according to the sample requirements of the Biomedical Sequencing Facility in Vienna, were 50 nt single-end sequencing was performed.

7.6.3 cDNA synthesis

For cDNA synthesis 500-1000 ng of TRIzol purified RNA (7.6.15) was first subjected to a DNase digest with either DNaseI or the TURBO DNA-free Kit (both Thermo), according to the manufacturer's instructions, followed by cDNA synthesis with the iScript cDNA Synthesis Kit (Bio-Rad, following the manufacturer's instructions). In contrast to this, RNA obtained with the RNeasy Plus Mini Kit was directly used for cDNA synthesis without prior DNase digest.

After cDNA synthesis, H_2O was added to give a final volume of 200 µl and the cDNA was stored at -20°C.

7.6.4 Cellular fractionation of keratinocytes

D3 to D6 differentiated keratinocytes (see 7.2.9) were washed twice with PBS and detached with 0.25% Trypsin-EDTA (Thermo Fisher Scientific) and incubation at 37°C for 4 minutes. Next, three volumes of DMEM+BCS were added, cells were scraped off with a cell scraper and singularized by resuspension and filtering through a 40 µm cell sieve (Falcon). After

pelleting (200 rcf, RT, 3 min), cells were resuspended in 1 ml PBS and the fractionation was performed with the Cytoplasmic and Nuclear RNA Purification Kit (Norgen Biotek) according to the manufacturer's instructions and with several variations (lifted cells protocol or tissue protocol, with Accutase instead of trypsin, with or without sucrose cushion during centrifugation, longer incubation times), here solely the pursued protocol for the depicted data is given. To this end, the protocol for cells growing in suspension was pursued, however the nuclear RNA pellet was washed once with 200 µl ice-cold lysis buffer (1 min, 4°C, full speed) before RNA extraction to remove residual cytoplasmic carry-over.

7.6.5 Chromatin Immunoprecipitation with DNA-sequencing (ChIP-Seq)

Sample preparation for ChIP-Seq was done together with the AG Rehli (University Hospital Regensburg) and Katrin Hartinger (master student under my supervision). Therefor keratinocytes with or without SPRR5 depletion were harvested on day 4 or day 0 (undifferentiated) of differentiation as described above (7.6.3 and 7.2.2) and formaldehyde crosslink was achieved by resuspending 10 million cells in 10 ml 1% (v/v) formaldehyde solution and incubation at room temperature for 10 minutes under constant rotation. Next, excess formaldehyde was quenched by addition of 1/10 volume 1.25 M glycine (5 min, RT, rotation) and crosslinked cells were pelleted (2,000 rcf, 4°C, 10 min), washed once with 10 ml ice-cold PBS supplemented with 1 mM AEBSF and 10 mM sodium butyrate (2,000 rcf, 4°C, 10 min) and transferred to a 1.5 ml reaction tube with 750 µl ice-cold PBS supplemented with 1 mM AEBSF and 10 mM sodium butyrate. Following centrifugation (6,000 rcf, 5 min, 4°C), the liquid was removed completely, the weight of the cell pellet was determined, samples were flash frozen in liquid nitrogen and stored at -80°C.

For chromatin sonication, crosslinked cells were thawed on ice and resuspended in 20 µl ChIP swelling buffer per 1 million crosslinked cells (the number of cells was determined by dividing their weight in mg by 4.5, which resulted in the millions of crosslinked cells as determined during preliminary experiments) but with a minimum of 500 µl swelling buffer per sample. Next, nuclei were extracted by incubation on ice for 10 minutes with occasional mixing and subsequent compression of the cells 13 times in a 2 ml douncer, pelleted (1,500 rcf, 5 min, 4°C) and resuspended in 1 µl ChIP RIPA buffer per 100 µg nuclei but with a minimum of 800 µl. Following incubation on ice for 10 minutes, the crosslinked DNA was sheared to fragments between 150 bp to 250 bp by sonication with the S220 Focused-ultrasonicator (Covaris) in combination with 12x24 mm sonication tubes (Covaris, P/N 520056) and the indicated settings (Table 31).

Table 31: Settings for sonication of chromatin

Duty cycle	20%
Mode	Freq sweeping
Intensity	8
Cycles of burst	200
Water level	15
Time	30 min

After centrifugation (15,870 rcf, 5 min, 4°C), 20 µl of the supernatant were set aside, while from the remaining supernatant was transferred to a fresh reaction tube. A 5% (v/v) input sample was taken, and input and chromatin were flash-frozen and stored at -80°C. The 20 µl aliquot was mixed with 2 µl Proteinase K solution (20 µg/µl, Roche), incubated over-night at 65°C and after the addition of 10 µg RNase A (Qiagen) incubated at 37°C for 30 min. Following DNA extraction with the Monarch PCR & DNA Purification Kit (New England Biolabs, according to the manufacturer's instructions), efficient chromatin shearing was verified by agarose gel electrophoresis (7.6.6).

For chromatin immunoprecipitation, beads were washed three times with TE (pH 8.0) (850 rcf, 1 min, RT), resuspended in the initial volume ChIP-seq DB supplemented with 0.5% (w/v) BSA and 20 ng/ml glycogen and incubated at RT for 2 hours or overnight at 4°C. Next, the chromatin corresponding to 2 million cells was thawed on ice, 1.5 volumes of ChIP-Seq DB and 0.25 final volumes washed Sepharose CL-4B beads were added and this mixture was incubated under constant rotation at 4°C for two hours. After separation of the preclearing beads (16,000 rcf, 5 min, 4°C), 2.5 µg antibody per 200 µl supernatant was added and antibody binding was achieved overnight at 4°C under constant rotation.

The next day, 0.25 volumes of prepared nProteinA Sepharose 4 Fast Flow beads were added and the beads were allowed to bind the antibody for 3 hours at 4°C with constant rotation. Beads were washed (750 rcf, 5 min, 4°C) twice with 400 µl ice-cold ChIP-Seq WB I, ChIP-Seq WB II and ChIP-Seq WB III, followed by three washing steps with 400 µl ice-cold TE (pH 8). Eventually, DNA was eluted from the beads by addition of 110 µl ChIP-Seq elution buffer and incubation at room temperature for 20 minutes, and a second elution with 100 µl ChIP-Seq elution buffer. After separation of the beads (750 rcf, 3 min, RT) the supernatant was combined into a fresh reaction tube, brought to a final concentration of 250 mM sodium chloride and incubated overnight at 65°C to reverse the formaldehyde crosslinks. Likewise, the 5% input samples were mixed with 200 µl ChIP-Seq elution buffer, reverse crosslinked and from here on treated identical to the ChIP samples.

The following day, DNA was extracted by addition of 70 µg RNAse A (Qiagen) and incubation for one hour at 37°C, subsequent addition of 100 µg Proteinase K (Roche) and

incubation for another hour at 37°C followed by purification with the Monarch PCR & DNA Purification Kit (New England Biolabs, according to the manufacturer's instructions) and elution in 43 µl of the provided elution buffer.

Library preparation for ChIP-Seq was performed with the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs) in combination with the Index Primer Set 1+2 of the NEBNext Multiplex Oligos for Illumina (New England Biolabs). In general, the manufacturer's instructions were followed, however for adapter ligation a 1:10 dilution of the NEBNext adapter was used and Agencourt AMPure XP beads (Beckman Coulter) were used for PCR clean-up as well as for size selection of DNA inserts of approximately 200 bp. Furthermore, instead of the recommended 5 µl only 2 µl of each index primer and the universal PCR primer were utilized and DNA elution after PCR clean-up was achieved with 16 µl elution buffer from the Monarch PCR & DNA Purification Kit (New England Biolabs). Finally, the obtained libraries were quantified on the Qubit 2.0 Fluorometer in combination with the Qubit dsDNA HS Assay Kit (both Thermo Fisher Scientific) according to the manufacturer´s instruction. The size distribution was analyzed on the TapeStation in combination with the High Sensitivity D1000 ScreenTape and its corresponding reagents (Agilent) and libraries were pooled in an equimolar ratio. The resulting pool was sent for sequencing to the Biomedical Sequencing Facility in Vienna according to their sample requirements, were 65 nt single-end sequencing was performed on an Illumina HiSeq3000 platform.

7.6.6 DNA agarose gel electrophoresis

Depending on the expected DNA size, horizontal agarose gel electrophoresis was performed with 0.8 – 1.4% (w/v) agarose gels (dissolved in TAE buffer, supplemented with a final concentration of 0.5 µg/ml ethidium bromide). To this end, samples were mixed with 1/6 volume of DNA loading dye, loaded next to 5 µl of 1 kb Plus DNA Ladder (Thermo Fisher Scientific) and separation of DNA fragments was achieved by applying 90 V in gel chambers filled with TAE buffer. Gel documentation was done with the Transilluminator Quantum ST4 (PEQLAB).

7.6.7 Full transcriptome RNA sequencing

Library preparation as well as sequencing was performed by the "Kompetenzzentrum fuer Fluoreszente Bioanalytik" Regensburg with RNAs isolated from regenerated organotypic

epidermal tissue (7.6.14).

In general, library preparation and mRNA sequencing were carried out according to the Illumina TruSeq Stranded mRNA Sample Preparation Guide, the Illumina HiSeq 1000 System User Guide (Illumina), and the KAPA Library Quantification Kit - Illumina/ABI Prism User Guide (Kapa Biosystems), with minor modifications.

In brief, mRNA molecules were purified using oligo-dT probes immobilized on magnetic beads starting with 250 ng of total RNA supplied with ERCC spike ins[233]. Chemical fragmentation of the mRNA to an average insert size of 200-400 bases was performed using divalent cations under elevated temperature (94°C, 4 minutes). First strand cDNA was produced by reverse transcription with random primers. Actinomycin D was added to improve strand specificity by preventing spurious DNA-dependent synthesis. Blunt-ended second strand cDNA was synthesized using DNA Polymerase I, RNase H and dUTP nucleotides. The resulting cDNA fragments were adenylated at the 3' ends, the indexing adapters were ligated and subsequently specific cDNA libraries were created by PCR enrichment. The libraries were quantified using the KAPA SYBR FAST ABI Prism Library Quantification Kit. Equimolar amounts of each library were used for cluster generation on the cBot (TruSeq SR Cluster Kit v3). The sequencing run was performed on a HiSeq 1000 instrument using the indexed, 1x50 cycles single end protocol and TruSeq SBS v3 Reagents according to the Illumina HiSeq 1000 System User Guide. Image analysis and base calling resulted in .bcl files, which were converted into .fastq files by the CASAVA1.8.2 software (Bio Gist).

7.6.8 Generation of plasmids

The concentration of digested and purified vector and insert (see 7.6.13) was determined using the NanoDrop 1000 device (Thermo Fisher Scientific) and 50 ng vector were mixed with the designated insert in a 1:3 molar ratio with 1x T4 DNA ligase buffer (New England Biolabs) and 400 units T4 DNA ligase (New England Biolabs) in a 20 µl reaction. Ligation was achieved by incubation at room temperature for one hour or overnight at 4°C. Following heat inactivation (10 min, 65°C), the obtained plasmids were used for transformation of competent *Escherichia coli* cells (7.5.3).

7.6.9 Northern blot analysis

15 to 25 µg TRIzol purified RNA from keratinocytes at varying timepoints of differentiation as well as 5 µl RiboRuler High Range or Low Range RNA Ladder (Thermo Fisher Scientific)

were mixed with the same amount of 2x RNA loading dye, incubated at 65°C for 10 min and subsequently placed on ice. Ethidium bromide was added to a final concentration of 40 µg/ml and the samples were loaded onto a formaldehyde agarose gel (1.2% agarose dissolved in RNA gel buffer) and separated by horizontal gel electrophoresis (running buffer: RNA gel buffer, 70 V) until the blue dye was 3 cm above the end of the gel. After shaking the gel for 30 min in 20x SSC, the RNA was transferred onto a trimmed piece of Amersham Hybond-N+ membrane (GE Healthcare) by upward capillary transfer using 20x SSC as transfer buffer and incubation overnight (detailed description of the setup can be found here[262]).

The next day, the RNA was crosslinked to the membrane with UV-light at 254 nm (auto crosslink function of the UV Stratalinker 2400 (Stratagene)) and successful RNA transfer as well as migration pattern of the RNA ladder was captured with the Transilluminator Quantum ST4 (PEQLAB). After prehybridization (one hour at 40-50°C with hybridization solution supplemented with 1 mg heat denaturated hering sperm DNA (Promega)) specific transcripts were detected by adding 20 pmol antisense DNA-oligos (see Table 21) that have been labeled in a T4 PNK reaction (Thermo Fisher Scientific) with 20 µCi ^{32}P (20 µl total reaction volume, according to the manufacturer's instructions), purified with a G-25 column (GE Healthcare, according to the manufacturer's instructions) and incubation overnight at 40-50°C under constant rotation.

The next day, excess radioactivity was washed away (10 min each, 40-50°C, twice with NB wash I and once with NB wash II) and a phosphorimager screen (Kodak) was used to accumulate the radioactive signal. The resulting radioactive signals were read with the Personal Molecular Imager (Bio-Rad) and analyzed with the Quantity One software (Bio-Rad). Sizes for the detected transcripts were determined by superimposing the ethidium bromide picture with the obtained phosphorimager signals.

7.6.10 PCR-based screening of potential SPRR5 knockout cell lines

Screening PCRs for potential SPRR5 knockout cell lines (7.2.8) were performed as nested PCRs utilizing the Taq DNA polymerase (New England Biolabs) and a first set of PCRs to detect genomic alterations (KO allele PCR) and a second to detect remaining WT alleles (WT allele PCR). For PCR1 12.5 µl cell lysate (see 7.2.8) was brought to a final concentration of 200 µM dNTPs, 0.2 µM forward and reverse primer (see below), 1x ThermoPol buffer (New England Biolabs) and 1.25 units Taq DNA polymerase in a 25 µl reaction. After conducting the thermal cycling program, 1 µl of this first PCR was used as a

template for the second PCR which was also done in a 25 µl scale with the same final concentrations as specified for PCR1. 15 µl from PCR2 were analyzed by DNA gel electrophoresis (7.6.5) and selected PCR products were purified with the NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel) according to the manufacturer's instructions and their sequence was determined by Sanger sequencing (Macrogen) to ensure specificity of the employed screening approach as well as to detect genomic alterations.

The localization of the utilized PCR primers and gRNAs can be seen in Figure 18 and the PCR conditions and utilized primers are listed below.

Table 32:Thermal cycling conditions for KO allele PCRs (left) and WT allele PCRs (right)

Step	Time		Step	Time	
95°C	5 min		95°C	5 min	
95°C	30 s	25 x	95°C	30 s	25 x
55°C	1 min		55°C	1 min	
68°C	3 min		68°C	1.5 min	
68°C	7 min		68°C	7 min	
10°C	∞		10°C	∞	

Table 33: Primer sequences for KO allele PCRs (left) and WT allele PCRs (right)

Name	Sequence (5' to 3')	Name	Sequence (5' to 3')
PCR 1		**PCR 1**	
KO_out_F1	CTGCATGTGTGTGTCGGTAC	WT_In_F1	CGTCCAAGCAGAAGTAAGCC
KO_out_R1	TTACTGTCGGGTGTGTGTGT	KO_out_R1	TTACTGTCGGGTGTGTGTGT
PCR 2		**PCR 2**	
KO_out_F2	CCAACTCTAAGAGAGGTAAGTATG	WT_In_F2	CCCCTCAGCTCCTGTTTTCA
KO_out_R2	GTATGCAGTGTTTGCATAGACTGTC	KO_out_R2	GTATGCAGTGTTTGCATAGACTGTC

7.6.11 Plasmid purification

Depending on the required amount of DNA vectors, plasmid purification was done with the NucleoBond Xtra Maxi EF (for endotoxin-free plasmid preparations), NucleoBond Xtra Midi or the NucleoSpin Plasmid Kit (Macherey-Nagel) following the manufacturer's instructions. The obtained DNA was dissolved in a suitable amount of H_2O, the concentration was determined with the NanoDrop 1000 (Thermo Fisher Scientific) and the sequence integrity was verified by restriction enzyme digest as well as Sanger sequencing of the included inserts (Macrogen, sequencing primers are listed in Table 18).

7.6.12 Polymerase chain reaction

For PCR amplification of DNA fragments designated for the generation of new plasmids, PCR was performed in a 50 µl scale using Phusion High-Fidelity DNA Polymerase (New England Biolabs) with 50 ng of DNA template and a final concentration of 1x HF buffer, 200 µM dNTPs, 0.5 µM of each primer (sequences are given in Table 20), 3% DMSO and 1 unit of Phusion DNA Polymerase in combination with the thermal cycling program given below.

Table 34: Thermal cycling program for PCR with the Phusion High-Fidelity DNA Polymerase

Step	Time	
98°C	30 s	
98°C	10 s	35 x
55-60°C	30 s	
72°C	30 s/kbp	
72°C	10 min	
4°C	∞	

Successful PCR amplification was verified by DNA agarose gel electrophoresis (7.6.5) and the desired PCR product was purified by either PCR clean-up (in case of only one PCR amplicon) or gel extraction, both done with the NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel) according to the manufacturer's instructions.

7.6.13 Restriction enzyme digest

For preparative restriction enzyme digests, 5 to 8 µg DNA vector (1 µg for analytical purposes) or the complete amount of purified PCR product was entirely digested with suitable restriction enzymes in the designated buffer according to the New England Biolabs guidelines and in case of vector backbone preparation for molecular cloning a dephosphorylation step with Antarctic phosphatase (New England Biolabs) was done according to the manufacturer's instructions. Following heat inactivation of the enzymes, the obtained DNA fragments were analyzed by DNA agarose gel electrophoresis (7.6.5) and desired fragments for molecular cloning were purified via gel extraction with the NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel) according to the manufacturer's instructions.

7.6.14 RNA extraction from organotypic tissue

The designated portion of organotypic tissue (7.2.6 and 7.2.7) was minced and transferred into a Lysing Matrix D tube (MP Biomedicals) that contained 800 µl RLT Plus buffer (from the RNeasy Plus Mini Kit, Qiagen) supplemented with 1% β-mercaptoethanol. Cell lysis and

homogenization was performed with the FastPrep-24 Instrument (MP Biomedicals) for 45 s at 6.5 m/s. After centrifugation (2 min, 4°C, 13,000 rcf), the supernatant was transferred to a QIAShredder tube (Qiagen) and additionally homogenized by centrifugation (2 min, 4°C, 13,000 rcf). The flow-through from this step was then subjected to RNA purification with the RNeasy Plus Mini Kit (Qiagen), following the manufacturer's instructions, except all centrifugation steps were done at 4°C. The RNA was eluted in 30 µl H_2O, quantified with the NanoDrop 1000 (Thermo Fisher Scientific) and stored at -80°C until further usage.

7.6.15 RNA extraction with TRIzol

Keratinocytes were washed with DPBS and directly lysed in a suitable volume of TRIzol Reagent (Thermo Fisher Scientific) and the RNA purification was performed according to the manufacturer's instructions but with an additional chloroform extraction step. The obtained RNA pellet was dissolved in an appropriate amount of H_2O (55°C, 900 rpm, 5 min), quantified with the NanoDrop 1000 (Thermo Fisher Scientific) and stored at -80°C.

7.6.16 RT-qPCR analysis

For RT-qPCR analysis, 7.5 µl of SsoFast EvaGreen Mix (Bio-Rad) or Takyon Mix (Eurogentec) were mixed with 4.5 µl H_2O, 1 µl Primermix (5 µM each primer, for sequences see Table 19), 2 µl cDNA and analyzed in a 96-well format using the Real-Time PCR Cycler CFX96 (Bio-Rad) in combination with the Bio-Rad CFX Manager 3.1 (Bio-Rad) or the Mx3000P (Agilent Technologies) in combination with the MxPro QPCR Software (Agilent Technologies). Samples were at least run in duplicates and specificity of each reaction was monitored using a melt curve analysis for each PCR product whereas the linear amplification for each Primermix was ensured by testing their amplification range with a serial cDNA dilution series in preliminary experiments.

Table 35: Thermal cycling program for RT-qPCR analysis

Step	Time	
95°C	3 min	
95°C 60°C 72°C plate read	15 s 30 s 30 s	40 x
95°C	10 s	
65°C to 95°C +0.5°C/step plate read	5s	melt curve

Sample wise fold changes were calculated for each gene of interest (goi) in reference to the control from the exported C_q-values using the $2^{-\Delta\Delta Cq}$ method[263] and L32 (where applicable) for normalization according to the following formula:

$$\text{Fold change} = 2^{-[(Cq\ sample(goi) - Cq\ control(goi)) - (Cq\ sample(L32) - Cq\ control\ (L32))]}$$

7.7 Protein biochemistry

7.7.1 BCA assay for protein quantification

Protein concentration was determined using the Pierce BCA protein assay kit (Thermo Fisher Scientific) according to the manufacturer's instructions in a 96-well format. Samples were measured in duplicates, analyzed at 562 nm and the protein concentration was determined using the standard curve obtained by plotting the blank-corrected A_{562} for the protein standards against their concentration.

7.7.2 Bradford assay for protein quantification

5x Roti-Quant (Carl Roth) was diluted 1:5 with water and 900 µl of this dilution were mixed with 100 µl prepared bovine serum albumin standard ranging from 0 to 150 µg/ml. 1 to 10 µl of protein sample were brought to a final volume of 10 µl with the respective protein lysis buffer and mixed with 990 µl 1x Roti-Quant. After 5 min at room temperature, the absorption at 595 nm was measured with a Nanophotometer (Implen) and the total protein amount for each sample was determined using the standard curve obtained by plotting the blank-corrected A_{595} for the protein standards against their total protein amount. The concentration of the protein lysate was obtained by dividing the total protein amount through the utilized volume of protein sample.

7.7.3 Mass spectrometry analysis of SPRR5

10 mg protein lysate from D5 differentiated keratinocytes was loaded onto a Vivaspin 100 kDa cutoff column (GE Healthcare) and centrifuged two hours at 4°C with 3000 rcf. The flow-through was mixed with 1/5 volume 5x laemmli buffer and separated onto a 15% SDS-PAGE with subsequent coomassie staining (7.7.5). The edges of the gel were removed and it was divided horizontally into 28 stripes based on the coomassie staining and each stripe was divided into ten equally sized pieces. Each gel piece was minced and transferred into a 2 ml micro tube (Eppendorf), washed for 30 min each with 950 µl 50 mM NH_4HCO_3,

50 mM NH_4HCO_3/acetonitrile (3/1), 10 mM NH_4HCO_3/acetonitrile (3/1), 10 mM NH_4HCO_3/ acetonitrile (1/1) and lyophilized. After reduction with 300 µl 1 mg/ml DTT dissolved in 100 mM NH_4HCO_3 (50°C, 1 hour), cysteines were alkylated with 200 µl 5 mg/ml Iodoacetamide in 50 mM NH_4HCO_3 (35 min, RT) and the gel pieces were washed and lyophilized again (see above). Next, proteins were subjected to an *in gel* tryptic digest overnight at 37°C with 1 µg Trypsin Gold mass spectrometry grade (Promega) in 50 mM NH_4HCO_3. Peptides were extracted twice with 150 µl 100 mM NH_4HCO_3, followed by one elution with 150 µl 100 mM NH_4HCO_3/ acetonitrile (2/1) and the obtained eluates were combined and lyophilized.

Further processing of samples, mass spectrometry measurements and protein identification was done by the mass spectrometry facility of Biochemistry I, University of Regensburg by Dr. Astrid Bruckmann and Eduard Hochmuth. In brief, data obtained from samples analyzed on the MaXiS mass spectrometer was transferred to MASCOT 2.5.1 using the Protein-Scape software 3.1.3 (Bruker Daltonics). MASCOT aligned the obtained data to the annotated proteins of the SWISS-PROT database.

For SRM measurements 100 fmol synthetic peptide (see Table 10) was spiked in before the trypsin digestion, all other steps were performed as described above but samples were analyzed on a QTRAP instrument. Data obtained from relative quantification of SPRR5 protein levels were first exported to Excel and ratios of spike in peptides were compared to measured SPRR5 peptides and multiplied with the known amount of 100 fmol peptide spike in. This amount of SPRR5 protein within the sample was then divided by the number of cells used for sample generation, to obtain the number of SPRR5 proteins within one cell.

7.7.4 Preparation of protein lysates from keratinocytes

Protein lysates were prepared by collecting the cells in RIPA buffer, incubation on ice for 15 minutes and centrifugation for 15 minutes at 4°C and full speed. The obtained supernatant was transferred into a new tube and this protein lysate was stored at -80°C until further use.

For mass spectrometry analysis, day 5 differentiated keratinocytes were detached and singularized as described in 7.6.2. After centrifugation (7 min, 4°C, 200 rcf), cells were washed once with ice-cold PBS (5 min, 4°C, 200 rcf) and transferred into a clean Eppendorf tube with 1 ml PBS. Cells were pelleted again (400 rcf, 4°C, 7 min) the PBS was completely removed and the pellet was snap frozen in liquid nitrogen. Following thawing on ice, cells were resuspended in protein lysis buffer and incubated on ice for 20 minutes. The protein lysate was obtained as the supernatant after centrifugation (15 min, 4°C, full speed).

7.7.5 SDS-PAGE analysis and Coomassie staining

Depending on the size of the protein of interest, proteins were separated on 10% or 15% SDS-polyacrylamide gels (composition see Table 36). To this end, protein samples were mixed with 1/5 volume of 5x laemmli buffer and denaturated at 95°C for 5 minutes. For size determination, 5 µl pre-stained Precision Plus Protein Standard Dual Color (Bio-Rad) was used and electrophoresis was performed with TGS at 100 V until the dye front reached the bottom of the gel.

Gels for mass spectrometry analysis were fixed with coomassie fixative (30 min, RT) washed with water (5 min, RT, 3 times) and stained with Bio-Safe Coomassie Stain (Bio-Rad) at 4°C overnight. Destaining with water was done until protein bands became visible and pictures were captured with the Odyssey Imaging System (LI-COR Biosciences) and analyzed with the Odyssey software (LI-COR Biosciences).

Table 36: Composition of SDS-PAGE gels

Stacking gel (4%)	Resolving gel (10%)	Resolving gel (15%)	Component
2.25 ml	3.2 ml	1.9 ml	Water
-	1.9 ml	1.9 ml	4x Resolving gel buffer
0.95 ml	-	-	4x Stacking gel buffer
0.5 ml	2.6 ml	3.9 ml	Acrylamid/Bis-solution 30% (37.5:1)
5 µl	4.5 µl	4.5 µl	TEMED
22.5 µl	45 µl	45 µl	10% APS

7.7.6 Western Blot analysis

7.5 to 30 µg total protein (7.7.4) from keratinocytes was separated via SDS-PAGE (7.7.5) and the proteins were subsequently transferred onto the Amersham Hybond-ECL membrane (GE Healthcare) by semi-dry blot (Bio-Rad system) using western blot transfer buffer and 13 V for one hour. Blocking was done in 5% milk powder in TBS-T for one hour at room temperature and primary antibodies (Table 6) were diluted in 5% milk powder in TBS-T and applied for one hour at room temperature or overnight at 4°C. After washing the membrane three times with TBS-T (5 min each, RT), secondary antibodies were diluted in TBS-T supplemented with 5% milk powder, added to the membrane and incubated for one hour at room temperature. Following three washing steps with TBS-T (5 min, RT), the signal was captured with the Odyssey Imaging System (LI-COR Biosciences) and analyzed with the Odyssey software (LI-COR Biosciences).

8 Publications

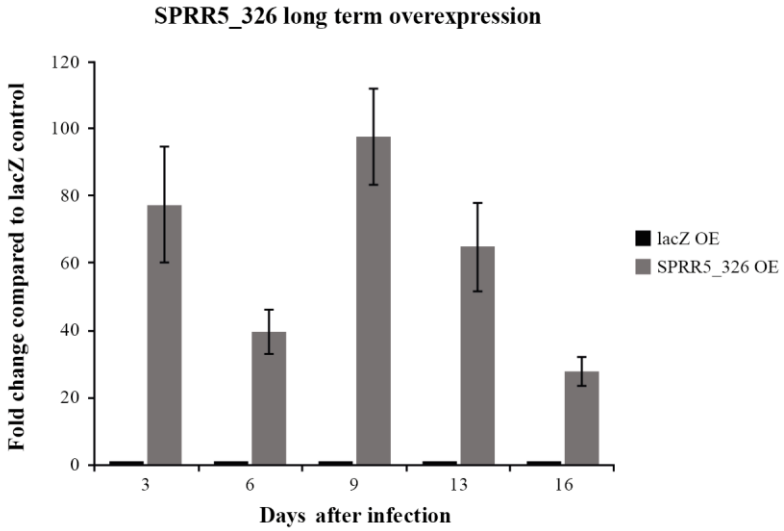
Ziegler, C., Graf, J., Faderl, S., Schedlbauer, J., Strieder, N., Förstl, B., Spang, R., Bruckmann, A., Merkl, R., Hombach, S., & Kretz, M. The long non-coding RNA LINC00941 and SPRR5 are novel regulators of human epidermal homeostasis Manuscript under review (EMBO Reports, 2018).

Ziegler, C. & Kretz, M. The More the Merrier—Complexity in Long Non-Coding RNA Loci. Front. Endocrinol. 8, (2017).

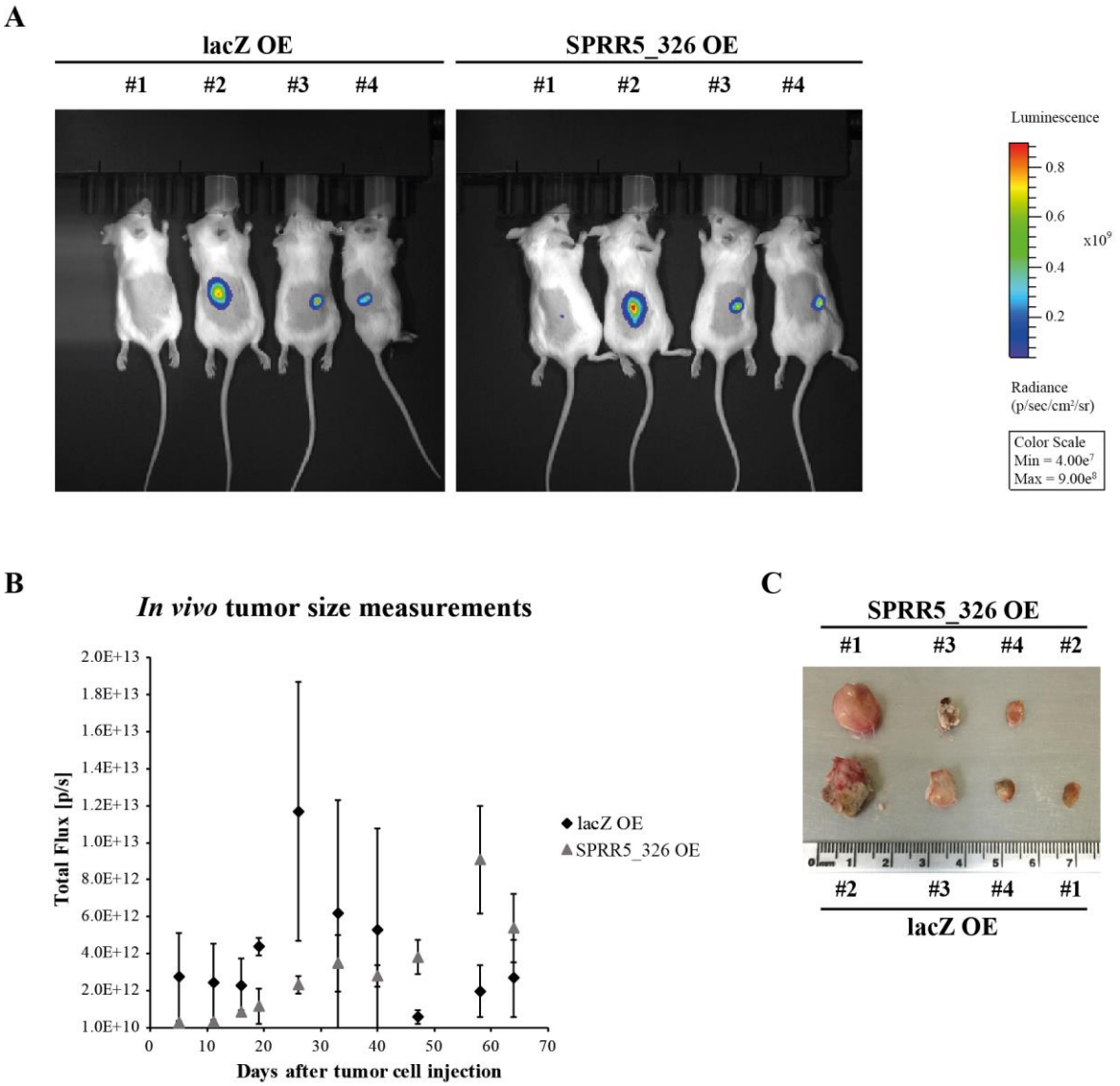
Dueck, A., Ziegler, C., Eichner, A., Berezikov, E. & Meister, G. microRNAs associated with the different human Argonaute proteins. Nucleic Acids Res. 40, 9850–9862 (2012).

9 Appendix

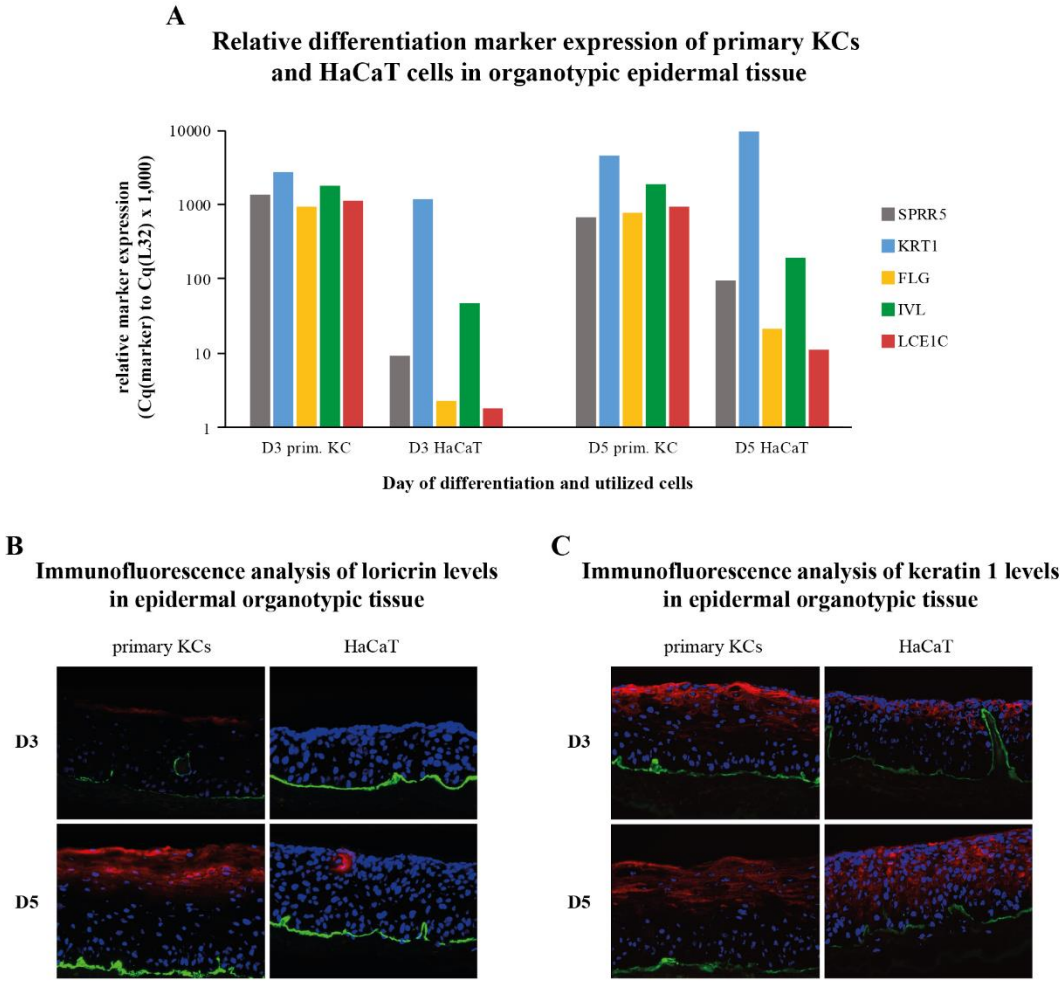
9.1 Supplementary Figures



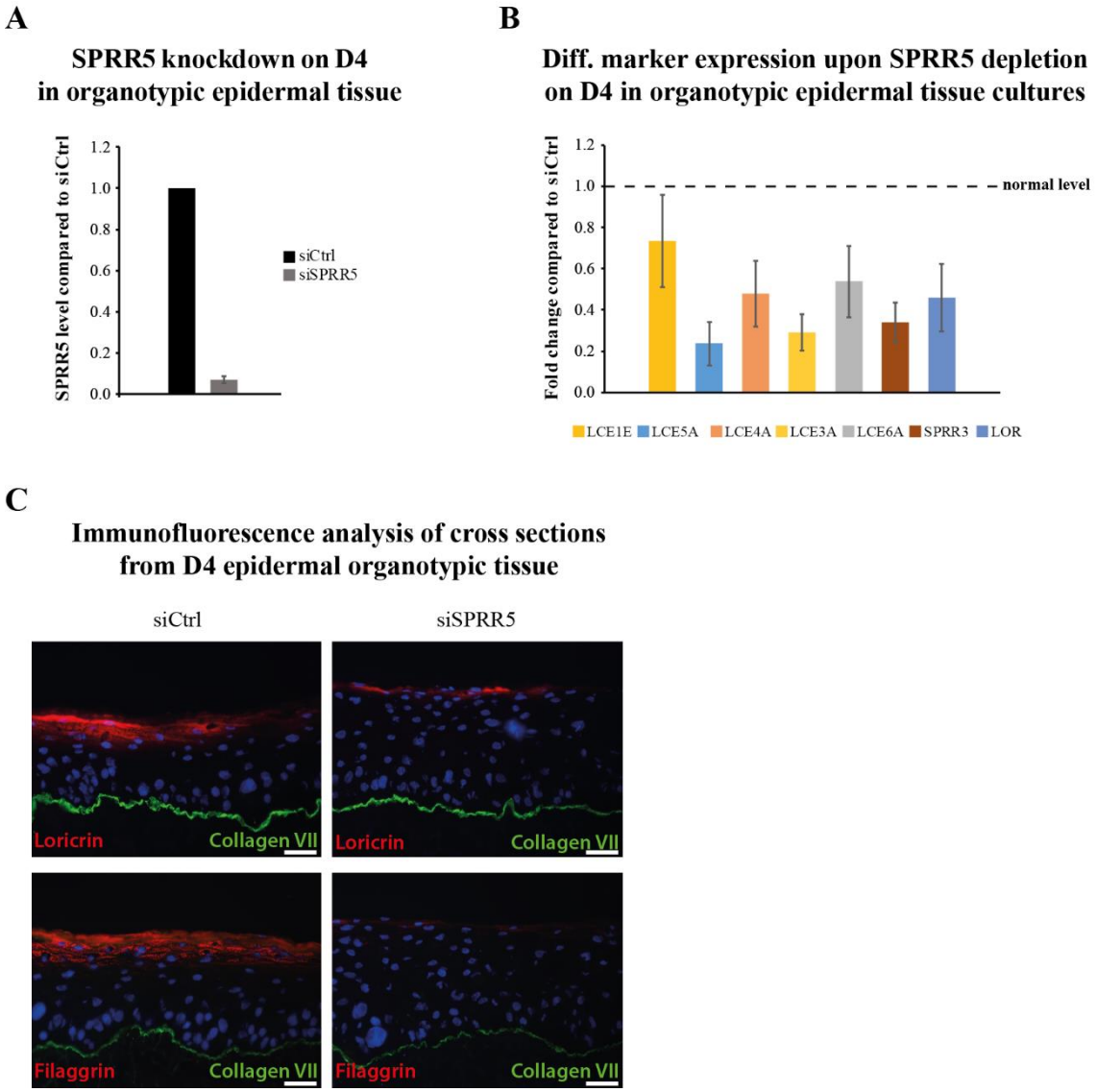
Supplementary Figure 1: Long term overexpression of SPRR5_326 in primary keratinocytes
SPRR5_326 levels in primary keratinocytes were analyzed by RT-qPCR at various timepoints after keratinocyte transduction (n=3). The obtained values were normalized to L32 and compared to the expression in lacZ overexpressing cells highlighting an adequate and long-lasting SPRR5_326 overexpression in primary keratinocytes. OE = overexpression



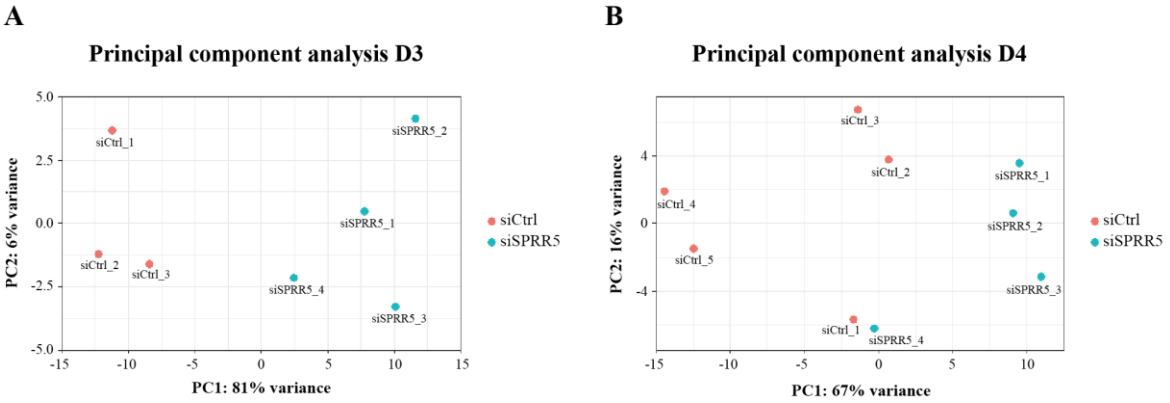
Supplementary Figure 2: *In vivo* luciferase measurements and explanted tumors
Exemplary picture of *in vivo* tumor size measurement on day 16 after tumor cell injection (A). (B) depicts the mean luciferase signals (Total flux) and the corresponding standard deviations for *in vivo* tumor size measurements of all biological replicates. Subfigure (C) shows the explanted tumors 85 days after tumor cell injection. OE = overexpression, p = photon, s = second



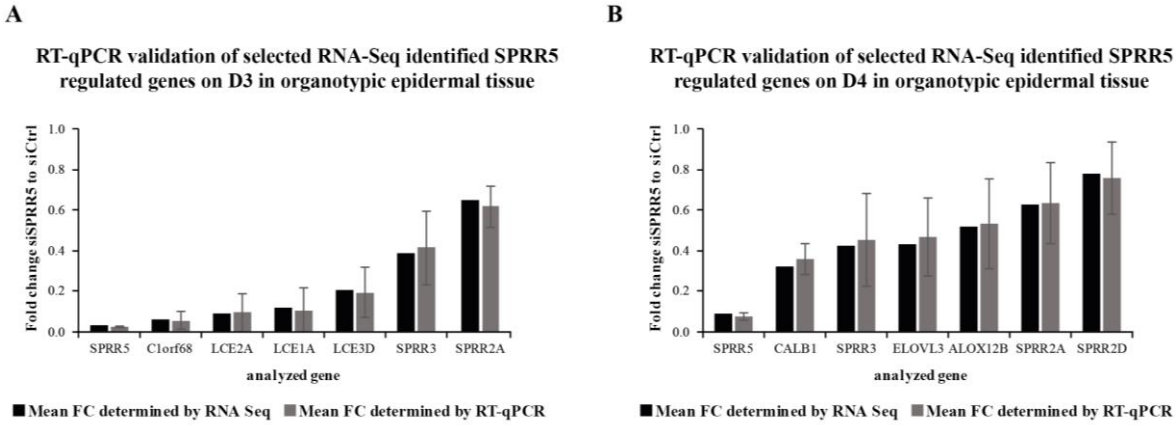
Supplementary Figure 3: HaCaT cells differentiate properly but slightly different than primary KCs
HaCaT cells exhibit slightly reduced amounts of differentiation marker transcripts compared to primary keratinocytes but differentiate properly (A). The relative expression of differentiation markers on day 3 and day 5 of differentiation was normalized to L32 expression as obtained by RT-qPCR (n=1). Exemplary pictures of immunofluorescence analysis from cross sections of the corresponding regenerated epidermal tissue on day 3 (top) or day 5 (bottom) reveals no detectable loricrin protein in HaCaT cells (B, loricrin is shown in red) but comparable amounts of the differentiation protein keratin 1 (C, keratin 1 is shown in red). Furthermore, collagen VII (green) is depicted for orientation as it separates the epidermis from the beneath lying dermis and nuclei are shown in blue. KC = keratinocyte, prim. = primary



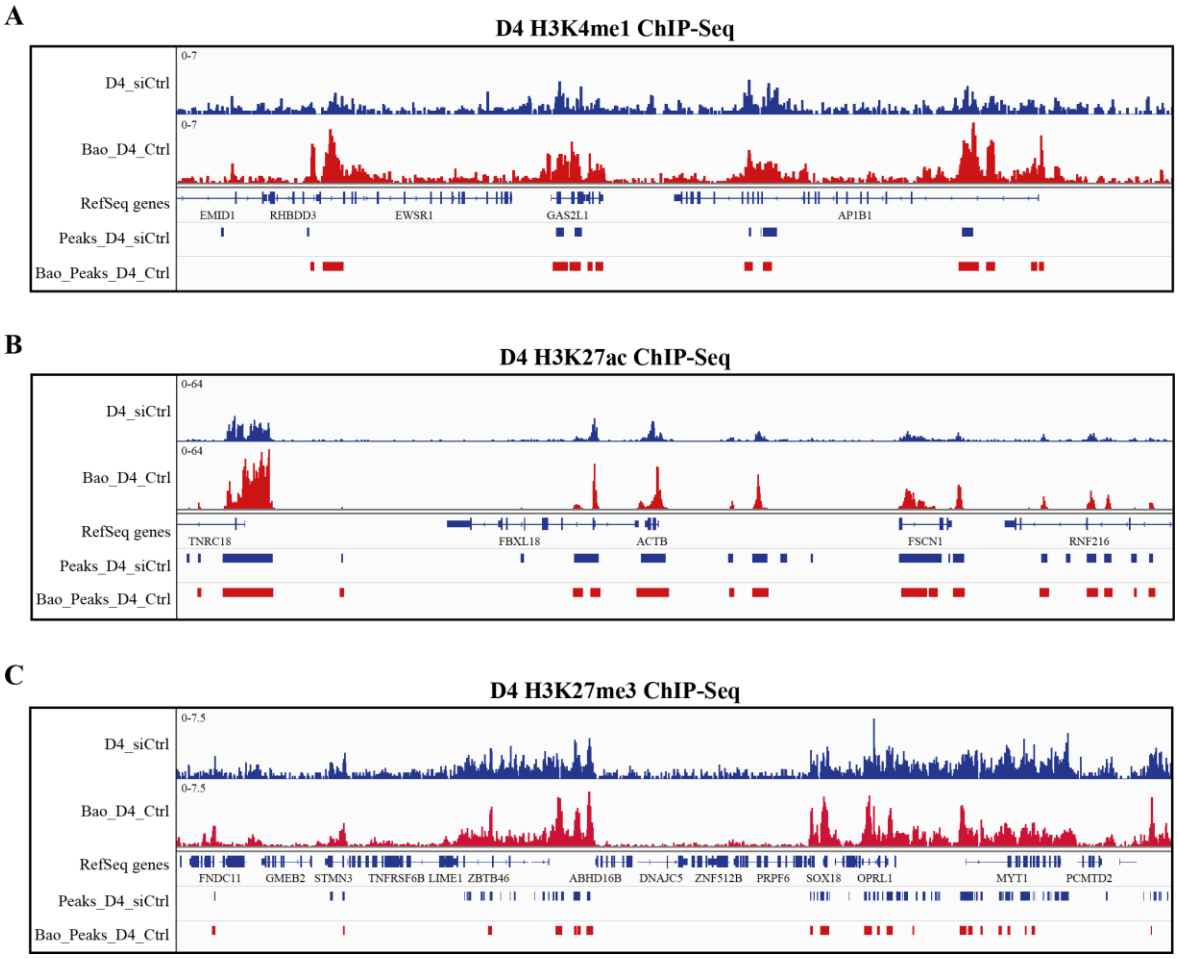
Supplementary Figure 4: SPRR5 controls epidermal tissue homeostasis on day 4 of differentiation
Efficient depletion of SPRR5 in organotypic epidermal tissue cultures (A) results in diminished expression of differentiation marker mRNAs (B) as obtained by RT-qPCR analysis (n=3-4). Exemplary pictures of immunofluorescence analysis of cross sections from the corresponding regenerated epidermal tissue (C) reveals reduced amounts of the differentiation proteins loricrin and filaggrin (shown in red) in SPRR5 depleted tissue. Collagen VII (green) is depicted for orientation as it separates the epidermis from the beneath lying dermis, nuclei are shown in blue and the scale bar (white) indicates 50 µm. Diff. = differentiation, Ctrl = control



Supplementary Figure 5: Principal component analysis of full transcriptome sequencing samples
Principal component analysis of the transcriptome upon SPRR5 depletion on D3 (A) and D4 (B) in regenerated organotypic epidermal tissue shows good clustering of control samples (siCtrl) and SPRR5 depleted tissue (siSPRR5) for both analyzed timepoints of differentiation. Ctrl = control, PC = principal component



Supplementary Figure 6: Validation of RNA-Seq results
DeSeq2 calculated fold changes between siSPRR5 depleted and control tissue from RNA-Seq analysis are plotted against fold changes obtained via RT-qPCR measurements for selected transcripts on day 3 of (A) or day 4 (B) of differentiation in epidermal organotypic tissue, exhibiting a decent correlation of the results (n=3-5). FC = fold change, Ctrl = control



Supplementary Figure 7: Comparison of ChIP-Seq reads and called peaks with results from Bao et al. Obtained mapped reads (top) and called peaks (bottom) for siCtrl ChIP-Seq samples on day 4 of differentiation (blue) were loaded into IGV and compared to already analyzed and published data from Bao et al. (red) for H3K4me1 (A), H3K27ac (B) and H3K27me3 (C), showing a good correlation between both datasets.

9.2 Sequences

9.2.1 SPRR5_326 sequence

5'AAGTTCCTTCTGGGCCAGCATCTTACACTCCACAGGCCAAGCAAAGCCCTGCTCACCAGTGCTTCTGGATGCAGCCACA
CTTCACACTTGCAGGAGAGGCAGCACCAGGCTGGCAGCTGGCGGCTTCCTCGAGCTCTACTCTTCTTAATTAGCAGCTGCA
GTTTCCATCTCAGACAAGGATCTGGTTCTAAACTATCTAATCATCTTTCCTGACATGCATCTACATTGCCCTGCTGGAAAC
ATGAGCGGCATGCACAGGCATCTTCTCCTACCACGCTTCCCCTCAGCTCCTGTTTTCAAACAATTAAAACAGAATTCCTCT
GGCA 3'

9.2.2 SPRR5 transcript sequence and genomic localization

5'AAACCCCTCCACACCTATCTGCTCCAGTCGCTGGTGAACCTTGTTGCAAGTCCAGCCCAGAATGTCTCAGCAGAAGCAG
AAGCAGTGTGCTCCCCCGCAGCAGTGCTGCCCCCCACCCCAGCAGCGCTGCCCCCCACCCCAGCAGTGCTGCCCCCCGCCT
CAACAGTGCTGCCCCCCACCCCAGCAGTGCTGCCCGCCGCCTCAACAATGCTGCCCCCCTCCCCAGCAGTGCTGTCCCCCA
CCCCAGCAGTGCTGCCCTCCACCTCAGCAGTACTGCCCCCCACCCCAACAGACCAAGCAGCCTTGCCAGCCCCCACCCAAG
TGCCAGGAGCCCTGTGCCCCCAAGTGCCCACCCCCTCAGCAGTGCCAGACGTCCAAGCAGAAGTAAGCCCTGGGCATGTGA
TCAGAGGCGAACCCCCAGGAAAAACAGAGCATGAAGTTCCTTCTGGGCCAGCATCTTACACTCCACAGGCCAAGCAAAGCC
CTGCTCACCAGTGCTTCTGGATGCAGCCACACTTCACACTTGCAGGAGAGGCAGCACCAGGCTGGCAGCTGGCGGCTTCCT
CGAGCTCTACTCTTCTTAATTAGCAGCTGCAGTTTCCATCTCAGACAAGGATCTGGTTCTAAACTATCTAATCATCTTTCC
TGACATGCATCTACATTGCCCTGCTGGAAACATGAGCGGCATGCACAGGCATCTTCTCCTACCACGCTTCCCCTCAGCTCC
TGTTTTCAAACAATTAAAACAGAATTCCTCTGGCA 3'

Table 37: Genomic localization of SPRR5 (hg38)

	Chromosome	Strand	Start	End
Exon1	1	+	152947206	152947248
Intron	1	+	152947249	152948536
Exon2	1	+	152948537	152949255

9.2.3 SPRR5 protein sequence and peptides for mass spectrometry

The protein sequence of the predicted human SPRR5 protein is given and the sequences of the four heavy labelled peptides for the SRM mass spectrometry approach are indicated in colors. Only the peptide indicated in green could be successfully detected via mass spectrometry and was used for protein quantification.

MSQQKQKQCAPPQQCCPPPQQRCPPPQQCCPPPQQCCPPPQQCCPPPQQCCPPPQQCCPP
PQQCCPPPQQYCPPPQQTKQPCQPPPKCQEPCAPKCPPPQQCQTSKQK

9.3 Lists of significantly altered genes upon SPRR5 depletion

Only those genes with -0.5 > $\log_2$(fold change) > 0.5 and an adjusted p-value < 0.05 are shown here.

9.3.1 SPRR5 regulated genes on day 3 in organotypic epidermis

Ensembl identifier	baseMean	log2FoldChange	p-value	adjusted p-value	gene symbol
ENSG00000283227	513.874134	-5.03141714	2.24E-114	4.11E-110	SPRR5
ENSG00000243627	12.1594968	-4.99622056	4.25E-08	1.50E-05	KCNE1
ENSG00000186207	41.8926237	-4.66514271	7.49E-13	1.10E-09	LCE5A
ENSG00000198854	342.418464	-4.03563816	1.76E-18	5.39E-15	C1orf68
ENSG00000187170	14.6175526	-3.67702734	4.47E-07	0.00010783	LCE4A
ENSG00000221878	37.2595988	-3.66566427	3.37E-14	6.18E-11	PSG7
ENSG00000124920	89.1559031	-3.66158625	6.28E-10	4.11E-07	MYRF
ENSG00000125571	149.659943	-3.57769005	2.90E-08	1.06E-05	IL37
ENSG00000180332	120.510814	-3.55744864	2.57E-22	1.18E-18	KCTD4
ENSG00000186226	660.621615	-3.50168296	4.71E-09	2.27E-06	LCE1E
ENSG00000187180	1414.49716	-3.49980149	1.17E-09	6.90E-07	LCE2C
ENSG00000187223	735.573059	-3.47860806	7.70E-09	3.28E-06	LCE2D
ENSG00000187173	940.33126	-3.46497331	3.94E-09	1.95E-06	LCE2A
ENSG00000235942	619.505669	-3.4485924	2.67E-12	3.49E-09	LCE6A
ENSG00000247011	9.7089836	-3.44278242	0.00011371	0.00993143	NA
ENSG00000176075	111.94385	-3.39795461	8.17E-11	7.35E-08	NA
ENSG00000159455	2178.89237	-3.34218831	2.58E-08	9.64E-06	LCE2B
ENSG00000243349	18.4093612	-3.31337165	8.16E-06	0.00121126	NA
ENSG00000240386	1084.01822	-3.28091094	6.72E-08	2.28E-05	LCE1F
ENSG00000244057	53.9550545	-3.26579291	5.34E-16	1.40E-12	LCE3C
ENSG00000172155	388.542911	-3.25807433	2.73E-07	7.27E-05	LCE1D
ENSG00000214491	31.9734555	-3.23706075	6.39E-12	7.82E-09	SEC14L6
ENSG00000177243	151.443432	-3.186702	3.07E-24	2.81E-20	DEFB103B
ENSG00000185962	315.045005	-3.15423796	7.45E-10	4.71E-07	LCE3A
ENSG00000215183	6.99525492	-3.13290739	0.00083908	0.04824319	MSMP
ENSG00000168952	25.7267698	-3.12873467	1.79E-10	1.43E-07	STXBP6
ENSG00000042062	55.7964831	-3.11053046	3.42E-15	7.84E-12	FAM65C
ENSG00000186844	1628.34182	-3.11029258	1.83E-07	5.08E-05	LCE1A
ENSG00000105143	8.01004631	-3.01050399	0.00048389	0.031472	SLC1A6
ENSG00000118137	11.2682561	-2.99693043	3.78E-06	0.00064225	APOA1
ENSG00000203786	5644.59336	-2.9743959	1.53E-07	4.59E-05	KPRP
ENSG00000263429	12.9688611	-2.95835945	5.43E-07	0.00012618	LINC00675
ENSG00000196734	2406.54552	-2.93286619	3.32E-07	8.45E-05	LCE1B
ENSG00000162399	8.61030677	-2.8017844	0.00037386	0.02587521	BSND
ENSG00000168703	1914.52298	-2.75023529	1.21E-14	2.46E-11	WFDC12
ENSG00000197084	2987.16501	-2.74213045	1.35E-07	4.28E-05	LCE1C
ENSG00000019102	13.1589975	-2.71164387	0.00010484	0.00938	VSIG2
ENSG00000273008	13.1133029	-2.62184407	0.00013062	0.01093888	NA
ENSG00000185966	2968.35942	-2.5496997	7.32E-09	3.20E-06	LCE3E
ENSG00000165799	1831.03649	-2.52981706	1.47E-08	5.62E-06	RNASE7
ENSG00000214313	183.629185	-2.510299	8.36E-07	0.00018478	AZGP1P1
ENSG00000136449	11.6208583	-2.45016881	4.97E-05	0.0052717	MYCBPAP
ENSG00000283376	30.2596033	-2.44775058	6.16E-08	2.13E-05	NA
ENSG00000233005	7.98194161	-2.44488673	0.00075574	0.04442631	NA
ENSG00000232679	25.8411909	-2.36776703	4.50E-06	0.0007233	NA
ENSG00000231672	40.6818906	-2.36135736	2.13E-13	3.55E-10	NA
ENSG00000163202	6984.72841	-2.27748273	3.58E-07	9.01E-05	LCE3D
ENSG00000160862	3790.91231	-2.26880303	2.97E-07	7.67E-05	AZGP1
ENSG00000172005	728.459707	-2.23362638	0.00061678	0.03783388	MAL
ENSG00000100302	43.4334226	-2.22032481	5.92E-05	0.00609651	RASD2
ENSG00000125355	24.4733055	-2.20829738	2.15E-06	0.00041404	TMEM255A
ENSG00000170054	43.8936574	-2.17245917	1.28E-07	4.12E-05	SERPINA9
ENSG00000224614	47.7805588	-2.1600126	4.13E-10	2.81E-07	NA
ENSG00000104327	30.4703162	-2.13004147	3.72E-10	2.73E-07	CALB1
ENSG00000102962	54.394796	-2.02846862	6.49E-09	3.03E-06	CCL22
ENSG00000094963	56.9429352	-1.98661688	1.41E-06	0.00028172	FMO2
ENSG00000174945	19.0782755	-1.98526116	3.81E-05	0.00423681	AMZ1
ENSG00000100167	12.688576	-1.97982956	0.00030271	0.02203167	SEPT3
ENSG00000145879	123.319464	-1.97385456	7.78E-05	0.00783644	SPINK7
ENSG00000131969	263.005535	-1.95755787	7.78E-13	1.10E-09	ABHD12B
ENSG00000181652	407.50275	-1.90093625	4.95E-05	0.0052717	ATG9B
ENSG00000169509	5322.98824	-1.85744307	3.49E-10	2.67E-07	CRCT1
ENSG00000164283	17.118721	-1.84289744	0.00085984	0.04921412	ESM1
ENSG00000185873	36.1534951	-1.78202614	0.00073671	0.04344696	TMPRSS11B
ENSG00000115488	39.2119949	-1.77503721	2.57E-06	0.00047187	NEU2
ENSG00000160360	484.526814	-1.77470658	5.93E-22	2.18E-18	GPSM1
ENSG00000073737	762.969632	-1.77399801	4.64E-23	2.84E-19	DHRS9
ENSG00000100292	2371.71075	-1.72362536	7.21E-07	0.00016318	HMOX1
ENSG00000183760	1235.17771	-1.71521881	3.78E-06	0.00064225	ACP7
ENSG00000110675	336.936545	-1.71202166	2.43E-05	0.00289689	ELMOD1

ENSG00000165495	43.6527979	-1.61326983	7.42E-05	0.00752237	PKNOX2
ENSG00000198691	44.5882514	-1.54304894	3.88E-05	0.00426125	ABCA4
ENSG00000183018	2637.04055	-1.52824071	4.35E-06	0.00070595	SPNS2
ENSG00000165091	20.5657198	-1.52422433	0.00020699	0.01601865	TMC1
ENSG00000164120	2401.66385	-1.51409686	4.64E-06	0.00074049	HPGD
ENSG00000237330	211.638197	-1.48760193	5.29E-11	5.39E-08	RNF223
ENSG00000244617	19016.2928	-1.48620801	0.00032795	0.02331385	ASPRV1
ENSG00000166105	180.75797	-1.47602632	7.13E-09	3.19E-06	GLB1L3
ENSG00000143536	3643.4723	-1.46502781	8.53E-05	0.00827934	CRNN
ENSG00000114854	38.5968371	-1.44218597	0.00031839	0.02290001	TNNC1
ENSG00000234478	31.821916	-1.4331129	1.83E-05	0.00227245	NA
ENSG00000197249	84.1911676	-1.42846432	1.55E-07	4.60E-05	SERPINA1
ENSG00000159899	180.478159	-1.42262795	1.66E-11	1.90E-08	NPR2
ENSG00000101443	22.2411063	-1.4185134	0.00071803	0.04289712	WFDC2
ENSG00000075673	7385.2923	-1.37920244	8.67E-07	0.00018826	ATP12A
ENSG00000261572	22.2935538	-1.378175	0.00047931	0.03128448	NA
ENSG00000163209	11418.8295	-1.36916956	4.58E-05	0.00491688	SPRR3
ENSG00000115590	479.246471	-1.36213215	9.87E-09	4.02E-06	IL1R2
ENSG00000197948	2320.26077	-1.34453354	1.12E-10	9.37E-08	FCHSD1
ENSG00000271573	22.2955124	-1.33763701	0.00027494	0.02048478	NA
ENSG00000154274	46.2106948	-1.33519099	1.20E-05	0.00170114	C4orf19
ENSG00000130751	164.744222	-1.31555802	1.47E-07	4.48E-05	NPAS1
ENSG00000178597	1622.73713	-1.29188729	0.00035677	0.02497517	PSAPL1
ENSG00000213904	91.873446	-1.28300954	2.08E-07	5.68E-05	LIPE-AS1
ENSG00000123689	113.314208	-1.27052753	9.51E-05	0.0087776	G0S2
ENSG00000204538	231.324728	-1.26407324	3.82E-07	9.40E-05	PSORS1C2
ENSG00000103056	4479.72834	-1.25036953	8.59E-05	0.00829472	SMPD3
ENSG00000096006	169.679239	-1.22337481	3.72E-05	0.00415999	CRISP3
ENSG00000106484	176.164652	-1.21240147	9.57E-07	0.00019953	MEST
ENSG00000204539	13611.3822	-1.20441971	5.01E-05	0.0052775	CDSN
ENSG00000185499	277.625812	-1.18962572	5.84E-07	0.00013393	MUC1
ENSG00000095637	45.0271788	-1.18077482	0.0005075	0.03243231	SORBS1
ENSG00000099822	45.3169922	-1.15953072	3.60E-05	0.00405442	HCN2
ENSG00000138678	163.113427	-1.14573701	4.11E-10	2.81E-07	GPAT3
ENSG00000167755	19407.6797	-1.13973624	7.89E-05	0.00790911	KLK6
ENSG00000276900	80.399488	-1.1330729	0.00015031	0.01219877	NA
ENSG00000158220	520.741592	-1.12915878	3.97E-06	0.00066114	ESYT3
ENSG00000196549	65.7502692	-1.12030302	1.44E-05	0.00189569	MME
ENSG00000185168	73.2856637	-1.10535611	0.00018183	0.01449962	NA
ENSG00000129437	119.49664	-1.10461539	1.21E-05	0.00170114	KLK14
ENSG00000072201	185.483973	-1.10063003	3.65E-09	1.91E-06	LNX1
ENSG00000179859	359.066158	-1.09805558	1.20E-07	3.93E-05	LOC284023
ENSG00000170961	283.020574	-1.09437712	2.98E-05	0.0034179	HAS2
ENSG00000088386	1553.16477	-1.08958191	6.60E-09	3.03E-06	SLC15A1
ENSG00000066735	33.5360524	-1.08833573	0.00026019	0.01955784	KIF26A
ENSG00000240891	124.053678	-1.08498199	0.00013831	0.01132459	PLCXD2
ENSG00000163803	219.164174	-1.08461901	3.32E-06	0.00057979	PLB1
ENSG00000236263	33.8403849	-1.07444381	0.00010322	0.00928034	NA
ENSG00000254966	57.8679958	-1.06431572	0.00086677	0.04937087	NA
ENSG00000119125	370.366937	-1.05281545	5.18E-07	0.00012173	GDA
ENSG00000166920	263.784621	-1.05218484	4.09E-06	0.00067654	C15orf48
ENSG00000130475	251.594059	-1.04896736	9.26E-05	0.00869217	FCHO1
ENSG00000105825	665.040185	-1.03391214	1.06E-08	4.21E-06	TFPI2
ENSG00000238266	78.7081231	-1.01851775	2.60E-06	0.00047246	LINC00707
ENSG00000151117	3315.16386	-1.01120508	2.81E-07	7.36E-05	TMEM86A
ENSG00000146674	407.192089	-1.00569462	9.97E-05	0.00909973	IGFBP3
ENSG00000144834	195.363505	-0.98733782	6.82E-08	2.28E-05	TAGLN3
ENSG00000036672	501.451051	-0.98722765	1.41E-07	4.38E-05	USP2
ENSG00000177628	2573.03642	-0.98609951	3.43E-11	3.71E-08	GBA
ENSG00000253125	174.046833	-0.98547045	0.00023626	0.01805531	NA
ENSG00000175352	57.4104737	-0.9796569	0.00057586	0.03592473	NRIP3
ENSG00000160766	626.793421	-0.95477444	1.64E-09	9.40E-07	GBAP1
ENSG00000122643	463.17592	-0.95122771	1.11E-09	6.77E-07	NT5C3A
ENSG00000173926	397.278437	-0.95074763	2.50E-05	0.00295054	MARCH3
ENSG00000159516	3222.29042	-0.90712328	1.25E-08	4.87E-06	SPRR2G
ENSG00000213492	69.1360923	-0.90653145	1.43E-05	0.00189569	NA
ENSG00000171476	7257.99242	-0.90402513	1.38E-05	0.00189569	HOPX
ENSG00000274276	66.3224885	-0.89832109	9.29E-05	0.00869217	CBSL
ENSG00000157368	280.223859	-0.88764079	5.45E-05	0.00567685	IL34
ENSG00000138623	630.023455	-0.87470739	8.19E-06	0.00121126	SEMA7A
ENSG00000184148	1596.9211	-0.87461564	8.72E-07	0.00018826	SPRR4
ENSG00000149328	276.145628	-0.84681181	0.0001201	0.01029329	GLB1L2
ENSG00000122042	756.411871	-0.82954603	9.35E-06	0.00136161	UBL3
ENSG00000204618	1046.37417	-0.82734807	2.15E-05	0.0025917	RNF39
ENSG00000135821	5089.88469	-0.82466367	8.24E-11	7.35E-08	GLUL
ENSG00000088826	549.149267	-0.82287093	2.51E-05	0.00295054	SMOX
ENSG00000035862	1356.79083	-0.82232271	4.70E-07	0.00011189	TIMP2
ENSG00000026025	4409.40616	-0.81166276	0.00041763	0.02775285	VIM
ENSG00000268307	347.770148	-0.8051815	2.54E-05	0.00297107	NA
ENSG00000108244	5044.55271	-0.80251717	3.84E-07	9.40E-05	KRT23
ENSG00000180667	690.793459	-0.79422615	0.00068434	0.04142399	YOD1
ENSG00000187210	176.932739	-0.79243759	0.00083211	0.04799303	GCNT1
ENSG00000173846	858.450962	-0.78566642	3.52E-09	1.90E-06	PLK3

ENSG00000116741	167.147564	-0.7826287	8.34E-05	0.00817244	RGS2
ENSG00000167779	2096.67669	-0.77862704	2.45E-09	1.36E-06	IGFBP6
ENSG00000150471	96.2327106	-0.77753359	0.00039483	0.02682092	ADGRL3
ENSG00000236824	152.63427	-0.77336052	0.00041693	0.02775285	BCYRN1
ENSG00000175315	2879.58924	-0.76722472	0.00038474	0.02642888	CST6
ENSG00000162040	346.101646	-0.7558579	0.00010233	0.00928034	HS3ST6
ENSG00000137033	314.684776	-0.74407009	0.00056366	0.03528336	IL33
ENSG00000091428	100.574973	-0.74133426	0.00027587	0.02048478	RAPGEF4
ENSG00000159871	3016.11459	-0.72397918	0.00023833	0.01808303	LYPD5
ENSG00000047346	742.845322	-0.72280876	2.32E-05	0.00278589	FAM214A
ENSG00000166828	132.101477	-0.72115269	0.00033537	0.02373809	SCNN1G
ENSG00000128510	3714.24773	-0.72095526	4.16E-06	0.00068091	CPA4
ENSG00000059804	152.528818	-0.71789709	8.15E-05	0.00803722	SLC2A3
ENSG00000148346	8492.37768	-0.71105145	0.00012571	0.010674	LCN2
ENSG00000100889	1015.4049	-0.70798278	1.35E-06	0.00027457	PCK2
ENSG00000103811	3398.2072	-0.70548201	1.70E-05	0.00213183	CTSH
ENSG00000172478	2777.20801	-0.70213032	1.47E-05	0.00192131	C2orf54
ENSG00000198853	492.367792	-0.6957467	1.60E-07	4.61E-05	RUSC2
ENSG00000246100	84.1336163	-0.69327602	0.00032676	0.02331385	NA
ENSG00000137962	865.311892	-0.68898797	0.00058799	0.03655729	ARHGAP29
ENSG00000006652	710.396428	-0.67617755	0.00012968	0.01091011	IFRD1
ENSG00000162852	457.580945	-0.67386283	9.27E-05	0.00869217	CNST
ENSG00000102312	951.884337	-0.66880518	7.33E-06	0.00112102	PORCN
ENSG00000102879	169.360979	-0.66609702	0.00018929	0.01496463	CORO1A
ENSG00000255501	2374.57461	-0.66282544	0.00040185	0.0270966	CARD18
ENSG00000069011	1329.72135	-0.65608318	0.00050572	0.03243231	PITX1
ENSG00000005001	866.47503	-0.64898823	9.53E-07	0.00019953	PRSS22
ENSG00000142623	523.197329	-0.64567236	1.41E-05	0.00189569	PADI1
ENSG00000189433	1066.9903	-0.6403917	1.68E-06	0.00033037	GJB4
ENSG00000101132	163.424112	-0.64031647	0.00019848	0.01555682	PFDN4
ENSG00000155324	828.778307	-0.63041973	2.41E-07	6.49E-05	GRAMD3
ENSG00000130818	488.155397	-0.62665279	8.85E-05	0.00849919	ZNF426
ENSG00000228903	260.712979	-0.6257139	0.0001282	0.01083593	NA
ENSG00000241794	2746.21293	-0.62238799	0.00010644	0.00947687	SPRR2A
ENSG00000164294	316.088318	-0.62084145	5.73E-05	0.00594201	GPX8
ENSG00000243137	149.609177	-0.62082158	0.00060455	0.03720812	PSG4
ENSG00000143845	310.389168	-0.61710272	5.25E-05	0.00550589	ETNK2
ENSG00000057657	1262.42699	-0.61685601	0.00028237	0.02079875	PRDM1
ENSG00000142227	841.058036	-0.60239327	9.67E-05	0.00886868	EMP3
ENSG00000143369	4379.1028	-0.58940462	1.86E-05	0.00228696	ECM1
ENSG00000164442	1598.79926	-0.58731186	7.96E-05	0.00793182	CITED2
ENSG00000158158	968.605702	-0.58448369	0.00024881	0.01877961	CNNM4
ENSG00000151498	889.302904	-0.57349128	3.89E-06	0.00065491	ACAD8
ENSG00000169715	590.515775	-0.57044187	0.00077482	0.04540225	MT1E
ENSG00000175121	3087.23372	-0.56371228	0.00044741	0.02941182	WFDC5
ENSG00000104783	248.813138	-0.56149299	0.00072765	0.04305114	KCNN4
ENSG00000135069	1348.93701	-0.55987074	0.00021479	0.01655238	PSAT1
ENSG00000177406	173.149081	-0.55931081	0.00030716	0.02226727	LOC100049716
ENSG00000136840	632.952461	-0.55001	3.28E-06	0.0005777	ST6GALNAC4
ENSG00000136943	6827.53223	-0.54727274	0.00060228	0.03719308	CTSV
ENSG00000070669	573.242	-0.54128676	0.00012569	0.010674	ASNS
ENSG00000261068	772.807209	-0.53007623	1.43E-05	0.00189569	NA
ENSG00000104267	4119.18146	-0.52801134	6.30E-06	0.00097119	CA2
ENSG00000177427	434.512944	-0.51431702	8.38E-05	0.00817244	MIEF2
ENSG00000150782	2052.10969	-0.50697654	3.00E-06	0.0005391	IL18
ENSG00000183696	2636.8153	-0.50288994	1.14E-05	0.00163169	UPP1
ENSG00000146477	663.870801	0.5002563	0.00017356	0.01396161	SLC22A3
ENSG00000179403	1541.51448	0.51352792	1.41E-06	0.00028172	VWA1
ENSG00000185507	1562.78856	0.51653243	5.41E-06	0.0008559	IRF7
ENSG00000143819	1077.19323	0.5215535	8.28E-06	0.00121485	EPHX1
ENSG00000172456	258.611804	0.52242949	0.00042631	0.02822702	FGGY
ENSG00000149260	163.380044	0.53586928	0.00069581	0.04197956	CAPN5
ENSG00000025708	11429.792	0.54897168	6.07E-06	0.00094395	TYMP
ENSG00000120149	215.075765	0.57370349	0.00080254	0.04672834	MSX2
ENSG00000158106	159.772803	0.60546325	0.000555	0.0348606	RHPN1
ENSG00000168016	581.446296	0.61378149	0.00039648	0.02683353	TRANK1
ENSG00000104722	393.712129	0.61994355	0.00027032	0.02023634	NEFM
ENSG00000139629	4693.40961	0.62605537	1.32E-05	0.00183658	GALNT6
ENSG00000205978	2940.29952	0.63629516	2.34E-06	0.0004422	NYNRIN
ENSG00000261150	2767.02193	0.64252751	4.21E-05	0.00457346	EPPK1
ENSG00000147883	8539.40571	0.64343688	0.0003378	0.02373809	CDKN2B
ENSG00000135218	1997.65758	0.66570355	0.0001334	0.01102686	CD36
ENSG00000185215	2133.26067	0.6664111	4.06E-08	1.46E-05	TNFAIP2
ENSG00000127954	2329.66337	0.68886045	1.40E-05	0.00189569	STEAP4
ENSG00000184371	159.796742	0.7033708	7.82E-06	0.00117629	CSF1
ENSG00000183486	764.814869	0.72060939	1.61E-07	4.61E-05	MX2
ENSG00000138606	979.872692	0.72163074	4.27E-05	0.00460813	SHF
ENSG00000105357	3950.88025	0.74732363	8.41E-11	7.35E-08	MYH14
ENSG00000125285	160.52706	0.7644877	0.00010297	0.00928034	SOX21
ENSG00000114812	527.398531	0.76776521	1.63E-05	0.00208952	VIPR1
ENSG00000158966	295.016632	0.76873403	1.48E-05	0.00192148	CACHD1
ENSG00000064655	66.4398786	0.88879888	0.00064292	0.03904556	EYA2
ENSG00000105523	101.777971	0.91976755	0.00050646	0.03243231	FAM83E

ENSG00000167972	53.6955493	0.9653959	0.00061969	0.03788557	ABCA3
ENSG00000138798	85.590101	0.98302478	4.12E-05	0.00450006	EGF
ENSG00000188613	170.404737	1.00732483	3.21E-06	0.00057168	NANOS1
ENSG00000243284	156.515732	1.09311829	1.22E-05	0.00170114	VSIG8
ENSG00000112394	62.4966976	1.14292158	0.00011959	0.01029329	SLC16A10
ENSG00000132854	295.377735	1.18899439	2.05E-05	0.00250158	KANK4
ENSG00000172867	6782.94035	1.21204235	3.78E-09	1.93E-06	KRT2
ENSG00000231574	47.5220821	1.21622261	1.65E-05	0.00209949	KCCAT211
ENSG00000159450	437.808533	1.28111414	9.72E-09	4.02E-06	TCHH
ENSG00000167157	127.69851	1.29487775	0.00029359	0.02145349	PRRX2
ENSG00000064886	36.875197	1.32706917	8.14E-05	0.00803722	CHI3L2
ENSG00000169116	193.343707	1.42100733	9.34E-07	0.00019923	PARM1
ENSG00000002745	16.2110806	1.80108235	0.00020048	0.0155802	WNT16

9.3.2 SPRR5 regulated genes on day 4 in organotypic epidermis

Ensembl identifier	baseMean	log2FoldChange	p-value	adjusted p-value	gene symbol
ENSG00000283227	513.874134	-3.52298464	4.90E-101	8.65E-97	SPRR5
ENSG00000241717	12.5709246	-2.882494	1.47E-05	0.00166099	NA
ENSG00000276241	15.8355602	-2.41705731	2.45E-05	0.00244803	NA
ENSG00000084674	14.5669241	-2.32229112	4.27E-05	0.00371816	APOB
ENSG00000124920	89.1559031	-2.32198659	5.54E-06	0.00085138	MYRF
ENSG00000110799	297.703255	-2.18280648	6.62E-07	0.00017723	VWF
ENSG00000100341	70.1883166	-1.94433734	0.00075354	0.0266826	PNPLA5
ENSG00000247011	9.7089836	-1.91520194	0.00172691	0.04866732	NA
ENSG00000101049	42.6212109	-1.82589092	0.00179577	0.0499739	SGK2
ENSG00000186912	20.5864765	-1.81720002	0.00139605	0.04177176	P2RY4
ENSG00000186207	41.8926237	-1.7345105	3.37E-06	0.00056793	LCE5A
ENSG00000204195	12.0344511	-1.69067923	0.00115233	0.03649044	AWAT1
ENSG00000104327	30.4703162	-1.64243628	1.90E-07	6.72E-05	CALB1
ENSG00000094963	56.9429352	-1.63069487	1.15E-06	0.00027123	FMO2
ENSG00000266274	9.94559428	-1.60897033	0.00141376	0.04205584	NA
ENSG00000073737	762.969632	-1.60730037	2.29E-24	2.02E-20	DHRS9
ENSG00000243627	12.1594968	-1.56556206	0.00035991	0.01605963	KCNE1
ENSG00000214491	31.9734555	-1.55158262	1.67E-05	0.00180649	SEC14L6
ENSG00000140678	89.1101808	-1.5125636	0.00082082	0.02838344	ITGAX
ENSG00000198691	44.5882514	-1.50592218	6.22E-07	0.00016901	ABCA4
ENSG00000215853	4219.94309	-1.46893694	3.88E-08	2.29E-05	RPTN
ENSG00000235091	14.426802	-1.45896818	0.0008342	0.02873369	NA
ENSG00000168952	25.7267698	-1.44687893	6.78E-05	0.00522861	STXBP6
ENSG00000178597	1622.73713	-1.43189091	6.63E-06	0.00096772	PSAPL1
ENSG00000019102	13.1589975	-1.41747737	0.00119659	0.03742243	VSIG2
ENSG00000167046	37.0673009	-1.41223095	2.79E-05	0.00267723	NA
ENSG00000214456	40.2181842	-1.40402265	0.00093385	0.03084324	PLIN5
ENSG00000276317	42.9526497	-1.39927405	0.00038227	0.01673211	NA
ENSG00000230102	10.080814	-1.39692844	0.00087014	0.02934239	LOC285389
ENSG00000185873	36.1534951	-1.38678281	4.16E-05	0.00370175	TMPRSS11B
ENSG00000183166	18.9273537	-1.37520096	0.00043251	0.01828334	CALN1
ENSG00000172164	178.681825	-1.34370582	4.25E-05	0.00371571	SNTB1
ENSG00000249307	14.2818237	-1.32812432	0.0007132	0.02556252	NA
ENSG00000082397	63.7529462	-1.32790744	1.13E-06	0.00027053	EPB41L3
ENSG00000118160	17.173707	-1.32127809	0.00057079	0.02210877	SLC8A2
ENSG00000255173	84.992939	-1.32089199	1.57E-05	0.00174229	NA
ENSG00000279693	53.0659157	-1.31826876	0.00014886	0.00878874	NA
ENSG00000283376	30.2596033	-1.30534995	0.00041646	0.01777497	NA
ENSG00000158014	27.7030009	-1.30521812	0.00121937	0.03783549	SLC30A2
ENSG00000007237	1307.95128	-1.29232872	1.20E-11	2.36E-08	GAS7
ENSG00000143520	34458.6098	-1.27006184	7.47E-05	0.00553187	FLG2
ENSG00000125355	24.4733055	-1.26336506	0.00062969	0.02349715	TMEM255A
ENSG00000163209	11418.8295	-1.23447676	2.87E-05	0.00274321	SPRR3
ENSG00000197046	33.1024134	-1.22527797	0.00021345	0.01113551	SIGLEC15
ENSG00000113739	339.737227	-1.2196126	0.00020539	0.01083364	STC2
ENSG00000119457	298.261118	-1.21562788	0.00055696	0.02180801	SLC46A2
ENSG00000112303	38.9028275	-1.21475454	2.60E-06	0.00049404	VNN2
ENSG00000119915	182.198162	-1.21353617	0.00103098	0.03324334	ELOVL3
ENSG00000230628	37.540418	-1.21110241	3.15E-06	0.00054073	NA
ENSG00000224614	47.7805588	-1.20744124	3.28E-05	0.00308577	NA
ENSG00000166105	180.75797	-1.19908626	5.81E-08	2.93E-05	GLB1L3
ENSG00000168703	1914.52298	-1.19412949	0.00012224	0.00779751	WFDC12
ENSG00000119125	370.366937	-1.18797255	1.47E-10	1.74E-07	GDA
ENSG00000277496	205.244401	-1.1806761	5.88E-06	0.00089486	NA
ENSG00000182261	917.255682	-1.16301754	4.26E-10	4.70E-07	NLRP10
ENSG00000271573	22.2955124	-1.16161623	7.34E-05	0.00553185	NA
ENSG00000108578	1498.92205	-1.16038456	6.08E-05	0.0048165	BLMH
ENSG00000205363	378.365549	-1.15055165	1.94E-12	4.89E-09	C15orf59
ENSG00000278952	324.699909	-1.14392128	0.00075239	0.0266826	NA
ENSG00000181577	44.3718918	-1.1436787	1.14E-05	0.00139273	C6orf223
ENSG00000167769	1712.69513	-1.132983	0.00011371	0.00752528	ACER1
ENSG00000224721	50.0125041	-1.13022085	0.00026459	0.01298703	LOC102724153
ENSG00000258689	26.9410759	-1.11654291	0.00053612	0.02119289	LINC01269
ENSG00000244057	53.9550545	-1.10561843	5.32E-05	0.00434881	LCE3C

ENSG00000136010	185.550825	-1.0962911	9.60E-09	6.52E-06	ALDH1L2
ENSG00000234478	31.821916	-1.09500955	4.30E-05	0.00372032	NA
ENSG00000183091	244.778047	-1.07046776	2.19E-05	0.00223298	NEB
ENSG00000108830	68.9619276	-1.06688844	0.00023834	0.01203292	RND2
ENSG00000135917	174.916841	-1.06102983	2.73E-05	0.0026533	SLC19A3
ENSG00000103740	401.663782	-1.05733956	1.52E-06	0.00033469	ACSBG1
ENSG00000188100	237.86449	-1.05712793	0.00070597	0.02546623	FAM25A
ENSG00000171476	7257.99242	-1.05542329	7.77E-09	5.49E-06	HOPX
ENSG00000103056	4479.72834	-1.05524626	0.00016093	0.00929276	SMPD3
ENSG00000221878	37.2595988	-1.05402921	0.00026779	0.01303551	PSG7
ENSG00000084710	105.343343	-1.05186365	8.04E-06	0.00108466	EFR3B
ENSG00000183479	678.815826	-1.04865851	0.00067885	0.02483493	TREX2
ENSG00000268601	99.0395579	-1.04662294	8.73E-07	0.00022049	NA
ENSG00000090512	495.387459	-1.04294039	1.81E-06	0.00038047	FETUB
ENSG00000185168	73.2856637	-1.04130514	6.63E-05	0.00514194	NA
ENSG00000188051	27.5558842	-1.03498169	0.00070649	0.02546623	TMEM221
ENSG00000184459	1714.29598	-1.0326541	5.58E-08	2.93E-05	BPIFC
ENSG00000126233	2400.69116	-1.03264763	0.00016385	0.00940019	SLURP1
ENSG00000143536	3643.4723	-1.02711876	0.00170811	0.04840647	CRNN
ENSG00000160200	129.506232	-1.02388358	2.12E-06	0.00042501	CBS
ENSG00000180332	120.510814	-1.02190088	4.20E-05	0.0037098	KCTD4
ENSG00000088053	18.7413241	-1.0203505	0.0009963	0.03236142	GP6
ENSG00000102962	54.394796	-1.01844101	7.16E-05	0.00547343	CCL22
ENSG00000088386	1553.16477	-1.00703759	1.08E-09	1.00E-06	SLC15A1
ENSG00000106302	151.894984	-0.99730188	0.00016616	0.00944664	HYAL4
ENSG00000254951	36.7381632	-0.98975661	0.00015779	0.00918673	LOC283299
ENSG00000173926	397.278437	-0.98723731	7.46E-07	0.00019388	MARCH3
ENSG00000151117	3315.16386	-0.98184685	1.34E-08	8.48E-06	TMEM86A
ENSG00000185499	277.625812	-0.97960625	4.11E-06	0.00066609	MUC1
ENSG00000189377	37.4349211	-0.97514591	0.00029016	0.01384258	CXCL17
ENSG00000128165	185.777317	-0.97048878	9.36E-08	4.24E-05	ADM2
ENSG00000136697	92.1102446	-0.9702897	1.75E-07	6.43E-05	IL1F10
ENSG00000139438	58.7409761	-0.9694445	3.28E-05	0.00308577	FAM222A
ENSG00000075673	7385.2923	-0.96418263	8.96E-05	0.00631686	ATP12A
ENSG00000159871	3016.11459	-0.96309555	2.77E-08	1.69E-05	LYPD5
ENSG00000253125	174.046833	-0.95584632	7.37E-05	0.00553185	NA
ENSG00000135678	1913.84923	-0.95079813	0.00012826	0.00797992	CPM
ENSG00000179477	4166.84704	-0.95054422	0.00019571	0.01044776	ALOX12B
ENSG00000136696	182.027696	-0.94817105	0.00016449	0.00940603	IL36B
ENSG00000130475	251.594059	-0.94472845	7.48E-05	0.00553187	FCHO1
ENSG00000254966	57.8679958	-0.94333673	0.0003731	0.01642081	NA
ENSG00000136695	3542.98192	-0.94297217	5.29E-09	4.25E-06	IL36RN
ENSG00000130751	164.744222	-0.93816482	7.34E-06	0.00104598	NPAS1
ENSG00000169026	130.643406	-0.93308035	0.00016321	0.00939375	MFSD7
ENSG00000099960	63.3117146	-0.93014997	3.41E-05	0.00313871	SLC7A4
ENSG00000184060	535.599272	-0.92198644	9.30E-06	0.00119916	ADAP2
ENSG00000186115	67.322133	-0.9209359	0.00022615	0.01154921	CYP4F2
ENSG00000131969	263.005535	-0.91065581	0.00010914	0.00733293	ABHD12B
ENSG00000188373	1813.12752	-0.90992978	1.48E-07	5.82E-05	C10orf99
ENSG00000263823	118.673433	-0.90867887	1.02E-05	0.00128423	NA
ENSG00000189090	68.9723657	-0.90863726	8.93E-06	0.00115975	FAM25G
ENSG00000240891	124.053678	-0.90530708	0.00024507	0.01226734	PLCXD2
ENSG00000162040	346.101646	-0.90240678	9.65E-08	4.26E-05	HS3ST6
ENSG00000186806	3712.46855	-0.90077543	2.68E-06	0.0005038	VSIG10L
ENSG00000173239	796.873851	-0.89903995	0.00068663	0.0250159	LIPM
ENSG00000103569	2065.16433	-0.8821354	6.24E-08	2.98E-05	AQP9
ENSG00000204539	13611.3822	-0.88210267	0.00072234	0.02575331	CDSN
ENSG00000105427	13390.0567	-0.88008507	0.00043658	0.01832405	CNFN
ENSG00000163221	1528.94304	-0.87512587	2.29E-07	7.36E-05	S100A12
ENSG00000182489	542.511261	-0.87484192	3.36E-07	9.73E-05	XKRX
ENSG00000204175	153.909115	-0.86950286	1.28E-06	0.00029333	GPRIN2
ENSG00000106327	46.7176389	-0.8684959	9.15E-05	0.00641395	TFR2
ENSG00000204421	3930.82629	-0.86441024	2.74E-06	0.00050485	LY6G6C
ENSG00000171812	33.874487	-0.85831757	0.00071923	0.0257262	COL8A2
ENSG00000213963	183.603614	-0.85562272	1.21E-05	0.00143864	LOC100130691
ENSG00000146192	82.9910417	-0.85542677	1.58E-05	0.00174631	FGD2
ENSG00000269855	322.702607	-0.85262452	0.00010854	0.00732036	RNF225
ENSG00000171711	56.6438482	-0.84723865	0.00032032	0.01484644	DEFB4A
ENSG00000169896	205.272045	-0.84485662	0.00021364	0.01113551	ITGAM
ENSG00000283167	126.832279	-0.84373593	0.00013337	0.00812877	NA
ENSG00000179388	816.375548	-0.84352611	5.09E-05	0.0042343	EGR3
ENSG00000169509	5322.98824	-0.83998888	0.0012205	0.03783549	CRCT1
ENSG00000140479	2270.3876	-0.83769404	4.91E-11	7.89E-08	PCSK6
ENSG00000106484	176.164652	-0.83218157	0.00011723	0.00768425	MEST
ENSG00000101577	273.071642	-0.83064922	7.62E-06	0.00105827	LPIN2
ENSG00000198948	926.376416	-0.82914287	0.00127608	0.03921444	MFAP3L
ENSG00000167759	5482.17759	-0.82851382	0.00058828	0.02235449	KLK13
ENSG00000204618	1046.37417	-0.8275039	1.19E-06	0.00027685	RNF39
ENSG00000276430	245.514883	-0.8260823	1.29E-12	4.56E-09	FAM25C
ENSG00000152137	4289.16382	-0.82342086	1.19E-07	4.89E-05	HSPB8
ENSG00000108309	523.927312	-0.82307847	7.73E-06	0.00105827	RUNDC3A
ENSG00000134242	262.166068	-0.82203027	2.76E-05	0.00266935	PTPN22
ENSG00000170426	3286.79961	-0.82104654	1.98E-06	0.00040613	SDR9C7

ENSG00000132170	84.7686778	-0.82051657	0.00077732	0.02714489	PPARG
ENSG00000170423	4001.51057	-0.81908496	9.87E-06	0.00125529	KRT78
ENSG00000137878	187.556643	-0.81366268	0.00167295	0.04781409	GCOM1
ENSG00000283270	57.945824	-0.80755779	0.00013063	0.00809878	NA
ENSG00000213904	91.873446	-0.80368946	0.00018044	0.01008969	LIPE-AS1
ENSG00000179148	2150.66911	-0.80318487	2.82E-06	0.00051346	ALOXE3
ENSG00000081181	276.498296	-0.79788461	0.00022481	0.01152167	ARG2
ENSG00000069535	56.1485512	-0.7968147	0.00162143	0.04702765	MAOB
ENSG00000073910	308.966697	-0.79592596	0.00018161	0.01009156	FRY
ENSG00000184702	2020.60014	-0.79377698	1.11E-10	1.40E-07	SEPT5
ENSG00000231672	40.6818906	-0.79302077	0.00044377	0.01840707	NA
ENSG00000188089	9747.47179	-0.7914321	7.71E-06	0.00105827	PLA2G4E
ENSG00000179859	359.066158	-0.78368578	1.19E-05	0.00143107	LOC284023
ENSG00000177243	151.443432	-0.77966286	0.00069951	0.0253804	DEFB103B
ENSG00000108839	249.557648	-0.77157434	3.16E-05	0.0030037	ALOX12
ENSG00000229035	318.201074	-0.76507981	4.38E-07	0.00012491	NA
ENSG00000274276	66.3224885	-0.76196596	0.00012136	0.0077695	CBSL
ENSG00000176194	1096.08785	-0.75842045	7.77E-05	0.00567344	CIDEA
ENSG00000148798	476.31571	-0.75719315	1.34E-07	5.37E-05	INA
ENSG00000162444	265.125694	-0.75701048	7.48E-11	1.02E-07	RBP7
ENSG00000204538	231.324728	-0.75095919	0.00063385	0.02351062	PSORS1C2
ENSG00000069667	1241.23344	-0.74635599	0.00015805	0.00918673	RORA
ENSG00000179913	895.464824	-0.74615557	3.69E-13	1.63E-09	B3GNT3
ENSG00000175121	3087.23372	-0.73731343	1.69E-07	6.34E-05	WFDC5
ENSG00000072954	213.970487	-0.73569194	0.0003835	0.01673211	TMEM38A
ENSG00000173221	933.472085	-0.73479442	0.00038594	0.01674194	GLRX
ENSG00000150471	96.2327106	-0.7285214	0.00048198	0.01966866	ADGRL3
ENSG00000092295	11319.7318	-0.71999895	6.52E-05	0.0050716	TGM1
ENSG00000269741	2122.38047	-0.71742276	0.00013511	0.00820414	KLK9
ENSG00000102886	999.640387	-0.7161523	1.14E-05	0.00139273	GDPD3
ENSG00000159958	96.4132669	-0.71257223	0.00120801	0.03764642	TNFRSF13C
ENSG00000255501	2374.57461	-0.71066529	1.62E-05	0.00177405	CARD18
ENSG00000166828	132.101477	-0.70978713	8.82E-05	0.00626092	SCNN1G
ENSG00000100889	1015.4049	-0.7086388	4.87E-08	2.77E-05	PCK2
ENSG00000188001	814.191268	-0.70849709	8.80E-06	0.00115196	TPRG1
ENSG00000149328	276.145628	-0.70655603	0.00023091	0.01175843	GLB1L2
ENSG00000135114	674.783418	-0.70112737	7.48E-06	0.00104957	OASL
ENSG00000111319	3699.13668	-0.69825182	3.20E-06	0.0005441	SCNN1A
ENSG00000143412	1993.46922	-0.69713954	0.00012628	0.00794106	ANXA9
ENSG00000167914	3408.25825	-0.69612722	0.00043644	0.01832405	GSDMA
ENSG00000158220	520.741592	-0.69127497	0.00128079	0.03929077	ESYT3
ENSG00000182040	87.0928916	-0.68645802	0.0002069	0.01085063	USH1G
ENSG00000069011	1329.72135	-0.68622372	3.36E-05	0.00312214	PITX1
ENSG00000179846	1091.93165	-0.68433173	1.56E-05	0.00174143	NKPD1
ENSG00000060982	910.805475	-0.68062986	0.00013311	0.00812877	BCAT1
ENSG00000186474	1175.97246	-0.67825105	1.21E-05	0.00143864	KLK12
ENSG00000203785	8623.65197	-0.67612988	1.68E-07	6.34E-05	SPRR2E
ENSG00000022567	482.102587	-0.67330246	0.00077688	0.02714489	SLC45A4
ENSG00000261175	105.730542	-0.67011847	7.83E-05	0.00569715	LOC102724344
ENSG00000072952	61.6556608	-0.67006417	0.00113981	0.03615887	MRVI1
ENSG00000108244	5044.55271	-0.66373969	1.77E-06	0.0003763	KRT23
ENSG00000241794	2746.21293	-0.66358261	2.72E-06	0.00050485	SPRR2A
ENSG00000275880	272.512208	-0.66338464	7.39E-05	0.00553185	NA
ENSG00000259581	70.9867833	-0.66334985	0.00108459	0.03478164	NA
ENSG00000280587	90.774473	-0.66047256	0.00150839	0.04442212	NA
ENSG00000197580	67.071225	-0.65858028	0.00139862	0.04177176	BCO2
ENSG00000180155	5430.49867	-0.65855937	5.86E-09	4.50E-06	LYNX1
ENSG00000129521	689.749124	-0.65757055	0.00026577	0.0129958	EGLN3
ENSG00000182378	2433.78009	-0.65696922	1.52E-05	0.00170768	PLCXD1
ENSG00000118402	3420.45053	-0.65495378	6.03E-05	0.00479662	ELOVL4
ENSG00000104055	940.039568	-0.65374843	6.38E-05	0.00503045	TGM5
ENSG00000188277	462.41795	-0.65172371	0.00090067	0.02997147	C15orf62
ENSG00000100368	121.162019	-0.64976468	0.00059992	0.02269937	CSF2RB
ENSG00000259153	280.466467	-0.64953206	3.13E-07	9.36E-05	NA
ENSG00000182749	4013.38672	-0.64923205	6.30E-09	4.64E-06	PAQR7
ENSG00000185052	90.0202632	-0.64635358	0.00022496	0.01152167	SLC24A3
ENSG00000171954	5727.90711	-0.64471021	0.00053736	0.02119437	CYP4F22
ENSG00000157368	280.223859	-0.64331379	0.00084254	0.02885217	IL34
ENSG00000103888	502.460984	-0.6426779	1.75E-06	0.00037611	CEMIP
ENSG00000237499	168.065878	-0.64130758	0.0008995	0.02997147	LOC100130476
ENSG00000100170	2707.53724	-0.63989566	8.46E-07	0.00021654	SLC5A1
ENSG00000036672	501.451051	-0.63961913	8.52E-05	0.0060931	USP2
ENSG00000014257	871.41059	-0.63937138	0.00013284	0.00812877	ACPP
ENSG00000133401	1142.31268	-0.63926113	0.00049888	0.0200804	PDZD2
ENSG00000070669	573.242	-0.63897125	2.10E-07	6.99E-05	ASNS
ENSG00000136840	632.952461	-0.6377598	9.51E-10	9.33E-07	ST6GALNAC4
ENSG00000180316	843.974492	-0.63609499	0.00043563	0.01832405	PNPLA1
ENSG00000236824	152.63427	-0.63376792	0.00171794	0.04856967	BCYRN1
ENSG00000198478	1292.46644	-0.63373919	0.00018132	0.01009156	SH3BGRL2
ENSG00000126878	636.418581	-0.63217066	4.35E-09	3.66E-06	AIF1L
ENSG00000145283	402.39556	-0.63041058	0.00012339	0.00781458	SLC10A6
ENSG00000105889	160.774856	-0.63018235	1.09E-06	0.00026798	STEAP1B
ENSG00000158786	706.216336	-0.62669783	0.00039025	0.01685977	PLA2G2F

ENSG00000255138	909.080055	-0.62488652	0.00017645	0.00992951	NA
ENSG00000197191	2342.43047	-0.62380953	0.00013341	0.00812877	CYSRT1
ENSG00000006625	5621.43316	-0.62189542	0.00094801	0.03113638	GGCT
ENSG00000169213	79.7631689	-0.62076173	0.00063064	0.02349715	RAB3B
ENSG00000198885	64.6885461	-0.6196484	0.00107902	0.03466594	ITPRIPL1
ENSG00000142677	2598.83165	-0.61928527	8.60E-06	0.00114233	IL22RA1
ENSG00000156675	4376.72631	-0.6189653	2.45E-07	7.74E-05	RAB11FIP1
ENSG00000108379	234.678786	-0.61739579	0.00025931	0.0127988	WNT3
ENSG00000196542	1252.27912	-0.61692836	0.00030791	0.01454766	SPTSSB
ENSG00000105289	227.067392	-0.61592177	2.43E-05	0.00243892	TJP3
ENSG00000196805	764.811964	-0.61452343	6.95E-07	0.00018326	SPRR2B
ENSG00000116741	167.147564	-0.61385043	0.00033072	0.01513961	RGS2
ENSG00000124429	3884.18843	-0.61369446	4.23E-05	0.00371571	POF1B
ENSG00000168140	891.968887	-0.61190753	0.00168478	0.047939	VASN
ENSG00000130707	1127.15772	-0.61109251	5.57E-05	0.004512	ASS1
ENSG00000163082	2179.70383	-0.60790471	0.00011509	0.00758834	SGPP2
ENSG00000133048	139.138601	-0.60769443	0.00014238	0.00856217	CHI3L1
ENSG00000170775	93.9171891	-0.60316721	0.00083883	0.02883609	GPR37
ENSG00000184368	102.103273	-0.60086148	0.00088137	0.02960056	MAP7D2
ENSG00000136153	1769.57186	-0.59802619	8.74E-06	0.00115196	LMO7
ENSG00000233381	101.047522	-0.59618351	0.00106314	0.03421809	NA
ENSG00000174502	185.746056	-0.59569578	0.0006064	0.02289536	SLC26A9
ENSG00000196407	1366.53002	-0.59445611	0.00011011	0.00736969	THEM5
ENSG00000166123	1721.06167	-0.59372136	1.06E-05	0.00131536	GPT2
ENSG00000101096	394.234529	-0.59264292	0.00080201	0.02778736	NFATC2
ENSG00000112033	5288.87242	-0.59256805	3.70E-12	8.17E-09	PPARD
ENSG00000064270	1076.5338	-0.59232183	0.00125854	0.03881043	ATP2C2
ENSG00000101849	1339.20421	-0.58916869	1.10E-07	4.62E-05	TBL1X
ENSG00000064300	198.676688	-0.58718227	0.00031482	0.01476491	NGFR
ENSG00000155324	828.778307	-0.58503049	5.02E-08	2.77E-05	GRAMD3
ENSG00000173212	263.096427	-0.58370961	0.00110841	0.03548122	MAB21L3
ENSG00000173083	1348.5552	-0.58258443	0.00019468	0.01044776	HPSE
ENSG00000188505	11806.869	-0.57978795	6.54E-06	0.00096351	NCCRP1
ENSG00000089723	1738.42677	-0.57938197	6.72E-06	0.00097367	OTUB2
ENSG00000167767	13435.985	-0.57929441	0.00016627	0.00944664	KRT80
ENSG00000141655	389.107419	-0.57823448	0.00043907	0.01834116	TNFRSF11A
ENSG00000169474	68783.4527	-0.57705692	3.09E-06	0.00054073	SPRR1A
ENSG00000158158	968.605702	-0.57446609	4.17E-05	0.00370175	CNNM4
ENSG00000197372	463.840954	-0.57369826	0.00156343	0.04593365	ZNF675
ENSG00000189051	612.398247	-0.57355735	0.00159433	0.04648819	RNF222
ENSG00000124102	135631.372	-0.57242148	5.97E-08	2.93E-05	PI3
ENSG00000135709	4706.17223	-0.56926207	4.73E-05	0.00401447	KIAA0513
ENSG00000168447	2064.89785	-0.56786773	0.0003693	0.01635451	SCNN1B
ENSG00000057657	1262.42699	-0.56527366	0.00014921	0.00878874	PRDM1
ENSG00000261040	943.978551	-0.56519878	7.23E-05	0.00550415	WFDC21P
ENSG00000189433	1066.9903	-0.56270636	1.36E-06	0.00030317	GJB4
ENSG00000135069	1348.93701	-0.56257812	2.70E-05	0.0026533	PSAT1
ENSG00000121316	2575.3047	-0.56142905	4.33E-09	3.66E-06	PLBD1
ENSG00000122643	463.17592	-0.56060577	5.35E-05	0.00435508	NT5C3A
ENSG00000138678	163.113427	-0.55910744	0.00057852	0.02222257	GPAT3
ENSG00000141574	154.127317	-0.55219749	0.00042616	0.0181015	SECTM1
ENSG00000144452	12011.8166	-0.55017316	0.00121441	0.03777935	ABCA12
ENSG00000071242	640.909899	-0.54861928	4.01E-06	0.00065609	RPS6KA2
ENSG00000122694	341.097576	-0.5456174	0.00023595	0.01194618	GLIPR2
ENSG00000134955	5481.98317	-0.54287887	0.00166857	0.04781409	SLC37A2
ENSG00000105825	665.040185	-0.54065298	0.00084044	0.02883609	TFPI2
ENSG00000135605	180.52257	-0.53967095	0.00038577	0.01674194	TEC
ENSG00000243137	149.609177	-0.53915114	0.00112103	0.03575545	PSG4
ENSG00000078900	324.900769	-0.53868345	0.00023951	0.01205746	TP73
ENSG00000066629	697.617564	-0.53635421	3.96E-06	0.00065466	EML1
ENSG00000129455	6500.21654	-0.53398002	2.80E-07	8.53E-05	KLK8
ENSG00000148154	1505.45963	-0.53346948	0.00124463	0.03844855	UGCG
ENSG00000170786	2071.38816	-0.53343487	8.20E-05	0.00594083	SDR16C5
ENSG00000100285	168.6928	-0.53268478	0.00079579	0.02762602	NEFH
ENSG00000197822	507.065989	-0.53186711	0.00031776	0.01481485	OCLN
ENSG00000166396	12059.4734	-0.52999968	1.69E-06	0.00036846	SERPINB7
ENSG00000204866	4603.72894	-0.52950877	1.69E-05	0.00180659	IGFL2
ENSG00000101846	204.440526	-0.52753945	0.00096846	0.03169006	STS
ENSG00000087128	919.180662	-0.52354628	4.32E-05	0.00372032	TMPRSS11E
ENSG00000155158	1367.09542	-0.52312054	0.00093756	0.03090791	TTC39B
ENSG00000101825	553.05157	-0.52242399	1.37E-05	0.00157713	MXRA5
ENSG00000205488	353.335067	-0.52084878	0.00163122	0.04709737	NA
ENSG00000184148	1596.9211	-0.52059334	0.00075979	0.0267691	SPRR4
ENSG00000189182	975.871337	-0.52010863	6.50E-05	0.0050716	KRT77
ENSG00000139433	39015.9653	-0.51977203	7.07E-06	0.00101613	GLTP
ENSG00000137563	2490.17643	-0.51934203	5.05E-06	0.00079459	GGH
ENSG00000136155	12503.0891	-0.51864667	0.00014637	0.00870823	SCEL
ENSG00000035862	1356.79083	-0.51598786	0.00027336	0.01323741	TIMP2
ENSG00000126903	1721.47116	-0.51578092	2.28E-07	7.36E-05	SLC10A3
ENSG00000145934	8356.26109	-0.51409535	0.00058365	0.02229914	TENM2
ENSG00000139988	4046.88867	-0.51345055	0.00094165	0.03098493	RDH12
ENSG00000088002	10616.9111	-0.50967355	3.98E-05	0.00358643	SULT2B1
ENSG00000153292	775.30613	-0.50919351	0.00041463	0.01773984	ADGRF1

ENSG00000137486	638.278236	-0.50879449	0.00015995	0.00926649	ARRB1
ENSG00000162390	1871.41858	-0.50854755	2.46E-06	0.00047238	ACOT11
ENSG00000100344	560.259883	-0.50806306	0.00047704	0.01955735	PNPLA3
ENSG00000070159	2367.72005	-0.50797589	2.04E-07	6.92E-05	PTPN3
ENSG00000160179	850.861419	-0.50570073	0.00133792	0.04076031	ABCG1
ENSG00000261068	772.807209	-0.50496022	3.72E-06	0.00061955	NA
ENSG00000244094	2183.01387	-0.50459319	3.15E-06	0.00054073	SPRR2F
ENSG00000153246	594.788669	-0.50291983	0.00032096	0.01484644	PLA2R1
ENSG00000237350	845.68632	-0.50188468	0.00022235	0.01145473	NA
ENSG00000173210	665.127643	-0.50010096	1.94E-07	6.72E-05	ABLIM3
ENSG00000162066	557.998688	0.5107202	3.11E-06	0.00054073	AMDHD2
ENSG00000094796	1066.08115	0.51451711	0.00139926	0.04177176	KRT31
ENSG00000147889	351.142108	0.51732486	0.00086186	0.02922193	CDKN2A
ENSG00000198715	1018.68266	0.52209751	8.64E-05	0.00615411	GLMP
ENSG00000137343	173.054374	0.52637175	1.73E-05	0.0018457	ATAT1
ENSG00000171621	753.517322	0.52645727	0.00170943	0.04840647	SPSB1
ENSG00000213366	150.786843	0.53325298	0.00011175	0.00743091	GSTM2
ENSG00000178467	150.567376	0.54270124	2.16E-05	0.00221527	P4HTM
ENSG00000133069	205.086057	0.55125401	0.00117235	0.03686021	TMCC2
ENSG00000035664	271.970871	0.56104127	0.00052741	0.0209421	DAPK2
ENSG00000115648	202.137305	0.58764872	0.00085486	0.029161	MLPH
ENSG00000114812	527.398531	0.59570349	5.26E-05	0.00432573	VIPR1
ENSG00000147883	8539.40571	0.60948091	0.00010532	0.00718343	CDKN2B
ENSG00000125772	620.355603	0.62112606	0.00010413	0.00713195	GPCPD1
ENSG00000139631	222.162215	0.62468355	7.74E-05	0.00567335	CSAD
ENSG00000135218	1997.65758	0.65823086	1.35E-05	0.0015689	CD36
ENSG00000111432	165.726492	0.66578083	0.00011878	0.00768425	FZD10
ENSG00000088756	221.992909	0.66817817	0.00137975	0.04175577	ARHGAP28
ENSG00000188883	190.415467	0.67495482	0.00161407	0.04690895	KLRG2
ENSG00000138606	979.872692	0.69340706	5.16E-06	0.00080003	SHF
ENSG00000182272	68.6146902	0.69606089	0.00047379	0.01951503	B4GALNT4
ENSG00000184545	169.577824	0.71013074	6.96E-05	0.00534467	DUSP8
ENSG00000128849	261.205746	0.71112757	7.36E-05	0.00553185	CGNL1
ENSG00000050438	71.4982862	0.71269616	0.00126217	0.0388545	SLC4A8
ENSG00000130600	625.665104	0.72151234	1.68E-05	0.00180659	H19
ENSG00000170454	1235.92488	0.7354675	6.41E-05	0.0050311	KRT75
ENSG00000129946	89.5935643	0.7370941	0.00051255	0.020537	SHC2
ENSG00000250722	660.406622	0.75102231	2.15E-05	0.00221527	SELENOP
ENSG00000165092	809.433746	0.75363186	5.98E-06	0.00089486	ALDH1A1
ENSG00000100767	255.830958	0.75510461	0.00093297	0.03084324	PAPLN
ENSG00000198832	228.097785	0.767709	4.37E-05	0.0037457	SELENOM
ENSG00000092096	51.9518257	0.80943674	0.00011186	0.00743091	SLC22A17
ENSG00000166405	40.1165785	0.81963125	0.00156492	0.04593365	RIC3
ENSG00000263639	37.0216264	0.82023467	5.96E-05	0.00478964	MSMB
ENSG00000126217	34.1288481	0.84702097	0.00170068	0.04831343	MCF2L
ENSG00000144821	185.151601	0.85247711	0.00055677	0.02180801	MYH15
ENSG00000016082	134.770823	0.88159942	7.89E-10	8.20E-07	ISL1
ENSG00000104081	158.583269	0.89080707	0.00013564	0.00820826	BMF
ENSG00000143595	24.1410857	0.93509011	0.00133239	0.04066221	AQP10
ENSG00000197635	835.905675	0.9418402	1.75E-11	3.09E-08	DPP4
ENSG00000186891	42.0667921	0.98970837	0.00019427	0.01044776	TNFRSF18
ENSG00000123612	20.1001669	1.00013431	0.00150011	0.04425202	ACVR1C
ENSG00000114013	32.2780116	1.06690818	0.00040786	0.01749244	CD86
ENSG00000006025	45.6921496	1.08577094	6.01E-05	0.00479662	OSBPL7
ENSG00000123095	153.782993	1.09323924	3.44E-05	0.00313871	BHLHE41
ENSG00000137673	30.9583265	1.15400175	0.00070764	0.02546623	MMP7
ENSG00000124875	26.0787286	1.19581501	0.00088417	0.02960056	CXCL6
ENSG00000185634	13.6525308	1.28827026	0.00158395	0.04633845	SHC4
ENSG00000071282	18.8189862	1.47380916	3.31E-05	0.00309227	LMCD1
ENSG00000159674	846.686296	1.55319853	4.41E-14	2.60E-10	SPON2
ENSG00000137507	160.788174	1.63903828	0.00113081	0.03593777	LRRC32
ENSG00000169116	193.343707	1.66484148	1.60E-12	4.72E-09	PARM1
ENSG00000162747	20.5353973	1.66494169	2.10E-06	0.00042501	FCGR3B
ENSG00000065371	12.549066	1.70006403	0.00047604	0.01955735	ROPN1
ENSG00000204136	9.428147	1.70582179	0.00012761	0.00797992	GGTA1P
ENSG00000105664	66.2751954	2.70167068	6.07E-07	0.00016752	COMP

9.4 List of Figures

9.5 List of Tables

10 References

1. Martin, M. T., Vulin, A. & Hendry, J. H. Human epidermal stem cells: Role in adverse skin reactions and carcinogenesis from radiation. *Mutat. Res. Mutat. Res.* **770,** 349–368 (2016).
2. Uy Gonzales, K. A. & Fuchs, E. Skin and Its Regenerative Powers: An Alliance between Stem Cells and Their Niche. *Dev. Cell* **43,** 387–401 (2017).
3. Abe, Y. & Tanaka, N. Roles of the Hedgehog Signaling Pathway in Epidermal and Hair Follicle Development, Homeostasis, and Cancer. *J. Dev. Biol.* **5,** (2017).
4. Blanpain, C. & Fuchs, E. Epidermal stem cells of the skin. *Annu. Rev. Cell Dev. Biol.* **22,** 339–373 (2006).
5. Kanitakis, J. Anatomy, histology and immunohistochemistry of normal human skin. *Eur. J. Dermatol. EJD* **12,** 390–399; quiz 400–401 (2002).
6. Millington, P. F. & Wilkinson, R. *Skin*. (Cambridge University Press, 2009).
7. *The biology of the skin*. (Parthenon Pub. Group, 2001).
8. Roméro-Graillet, C. *et al.* Ultraviolet B radiation acts through the nitric oxide and cGMP signal transduction pathway to stimulate melanogenesis in human melanocytes. *J. Biol. Chem.* **271,** 28052–28056 (1996).
9. MacNeil, S. Progress and opportunities for tissue-engineered skin. *Nature* **445,** 874–880 (2007).
10. Solanas, G. & Benitah, S. A. Regenerating the skin: a task for the heterogeneous stem cell pool and surrounding niche. *Nat. Rev. Mol. Cell Biol.* **14,** 737–748 (2013).
11. Candi, E., Schmidt, R. & Melino, G. The cornified envelope: a model of cell death in the skin. *Nat. Rev. Mol. Cell Biol.* **6,** 328–340 (2005).
12. Baroni, A. *et al.* Structure and function of the epidermis related to barrier properties. *Clin. Dermatol.* **30,** 257–262 (2012).
13. Potten, C. S. & Morris, R. J. Epithelial stem cells in vivo. *J. Cell Sci. Suppl.* **10,** 45–62 (1988).
14. Doupé, D. P. & Jones, P. H. Interfollicular epidermal homeostasis: dicing with differentiation: Interfollicular epidermal homeostasis. *Exp. Dermatol.* **21,** 249–253 (2012).
15. Clayton, E. *et al.* A single type of progenitor cell maintains normal epidermis. *Nature* **446,** 185–189 (2007).
16. Hennings, H. & Holbrook, K. A. Calcium regulation of cell-cell contact and differentiation of epidermal cells in culture. An ultrastructural study. *Exp. Cell Res.* **143,** 127–142 (1983).
17. Hennings, H. *et al.* Calcium regulation of growth and differentiation of mouse epidermal cells in culture. *Cell* **19,** 245–254 (1980).
18. Steven, A. C. & Steinert, P. M. Protein composition of cornified cell envelopes of epidermal keratinocytes. *J. Cell Sci.* **107 (Pt 2),** 693–700 (1994).
19. Bergboer, J. G. M. *et al.* Psoriasis Risk Genes of the Late Cornified Envelope-3 Group Are Distinctly Expressed Compared with Genes of Other LCE Groups. *Am. J. Pathol.* **178,** 1470–1477 (2011).
20. Fuchs, E. Epidermal differentiation: the bare essentials. *J. Cell Biol.* **111,** 2807–2814 (1990).

21. Warhol, M. J., Roth, J., Lucocq, J. M., Pinkus, G. S. & Rice, R. H. Immuno-ultrastructural localization of involucrin in squamous epithelium and cultured keratinocytes. *J. Histochem. Cytochem. Off. J. Histochem. Soc.* **33,** 141–149 (1985).
22. Sahle, F. F., Gebre-Mariam, T., Dobner, B., Wohlrab, J. & Neubert, R. H. H. Skin Diseases Associated with the Depletion of Stratum Corneum Lipids and Stratum Corneum Lipid Substitution Therapy. *Skin Pharmacol. Physiol.* **28,** 42–55 (2015).
23. Steinert, P. M., Cantieri, J. S., Teller, D. C., Lonsdale-Eccles, J. D. & Dale, B. A. Characterization of a class of cationic proteins that specifically interact with intermediate filaments. *Proc. Natl. Acad. Sci. U. S. A.* **78,** 4097–4101 (1981).
24. Kalinin, A., Marekov, L. N. & Steinert, P. M. Assembly of the epidermal cornified cell envelope. *J. Cell Sci.* **114,** 3069–3070 (2001).
25. Hitomi, K. Transglutaminases in skin epidermis. *Eur. J. Dermatol. EJD* **15,** 313–319 (2005).
26. Nemes, Z. & Steinert, P. M. Bricks and mortar of the epidermal barrier. *Exp. Mol. Med.* **31,** 5–19 (1999).
27. Bikle, D. D., Xie, Z. & Tu, C.-L. Calcium regulation of keratinocyte differentiation. *Expert Rev. Endocrinol. Metab.* **7,** 461–472 (2012).
28. Lee, S. E. & Lee, S. H. Skin Barrier and Calcium. *Ann. Dermatol.* **30,** 265–275 (2018).
29. Pillai, S., Bikle, D. D. & Elias, P. M. 1,25-Dihydroxyvitamin D production and receptor binding in human keratinocytes varies with differentiation. *J. Biol. Chem.* **263,** 5390–5395 (1988).
30. Bikle, D. D. Vitamin D regulated keratinocyte differentiation. *J. Cell. Biochem.* **92,** 436–444 (2004).
31. Pillai, S., Bikle, D. D., Su, M. J., Ratnam, A. & Abe, J. 1,25-Dihydroxyvitamin D3 upregulates the phosphatidylinositol signaling pathway in human keratinocytes by increasing phospholipase C levels. *J. Clin. Invest.* **96,** 602–609 (1995).
32. Eckert, R. L. *et al.* AP1 Transcription Factors in Epidermal Differentiation and Skin Cancer. *J. Skin Cancer* **2013,** 1–9 (2013).
33. Botchkarev, V. A. Integration of the Transcription Factor-Regulated and Epigenetic Mechanisms in the Control of Keratinocyte Differentiation. *J. Investig. Dermatol. Symp. Proc.* **17,** 30–32 (2015).
34. Lopez-Pajares, V., Yan, K., Zarnegar, B. J., Jameson, K. L. & Khavari, P. A. Genetic Pathways in Disorders of Epidermal Differentiation. *Trends Genet. TIG* **29,** 31–40 (2013).
35. Sen, G. L. *et al.* ZNF750 is a p63 target gene that induces KLF4 to drive terminal epidermal differentiation. *Dev. Cell* **22,** 669–677 (2012).
36. Lopez-Pajares, V. *et al.* A LncRNA-MAF:MAFB transcription factor network regulates epidermal differentiation. *Dev. Cell* **32,** 693–706 (2015).
37. Teng, A., Nair, M., Wells, J., Segre, J. A. & Dai, X. Strain-dependent perinatal lethality of Ovol1-deficient mice and identification of Ovol2 as a downstream target of Ovol1 in skin epidermis. *Biochim. Biophys. Acta* **1772,** 89–95 (2007).
38. Yu, Z. *et al.* The Grainyhead-like epithelial transactivator Get-1/Grhl3 regulates epidermal terminal differentiation and interacts functionally with LMO4. *Dev. Biol.* **299,** 122–136 (2006).

39. Romano, R.-A. *et al.* ΔNp63 knockout mice reveal its indispensable role as a master regulator of epithelial development and differentiation. *Dev. Camb. Engl.* **139,** 772–782 (2012).
40. Truong, A. B., Kretz, M., Ridky, T. W., Kimmel, R. & Khavari, P. A. p63 regulates proliferation and differentiation of developmentally mature keratinocytes. *Genes Dev.* **20,** 3185–3197 (2006).
41. Yang, A. *et al.* p63 is essential for regenerative proliferation in limb, craniofacial and epithelial development. *Nature* **398,** 714–718 (1999).
42. Koster, M. I. p63 in skin development and ectodermal dysplasias. *J. Invest. Dermatol.* **130,** 2352–2358 (2010).
43. Soares, E. & Zhou, H. Master regulatory role of p63 in epidermal development and disease. *Cell. Mol. Life Sci.* **75,** 1179–1190 (2018).
44. Bao, X. *et al.* A novel ATAC-seq approach reveals lineage-specific reinforcement of the open chromatin landscape via cooperation between BAF and p63. *Genome Biol.* **16,** (2015).
45. LeBoeuf, M. *et al.* Hdac1 and Hdac2 act redundantly to control p63 and p53 functions in epidermal progenitor cells. *Dev. Cell* **19,** 807–818 (2010).
46. Perdigoto, C. N., Valdes, V. J., Bardot, E. S. & Ezhkova, E. Epigenetic Regulation of Epidermal Differentiation. *Cell. Mol. Life Sci. CMLS* **69,** 2161–2172 (2012).
47. Ho, L. & Crabtree, G. R. Chromatin remodelling during development. *Nature* **463,** 474–484 (2010).
48. Margueron, R. *et al.* Ezh1 and Ezh2 Maintain Repressive Chromatin through Different Mechanisms. *Mol. Cell* **32,** 503–518 (2008).
49. Ezhkova, E. *et al.* Ezh2 Orchestrates Gene Expression for the Stepwise Differentiation of Tissue-Specific Stem Cells. *Cell* **136,** 1122–1135 (2009).
50. Agger, K. *et al.* UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **449,** 731–734 (2007).
51. Gdula, M. R. *et al.* Remodeling of three-dimensional organization of the nucleus during terminal keratinocyte differentiation in the epidermis. *J. Invest. Dermatol.* **133,** 2191–2201 (2013).
52. Mischke, D., Korge, B. P., Marenholz, I., Volz, A. & Ziegler, A. Genes encoding structural proteins of epidermal cornification and S100 calcium-binding proteins form a gene complex ('epidermal differentiation complex') on human chromosome 1q21. *J. Invest. Dermatol.* **106,** 989–992 (1996).
53. Kypriotou, M., Huber, M. & Hohl, D. The human epidermal differentiation complex: cornified envelope precursors, S100 proteins and the 'fused genes' family. *Exp. Dermatol.* **21,** 643–649 (2012).
54. de Guzman Strong, C. *et al.* A milieu of regulatory elements in the epidermal differentiation complex syntenic block: implications for atopic dermatitis and psoriasis. *Hum. Mol. Genet.* **19,** 1453–1460 (2010).
55. Eckert, R. L. *et al.* S100 proteins in the epidermis. *J. Invest. Dermatol.* **123,** 23–33 (2004).
56. Gibbs, S. *et al.* Molecular characterization and evolution of the SPRR family of keratinocyte differentiation markers encoding small proline-rich proteins. *Genomics* **16,** 630–637 (1993).

57. Carregaro, F., Stefanini, A. C. B., Henrique, T. & Tajara, E. H. Study of small proline-rich proteins (SPRRs) in health and disease: a review of the literature. *Arch. Dermatol. Res.* **305,** 857–866 (2013).
58. Fischer, D. F. & Backendorf, C. Promoter analysis in the human SPRR gene family. *Methods Mol. Biol. Clifton NJ* **289,** 303–314 (2005).
59. Cabral, A. *et al.* SPRR4, a novel cornified envelope precursor: UV-dependent epidermal expression and selective incorporation into fragile envelopes. *J. Cell Sci.* **114,** 3837–3843 (2001).
60. Vermeij, W. P. & Backendorf, C. Skin Cornification Proteins Provide Global Link between ROS Detoxification and Cell Migration during Wound Healing. *PLOS ONE* **5,** e11957 (2010).
61. Mizuguchi, Y. *et al.* SPRR2A enhances p53 deacetylation through HDAC1 and down regulates p21 promoter activity. *BMC Mol. Biol.* **13,** 20 (2012).
62. Vermeij, W. P., Alia, A. & Backendorf, C. ROS Quenching Potential of the Epidermal Cornified Cell Envelope. *J. Invest. Dermatol.* **131,** 1435–1441 (2011).
63. Vermeij, W. P. *et al.* Proteomic identification of in vivo interactors reveals novel function of skin cornification proteins. *J. Proteome Res.* **11,** 3068–3076 (2012).
64. Dickinson, M. E. *et al.* High-throughput discovery of novel developmental phenotypes. *Nature* **537,** 508–514 (2016).
65. Blake, J. A. *et al.* Mouse Genome Database (MGD)-2017: community knowledge resource for the laboratory mouse. *Nucleic Acids Res.* **45,** D723–D729 (2017).
66. Green, C. L. & Khavari, P. A. Targets for molecular therapy of skin cancer. *Semin. Cancer Biol.* **14,** 63–69 (2004).
67. Skin Cancer Treatment. *National Cancer Institute* Available at: https://www.cancer.gov/types/skin/hp/skin-treatment-pdq#section/all. (Accessed: 11th May 2018)
68. Saladi, R. N. & Persaud, A. N. The causes of skin cancer: A comprehensive review. *Drugs Today* **41,** 37 (2005).
69. *World cancer report 2014*. (International Agency for Research on Cancer, 2014).
70. Swetter, S. M. Dermatological perspectives of malignant melanoma. *Surg. Clin. North Am.* **83,** 77–95 (2003).
71. Didona, D., Paolino, G., Bottoni, U. & Cantisani, C. Non Melanoma Skin Cancer Pathogenesis Overview. *Biomedicines* **6,** (2018).
72. Jr, V. T. D., Lawrence, T. & Rosenberg, S. A. *Cancer: Principles & Practice of Oncology: Annual Advances in Oncology*. (Lippincott Williams & Wilkins, 2012).
73. Xie, J. *et al.* A role of PDGFRα in basal cell carcinoma proliferation. *Proc. Natl. Acad. Sci.* **98,** 9255–9259 (2001).
74. Yanagi, T., Kitamura, S. & Hata, H. Novel Therapeutic Targets in Cutaneous Squamous Cell Carcinoma. *Front. Oncol.* **8,** (2018).
75. Ashford, B. G. *et al.* Reviewing the genetic alterations in high-risk cutaneous squamous cell carcinoma: A search for prognostic markers and therapeutic targets. *Head Neck* **39,** 1462–1469 (2017).
76. Lazarov, M. *et al.* CDK4 coexpression with Ras generates malignant human epidermal tumorigenesis. *Nat. Med.* **8,** 1105–1114 (2002).

77. Dajee, M. *et al.* NF-κB blockade and oncogenic Ras trigger invasive human epidermal neoplasia. *Nature* **421,** 639–643 (2003).
78. Seitz, C. S., Lin, Q., Deng, H. & Khavari, P. A. Alterations in NF-kappaB function in transgenic epithelial tissue demonstrate a growth inhibitory role for NF-kappaB. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 2307–2312 (1998).
79. Maurelli, R. *et al.* The role of oncogenic Ras in human skin tumorigenesis depends on the clonogenic potential of the founding keratinocytes. *J. Cell Sci.* **129,** 1003–1017 (2016).
80. Takaoka, M. *et al.* Ha-Ras(G12V) induces senescence in primary and immortalized human esophageal keratinocytes with p53 dysfunction. *Oncogene* **23,** 6760–6768 (2004).
81. Pierceall, W. E., Goldberg, L. H., Tainsky, M. A., Mukhopadhyay, T. & Ananthaswamy, H. N. Ras gene mutation and amplification in human nonmelanoma skin cancers. *Mol. Carcinog.* **4,** 196–202 (1991).
82. Gannon, O. M., Merida de Long, L., Endo-Munoz, L., Hazar-Rethinam, M. & Saunders, N. A. Dysregulation of the repressive H3K27 trimethylation mark in head and neck squamous cell carcinoma contributes to dysregulated squamous differentiation. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **19,** 428–441 (2013).
83. Scott, R. E. *et al.* Human squamous carcinoma cells express complex defects in the control of proliferation and differentiation. *Am. J. Pathol.* **133,** 374–380 (1988).
84. Ratushny, V., Gober, M. D., Hick, R., Ridky, T. W. & Seykora, J. T. From keratinocyte to cancer: the pathogenesis and modeling of cutaneous squamous cell carcinoma. *J. Clin. Invest.* **122,** 464–472 (2012).
85. Konicke, K. *et al.* The microRNA landscape of cutaneous squamous cell carcinoma. *Drug Discov. Today* **23,** 864–870 (2018).
86. Mizrahi, A. *et al.* Alterations of microRNAs throughout the malignant evolution of cutaneous squamous cell carcinoma: the role of miR-497 in epithelial to mesenchymal transition of keratinocytes. *Oncogene* **37,** 218–230 (2018).
87. Fatica, A. & Bozzoni, I. Long non-coding RNAs: new players in cell differentiation and development. *Nat. Rev. Genet.* **15,** 7–21 (2013).
88. Kretz, M. *et al.* Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature* **493,** 231–235 (2012).
89. Lee, C. S. *et al.* Cancer-Associated Long Noncoding RNA SMRT-2 Controls Epidermal Differentiation. *J. Invest. Dermatol.* (2018). doi:10.1016/j.jid.2018.01.003
90. Piipponen, M. *et al.* Long Noncoding RNA PICSAR Promotes Growth of Cutaneous Squamous Cell Carcinoma by Regulating ERK1/2 Activity. *J. Invest. Dermatol.* **136,** 1701–1710 (2016).
91. Yildirim, E. *et al.* Xist RNA Is a Potent Suppressor of Hematologic Cancer in Mice. *Cell* **152,** 727–742 (2013).
92. Djebali, S. *et al.* Landscape of transcription in human cells. *Nature* **489,** 101–108 (2012).
93. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* **431,** 931–945 (2004).
94. Matera, A. G., Terns, R. M. & Terns, M. P. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat. Rev. Mol. Cell Biol.* **8,** 209–220 (2007).

95. Okamura, K. & Lai, E. C. Endogenous small interfering RNAs in animals. *Nat. Rev. Mol. Cell Biol.* **9,** 673–678 (2008).
96. Bartel, D. P. MicroRNAs: Target Recognition and Regulatory Functions. *Cell* **136,** 215–233 (2009).
97. Farazi, T. A., Juranek, S. A. & Tuschl, T. The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members. *Development* **135,** 1201–1214 (2008).
98. Derrien, T. *et al.* The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Res.* **22,** 1775–1789 (2012).
99. Jia, H. *et al.* Genome-wide computational identification and manual annotation of human long noncoding RNA genes. *RNA* **16,** 1478–1487 (2010).
100. Cabili, M. N. *et al.* Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* **25,** 1915–1927 (2011).
101. Mercer, T. R., Dinger, M. E. & Mattick, J. S. Long non-coding RNAs: insights into functions. *Nat. Rev. Genet.* **10,** 155–159 (2009).
102. Ingolia, N. T., Lareau, L. F. & Weissman, J. S. Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity and Dynamics of Mammalian Proteomes. *Cell* **147,** 789–802 (2011).
103. Guttman, M. *et al.* lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* **477,** 295–300 (2011).
104. Anderson, D. M. *et al.* A Micropeptide Encoded by a Putative Long Noncoding RNA Regulates Muscle Performance. *Cell* **160,** 595–606 (2015).
105. Cooper, C. *et al.* Steroid receptor RNA activator bi-faceted genetic system: Heads or Tails? *Biochimie* **93,** 1973–1980 (2011).
106. Lanz, R. B. *et al.* A Steroid Receptor Coactivator, SRA, Functions as an RNA and Is Present in an SRC-1 Complex. *Cell* **97,** 17–27 (1999).
107. Hubé, F., Velasco, G., Rollin, J., Furling, D. & Francastel, C. Steroid receptor RNA activator protein binds to and counteracts SRA RNA-mediated activation of MyoD and muscle differentiation. *Nucleic Acids Res.* **39,** 513–525 (2011).
108. Gascoigne, D. K. *et al.* Pinstripe: a suite of programs for integrating transcriptomic and proteomic datasets identifies novel proteins and improves differentiation of protein-coding and non-coding genes. *Bioinformatics* **28,** 3042–3050 (2012).
109. Banfai, B. *et al.* Long noncoding RNAs are rarely translated in two human cell lines. *Genome Res.* **22,** 1646–1657 (2012).
110. Nelson, B. R. *et al.* A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. *Science* **351,** 271–275 (2016).
111. Chakraborty, S., Deb, A., Maji, R. K., Saha, S. & Ghosh, Z. LncRBase: An Enriched Resource for lncRNA Information. *PLoS ONE* **9,** e108010 (2014).
112. Rinn, J. L. & Chang, H. Y. Genome Regulation by Long Noncoding RNAs. *Annu. Rev. Biochem.* **81,** 145–166 (2012).
113. Stein, J. M. The effect of adrenaline and of alpha- and beta-adrenergic blocking agents on ATP concentration and on incorporation of 32Pi into ATP in rat fat cells. *Biochem. Pharmacol.* **24,** 1659–1662 (1975).

114. Yin, Q.-F. *et al.* Long Noncoding RNAs with snoRNA Ends. *Mol. Cell* **48,** 219–230 (2012).
115. Amaral, P. P. *et al.* Complex architecture and regulated expression of the Sox2ot locus during vertebrate development. *RNA* **15,** 2013–2027 (2009).
116. Quinn, J. J. & Chang, H. Y. Unique features of long non-coding RNA biogenesis and function. *Nat. Rev. Genet.* **17,** 47–62 (2015).
117. Niemczyk, M. *et al.* Imprinted Chromatin around DIRAS3 Regulates Alternative Splicing of GNG12-AS1, a Long Noncoding RNA. *Am. J. Hum. Genet.* **93,** 224–235 (2013).
118. Kornienko, A. E. *et al.* Long non-coding RNAs display higher natural expression variation than protein-coding genes in healthy humans. *Genome Biol.* **17,** (2016).
119. Ziegler, C. & Kretz, M. The More the Merrier—Complexity in Long Non-Coding RNA Loci. *Front. Endocrinol.* **8,** (2017).
120. Rinn, J. L. *et al.* Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. *Cell* **129,** 1311–1323 (2007).
121. Gupta, R. A. *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464,** 1071–1076 (2010).
122. Tsai, M.-C. *et al.* Long Noncoding RNA as Modular Scaffold of Histone Modification Complexes. *Science* **329,** 689–693 (2010).
123. Martianov, I., Ramadass, A., Serra Barros, A., Chow, N. & Akoulitchev, A. Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature* **445,** 666–670 (2007).
124. Feng, J. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes Dev.* **20,** 1470–1484 (2006).
125. Lee, J. T. Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome. *Genes Dev.* **23,** 1831–1842 (2009).
126. Bartolomei, M. S., Zemel, S. & Tilghman, S. M. Parental imprinting of the mouse H19 gene. *Nature* **351,** 153–155 (1991).
127. Sleutels, F., Zwart, R. & Barlow, D. P. The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* **415,** 810–813 (2002).
128. Flynn, R. A. & Chang, H. Y. Long Noncoding RNAs in Cell-Fate Programming and Reprogramming. *Cell Stem Cell* **14,** 752–761 (2014).
129. Derrien, T. *et al.* The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* **22,** 1775–1789 (2012).
130. Wang, K. C. & Chang, H. Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **43,** 904–914 (2011).
131. Kopp, F. & Mendell, J. T. Functional Classification and Experimental Dissection of Long Noncoding RNAs. *Cell* **172,** 393–407 (2018).
132. Brown, C. J. *et al.* A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* **349,** 38–44 (1991).
133. da Rocha, S. T. & Heard, E. Novel players in X inactivation: insights into Xist-mediated gene silencing and chromosome conformation. *Nat. Struct. Mol. Biol.* **24,** 197–204 (2017).

134. Bonasio, R. & Shiekhattar, R. Regulation of Transcription by Long Noncoding RNAs. *Annu. Rev. Genet.* **48,** 433–455 (2014).
135. Hamada, F. N. Global regulation of X chromosomal genes by the MSL complex in Drosophila melanogaster. *Genes Dev.* **19,** 2289–2294 (2005).
136. Kung, J. T. Y., Colognori, D. & Lee, J. T. Long Noncoding RNAs: Past, Present, and Future. *Genetics* **193,** 651–669 (2013).
137. Yang, G., Lu, X. & Yuan, L. LncRNA: A link between RNA and cancer. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* **1839,** 1097–1109 (2014).
138. Faghihi, M. A. *et al.* Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biol.* **11,** R56 (2010).
139. Cesana, M. *et al.* A Long Noncoding RNA Controls Muscle Differentiation by Functioning as a Competing Endogenous RNA. *Cell* **147,** 358–369 (2011).
140. Salviano-Silva, A., Lobo-Alves, S. C., Almeida, R. C. de, Malheiros, D. & Petzl-Erler, M. L. Besides Pathology: Long Non-Coding RNA in Cell and Tissue Homeostasis. *Non-Coding RNA* **4,** 3 (2018).
141. Klattenhoff, C. A. *et al.* Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. *Cell* **152,** 570–583 (2013).
142. Grote, P. *et al.* The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. *Dev. Cell* **24,** 206–214 (2013).
143. Ounzain, S. *et al.* CARMEN, a human super enhancer-associated long noncoding RNA controlling cardiac specification, differentiation and homeostasis. *J. Mol. Cell. Cardiol.* **89,** 98–112 (2015).
144. Ramos, A. D. *et al.* The Long Noncoding RNA Pnky Regulates Neuronal Differentiation of Embryonic and Postnatal Neural Stem Cells. *Cell Stem Cell* **16,** 439–447 (2015).
145. Liu, S. J. *et al.* Single-cell analysis of long non-coding RNAs in the developing human neocortex. *Genome Biol.* **17,** 67 (2016).
146. Kretz, M. *et al.* Suppression of progenitor differentiation requires the long noncoding RNA ANCR. *Genes Dev.* **26,** 338–343 (2012).
147. Hombach, S. & Kretz, M. The non-coding skin: Exploring the roles of long non-coding RNAs in epidermal homeostasis and disease: Review essay. *BioEssays* **35,** 1093–1100 (2013).
148. Leucci, E., Coe, E. A., Marine, J.-C. & Vance, K. W. The emerging role of long non-coding RNAs in cutaneous melanoma. *Pigment Cell Melanoma Res.* **29,** 619–626 (2016).
149. Flockhart, R. J. *et al.* BRAFV600E remodels the melanocyte transcriptome and induces BANCR to regulate melanoma cell migration. *Genome Res.* **22,** 1006–1014 (2012).
150. Cai, B. *et al.* BANCR contributes to the growth and invasion of melanoma by functioning as a competing endogenous RNA to upregulate Notch2 expression by sponging miR-204. *Int. J. Oncol.* **51,** 1941–1951 (2017).
151. Li, R. *et al.* Long non-coding RNA BANCR promotes proliferation in malignant melanoma by regulating MAPK pathway activation. *PloS One* **9,** e100893 (2014).
152. Danis, J., Göblös, A., Bata-Csörgő, Z., Kemény, L. & Széll, M. PRINS Non-Coding RNA Regulates Nucleic Acid-Induced Innate Immune Responses of Human Keratinocytes. *Front. Immunol.* **8,** (2017).

153. Sonkoly, E. *et al.* Identification and characterization of a novel, psoriasis susceptibility-related noncoding RNA gene, PRINS. *J. Biol. Chem.* **280,** 24159–24167 (2005).
154. Ridky, T. W., Chow, J. M., Wong, D. J. & Khavari, P. A. Invasive three-dimensional organotypic neoplasia from multiple normal human epithelia. *Nat. Med.* **16,** 1450–1455 (2010).
155. Khavari, P. A. Modelling cancer in human skin tissue. *Nat. Rev. Cancer* **6,** 270–280 (2006).
156. Reuter, J. A. *et al.* Modeling Inducible Human Tissue Neoplasia Identifies an Extracellular Matrix Interaction Network Involved in Cancer Progression. *Cancer Cell* **15,** 477–488 (2009).
157. Lachmann, A. *et al.* Massive Mining of Publicly Available RNA-seq Data from Human and Mouse. (2017). doi:10.1101/189092
158. Kouwenhoven, E. N. *et al.* Transcription factor p63 bookmarks and regulates dynamic enhancers during epidermal differentiation. *EMBO Rep.* **16,** 863–878 (2015).
159. Barrett, T. *et al.* NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* **41,** D991–D995 (2013).
160. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30,** 207–210 (2002).
161. Zerbino, D. R. *et al.* Ensembl 2018. *Nucleic Acids Res.* **46,** D754–D761 (2018).
162. Hon, C.-C. *et al.* An atlas of human long non-coding RNAs with accurate 5′ ends. *Nature* **543,** 199–204 (2017).
163. Gao, W., Chan, J. Y.-W. & Wong, T.-S. Long Non-Coding RNA Deregulation in Tongue Squamous Cell Carcinoma. *BioMed Res. Int.* **2014,** 1–10 (2014).
164. Kong, L. *et al.* CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* **35,** W345-349 (2007).
165. Sun, K. *et al.* iSeeRNA: identification of long intergenic non-coding RNA transcripts from transcriptome sequencing data. *BMC Genomics* **14,** S7 (2013).
166. Lin, M. F., Jungreis, I. & Kellis, M. PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. *Bioinformatics* **27,** i275–i282 (2011).
167. Sendoel, A. *et al.* Translation from unconventional 5′ start sites drives tumour initiation. *Nature* **541,** 494–499 (2017).
168. Steinert, P. M., Candi, E., Kartasova, T. & Marekov, L. Small proline-rich proteins are cross-bridging proteins in the cornified cell envelopes of stratified squamous epithelia. *J. Struct. Biol.* **122,** 76–85 (1998).
169. Gerber, P. A. *et al.* The top skin-associated genes: a comparative analysis of human and mouse skin transcriptomes. *Biol. Chem.* **395,** 577–591 (2014).
170. Lay, K., Kume, T. & Fuchs, E. FOXC1 maintains the hair follicle stem cell niche and governs stem cell quiescence to preserve long-term tissue-regenerating potential. *Proc. Natl. Acad. Sci. U. S. A.* **113,** E1506-1515 (2016).
171. Bin, L. *et al.* Forkhead Box C1 Regulates Human Primary Keratinocyte Terminal Differentiation. *PloS One* **11,** e0167392 (2016).
172. Beck, M. *et al.* The quantitative proteome of a human cell line. *Mol. Syst. Biol.* **7,** 549 (2011).

173. Choi, M. & Lee, C. Immortalization of Primary Keratinocytes and Its Application to Skin Research. *Biomol. Ther.* **23,** 391–399 (2015).
174. Stöppler, H., Hartmann, D.-P., Sherman, L. & Schlegel, R. The Human Papillomavirus Type 16 E6 and E7 Oncoproteins Dissociate Cellular Telomerase Activity from the Maintenance of Telomere Length. *J. Biol. Chem.* **272,** 13332–13337 (1997).
175. Boukamp, P. *et al.* Normal keratinization in a spontaneously immortalized aneuploid human keratinocyte cell line. *J. Cell Biol.* **106,** 761–771 (1988).
176. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339,** 819–823 (2013).
177. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8,** 2281–2308 (2013).
178. Sen, G. L., Reuter, J. A., Webster, D. E., Zhu, L. & Khavari, P. A. DNMT1 maintains progenitor function in self-renewing somatic tissue. *Nature* **463,** 563–567 (2010).
179. Botchkarev, V. A., Gdula, M. R., Mardaryev, A. N., Sharov, A. A. & Fessing, M. Y. Epigenetic Regulation of Gene Expression in Keratinocytes. *J. Invest. Dermatol.* **132,** 2505–2521 (2012).
180. Cavazza, A. *et al.* Dynamic Transcriptional and Epigenetic Regulation of Human Epidermal Keratinocyte Differentiation. *Stem Cell Rep.* **6,** 618–632 (2016).
181. Sloan, C. A. *et al.* ENCODE data at the ENCODE portal. *Nucleic Acids Res.* **44,** D726–D732 (2016).
182. ATAC-seq Data Standards and Prototype Processing Pipeline – ENCODE. Available at: https://www.encodeproject.org/atac-seq/. (Accessed: 1st May 2018)
183. Calo, E. & Wysocka, J. Modification of enhancer chromatin: what, how and why? *Mol. Cell* **49,** (2013).
184. Lawrence, M., Daujat, S. & Schneider, R. Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet. TIG* **32,** 42–56 (2016).
185. Histone ChIP-seq Data Standards and Processing Pipeline – ENCODE. Available at: https://www.encodeproject.org/chip-seq/histone/. (Accessed: 1st May 2018)
186. Oh, I. Y. *et al.* Regulation of the Dynamic Chromatin Architecture of the Epidermal Differentiation Complex Is Mediated by a c-Jun/AP-1-Modulated Enhancer. *J. Invest. Dermatol.* **134,** 2371–2380 (2014).
187. Poterlowicz, K. *et al.* 5C analysis of the Epidermal Differentiation Complex locus reveals distinct chromatin interaction networks between gene-rich and gene-poor TADs in skin epithelial cells. *PLoS Genet.* **13,** (2017).
188. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38,** 576–589 (2010).
189. Xu, N. *et al.* MicroRNA-125b Down-regulates Matrix Metallopeptidase 13 and Inhibits Cutaneous Squamous Cell Carcinoma Cell Proliferation, Migration, and Invasion. *J. Biol. Chem.* **287,** 29899–29908 (2012).
190. Justus, C. R., Leffler, N., Ruiz-Echevarria, M. & Yang, L. V. In vitro Cell Migration and Invasion Assays. *J. Vis. Exp. JoVE* (2014). doi:10.3791/51046
191. Ji, N. & Oudenaarden, A. van. *Single molecule fluorescent in situ hybridization (smFISH) of C. elegans worms and embryos*. (WormBook, 2005).

192. Adelman, K. & Lis, J. T. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat. Rev. Genet.* **13,** 720–731 (2012).
193. Segre, J. A., Bauer, C. & Fuchs, E. Klf4 is a transcription factor required for establishing the barrier function of the skin. *Nat. Genet.* **22,** 356–360 (1999).
194. Patel, S., Xi, Z. F., Seo, E. Y., McGaughey, D. & Segre, J. A. Klf4 and corticosteroids activate an overlapping set of transcriptional targets to accelerate in utero epidermal barrier acquisition. *Proc. Natl. Acad. Sci.* **103,** 18668–18673 (2006).
195. Iyer, M. K. *et al.* The landscape of long noncoding RNAs in the human transcriptome. *Nat. Genet.* **47,** 199–208 (2015).
196. Frey, U. H., Bachmann, H. S., Peters, J. & Siffert, W. PCR-amplification of GC-rich regions: 'slowdown PCR'. *Nat. Protoc.* **3,** 1312–1317 (2008).
197. Cabral, A. *et al.* Structural Organization and Regulation of the Small Proline-rich Family of Cornified Envelope Precursors Suggest a Role in Adaptive Barrier Function. *J. Biol. Chem.* **276,** 19231–19237 (2001).
198. Hezroni, H. *et al.* A subset of conserved mammalian long non-coding RNAs are fossils of ancestral protein-coding genes. *Genome Biol.* **18,** 162 (2017).
199. Housman, G. & Ulitsky, I. Methods for distinguishing between protein-coding and long noncoding RNAs and the elusive biological purpose of translation of long noncoding RNAs. *Biochim. Biophys. Acta* **1859,** 31–40 (2016).
200. Staab, J. F., Ferrer, C. A. & Sundstrom, P. Developmental expression of a tandemly repeated, proline-and glutamine-rich amino acid motif on hyphal surfaces on Candida albicans. *J. Biol. Chem.* **271,** 6298–6305 (1996).
201. Human Small Proline-Rich Protein 2B (SPRR2B) Protein (Myc-DYKDDDDK Tag), Recombinant | ABIN2732228. Available at: https://www.antibodies-online.com/protein/2732228/Small+Proline-Rich+Protein+2B+SPRR2B+protein+Myc-DYKDDDDK+Tag/. (Accessed: 15th May 2018)
202. Bassett, A. R. *et al.* Considerations when investigating lncRNA function in vivo. *eLife* **3,** (2014).
203. Lee, M.-H., Padmashali, R. & Andreadis, S. T. JNK1 Is Required for Lentivirus Entry and Gene Transfer. *J. Virol.* **85,** 2657–2665 (2011).
204. Weaver, J. C. Electroporation: a general phenomenon for manipulating cells and tissues. *J. Cell. Biochem.* **51,** 426–435 (1993).
205. Eckert, R. L. *et al.* Keratinocyte Survival, Differentiation, and Death: Many Roads Lead to Mitogen-Activated Protein Kinase. *J. Investig. Dermatol. Symp. Proc.* **7,** 36–40 (2002).
206. Canver, M. C. *et al.* Characterization of Genomic Deletion Efficiency Mediated by Clustered Regularly Interspaced Palindromic Repeats (CRISPR)/Cas9 Nuclease System in Mammalian Cells♦. *J. Biol. Chem.* **289,** 21312–21324 (2014).
207. Mali, P. *et al.* RNA-Guided Human Genome Engineering via Cas9. *Science* **339,** 823–826 (2013).
208. Kent, W. J. *et al.* The Human Genome Browser at UCSC. *Genome Res.* **12,** 996–1006 (2002).
209. Abhishek, S. & Palamadai Krishnan, S. Epidermal Differentiation Complex: A Review on Its Epigenetic Regulation and Potential Drug Targets. *Cell J.* **18,** 1–6 (2016).

210. Kawaji, H. *et al.* The FANTOM web resource: from mammalian transcriptional landscape to its dynamic regulation. *Genome Biol.* **10,** R40 (2009).
211. Yee, J.-K. Off-target effects of engineered nucleases. *FEBS J.* **283,** 3239–3248 (2016).
212. Jamal, M. *et al.* Keeping CRISPR/Cas on-Target. *Curr. Issues Mol. Biol.* **20,** 1–12 (2016).
213. Kim, M. *et al.* Passage-dependent accumulation of somatic mutations in mesenchymal stromal cells during in vitro culture revealed by whole genome sequencing. *Sci. Rep.* **7,** 14508 (2017).
214. Xiong, X., Boyett, J. M., Webster, R. G. & Stech, J. A stochastic model for estimation of mutation rates in multiple-replication proliferation processes. *J. Math. Biol.* **59,** 175–191 (2009).
215. Smits, J. P. H. *et al.* Immortalized N/TERT keratinocytes as an alternative cell source in 3D human epidermal models. *Sci. Rep.* **7,** 11838 (2017).
216. Slaymaker, I. M. *et al.* Rationally engineered Cas9 nucleases with improved specificity. *Science* **351,** 84–88 (2016).
217. Ji, Z., Song, R., Regev, A. & Struhl, K. Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. *eLife* **4,** e08890 (2015).
218. Kawashima, H. *et al.* A novel steroid receptor co-activator protein (SRAP) as an alternative form of steroid receptor RNA-activator gene: expression in prostate cancer cells and enhancement of androgen receptor activity. *Biochem. J.* **369,** 163–171 (2003).
219. Sen, G. L., Webster, D. E., Barragan, D. I., Chang, H. Y. & Khavari, P. A. Control of differentiation in a self-renewing mammalian tissue by the histone demethylase JMJD3. *Genes Dev.* **22,** 1865–1870 (2008).
220. Driskell, I. *et al.* The histone methyltransferase Setd8 acts in concert with c-Myc and is required to maintain skin. *EMBO J.* **31,** 616–629 (2012).
221. Xiang, J.-F. *et al.* Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. *Cell Res.* **24,** 513–531 (2014).
222. Brameier, M., Krings, A. & MacCallum, R. M. NucPred--predicting nuclear localization of proteins. *Bioinforma. Oxf. Engl.* **23,** 1159–1160 (2007).
223. Kosugi, S. *et al.* Six classes of nuclear localization signals specific to different binding grooves of importin alpha. *J. Biol. Chem.* **284,** 478–485 (2009).
224. Nguyen Ba, A. N., Pogoutse, A., Provart, N. & Moses, A. M. NLStradamus: a simple Hidden Markov Model for nuclear localization signal prediction. *BMC Bioinformatics* **10,** 202 (2009).
225. Lin, J. & Hu, J. SeqNLS: nuclear localization signal prediction based on frequent pattern mining and linear motif scoring. *PloS One* **8,** e76864 (2013).
226. Liu, B. *et al.* A cytoplasmic NF-κB interacting long noncoding RNA blocks IκB phosphorylation and suppresses breast cancer metastasis. *Cancer Cell* **27,** 370–381 (2015).
227. Amelio, I. *et al.* miR-24 triggers epidermal differentiation by controlling actin adhesion and cell migration. *J Cell Biol* **199,** 347–363 (2012).
228. Li, C.-X. *et al.* H19 lncRNA regulates keratinocyte differentiation by targeting miR-130b-3p. *Cell Death Dis.* **8,** e3174 (2017).

229. Wang, X. *et al.* miR-378b Promotes Differentiation of Keratinocytes through NKX3.1. *PLOS ONE* **10,** e0136049 (2015).
230. Barbollat-Boutrand, L. *et al.* MicroRNA-23b-3p regulates human keratinocyte differentiation through repression of TGIF1 and activation of the TGF-ß-SMAD2 signalling pathway. *Exp. Dermatol.* **26,** 51–57 (2017).
231. Hildebrand, J. *et al.* A comprehensive analysis of microRNA expression during human keratinocyte differentiation in vitro and in vivo. *J. Invest. Dermatol.* **131,** 20–29 (2011).
232. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30,** 2114–2120 (2014).
233. External RNA Controls Consortium. Proposed methods for testing and selecting the ERCC external RNA controls. *BMC Genomics* **6,** 150 (2005).
234. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29,** 15–21 (2013).
235. Kolde, R. Pheatmap: pretty heatmaps. *R Package Version* **61,** (2012).
236. Hartley, S. W. & Mullikin, J. C. QoRTs: a comprehensive toolset for quality control and data processing of RNA-Seq experiments. *BMC Bioinformatics* **16,** 224 (2015).
237. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14,** 178–192 (2013).
238. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29,** 24–26 (2011).
239. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** (2014).
240. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* **57,** 289–300 (1995).
241. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4,** 44–57 (2009).
242. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37,** 1–13 (2009).
243. Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14,** 128 (2013).
244. Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44,** W90–W97 (2016).
245. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, YEAR).
246. Raney, B. J. *et al.* Track data hubs enable visualization of user-defined genome-wide annotations on the UCSC Genome Browser. *Bioinformatics* **30,** 1003–1005 (2014).
247. Afgan, E. *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.* **44,** W3–W10 (2016).
248. Andrews, S. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. (Accessed: 25th April 2018)

249. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10,** R25 (2009).
250. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9,** 357–359 (2012).
251. Barnett, D. W., Garrison, E. K., Quinlan, A. R., Stromberg, M. P. & Marth, G. T. BamTools: a C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics* **27,** 1691–1692 (2011).
252. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinforma. Oxf. Engl.* **25,** 2078–2079 (2009).
253. Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.* **9,** R137 (2008).
254. Feng, J., Liu, T., Qin, B., Zhang, Y. & Liu, X. S. Identifying ChIP-seq enrichment using MACS. *Nat. Protoc.* **7,** 1728–1740 (2012).
255. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26,** 841–842 (2010).
256. Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31,** 166–169 (2015).
257. Tamura, K. & Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10,** 512–526 (1993).
258. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **33,** 1870–1874 (2016).
259. Faustino-Rocha, A. *et al.* Estimation of rat mammary tumor volume using caliper and ultrasonography measurements. *Lab Anim.* **42,** 217–224 (2013).
260. Tomayko, M. M. & Reynolds, C. P. Determination of subcutaneous tumor size in athymic (nude) mice. *Cancer Chemother. Pharmacol.* **24,** 148–154 (1989).
261. Corces, M. R. *et al.* An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* **14,** 959–962 (2017).
262. Streit, S., Michalski, C. W., Erkan, M., Kleeff, J. & Friess, H. Northern blot analysis for detection and quantification of RNA in pancreatic cancer cells and tissues. *Nat. Protoc.* **4,** 37–43 (2009).
263. Livak, K. J. & Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods* **25,** 402–408 (2001).

11 Acknowledgements

This thesis could not have been completed without the help and support of many people.

MANY THANKS…

… to my supervisor PD Dr. Markus Kretz for the opportunity to start my PhD project in his laboratory. I am very grateful for all the inspiring scientific discussions, his positive attitude and encouragement and last but not least his appreciation along this journey.

… to all members of my thesis examination committee, especially Prof. Dr. Gunter Meister who supported and accompanied my whole scientific career.

… to all my collaborators. Prof. Dr. Rainer Merkl for the phylogenetic analysis, the whole AG Rehli for their welcoming way as well as their support during the epigenetic studies and Dr. Astrid Bruckmann and Eduard Hochmuth for the mass spectrometry analysis and data evaluation.

… to Prof. Dr. Paul Khavari and Dr. Zurab Siprashvili for the unique opportunity and the excellent supervision during my time at Stanford. Moreover, I am grateful to Dr. Cari Lee, Angela Mah & Rajani Shenoy for supplying several cells and reagents for my experiments.

… to all members of the AG Kretz (present and alumni). The unconditional support of Bianca for all matters, her always positive attitude and sympathetic ear during countless discussions made every challenge feasible. Moreover, I am grateful to my fellow PhD companion Johannes for discussions, reading this manuscript and a lot of support along the way. Also, many thanks to Sonja for scientific discussions and inspirations.

… to the whole AG Meister and AG Medenbach, especially Gerhard and Norbert for their endless help and support for all bioinformatic obstacles. Furthermore, Johannes and Hung who made the lab-life more enjoyable on numerous occasions.

… to my family, not only for their financial support along this long and sometimes challenging journey but also for their unconditional love.

As with many things in life, the most important things can be found at the very end. Along this line I want to thank my wonderful wife Nathalie at last, for always helping me through the tough times but also sharing the good times with me. I am very grateful for having you at my side during all the adventures that we encountered so far and I am looking forward to all the future endeavors that our life might offer!