IRF2 is a master regulator of human keratinocyte stem cell fate

Nicolas Mercado1, Gabi Schutzius1, Christian Kolter1, David Estoppey1, Sebastian Bergling1, Guglielmo Roma1, Caroline Gubser Keller1, Florian Nigsch1, Adrian Salathe1, Remi Terranova2, John Reece-Hoyes3, John Alford3, Carsten Russ3, Judith Knehr1, Dominic Hoepfner1, Alexandra Aebi1, Heinz Ruffner1, Tanner C. Beck4, Sajjeev Jagannathan4, Calla M. Olson4, Hadley E. Sheppard5, Selma Z. Elsarrag5, Tewis Bouwmeester1, Mathias Frederiksen1, Felix Lohmann6,7*, Charles Y. Lin4,5,7* & Susan Kirkland1,7*

Resident adult epithelial stem cells maintain tissue homeostasis by balancing self-renewal and differentiation. The stem cell potential of human epidermal keratinocytes is retained in vitro but lost over time suggesting extrinsic and intrinsic regulation. Transcription factor-controlled regulatory circuitries govern cell identity, are sufficient to induce pluripotency and transdifferentiate cells. We investigate whether transcriptional circuitry also governs phenotypic changes within a given cell type by comparing human primary keratinocytes with intrinsically high versus low stem cell potential. Using integrated chromatin and transcriptional profiling, we implicate IRF2 as antagonistic to stemness and show that it binds and regulates active cis-regulatory elements at interferon response and antigen presentation genes. CRISPR-KD of IRF2 in keratinocytes with low stem cell potential increases self-renewal, migration and epidermis formation. These data demonstrate that transcription factor regulatory circuitries, in addition to maintaining cell identity, control plasticity within cell types and offer potential for therapeutic modulation of cell function.
issue stem cells (SCs) maintain homeostasis and repair damage throughout life. Stem cells reside in specialized niches which preserve and balance stem cell renewal with delivery of differentiated cells. The best understood adult tissue stem cell, the hematopoietic stem cell (HSC), resides in the bone marrow niche and supplies all blood cell lineages. HSCs are distinct cell types which can be isolated and shown to renew the entire blood system experimentally and clinically with transplantation therapies. The HSC paradigm drove the search for similarly rare “hard-wired” epithelial stem cells. However, recently the existence of equivalent slow cycling quiescent epithelial stem cells has been questioned, leading to the proposal that shifting focus from stem cell phenotype to function would better enable advances in epithelial stem cell biology.

The Epidermis, the outermost layer of skin comprised mainly of keratinocytes, acts as a physical barrier and is replaced every few weeks by resident stem cells residing in the basal layer. Cultivation of human keratinocytes in vitro was first achieved using 3T3 feeders and is now enabled with specialist serum-free culture media. Epidermal stem cell function is retained in vitro, albeit with a high degree of heterogeneity suggesting some intrinsic regulation, and used therapeutically in burns patients. Furthermore, in vitro expanded transgenic epidermal stem cells were recently used to replace an entire epidermis in life-threatening epidermolysis bullosa, supporting the existence of a long-lived population of self-renewing epithelial stem cells.

Recent evidence demonstrates that the integration of niche signals with transcriptional circuitries specifies stem cell fate decisions, arguing for extrinsic and intrinsic maintenance of stemness. A role for epigenetic factors and individual TFs has been demonstrated in epithelial stem cell fate decisions and transcriptional enhancer clusters that specify cell identity, termed super-enhancers, have been shown to be involved in lineage decisions in hair follicle stem cells. As transcription factors (TFs) have a pivotal role in mediating cell state, exemplified in the discovery of induced pluripotent stem cells, we set out to identify the transcriptional landscape specifying the epithelial stem cell state.

TFs recruit co-activators like BRD4, mediator, and histone acetyltransferases to genes in order to drive their expression. We and others have shown that in any given cell type, cell identity is specified by a highly interconnected regulatory network of TFs termed core regulatory circuitry (CRC). CRC TFs often physically interact at the protein–protein level, are regulated by large enhancers or super-enhancers, and tend to form highly interconnected regulatory circuits by binding to the promoters and enhancers of other CRC TFs. Although such transcriptional circuitries define cell identity, they have not been used to differentiate cell states within a single cell type.

Here, we define the transcriptional network maintaining the human epidermal stem cell state using comprehensive transcriptional and epigenomic profiling, computational circuitry analysis and genetic validation. To generate cell numbers necessary for such an unbiased analysis, we employed replication-induced loss of stem cell function as our model system, supported by previous studies showing concomitant replicative senescence and differentiation in this in vitro model and in vivo in aged mouse gingival epithelium. Multiple phenotypic read-outs confirm high and low stem cell potential in our populations. Epidermal stem cells are more clonogenic, migratory, adherent to extracellular matrix (ECM), and better able to form a stratified epidermis in 3D skin models. We delineate chromatin-mediated cis-regulation associated with loss of stem cell function and identify putative TFs acting as master regulators of keratinocyte stem cell function, several of which are validated using CRISPR-mediated genetic loss of function. Most significantly, we show that loss of IRF2 induces a rapid, dramatic, and sustained enhancement of stem cell function in human keratinocytes.

CRC analysis identifies transcriptional master regulators in cell state transitions supporting the concept of a “hard-wired” stem cell circuitry where interconnectivity of TF network maintains stem cell function. Clearly, environmental factors will be expected to intersect with this circuitry to define the final ability of stem cells to retain their self-renewal state. Such understanding opens up possibilities for therapeutic intervention in epithelial regeneration.

**Results**

*Keratinocytes steadily lose stem cell potential in vitro.* Primary human keratinocytes lose stem cell function over time in vitro, and were characterized at early passage 7 as High Stem Cell Potential–Human Keratinocytes (HSCP-HKs) and after further passaging until passage 15 as Low Stem Cell Potential–Human Keratinocytes (“LSCP-HKs”) (Fig. 1a). HSCP-HKs generated significantly more and larger colonies (Fig. 1b), migrated faster (Fig. 1c), and attached either more quickly (Collagen I) or in larger numbers (Collagen IV, Fibronectin and Laminin V) than LSCP-HKs (Fig. 1d), supporting higher stem-cell function. Serialpassaging of primary cells induces senescence, which we reproduced in our system as confirmed by an increase in Senescence-Associated Beta-Galactosidase (SA-β-Gal) staining in LSCP-HK cells (Fig. 1e). Transcriptomic analysis (RNA-seq) and comparison of LSCP-HKs with HSCP-HKs shows enhanced expression of 178 genes and decreased expression of 416 genes in LSCP-HKs (Fig. 1f, Supplementary Data 1). Gene set analysis (GSA) of the HSCP-HK transcriptome shows enrichment of genes categorized as “Cell Cycle/Mitotic”, “DNA Replication”, and “Cell Adhesion Molecules” (Fig. 1g, Supplementary Fig. 1a, Supplementary Data 2), exemplified by known cell cycle regulators (CCNA2, AURKB, E2F1, and UBE2C) and supporting their higher proliferation and ECM-binding affinity (Fig. 1g, Supplementary Fig. 1a, Supplementary Data 2). LSCP-HKs were enriched for categories related to “Immune Response” and “Inflammatory Response” with upregulation of CXCL8, TNF, NFKBIZ, PTGS2, S100A9, and IL1B (Fig. 1g, Supplementary Fig. 1a, Supplementary Data 2), a transcriptome with similarities to other senescent phenotypes in line with the increased SA-β-Gal staining. Reduced cell proliferation in LSCP-HKs accompanied decrease in cell cycle promoting genes, CCNA2 and UBE2C, and concomitant increase in negative regulators such as CDKN2A (p16INK4a) but not CDKN1A (p21WAF1), as previously reported in passaging induced senescence. Generation of reactive oxygen species and ECM degradation was inferred by upregulation of anti-oxidant response genes such as NQO1, SOD2, GPX2, and PTGS2 and by matrix metalloproteinases (MMP1, -3, -9) all of which are again associated with stress-induced premature senescence (Supplementary Fig. 1c). In addition, an increase in pro-inflammatory chemokines and cytokines including CXCL8, TNF, IL1A, and IL1B suggests a possible senescent associated secretory phenotype. Finally, FOXM1, a member of the Forkhead superfamily of TFs, previously associated with high proliferative potential in keratinocytes, was decreased in LSCP-HKs (Supplementary Fig. 1c).

Metacore TF connectivity analysis showed an association of LSCP-HKs with pro-inflammatory (NF-κB family) and Interferon signaling TFs (STAT1 and IRF1), and HSCP-HKs with pluripotency-associated TFs (NANOG, Oct-3/4, and SOX2), cell cycle regulators (E2F1 and E2F4) and p63, a TF regulator of epithelial stem cell compartments (Fig. 1h).
BRD4 is dynamically distributed between HSCP-HK and LSCP-HK. Differential gene expression and stem cell function between HSCP-HKs and LSCP-HKs suggests a change in cell state which is imposed by a cohort of specific TFs and epigenetically maintained\(^\text{15,32}\). We contrasted chromatin landscapes between HSCP-HKs and LSCP-HKs by genome-wide profiling of H3K27ac, a histone mark found at active promoters and enhancers\(^\text{12,33–35}\), and of BRD4, a chromatin co-activator directly associated with active transcription elongation\(^\text{36}\), as well as of chromatin accessibility to identify nucleosome free cis-regulatory elements likely to harbor TF binding sites\(^\text{37}\). Using H3K27ac, we defined 24,614 discrete regions of active chromatin present in at least one sample in HSCP-HKs or LSCP-HKs. Although unbiased clustering of H3K27ac occupancy profiles segregated HSCP-HK and LSCP-HK replicates (\(n = 3\)), overall the active chromatin landscapes were highly similar.
(Supplementary Fig. 2a) and likely reflect keratinocyte identity. This suggests that stem cell potential arises from subtle alterations of chromatin landscape as opposed to de novo formation of euchromatin or heterochromatin.

Previously, we demonstrated that TFs can redistribute the transcriptional co-activator BRD4 across chromatin landscapes to reshape both chromatin and gene expression. Across active genes (n = 7278) in both HSCP-HK and LSCP-HK, we quantified BRD4 total occupancy (area under the curve, AUC) at all promoter and proximal enhancer elements (<50 kb from the transcription start site (TSS)). Ranking active genes by the change in BRD4 occupancy between HSCP-HK and LSCP-HK (Fig. 2a), revealed a strong association with changes in chromatin acetylation and gene expression both globally (Fig. 2a) and at individual exemplary genes (Fig. 2b, c). Notably, at genes where BRD4 was redistributed, H3K27ac chromatin was present in both HSCP-HKs and LSCP-HKs (Fig. 2c) suggesting that despite similarities in active chromatin landscapes, redistribution of transcriptional co-activators like BRD4 may play an important role in modulating stem cell potential.

Circuitry as potential TF master regulators. To quantify the role of individual TFs in reshaping the keratinocyte active chromatin landscape, we prioritized a cohort of TFs possessing features of CRC TFs (Fig. 3a). Actively expressed TFs were filtered for the presence of a known binding motif regulation by an enhancer element, and evidence of protein–protein interaction with other TFs. This resulted in 60 remaining TFs that via unbiased Markov Chain Linkage clustering segregated into six distinct clusters with two predominating clusters. Compared to other actively expressed TFs, these 60 displayed higher regulatory connectivity (binding to each other’s regulatory regions) (Supplementary Fig. 3a–c). Interestingly, these two large clusters contained TFs with higher inter-cluster regulatory connectivity (Supplementary Fig. 3d) and with biological functions consistent with thematic HSCP-HK or LSCP-HK gene expression and Metacore signatures (e.g., growth/proliferation for HSCP-HK and inflammatory signaling for LSCP-HK) (Fig. 3a). Based on these properties of enhancer regulation, high regulatory, and protein–protein interactions, and HSCP-HK and LSCP-HK associated functions, we hypothesized that these two clusters represent the CRC TFs that underlie stem cell potential in keratinocytes.

Using motif enrichment finding with FIMO, CRC TFs were assigned to putative nucleosome free regions (as defined by ATAC-seq). Clustering of CRC TF motif occupancy revealed modules with similar or contrasting binding patterns (Supplementary Fig. 3e). Interestingly, IRF TFs showed a highly dissimilar pattern of occupancy to a set of basic helix loop helix (bHLH) TFs (MYC, MAX, SREBF1/2, and BHLHE40) and a set of promoter associated housekeeping factors (E2F4, MAZ, YY1, SP1, TFAP2A, KLF5, and KLF13), both of which associated with the HSCP-HK cluster 1. These data suggested that the HSCP-HK and LSCP-HK associated clusters of TFs have distinct targets and gene-expression programs.

To measure the effect of individual CRC TFs on the keratinocyte chromatin landscape, we quantified changes in BRD4 occupancy proximal to predicted nucleosome free binding sites as a measure of TF outward effect (BRD4 OUT Degree) (Fig. 3b, c). We next ranked all 60 CRC TFs by change in BRD4 OUT degree between HSCP-HK and LSCP-HK. This ranking revealed a gradient of TF activity delineating BRD4 changes between HSCP-HKs and LSCP-HKs. TFs regulating growth and proliferation such as MYC, MAZ, and E2F4 drove increased BRD4 in HSCP-HKs and conversely, pro-inflammatory TFs including the IRF family drove increased BRD4 in LSCP-HKs. Based on the ability of these CRC TFs to drive BRD4 re-localization between HSCP-HKs and LSCP-HKs, we hypothesized that these TFs are not only associated with HSCP-HK versus LSCP-HK identity, but are functionally required for maintenance of keratinocyte stem cell function (Fig. 3d).

CRISPR-Cas9 screen identifies function modulating TFs. To validate the role of CRC predicted TFs in keratinocyte stem cell function, we performed a pooled CRISPR–Cas9 screen and sampled the population over time. Keratinocytes with greater stem cell properties proliferate faster and survive longer in vitro, therefore single-guide RNA (sgRNAs) targeting genes promoting stem cell function should be quickly lost from the pool while sgRNAs to genes antagonistic to stem cell function should persist. The screen was performed in neonatal keratinocytes which, unlike adult keratinocytes, were sufficiently robust for infection and selection of a Cas9 expressing population essential for such a pooled screen (Supplementary Fig. 4). The pool contained 2698 sgRNAs targeting 540 genes, including 34 out of the 60 putative TF regulators from the CRC analysis (Supplementary Fig. 4b), along with predicted stem cell regulators and controls. After infection at day 0 and selection with puromycin for 3 days, cell pellets were collected for DNA sequencing at days 12, 19, 28, 33, 38, 43, and 47 post infection (Fig. 4a). sgRNA barcode counts over time showed a rapid decrease with most barcodes lost at day 38 (Supplementary Fig. 5a, b) suggesting that Cas9 expressing keratinocyte proliferation was limited regardless of the effects of editing. Evidence from positive control sgRNAs targeting essential genes supported this suggestion showing only modest effects at days 12 and 19 (Supplementary Fig. 5c). Therefore, negative effects of editing on proliferation were evaluated early (days 12 and 19) while increased survival and cell proliferation were assessed later (days 38 and 43) (Fig. 4b, Supplementary Data 3). When evaluating the top seven most strongly represented CRC-associated TFs by significance (redundant siRNA activity (RAS) Dnne/Up) and magnitude (Q1/Q3), we observed an early
“drop-out” of HSCP-HK TFs (MYC, E2F4, MAZ, TEAD1, and NFE2L2), indicative of a role in maintaining stem cell function (Fig. 4b). E2F4 and MYC, within the top three CRC candidates, are known positive regulators of keratinocyte proliferation⁴⁵,⁴⁶ (Fig. 3d). MYC was amongst the top TFs whose editing drastically reduced cell proliferation (from day 12 to day 28) whereas a “neutral” CRC TF, SNAI2, was shown to be among the top TFs from day 19 onwards (Supplementary Fig. 5d, e). SNAI2 is known to repress expression of differentiation genes⁴⁷. At days 38 and 43, we also found an overrepresentation of sgRNAs to LSCP-HK TFs (IRF6, RELA, STAT1, ERF, TGIF1, VDR, RUNX1, ESSRA, and IRF2) implicating these CRC-TFs in the loss of stem cell phenotype with time in vitro (Fig. 4b). Of particular interest was the strong enrichment of sgRNAs targeting IRF2 from day 28 onward.

Fig. 2 Dynamic redistribution of BRD4 between HSCP-HKs and LSCP-HKs. **a** Ranking of all active genes by change in BRD4 occupancy at their promoter and proximal enhancers between HSCP-HK and LCSP-HK, with fold change in BRD4 occupancy shown in top panel. In middle and bottom panels, fold change in H3K27ac and mRNA levels are shown using a binned average (100 genes per bin), respectively. Error bars represent 95% confidence intervals (CI) of the average as determined by resampling with replacement (1000 permutations). A trend line (black) is overlaid by lowest locally weighted regression. **b** Transcript levels of RBP1 and ZFP42 based on RNA-seq (FPKM) comparing HSCP-HK to LSCP-HK. Error bars represent standard deviation. **c** Chromatin occupancy of BRD4 and H3K27ac (ChIP-seq) as well as chromatin accessibility (ATAC-seq) at RBP1 and ZFP42 genomic loci as read density in units of reads per million per base pair (rpm/bp). Individual HSCP-HK (blue) and LSCP-HK (red) replicates are shown as shaded traces. The average of replicates is drawn as a solid line.
Calculating change in BRD4 IN and OUT degree for each TF

For a given TFi:

Enhancer1
Gene1
TFi Gene
Other TF

Enhancer2
Gene3
Gene2
Enhancer3

Log2 change in BRD4 OUT degree
HSCP-HK/LSCP-HK
−0.05
0.00
0.05
0.10
0.15

IRF2
IRF9
ERF
SREBF1
IRF7
SREBF2
IRF6
IRF1
STAT1
ETS2
RFX5
RXRA
IRF3
ELF1
TCF7L1
MAFF
USF2
VDR
RARG
ESRRA
HBP1
NFKB2
RUNX1
TP63
ETS1
STAT6
FOXM1
TGIF1
KLF13
RELA
SMAD1
ATF4
NFKB1
USF1
SOX9
BHLHE40
SNAI2
SOX4
NR3C1
STAT3
NFE2L2
JUND
JUN
TEAD1
TCF12
SMAD3
MAX
BACH1
SRF
FOSL1
KLF5
SP3
MAZ
POU2F2
TFAP2A
MYC
EGR1
E2F4
YY1

Building keratinocyte TF core regulatory circuitry using the union of HSCP-HK and LSCP-HK landscapes

Active genes in both > 10 FPKM and H3K27ac and BRD4 at the TSS (+/− 1kb)
Annotated human TFs with position weight matrices (Jolma 2013, Jaspar, Transfac)

7278
126
24,614

Discreet active cis-regulatory regions in both (H3K27ac)
H3K27ac MACS-enriched peaks
rpm/bp
Ac
Ac
TF1
TF2
TF3
1.0
1.0
rpm/bp
1.0
1.0
rpm/bp
2.0
2.0

UBA6

(1) Active cis-regulatory regions defined as the union of discreet H3K27ac regions in HSCP-HK and LSCP-HK
(2) ATAC-seq subpeaks defined as the union of regions between HSCP-HK and LSCP-HK
H3K27acATAC
(3) Candidate nucleosome free region defined by ATAC-seq (D)
ΔBRD4 used to quantify change in TF activity

Map connectivity to active chromatin landscape using ATAC-seq peaks as NFRs
Filter and cluster by protein–protein interaction network (String dB and MCL clustering)
Keratinocyte TFs with P-P interactions (n = 60)

Experimental p–p interaction
Predicted p–p interaction

Cluster 1
HSCP-HK

Cluster 2
LSCP-HK

Cluster 3

Cluster 4

Cluster 5

Cluster 6

Cluster 8

Calculating change in BRD4 IN and OUT degree for each TF
For a given TFi:

BRD4 IN degree: BRD4 occupancy at cis-regulatory regions of TFi

BRD4 OUT degree: BRD4 occupancy at cis-regulatory regions of TFi binding sites at other cis-regulatory regions

Fig. 3 Changes in TF core regulatory circuitry between HSCP-HKs and LSCP-HKs. a Schematic showing method to identify potential CRC TFs. The union of active genes and H3K27ac sites in both HSCP-HKs and LSCP-HKs was formed in order to create the overall active cis-regulatory landscape for HKs. Subsequently, trans factors and cis-regulatory elements within the HK regulatory landscape showing LSCP or HSCP specific activity were identified, resulting in network of 60 candidate CRC TFs. Each TF node is colored based on Markov Chain Linkage clustering. Solid and dashed edges represent experimental and predicted protein–protein interactions, respectively, as determined by the StringDB protein–protein interaction database. b Tracks of BRD4 and H3K27ac ChiP-seq as well as chromatin accessibility (ATAC-seq) signal at UBA6 genomic locus with schematic detailing how ATAC-seq peaks are used for TF motif searching followed by quantification of proximal changes in BRD4. c Schematic showing how BRD4 IN and OUT degree are calculated for each CRC TF. d Ranking of the fold change in BRD4 OUT degree for all CRC TFs in HSCP-HK over LSCP-HK. Error bars represent 95% CI of the mean as determined by resampling with replacement (1000 permutations)
onwards (Supplementary Fig. 5e) as IRF2 was predicted as the top CRC-TF regulating LSCP-HKs (Fig. 3d). IRF2 knockdown (KD) in mice leads to epidermal hyperplasia suggesting a physiological role for IRF2 in promoting a more differentiated, less proliferative epidermis48. Furthermore, IRF6 and RELA have been shown to negatively regulate keratinocyte proliferation49,50. Besides TFs, KRT5 editing was shown to quickly reduce keratinocyte proliferation, confirming its essential role in basal keratinocytes51 (Fig. 4b). CDKN1A (p21) and MAPK14 (p38α) editing was shown to increase cell proliferation and could have been counteracting senescence associated cell cycle arrest52,53 possibly related to DNA damage induced by CRISPR–Cas954 (Fig. 4b, days 38 and 43). In addition to IRF2, BET bromodomain protein BRD2 knockout was strongly associated with enhanced cell proliferation (Fig. 4b, Supplementary Fig. 5e).

Some top CRC-predicted TF regulators did not validate in the screen, possibly due to technical limitations and variability in CRISPR editing efficiency or assumptions around TF occupancy used for the CRC analysis. For example, only 3 of the 5 sgRNA guides added against our top hits BRD2 and IRF2 remained at the end of the screen (Supplementary Fig. 5e). IRF2 sgRNAs (#2, #3, #4) enhanced keratinocyte migration while sgRNAs #1, #2, #3, and #4 enhanced clonogenicity, confirming IRF2 as a negative regulator of stem cell function (Supplementary Fig. 6b, c). Interestingly, although all BRD2 sgRNA enhanced the clonogenic potential of keratinocytes (Supplementary Fig. 6c) no effect on migration was observed (Supplementary Fig. 6b). One explanation could be a time dependent increase in stem cell function only captured in the longer clonogenic assay. The BRD2 findings are interesting in light of recent studies describing its role in keratinocyte inflammatory responses55 and warrants further research. However, for this study we decided to focus on IRF2 as its KD led to robust and reproducible phenotypic changes in keratinocytes.
CRISPR–Cas9 editing of TFs modulates stem cell function. To orthogonally validate the role of select TF candidates, we repeated genetic loss of function studies in adult keratinocytes. Editing of YY1 and SNAI2 in HSCP-HKs resulted in a loss of clonogenicity and migration (Fig. 5a). Further physiological relevance was evaluated in a 3D human dermo-epidermal skin model that tests the ability of keratinocytes to form a stratified epidermis when combined with an artificially constructed dermal layer56, a model previously shown to link keratinocyte stemness, delayed replicative senescence and increased epidermal thickness44. YY1 KD and SNAI2 KD in adult keratinocytes significantly impaired the formation of an epidermis and failed to maintain an undifferentiated basal layer (Fig. 5b, Supplementary Fig. 7a, b). Furthermore, YY1-KD and SNAI2-KD keratinocytes failed to migrate under the dermis as seen with control keratinocytes (Fig. 5b, Supplementary Fig. 7a, b), confirming both YY1 and SNAI2 as master regulators of keratinocyte stem cell function. To determine whether either YY1 or SNAI2 are sufficient to induce keratinocyte stem cell function would require their individual overexpression but this was not performed in this study.

IRF2 edited keratinocytes took longer to display phenotypic changes and were passaged until their proliferation rate surpassed controls. IRF2-KD keratinocytes showed increased clonogenicity and migration (Fig. 5c) and formed a thicker, more cellular controls. IRF2-KD cells were highly migratory generating an epidermis beneath the dermal construct which prevented statistical analysis of epidermal thickness (Fig. 5d, Supplementary Fig. 7d). To further characterize IRF2-KD keratinocytes in the clonogenic assay, we quantified Holoclones (formed by cells with highest proliferation and lifespan) and Meroclones (a transitional stage between holoclones and paraclones)37. IRF2 KD keratinocytes form more holoclones (Fig. 5c), reflecting greater stem cell function and self-renewal capacity9.

To evaluate IRF2 specificity, we edited another family member, IRF9, in HSCP-HKs. IRF9-KD cells showed no increased clonogenicity or migration (Supplementary Fig. 7c) or epidermal thickness in the 3D model but a better delineation of the undifferentiated basal cell layer (Supplementary Fig. 7a, b).

IRF2 knockout restores HSCP-HK gene-expression programs. As IRF2 KD in HSCP-HK cells prevented loss of stem cell function over time in vitro, we next asked whether IRF2 KD in LSCP-HK cells could restore stem cell function and transcriptional signature of HSCP-HKs (also edited for comparison). IRF2 KD induced global transcriptional changes in LSCP-HKs toward those of HSCP-HKs (Supplementary Fig. 8a–c), illustrated using genome-wide clustering (Fig. 6a). Cluster 1 showed genes upregulated in LSCP-HKs, compared to HSCP-HKs, and downregulated by IRF2 KD toward the expression in HSCP-HKs. The top GO terms for Cluster 1 are “Keratinization” and “Cornification”, processes involved in terminal differentiation, cell death and the formation of a cross-linked insoluble barrier, suggesting that IRF2 editing is associated with keratinocyte differentiation (Supplementary Fig. 9e, Supplementary Fig. 10). Two important genes in these processes, involucrin and SPPR1A (small proline rich protein 1A), components of the insoluble cross-linked envelope critical for skin barrier formation, were reduced by IRF2 KD in LSCP-HKs back to HSCP-HK levels Supplementary Fig. 9e). Interestingly, most of the genes in the subcluster 3, which were reversed back to HSCP-HKs (28 genes), are involved in GO term “Keratinocyte Differentiation” (Supplementary Fig. 10).

Cluster 2 is the largest cluster and characterized by genes downregulated in LSCP-HKs compared to HSCP-HKs and upregulated by IRF2 KD in LSCP-HK to the levels of HSCP-HKs (Fig. 6a). The top GO term associated with this cluster is “Cell Cycle”, as cell cycle-associated genes were restored to HSCP-HK levels after IRF2 editing in LSCP-HKs, suggesting IRF2 reduces cell proliferation after serial passaging (Fig. 6a, Supplementary Fig. 9e). These genes were also increased in HSCP-HKs with IRF2 KD, suggesting that cell cycle potential can be further increased even in low passage keratinocytes. Further sub-clustering showed that expression of most cell cycle associated genes was upregulated towards HSCP-HK levels following IRF2 KD in LSCP-HKs (Supplementary Fig. 11). Interestingly, CDKN2A (p16INK4a) was the only cell cycle associated gene linked with cluster 1, hence upregulated in LSCP-HKs and restored to HSCP-HK levels with IRF2 editing (Fig. 6a), supporting a reversal of senescence-associated cell cycle arrest.

Cluster 3 showed genes largely unchanged by IRF2 editing in both LSCP- and HSCP-HKs. Cluster 4 showed genes upregulated by IRF2 KD both in LSCP- and HSCP-HKs.
**Fig. 5** Loss of IRF2 induces stem cell potential whereas loss of YY1 or SNAI2 reduces it.  

**a** Assessment of CRISPR–Cas9 mediated knockdown of YY1 (left panel) or SNAI2 (right panel) on colony formation and migration (as in Fig. 1) in HSCP-HK compared to control cells, with protein levels monitored by Western blot analysis utilizing β-actin protein level as a loading control.  

**(b)** Comparison of SNAI2 KD and YY1 KD to control cells in their functional ability to generate an epidermis in human dermo-epidermal 3D model. # indicates Epidermis layer. Epidermis thickness (µm) was quantified for each condition and differences assessed using One-way ANOVA with Holm–Šidák multiple comparisons test. ****p < 0.0001. Scale bar = 167 µm. Means ± S.D. of n = 4 biological replicates.  

**(c)** Assessment of colony formation and migration in LSCP-HK cells with CRISPR–Cas9 mediated knockdown of IRF2 compared to control cells, with protein levels monitored by Western blot analysis utilizing β-actin protein level as a loading control.  

**(d)** Epidermis formation in human dermo-epidermal 3D model comparing IRF2 KD cells to controls. # indicates epidermis layer. Epidermis thickness (µm) was quantified for each condition (means ± S.D. of n = 3 biological replicates for CT and n = 1 for IRF2 KD). Scale bar = 167 µm.  

**(e)** Comparison of ability to clonally expand aHPEK-IRF2-KD versus control cells, where individual keratinocytes either fail to form a colony or form a colony belonging to three broad morphological types termed holo-, mero-, and paraclones (pictured, scale bar = 1.1 mm). Total colonies of each morphological type were counted and plotted as means ± S.D. of n = 4 biological replicates. *p < 0.05, ***p < 0.001, ****p < 0.0001 (Ordinary One-way ANOVA with Holm–Šidák multiple comparisons test).
IRF2 is best known as an antagonist to interferon mediated antiviral gene activation by IRF1 but is also associated with gene activation, as recapitulated in our data. GO term “Interferon Signaling” was strongly associated with LSCP-HKs as compared to HSCP-HKs, implicating interferon signaling in stem cell function (Supplementary Fig. 12). However, IRF2 KD in LSCP-HKs only partially reduced the genes ascribed to this GO term back to HSCP-HK levels, implying that interferon signaling is not the dominant driver of loss of stem cell potential (Supplementary Fig. 12). Inflammatory response-related genes are strongly associated with LSCP-HKs as compared to HSCP-HKs and also with LSCP-HKs CT versus LSCP-HKs IRF2 KD (Supplementary Fig. 13). Several of these genes were upregulated and reversed by IRF2 KD to HSCP-HK levels such as IL1B, KLK5, LCE3D, SPRR2B, KRT1, FLG, SPRR2A, SPRR1B, SPRR2G, TGM1, CDSN, SPRR1A, IVL, SPRR2D, PI3, SPRR2E, DSG1, SPRR3, KRT80, LCE3E, CASP14, KRT75, SPINK5, CDKN2A, CENPE, TOP2A, E2F2, PLK1, AURKB, NCAPH, NCAPG, BIRC5, RBL1, NEK2, CDC6, BUB1B, POLE2, NDC80, CDC25C, CCNB2, CDKN2C, PRIM1, ORC1, MCM10, MAD2L1, BUB1, CCNA2, ESPL1, CDC45, E2F1, CDK1, CCNE2, CLUSTER 2 (n = 5622), CLUSTER 1 (n = 3705), CLUSTER 4 (n = 1857), CLUSTER 3 (n = 1595), CTL, Cell cycle, Keratinization, Stem cell potential.
S100A8, and PTGS2 (Supplementary Fig. 13). Although TNF was increased in LSCP-HKs, IRF2 KD had no effect on TNF gene expression and only partially inhibited CXCL8. As shown in Supplementary Fig. 1c, various senescence-associated genes were either upregulated or downregulated in LSCP-HK compared to HSCP-HK as anticipated from their enhanced SA-Gal staining. IRF2 KD restored the expression of many of these genes, including AURKA, AURKB, CCNA2, CCNB1, CDKNA2A (p16INK4), FOXM1, IL1B, MMP3, PCNA, PTGS2, and UBE2C, strongly suggesting a function as a positive regulator of keratinocyte senescence. Our data showed a strong association between genes upregulated in LSCP-HK cells (compared with HSCP-HKs) and those upregulated in psoriatic compared to normal skin (Supplementary Fig. 14). Furthermore, we detected an equally strong association when comparing differential expression between LSCP-HKs CT and LSCP-HKs IRF2 KD with genes expressed in psoriatic skin, suggesting that many psoriasis-related genes were upregulated by serial passaging and reversed by IRF2 KD to low-passage-HSCP-HK levels (Supplementary Fig. 13).

As mentioned before, IRF2 can act both as an inhibitor and activator of gene expression. IRF2 binding was predicted to be increased in the regulatory regions of various genes (Fig. 3d, Supplementary Data 4) in LSCP-HKs compared to HSCP-HKs, and among these putative target genes we found expression to be both upregulated and downregulated when comparing LSCP-HK and HSCP-HK (Supplementary Fig. 15 left), arguing that an increased binding of IRF2 either activated or inhibited expression of a specific gene (Supplementary Fig. 15 left). The resulting genome-wide hierarchical clustering revealed two large clusters (2 and 3) and two smaller ones (1 and 4) where genes in cluster 2 tended to be upregulated in LSCP-HK but reduced by IRF2 KD. Based on this clustering, our prediction would be that PTGS2, NFKBIZ, and SOX4 are directly upregulated by IRF2. Similarly, cluster 3 genes tended to be downregulated in LSCP-HKs and increased by IRF2 KD. In this case, our prediction would be that AURKA and AURKB are directly inhibited by IRF2 thus linking IRF2 to senescence-related genes (Supplementary Fig. 15 right).

CRISPR-KO of IRF2 in LSCP-HKs restores stem cell function. Having established that IRF2 KD in LSCP-HKs reverses the genome-wide gene signature toward that of HSCP-HKs, we evaluated its effect on stem cell function. IRF2 KD by CRISPR–Cas9 editing was confirmed in HSCP-HKs and LSCP-HKs by Western blot (Fig. 7a). The senescence-associated protein p16 showed significantly higher expression in LSCP-HKs than HSCP-HKs as expected (Fig. 7a) and IRF2 KD significantly reduced p16 expression in both populations (Fig. 7a) supporting a reversal of the senescence phenotype suggested by transcriptional profiling. IRF2 editing in LSCP-HK restored clonogenicity towards a HSCP phenotype (Fig. 7a, b), indicating a functional effect of the induction of cell cycle genes (Fig. 6). The migration assay also showed a restoration of HSCP-HK phenotype in IRF2 KD LSCP-HK and a further enhanced migration potential in IRF2 KD HSCP-HK over control cells (Fig. 7c). Interestingly, β-galactosidase staining was also reduced in LSCP-HKs by IRF2 editing, again suggesting a role as a TF involved in cellular senescence (Fig. 7d) (supporting data from Fig. 6, Supplementary Fig. 11). IRF2 KD in LSCP-HK cells also restored the ability to form a thicker and more cellular epidermis in the 3D human skin model (Fig. 7e, Supplementary Fig. 16a, b). Again HSCP-HKs establish a clear undifferentiated basal layer (Fig. 7f) which is less convincing using LSCP-HKs where this layer is disorganized with few cells (Fig. 7f). IRF2 editing of LSCP-HKs results in less differentiation, a thicker epidermis and increased cellularity in the basal layer of 3D skin models (Fig. 7f, Supplementary Fig. 16a, b).

Discussion

While epithelial stem cell biology has advanced significantly over the last few decades, key questions remain as to the nature, plasticity and regulation of adult epithelial stem cells. Further insights are needed to guide successful therapeutic intervention to prevent epithelial damage or promote epithelial regeneration in disease or ageing. This study delineates the transcriptional circuitry imposing epidermal stem cell function, demonstrating that transcriptional regulatory circuits, previously shown to regulate cell identity, also govern the more nuanced phenotypic differences between human keratinocytes with low and high stem cell potential.

Using an unbiased approach based on stem cell function rather than marker expression, we exploited the loss of epidermal stem cell function over time in vitro for a high yield of keratinocytes with differential stem cell function. Confirming that replicative senescence paralleled increases in differentiation markers in our model, we employed transcriptional circuitry analysis to identify TFs predicted to modulate stem cell function, followed by validation thereof using a CRISPR-Cas9 screen. The validity of our screening approach was demonstrated with the inclusion of guides to known TF regulators of stem cell fate such as SNAI2, which were rapidly depleted from the pool. In addition, we successfully identified a number of TF regulators of epidermal stem cell fate. An outstanding and unexpected finding from this screen was the persistence and (very) strong enrichment of IRF2 edited cells in the final pool, confirming the utility of our transcriptional circuitry analysis and suggesting that IRF2 was significantly antagonistic to epidermal stem cell function.
We demonstrate that IRF2 is antagonistic to adult human epidermal stem cell function as IRF2 edited keratinocytes exhibited enhanced clonogenicity and migratory characteristics. Importantly, IRF2 KD cells retained their keratinocyte identity demonstrated by their formation of a stratified squamous epidermis in a human 3D skin model which was thicker and less differentiated than control cells. Furthermore, IRF2 editing could restore stem cell function and proliferative capacity in high passage keratinocytes that had undergone replicative senescence. These findings support a major role for IRF2 in driving loss of keratinocyte stem cell function although this was not directly demonstrated with overexpression experiments. Interestingly,
IRF2 expression showed only minimal change upon serial passaging of cells, which suggests that posttranscriptional regulation of IRF2 activity or differential interaction with other proteins might drive IRF2’s chromatin-directed functions.

IRF2 has been broadly associated with inflammatory diseases of the skin including psoriasis, but the relative contributions of epithelial and immune cells are difficult to discern with IRF2 KO models. IRF2 regulates pro-inflammatory response in macrophage data and suggest that IRF2 could also impair epidermal function critically required to maintain the essential skin barrier. Of note, a recent analysis of adult and fetal proerythroblast cells showed enrichment of IRF2 at adult-specific active promoters of genes associated with type I interferon and interferon gamma signaling. This implies that IRF2 may have a broader role in “aging” and specifying cell state between fetal and adult cells. Our data suggest that IRF2 drives type I interferon and gamma-interferon signaling in keratinocytes but the extent to which these pathways drive the loss of stem cell function requires further study. However, there is a growing appreciation of a role for inflammation in regulating stem cell behavior, most notably described in HSCs which are driven out of quiescence by interferon signaling, but also observed in intestinal stem cell activation. Importantly, low level systemic inflammation is associated with chronic diseases of ageing and may play a role in castration, supporting the importance of further mechanistic studies to define the role of IRF2 in epithelial biology.

There is considerable evidence that a cell’s gene-expression program is controlled by a limited and cell-specific number of TFs. Such core TFs are found enriched at super-enhancers, large clusters of transcriptional enhancers, and they form interconnected, auto regulatory and feedforward circuits to impose cell identity. TF regulatory circuits can induce pluripotency and transdifferentiate cell types. Indeed, a recent study shows in vivo transdifferentiation of mesenchymal cells into keratinocytes using TP63, GRHL2, TFPAP2A and MYC to enable healing of large skin wounds. Studies underscore the validity of TF-driven therapeutic approaches and interestingly two of the TFs used to convert mesenchymal to epithelial skin lineage, TFPAP2A and MYC, were in our leading edge for TFs associated with high stem cell potential.

This study adds to a growing understanding of the importance of TF networks in health and disease. Our findings reveal that, in addition to specifying cell identity, transcriptional networks define cell state within a single lineage thereby adding an additional layer of “fine-tuned” regulation of cellular plasticity. Refinement of the computational and chromatin profiling technology used in this study, particularly applied to low cell numbers, holds promise for the discovery of significant targets for disease-modifying small molecule therapeutics. Once considered “undruggable”, TFs are now an intense focus of chemistry efforts to enable small molecule regulation of their activity for therapeutic benefit. Although the current focus for such efforts is cancer, the ability to identify dominant TF regulators of cell state broadens potential therapeutic utility. Small molecules against key cell fate regulators could enable manipulation of cell plasticity for therapeutic benefit in tissue engineering, regenerative medicine and other disease settings.

**Methods**

**Human keratinocyte culture and serial passaging.** All human cells were purchased from approved commercial suppliers who had obtained the necessary patient consent for skin donation for research purposes. The cells were obtained and handled in compliance with the Swiss Human Research Act. Normal human neonatal epidermal keratinocyte progenitor cells (nHPEKs, Lot#EB1110044) and adult normal human epidermal keratinocyte progenitor cells (aHPEKs, Lots #ES1303277, and #MC1511045 (both abdominal)) were purchased from CELLnTEC (Bern, Switzerland) at Passage 2 (P2). HPEKs were expanded until Passage 4 (P4) and pellets of 1 x 10^6 cells were cryopreserved in liquid nitrogen for later use. HPEKs were cultured from P4 by seeding at an initial density of 5.3 x 10^4 cells/cm^2 using CatT-Prime media (CELLnTEC, cat#CnT-PR). When cells reached 70–85% confluence, they were detached from the cell culture dish using Accutase Cell Detachment Solution (CELLnTEC). For serial passaging experiments, aHPEKs were grown and passaged about 13 times. aHPEKs were collected for various assays and were considered as “HSCP-HKs” when fast-dividing (between Passages 7–9) whereas aHPEKs that were slow-dividing (Passages 13–17) were considered as “LSCP-HKs”.

**Clonogenic assay.** HPEKs were seeded at 1000–2000 cells/well in 6-well plates (in triplicates) with 5 ml of EpiLife Medium (ThermoFisher Scientific, Zug, Switzerland) with no BPE and 75–85% of human keratinocyte growth supplements (HKGS, ThermoFisher Scientific). 80% of the medium was replaced every 3 or 4 days. After 8–12 days plates were washed with phosphate-buffered saline (PBS) and colonies fixed with 10% neutral buffered formalin solution for 20 min. After washing with PBS, colonies were stained with 0.5% (w/v) crystal violet (50% of 1% crystal violet and 50% Ethanol) for 45 min. Excess crystal violet was removed by washing with water and then dishes were dried. Images of stained colonies were collected for various assays and were considered as “HSCP-HKs” when fast-dividing (between Passages 7–9) whereas aHPEKs that were slow-dividing (Passages 13–17) were considered as “LSCP-HKs”.

**ECM-binding assay.** In total, 96-well NUNC plates were coated with different ECM for two days at 4°C: 25 μg ml⁻¹ Collagen I (Corning, WVR, Vienna, Austria, cat#354236) (0.02 N acetic acid), 25 μg ml⁻¹ Fibronectin (Gibco, cat#PHE00203) and collagen I (in PBS) or 10 μg ml⁻¹ Laminin V (BioLamina, cat#L1111-02) (in PBS with Ca²⁺/Mg²⁺⁻). Unbound Collagen I and Fibronectin were removed with PBS and excess Laminin V was removed with PBS with Ca²⁺/Mg²⁺⁻. Cells were then incubated with CatT-PR media before applying 15 μM of Resazarin for 75 min. Fluorescence was read at 560/590 nm, sensitivity 40 (Synergy HT from Biotek + Bistack and Gen5 software).
Migration assay. HPEKs were detached and resuspended in CnT-FTAL5 medium (CELLnTEC, Bern, Switzerland, Lot#ES1303222) were expanded in DMEM/F-12 (CELLnTEC, Bern, Switzerland, Lot#ES1303222) were expanded in DMEM/F-12 − L-ascorbic acid (Sigma-Aldrich) was changed every 2 days.

Human dermo-epidermal 3D models. Human primary dermal fibroblasts (CELLnTEC, Bern, Switzerland, Lot#ES1303222) were expanded in DMEM/F-12 (CELLnTEC, Bern, Switzerland, Lot#ES1303222) were expanded in DMEM/F-12 from the cell inserts and the wells and HPEKs, resuspended in CnT-PR intermediate cell layer by using a 96-pin wound making tool (ESSEN BIO). After an additional washing step, 100 μl Assay medium was added and the plates were transferred into the Incucyte ZOOM (Essen Bioscience, cat#9600-0012) for wound area measurement overnight. Wound confluence was monitored and analyzed with the Incucyte ZOOM Live-Cell Imaging System and software according to the supplier’s protocols.

Human dermo-epidermal 3D models. Human primary dermal fibroblasts (CELLnTEC, Bern, Switzerland, Lot#ES1303222) were expanded in DMEM/F-12 with GlutaMAX™ supplemented with 10% fetal calf serum (ThermoFisher Scientific, Zug, Switzerland) up to passage 6 before used in the dermo-epidermal model. The dermal part was prepared with the RAFT kit (LONZA, Visp, Switzerland, cat#1016-0135) and the epidermal part with the HPEKs were covered with medium. The stratification process at the air–liquid interface was achieved by removing the medium level in the former just reached the bottom of the insert and into the wells, and models were incubated for 3 days whereby the medium level in the former just reached the bottom of the insert and into the wells, and models were incubated for 3 days whereby.

Western blot analysis. Protein extracts were prepared using RIPA buffer (Sigma: 150 mM NaCl, 1.0% IGEPAL® CA-630, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate (SDS), and 50 mM Tris, pH 8.0) completed with Protease Inhibitor (PI) Cocktail tablet (Sigma). Protein concentration was determined with the Pierce BCA protein assay kit (Thermo Scientific). Protein extracts (20–40 μg) were fractionated by SDS polyacrylamide gel electrophoresis on 4–12% Bis–Tris precast polyacrylamide gels (Invitrogen, Paisley, UK) and transferred to a nitrocellulose membrane (Invitrogen). Blots were visualized by chemiluminescence (ECL Plus, GE Healthcare, Hatfield, UK). Protein quantification was performed using ImageJ and statistical significance was assessed using GraphPad Prism 8. Anti-IRF2 (Abcam, cat#ab124744, 1:500 dilution), anti-IRF3 (Cell Signaling, cat#76684, 1:500 dilution), anti-Y1I (Cell Signaling, cat#2185, 1:500 dilution), anti-SNAI2 (Abcam, cat#ab71988, 1:200 dilution), and anti-p16 (CDKN2A, Abcam, cat#ab54210, 1:500 dilution) antibodies were used.

CRISPR-Cas9 editing with lentiviral constructs. Cas9 gene encoding the S. pyogenes CRISPR associated protein 9 RNA-guided DNA endonuclease was cloned under control of the human cytomegalovirus promoter into a lentiviral construct derived from pLenti5 (V#99610, Invitrogen) carrying a bipartite resistance cassette. Upon packaging, the active virus was used to transduce the control P6 nHPEKs grown in CnT-Prime media (CELLnTEC) (Supplementary Fig. 4a). After 7 days of bacterial at 0.625 μg/ml, 4 ml of bacterial cells were further grown for 2 days in CnT-Prime media, analyzed for Cas9 expression, cryopreserved (into pellets of 2 × 10^6 cells/l) in liquid nitrogen for later use and assessed for editing using a sgRNA against PIG-A (Phosphatidylinositol Glycan Class A) with the following sequence: 5′-TGGCGT GGAAGAGGATCAT-3′. For editing assessment nHPEKs-Cas9 were infected with the PIG-A sgRNA and a control at a multiplicity of infection of 1. Cells were maintained with puromycin selection until day 4 and transduction efficiency was assessed by flow cytometry using the red fluorescent protein (RFP) reporter encoded on the lentiviral construct. If >90% RFP-positive cells were measured, expansion of cells continued without further puromycin selection. Gene-editing efficiency was assessed using the software TIDE (Tracking of Indels by DEcomposition).

CRISPR-Cas9 mini-pool screen with lentiviral constructs. For the pooled sgRNA library, 2689 sgRNA sequences were selected for 540 genes. The library was constructed using chip-based synthesis (Costs and CRISPR associated protein 9 RNA-guided DNA endonuclease Cas9 were packaged into lentiviral particles by growing HEK293T cells in T150 flasks (Corning, cat#33113). For each flasks, 2.1 × 10^6 cells were transfected 24 h after plating using 510.3 ml of TransIT reagent (Mirus, Madison, WI, cat#MIR2300), diluted 1:8 ml of Opti-MEM and 7.6 μg of the sgRNA libraries and 94.5 mg of lentiviral packaging mix (cat#CPCP-K2A, Cellscitec, psPAX2 and MD2 plasmids that encode Gag/Pol and VSV-G, respectively). Seventy-two hours post transfection, lentivirus was harvested, aliquoted, and frozen at −80°C. Viral titer was measured by feeding virus-activated cell sorting into HCT116 cells and was typically in the range of 5 × 10^8TU/ml.

For the screen, nHPEKs-Cas9 were expanded in CnT-Prime media to 3 × 10^6 cells in T300 flasks (TPP, Trasadingen, Switzerland) and transduced with the lentiviral sgRNA library (described in detail above) with a coverage of 5 sgRNAs/gene and a MOI of 0.5 aiming for coverage of on average 1000 cells/sgRNA. On day 0, each flasks was infected with the lentivirus pool supplemented with 1 μg/ml−1 polybrene. After 24 h the culture media was replaced with fresh media containing 1 μg/ml−1 puromycin. Seventy-two hour after puromycin addition, cells were detached using Accutase and plated into new flasks at 4.5 × 10^3 cells per well. Selection efficiencies determined by flow cytometry were maintained in culture, split as needed to ensure confluence did not exceed 90% and at least 5 × 10^6 cell per well were maintained up to 5 days after each split. Cells were collected at days 12, 19, 28, 33, 38, 43, and 47 post infection.

Genomic DNA from cells was isolated using the PureLink Genomic DNA Mini Kit (Invitrogen, cat#K1820-02) and quantified using PicoGreen (Invitrogen, cat#P11496) following manufacturer’s instructions. Illumina sequencing libraries were generated using PCR amplification with primers specific to the genome integrated lentiviral backbone sequence. A total of eight 500 ng independent and unique sgRNA transduced pools were sequenced with an average depth of coverage of 4000 fold across all genes. The sgRNA library was sequenced onto an Illumina HiSeq 2500 (Illumina, San Diego, CA, USA). Sequencing libraries were loaded onto the flow cell with one lane per sample. For each sample, a unique 5′ and 3′ index was used to demultiplex following sequencing. For gene targeting, 5′-TGGCGT GGAAGAGGATCAT-3′, 5′-GAGATGGTCAGACACGCGTGGAC-3′, and a 1 × 11b index read, using the a custom indexing primer 6430 (5′-GATCTTGAGACAGATTGACGATGTTCTTCCTTGAC-3′) and an Illumina sequencing library, a 1 × 150 paired-end read (10 bp barcode used for data demultiplexing following sequencing), 0.5 μm dNTPs (Clontech, cat#4030), 1× Titanium Taq DNA polymerase and buffer (Clontech, #693242), PCR cycling conditions were as follows: 1 × 98°C for 5 min; 28 × 95°C for 30 s, 65°C for 30 s, 72°C for 30 s; 1 × 72°C for 5 min. The resulting Illumina libraries were purified using 1.8× SPRI AMPure XL beads (Beckman Coulter, #A63882) following the manufacturer’s recommendations and quantified by using primers specific to the Illumina sequences using standard methods. Illumina sequencing libraries were then pooled and sequenced with a HiSeq 2500 instrument (Illumina) with 1 × 300 reads, using a custom read 1 sequencing primer 5′-TCGCCATATGTTCTTCCTTGAC-3′ and using Illumina sequencing library, a 1 × 100 paired-end read (10 bp barcode used for data demultiplexing following sequencing) resulting in an average of approximately 600 reads per sgRNA.

Raw sequencing reads were aligned to the appropriate library using Bowtie5 allowing for no mismatches and counts were generated. To assess effects on proliferation, the fold change of the unsorted cell population compared to the input library was generated. For gene-based hit calling, RSA6 and average or maximal fold changes were calculated across all reagents for a given gene.

CRISPR-Cas9 editing using electroporation. Two sgRNA per gene were designed on a DNA sequence flanking an exon and an intron as close as possible to TSS using the Design sgRNA for CRISPRko from the Broad Institute. A guide targeting an exon and the other an intron were chosen based on the Combined Rank
provided by the design tool with minimal predicted off-target binding activity. Guides were constructed by Integrated DNA Technologies (IDT, Coralville, IA). 

Supplementary Table 1 shows the sequences of all gRNAs designed. A negative control gRNA was purchased from IDT (Alt-R® CRISPR-Cas9 Negative Control crRNA #1).

Recombinant Cas9 was produced in Escherichia coli containing two nuclear localization sequences and a His6 tag and purified with NTA chromatography, followed by size-exclusion chromatography. Two gRNAs per gene were initially designed and tested. gRNA #1 was validated (Benjamini Hochberg method). Enrichments were calculated using a hypergeometric overrepresentation test. The Benjamini-Hochberg method was performed all of the analyses using two gene rankings, where the first is by fold-change only and the second is by fold-change multiplied by the negative log10 adjusted P value; the latter version puts more weight on genes that move significantly. Within a contrast, results were adjusted using the Benjamini Hochberg method.

Quantification and statistical analysis. The gene-set database was compiled from multiple sources including Reactome, NCBI Biosystems, and Gene Ontology. Enrichments were calculated using a hypergeometric overrepresentation test. Gene counts were divided by the total number of mapped reads for each sample (using the standard consecutive enzymatic steps of end-polishing, dA-addition, and adaptor ligation). After a quick PCR (Qiagen) using Qubit Nucleic Acid Quantitation Kit, Samples were sequenced in paired-end sequencing technology. Each experimental condition was performed in triplicate. Reads were filtered using the Illumina TruSeq RNA Sample Preparation Kit v2 (Illumina), and input DNAs by the standard consecutive enzymatic steps of end-polishing, dA-addition, and adaptor ligation. Quantitative PCR (QPCR) reactions were carried out in triplicate on the original chromatin volume allowed quantitation of the total chromatin yield. An aliquot of chromatin (30 μg) was precleared with protein A agarose beads (Invitrogen). Genomic DNA regions of interest were isolated using 4 μg of antibody against BRD4 (Bethyl, cat# A301-985A100, Lot A301-985A100-6). Complexes were separated and DNA extracted using the MiniElute PCR purification kit (Qiagen). Illumina sequencing libraries were prepared from the ChIP and Input DNAs by the standard consecutive enzymatic steps of end-polishing, da-addition, and adaption ligation. After a final PCR amplification step, the resulting DNA libraries were quantified and sequenced on Illumina’s NextSeq 500 (75 nt reads, single end).

The raw ChIP-seq data for each of the above conditions is available in Supplementary Data 1.

Chromatin immunoprecipitation of H3K27ac and sequencing. Approximately, 4 × 10^6 cells were cross-linked with 10% formaldehyde at RT for 15 min. After quenching, cells were washed with PBS and centrifuged at 1,500 × g for 5 min. RIN values were reported in Supplementary Data 1. ChIP-Seq was performed by Active Motif Inc. (Carlsbad, CA, USA) as follows. Frozen cells (9 × 10^6) were thawed and fixed with 1% formaldehyde for 15 min and quenched with 0.125 M glycine. Chromatin was isolated by the lysis buffer, followed by digestion with a Dounce homogenizer. Lysates were nucleated and 1 million cells were used for each replicate. Cells were incubated on ice for 10 min in a crosslinking buffer (0.1 M NaCl, 0.5 mM EDTA, 10 mM Tris pH 8.0). Fixed nuclei were collected by centrifugation (14,000 rpm, 4 °C, 30 s), washed with Wash Buffer C (with PEG) (trucHiP Chromatin Shearing Kit, Covaris, Brighton, UK, cat#SF20154) and re-suspended in Shearing buffer (D3 plus PEG and 100 mM PMSE) (trucHiP Chromatin Shearing Kit, Covaris). Nuclei were vortexed for 3 rounds of 10 s and left 15 min in ice before freezing in −80 °C for later use. After thawing, 130 μl nuclei suspensions were transferred into microTUBE (Covaris) and sonicated using a Covaris E220 (10 min: duty cycle 2%, Intensity 5, Cycles per Burst 200). Samples were collected into 1.5 ml tubes, centrifuged at 10,000xg for 10 min at 4 °C and supernatants transferred into 2 ml tubes. DNA concentrations were calculated after DNA isolation using Quiaquick PCR (Qiagen) using Qubit Nucleic Acid Quantitation Kit (Invitrogen). For ChIP a modified protocol of the Magna ChiP A/G (Merck, cat#17-10085) protocol was used. In brief 5 μg of mice anti-H3K27ac (Active Motif, cat#339385S) or mAb IgG1 Isotype Control (Cell Signaling, cat#5415) were coupled with Protein A/G beads at 4 °C for 4 h. Coupled Ab-beads were washed 5 μg of samples and incubated overnight at 4 °C with rotation. Ab-beads were washed as per manufacturer’s instructions and protein/DNA complexes eluted in Elution buffer (30 mM NaCl and 1% SDS in D2H2O) for 4 h at 65 °C. Reverse-cross-linking was performed by adding RNAse for 30 min at 37 °C and 5 mM EDTA, 20 μM Tris pH 8.0, 1 μg μg ml−1 Proteinase K and heat for 30 min at 65 °C, and ChIP DNA was purified by phenol-chloroform extraction and ethanol precipitation. Quantitative PCR (QPCR) reactions were carried out in triplicate on specific genomic regions using SYBR Green Supermix (Bio-Rad). The resulting signals were normalized for primer efficiency by carrying out qPCR for each primer pair using input DNA. Illumina sequencing libraries were sequenced on the HiSeq 2500 platform distributed on the NanoDrop spectrophotometer. Extrapolation to the original chromatin volume allowed quantitation of the total chromatin yield. An aliquot of chromatin (30 μg) was precleared with protein A agarose beads (Invitrogen). Genomic DNA regions of interest were isolated using 4 μg of antibody against BRD4 (Bethyl, cat# A301-985A100, Lot A301-985A100-6). Complexes were separated, washed from the elutes with SDS buffer, and subjected to RNAse and proteinase K treatment. Crosslinking was reversed by incubating at 65 °C, and ChIP DNA was purified by phenol-chloroform extraction and ethanol precipitation. Quantitative PCR (QPCR) reactions were carried out in triplicate on specific genomic regions using SYBR Green Supermix (Bio-Rad). The resulting signals were normalized for primer efficiency by carrying out qPCR for each primer pair using input DNA. Illumina sequencing libraries were sequenced on the HiSeq 2500 platform. The raw ChIP-seq data are available in the NCBI GEO database.

Chromatin immunoprecipitation of BRD4 and sequencing. ChIP-Seq was performed by Active Motif Inc. (Carlsbad, CA, USA) as follows. Frozen cells (9 × 10^6) were thawed and fixed with 1% formaldehyde for 15 min and quenched with 0.125 M glycine. Chromatin was isolated by the lysis buffer, followed by digestion with a Dounce homogenizer. Lysates were nucleated and 1 million cells were used for each replicate. Cells were incubated on ice for 10 min in a crosslinking buffer (0.1 M NaCl, 0.5 mM EDTA, 10 mM Tris pH 8.0). Fixed nuclei were collected by centrifugation (14,000 rpm, 4 °C, 30 s), washed with Wash Buffer C (with PEG) (trucHiP Chromatin Shearing Kit, Covaris, Brighton, UK, cat#SF20154) and re-suspended in Shearing buffer D3 (plus PEG and 100 mM PMSE) (trucHiP Chromatin Shearing Kit, Covaris). Nuclei were vortexed for 3 rounds of 10 s and left 15 min in ice before freezing in −80 °C for later use. After thawing, 130 μl nuclei suspensions were transferred into microTUBE (Covaris) and sonicated using a Covaris E220 (10 min: duty cycle 2%, Intensity 5, Cycles per Burst 200). Samples were collected into 1.5 ml tubes, centrifuged at 10,000xg for 10 min at 4 °C and supernatants transferred into 2 ml tubes. DNA concentrations were calculated after DNA isolation using Quiaquick PCR (Qiagen) using Qubit Nucleic Acid Quantitation Kit (Invitrogen). For ChIP a modified protocol of the Magna ChiP A/G (Merck, cat#17-10085) protocol was used. In brief 5 μg of mice anti-H3K27ac (Active Motif, cat#339385S) or mAb IgG1 Isotype Control (Cell Signaling, cat#5415) were coupled with Protein A/G beads at 4 °C for 4 h. Coupled Ab-beads were washed 5 μg of samples and incubated overnight at 4 °C with rotation. Ab-beads were washed as per manufacturer’s instructions and protein/DNA complexes eluted in Elution buffer (30 mM NaCl and 1% SDS in D2H2O) for 4 h at 65 °C. Reverse-cross-linking was performed by adding RNAse for 30 min at 37 °C and 5 mM EDTA, 20 μM Tris pH 8.0, 1 μg μg ml−1 Proteinase K and heat for 30 min at 65 °C, and ChIP DNA was purified by phenol-chloroform extraction and ethanol precipitation. Quantitative PCR (QPCR) reactions were carried out in triplicate on specific genomic regions using SYBR Green Supermix (Bio-Rad). The resulting signals were normalized for primer efficiency by carrying out qPCR for each primer pair using input DNA. Illumina sequencing libraries were sequenced on the HiSeq 2500 platform. The raw ChIP-seq data are available in the NCBI GEO database.
from (https://sites.google.com/site/anshulkundaje/projects/blacklists).

Downloaded directly from the UCSC genome browser from their table browser (ftp://igenome:G3nom3s4u@ussd-ftp.illumina.com/Homo_sapiens/UCSC/hg19/site/anshulkundaje/projects/blacklists).

1 mM EDTA) with 0.1% SDS and DNA was sheared at 4°C using a covaris instrument. The chromatin was then cleared by centrifugation at 10000xg for 10 min and incubated with 5 μl of Dynabeads Protein A (Thermo Fisher Scientific, Waltham, MA) for one hour to pre-clear the lysate. After collecting the incubate, the supernatant was incubated with end-over-end rotation overnight at 4°C with Dynabeads Protein A magnetic beads prebound with a HA antibody (Cell Signaling, cat#3724, 0.2 μg). The HA antibody was bound to the beads by incubating the beads with the antibody in 200 μl of bead binding buffer (TE, 0.05% SDS and 1% Triton X-100) at 4°C for 2 hours. Beads were washed once with low salt wash buffer (250 mM NaCl, 10 mM Tris-HCl, 1 mM EDTA, 0.1% SDS, 0.1% Na-deoxycholate, 0.1% Triton X-100), once with wash buffer containing 500 mM NaCl, once with LiCl wash buffer (20 mM Tris pH8.0, 1 mM EDTA, 250 mM LiCl, 0.5% NP-40, 0.5% Na-deoxycholate), once with TE, 0.1% Triton X-100 and twice with ice cold 10 mM Tris-HCl. The beads and the input were then incubated at 37°C for 12 h with the Tagment DNA enzyme (illuminata, San Diego, CA) following tagmentation, the beads were washed twice with the low salt wash buffer and DNA was eluted TE Buffer, 250 mM NaCl and 0.3% SDS. Cross-links were reversed by incubation first at 55°C and continued with protein digestion with the addition of Proteinase K for 10 h at 64°C. DNA was purified using DNA Clean & Concentrator-5 (Zymo Research, Irvine, CA).

Library preparations were performed as described with minor changes⁷⁹. The amplified libraries were cleaned, and size selected using AMPure XP (Beckman Coulter, Indianapolis, IN). The libraries were sequenced using NextSeq® 500/550 High Output Kit v2 (75 cycles) (Illumina) in a NextSeq 550 (Illumina).

ATAC-seq processing. ATAC-seq (Assay for Transposable-Accessible Chromatin) was performed as described with minor changes. PCR products that were sheared in 6-well plates, washed out with 500 μl of PBS (on plate) and detached using Accucast (CELLTETC). Cells were centrifuged at 500g and cell pellets placed on ice. Cells were resuspended in 1 ml lysis buffer (10 mM Tris-HCl pH7.5, 10 mM NaCl, 3 mM MgCl2, 0.1% IGEPAL CA-630, 1 mM PMSF and PI cocktail). Cell lysates were split into 2 x 1.5 ml tubes (for an equivalent of 1 x 10⁸ HPEKs per tube) and DNA was sheared at 4°C using a covaris instrument. The chromatin was then cleared by centrifugation at 37°C for 30 min. The reaction was stopped using 50 μl Stop Buffer (30 mM EDTA, pH8.0 and 300 mM NaCl) at 55°C. Cross-links were reversed by incubation first at 55°C and continued with protein digestion with the addition of Proteinase K for 10 h at 64°C. DNA was purified using DNA Clean & Concentrator-5 (Zymo Research, Irvine, CA).

Library preparations were performed as described with minor changes⁷⁹. The amplified libraries were cleaned, and size selected using AMPure XP (Beckman Coulter, Indianapolis, IN). The libraries were sequenced using NextSeq® 500/550 High Output Kit v2 (75 cycles) (Illumina) in a NextSeq 550 (Illumina).

ChIP-seq analysis. All paired-end ChIP-seq datasets were aligned using Bowtie2 (version 2.2.8) to build version NCBI37/HG19 with the following parameters: -end-to-end --sensitive --no-unal --no-discordant --mm --net-stiller.
To process aligned ATAC-seq data, we used the RIESLING pipeline³⁹ developed jointly with the Gordon laboratory. Briefly, aligned bam files were filtered for low quality reads, duplicate reads were removed, reads were sorted using samtools, mitochondrial reads were removed, and further filtering against the ENCODE blacklist. Peaks were called using MACS1.4.2 which does not take into account paired end fragment information. A threshold of p value < 1e-9 was used to call enriched regions.

RNA-seq data processing. For integrated chromatin analysis, RNA-seq data were repro-grammed to align to and quantify the HG19 RefSeq transcriptome. Fastq files were aligned to the HG19 genome using HiSat2 (version 2.0.4) with default parameters. Transcripts were quantified and FPKM values were generated using cuffquant and cuffnorm from the cufflinks suite (version 2.2.1) (http://cole-trapnell-lab.github.io/cufflinks/). The HG19 RefSeq gene IDs were filtered against the ENCODE blacklist and joint core Interactome enrichment tool was used (https://clarivate.com/products/metacore/) to predict TF connectivity analysis.

Calculating read density. We calculated the normalized read density of a ChiP-seq or ATAC-seq dataset in any genomic region using the Bamlidiprocessor (version 1.0) read density calculator (https://github.com/BradnerLab/pipeline/wiki/bamlidiprocessor). Briefly, reads aligning to the region were extended to 200 bp and the density of reads as base pair (bp) was calculated. The density of reads in each region was normalized to the total number of million mapped reads producing read density in units of reads per million mapped reads per bp (rpm/bp). The AUC of read occupancy in each region was reported simply as the number of reads divided by the total number of million mapped reads producing an AUC measurement in total reads per million (rpm).

Mapping of active chromatin landscapes. We used H3K27ac ChiP-seq to define the active chromatin cis-regulatory landscape across human keratinocytes. Briefly, we took the union of all H3K27ac enriched regions in any single H3K27ac dataset in either HSCP-HK or LSCP-HK samples (n of 3 each). This yielded 24,614 discreet H3K27ac regions in the genome. To compare the similarity of H3K27ac at these regions between HSCP-HK and LSCP-HK samples, at each region on a per sample basis, we first quantified H3K27ac AUC (rpm). We next normalized on a per sample basis H3K27ac signal to the median region AUC. Similarity between samples was calculated using a Pearson correlation and samples were clustered based on a correlation distance matrix (1 – correlation) (Supplementary Fig. 2). H3K27ac RefSeq genes were considered active if they possessed an H3K27ac enriched region within the ± 1 kb of their TSS in at least one H3K27ac dataset and their expression was greater than or equal to 10 FPKM in at least one sample. Using these criteria, we defined 7278 active HG19 RefSeq transcripts. As genes often have more than one TSS, these 7278 active transcripts corresponded to 5846 common gene names.

We sought to examine relationships between H3K27ac, BRD4, and mRNA levels at active genes between HSCP-HK and LSCP-HK. First, we quantified H3K27ac and BRD4 AUC occupancy at all active genes. For each gene, we assigned H3K27ac and BRD4 occupancy within the ±50 kb window flanking the TSS. All active genes were ordered based on the difference in occupancy between HSCP-HK and LSCP-HK. Using bins of 100 genes, the average log. fold change in H3K27ac AUC or mRNA levels was plotted (Fig. 2a). Error bars represent the 95% confidence interval of the mean determined by empirical resampling with replacement of 1000 permutations.

To compactly display ChiP-seq and ATAC-seq signal at individual genomic loci, we used a simple meta representation. For all samples within a group, ChiP-seq or ATAC-seq signal is smoothed using a simple spline function and plotted as a translucent shape in units of rpm per bp. Darker regions indicate regions with signal in more samples. An opaque line is plotted and gives the average signal across all samples in a group.

Core transcriptional regulatory circuitry analysis. We first identified TFs in keratinocytes that were actively expressed and regulated by a proximal H3K27ac defined active cis-regulatory element (n = 126). We next filtered these TFs for protein–protein interactions as defined by the STRING protein–protein interaction database⁵⁹. Protein–protein interactions were defined in STRING as having a medium confidence >0.400 interaction defined by all active interaction sources. Markov Chain Linkage with an inflation parameter was then used to cluster TFs by protein–protein interaction. These resulted in 60 remaining TFs across 6 clusters (Fig. 3a).

To define the transcriptional circuitry of keratinocytes and regulatory connectivity of keratinocyte KTFs, we developed the CRC software package, an updated version of the COLTRON CRC software initially reported in ref.16. Using the 60 TFs previously defined, we defined a regulatory interaction as a TF binding to a nucleosome free region inside an active cis-regulatory element. cis-regulatory elements were previously defined by H3K27ac and nucleosome free regions are defined as ATAC-seq peaks co-localized within H3K27ac defined cis-regulatory regions. Within these regions, the CRC software uses FIMO⁶⁰ to find enriched (p value < 1e−5) TF motif occurrences. CRC defines TF motifs using a custom
HKs can be found under: GSE135675. H3K27ac and BRD4 ChIP-seq in HSCP-HKs and the difference in distributions was assessed using a one-sided Wilcoxon Rank-Sum test (*p value < 1e-5). Within either cluster 1 or cluster 2, the total maximum possible connectivity for TFs within the cluster and for TF-TF interactions with other actively expressed TFs was assessed (Supplementary Fig. 3d). The statistical significance of the difference in distributions was assessed using a one-sided Wilcoxon Rank-Sum test (*p value < 1e-3).

IRF2 binding analysis. ChIP-seq was performed in duplicate for the HA antibody in either HSCP-HKs engineered with doxycycline inducible IRF2-HA or parental unengineered HSCP-HKs. HA peaks were called versus whole-cell extract background using MACS4.2. High confidence IRF2 peaks were determined as those HA peaks found in both IRF2-HA expressing cells and absent in the parental unengineered cells. This resulted in 4135 peaks of which 2068 overlapped an ATAC-seq peak in at least one of the samples.

For both IRF2 peaks with and without overlapping ATAC-seq, we performed de novo motif finding using the meme-chip software (version 4.11.4) with the following parameters (-meme -nmotifs 5 -skip -db VertebratePWMs.txt). For the -db argument, the same table of vertebrate TF position weight matrices used in the transcriptional regulatory circuitry analysis was supplied (https://github.com/linlabcode/CRC/blob/master/erc/annotation/VertebratePWMs.txt).

For both sets of regions, the top identified motif is reported.

BRD4 out degree was calculated as before, but only on edges overlapping a high confidence IRF2 binding site. The average change in BRD4 out degree for these IRF2 ChIP-validated edges is reported and compared to IRF2 predicted edges and all edges in the cis-regulatory landscape. Error bars represent the standard error of the mean (Fig. 6d).

We used an approach previously developed[2] to quantify IRF2 signal proximal to each gene. Briefly, background subtracted IRF2 occupancy (AUC) was calculated at all 2086 high confidence regions. Genes were ranked by their combined promoter (±1 kb TSS) and distal IRF2 occupancy. Distal IRF2 occupancy represents the AUC of IRF2 peaks within ±50 kb of a gene’s TSS that are not overlapping another gene’s TSS. IRF2 signal was detected at 855 unique genes. These were ranked by cumulative promoter + distal IRF2 AUC (Fig. 6f).

Gene-set enrichment analysis was used on these 855 genes ranked by IRF2 AUC to identify signatures with leading edge enrichment (pulled from the mSigDB C2 set). Log fold expression changes upon IRF2 KD or between HSCP-HKs and LSCP-HKs were calculated at these 855 genes and for all 100 genes for which there was evidence of IRF2 binding. The statistical significance of the difference between distributions was tested using a two-tailed t test (Fig. 6f). Finally, the top 100 IRF2 targets were queried using the mSigDB investigate gene sets tool against the C2 curated gene sets. The significance of enrichment for the top ten gene sets are shown in Fig. 6g.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
High throughput sequencing data that support the findings of this study have been deposited in NCBI GEO under the overall accession code: GSE153680 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE153680) ATAC-seq in HSCP-HKs and LSCP-HKs can be found under: GSE135675. H3K27ac and BRD4 ChIP-seq in HSCP-HKs and LSCP-HKs can be found under: GSE135677. RNA-seq of baseline HSCP-HKs and LSCP-HKs can be found under: GSE135679. RNA-seq expression level count data for IRF2 KD experiments are available in Supplementary Data Files 5 and 6. Read depth for all samples is provided in Supplementary Data 7.

Code availability
All manuscript specific custom analysis scripts and code can be found at: https://github.com/charleslin/keratinocyte_scripts. The transcriptional core regulatory circuitry analysis code can be found at https://github.com/linlabcode/CRC.

Received: 9 November 2018; Accepted: 14 September 2019; Published online: 14 October 2019

References
1. Scadden, D. T. Nice neighborhood: emerging concepts of the stem cell niche. Cell 157, 41–50 (2014).
2. Eaves, C. J. Hematopoietic stem cells: concepts, definitions, and the new reality. Blood 125, 2605–2613 (2015).
3. Morris, R. J. & Potten, C. S. Slowly cycling (label-retaining) epidermal cells behave like clonogenic stem cells in vitro. Cell Proliﬁr. 27, 279–289 (1994).
4. Clevers, H. The intestinal crypt, a prototype stem cell compartment. Cell 154, 274–284 (2013).
5. Clevers, H. & Watt, F. M. Defining Adult Stem Cells by Function, Not by Phenotype. Annu Rev Biochem, https://doi.org/10.1146/annurev-biochem-062917-012341 (2018).
6. Gonzales, K. A. U. & Fuchs, E. Skin and Its Regenerative Powers: An Alliance between Stem Cells and Their Niche. Dev. Cell 43, 387–401 (2017).
7. Watt, F. M. Mammalian skin cell biology: at the interface between laboratory and clinical. Sci. (New Y., N. Y.) 346, 937–940 (2014).
8. Rheinwald, J. G. & Green, H. Serial cultivation of strains of human epidermal keratinocytes: the formation of keratinizing colonies from single cells. Cell 6, 331–343 (1975).
9. Hirsch, T. et al. Regulation of the entire human epidermis using transgenic stem cells. Nature 551, 327–332 (2017).
10. Adam, R. C. et al. Temporal Layering of Signaling EffectorS Drives Chromatin Remodeling during Hair Follicle Stem Cell Lineage Progression. cell stem cell 22, 398–413.e398 (2018).
11. Donati, G. & Watt, F. M. Stem cell heterogeneity and plasticity in epithelia. cell stem cell 16, 665–476 (2015).
12. Hnisz, D. et al. Super-enhancers in the control of cell identity and disease. Cell 155, 934–947 (2013).
13. Adam, R. C. et al. Pioneer factors govern super-enhancer dynamics in stem cell plasticity and lineage choice. Nature 521, 366–370 (2015).
14. Takahashi, K. & Yamanka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126, 663–676 (2006).
15. Lee, T. I. & Young, R. A. Transcriptional regulation and its misregulation in disease. Cell 152, 1237–1251 (2013).
16. Lin, C. Y. et al. Active medioloblastoma enhancers reveal subgroup-specific cellular origins. Nature 530, 57–62 (2016).
17. Saint-Andre, V. et al. Models of human core transcriptional regulatory circuitries. Genome Res. 26, 385–396 (2016).
18. Panne, D., Maniatis, T. & Harrison, S. C. An atomic model of the interferon-beta enhanceosome. Cell 129, 1111–1123 (2007).
19. Rodda, D. J. et al. Transcriptional regulation of nanog by OCT4 and SOX2. J. Biol. Chem. 280, 24731–24737 (2005).
20. Whyte, W. A. et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell 153, 307–319 (2013).
21. Boyer, L. A. et al. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122, 947–956 (2005).
22. Jiang, D. H. et al. A transcriptional roadmap to the senescence and differentiation of human oral keratinocytes. J. Gerontol. A Biol. Sci. Med. Sci. 70, 20–32 (2015).
23. Nanba, D. et al. Cell motion predicts human epidermal stemness. j. cell biol. 209, 305–315 (2015).
24. Chermnykh, E., Kalabusheva, E. & Vorotelyak, E. Extracellular Matrix as a Regulator of Epidermal Stem Cell Fate. Int J Mol Sci 19, https://doi.org/10.3390/ijms19010038 (2018).
25. Hayflick, L. & Moorhead, P. S. The serial cultivation of human diploid cell strains. Exp. Cell Res. 25, 585–621 (1961).
26. Toutoufarie, M., Bausens, E. & Debaq-Czainiaux, F. The impact of cellular senescence in skin ageing: A notion of mosaic and therapeutic strategies. Biochem Pharmacol. 142, 1–12 (2017).
27. Kim, R. H. et al. Regulation of p53 during senescence in normal human keratinocytes. Aging Cell 14, 838–846 (2015).
28. Andriani, G. A. et al. Whole Chromosome Instability induces senescence and promotes SASP. Sci. Rep. 6, 35218 (2016).
29. Smirnov, A. et al. FOXX1 regulates proliferation, senescence and oxidative stress in keratinocytes and cancer cells. Aging 8, 1384–1397 (2016).
30. Bessarabova, M., Ishkin, A., Jебале, L., Nikolskaya, T. & Nikolsky, Y. Knowledge-based analysis of proteomics data. BMC Bioinform. 13, S13 (2012).
31. Pellegrini, G. et al. p63 identifies keratinocyte stem cells. Proc. Natl. Acad. Sci. USA 98, 3156–3161 (2001).
32. Graf, T. & Enver, T. Forcing cells to change lineages. Nature 462, 587–594 (2009).
33. Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57–74 (2012).
34. Ernst, J. & Kellis, M. ChromHMM: automating chromatin-state discovery and characterization. Nat. methods 9, 215–216 (2012).
35. Roadmap Epigenomics, C. et al. Integrative analysis of 111 reference human genomes.
36. Winter, G. E. et al. BET Bromodomain Proteins Function as Master Transcription Elongation Factors Independent of CDK9 Recruitment. Mol. cell 67, 5–18 e19 (2017).
37. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. methods 10, 1213–1218 (2013).
38. Brown, J. D. et al. NF-kappaB Directs Dynamic Super Enhancer Formation in Inflammation and Atherosclerosis. Molecular cell, https://doi.org/10.1016/j.molcel.2014.08.024 (2014).
39. Mathelier, A. et al. JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic acids Res. 42, D142–D147 (2014).
40. Matys, V. et al. TRANSFAC and its module TRANTCSp: transcriptional gene regulation in eukaryotes. Nucleic acids Res. 34, D108–D110 (2006).
41. Jelma, A. et al. DNA-binding specificities of human transcription factors. Cell 152, 327–339 (2013).
42. Szklarczyk, D. et al. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. Nucleic acids Res. 45, D362–D368 (2017).
43. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. Bioinformatics 27, 1017–1018 (2011).
44. Jobeli, L. et al. Selenium preserves keratinocyte stemness and delays senescence by maintaining epidermal adhesion. Aging (Albany NY) 9, 2302–2315 (2017).
45. Wang, D., Russell, J. L. & Johnson, D. G. E2F4 and E2F1 have similar transcriptional attenuator of interferon-alpha/beta signaling. Proc. Natl. Acad. Sci. USA 113, 79–84 (2016).
46. Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic Screens in Human Cells Using the CRISPR-Cas9 System. Cell 162, 343–353 (2015).
47. Hoffmann, G. R. et al. Functional epigenetics approach identifies BRM/SMARC2 as a critical synthetic lethal target in BRG1-deficient cancers. Proc. Natl. Acad. Sci. USA 111, 3128–3133 (2014).
48. 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).
49. König, R. et al. A probability-based approach for the analysis of large-scale RNAi screens. Nat. Methods 4, 847 (https://www.nature.com/articles/nmeth1089#supplemental-information (2007).
50. Meierk, K. L. et al. The pIDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. Proc. Natl. Acad. Sci. USA 108, 3665–3670 (2011).
51. Rott, A. et al. CRISPRi: a potent and flexible system for force-directed genome editing. Cell Res. 25, 119–129 (2015).
52. König, R. et al. A probability-based approach for the analysis of large-scale RNAi screens. Nat. Methods 4, 847 (https://www.nature.com/articles/nmeth1089#supplemental-information (2007).
53. Meierk, K. L. et al. The pIDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. Proc. Natl. Acad. Sci. USA 108, 3665–3670 (2011).
54. Schmidt, C., Rendeiro, A. F., Shef, N. C. & Rock, C. ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors. Nat. Methods 12, 963–965 (2015).
55. Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: A Method for Assessing Chromatin Accessibility Genome-Wide. Curr. Protoc Mol. Biol 109, 21.29.21–29 (2015).
56. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9, R137 (2008).
57. Semenkovich, N. P. et al. Impact of the gut microbiota on enhancer accessibility in gut intraepithelial lymphocytes. Proc. Natl. Acad. Sci. USA 113, 14805–14810 (2016).
58. Zed, R. et al. Enhancer invasion shapes MYCN-dependent transcriptional amplification in neuroblastoma. Nat. Genet. 50, 515–523 (2018).

Acknowledgements

We thank Philipp Hoppe, Lapa Morelli, and Daniela Stauffer for help with CRISPR guide design and editing methodology. Walter Carbone for high throughput sequencing and...
informatics and Alicia Lindeman for high throughput sequencing of CRISPR libraries. Nicole Renaud and Florian Kiefer for support in developing the chromatin analysis workflow. Kayhangabriel Akyel and Oleg Iartchouk for enabling the BRD4 ChIP analysis at Active Motif. William Tschantz for supply of Cas9 protein. Svenja Ley for histology and staining on the 3D human skin model.

Author contributions
F.L., N.M., C.Y.L. and S.K. conceptualized the study, designed, and analyzed the experiments: J.R-H. and J.A. generated the reagents for the pooled CRISPR screen: N.M., D.E., D.H., C.R. and C.K. designed and executed the CRISPR-Cas9 mini-pool screen: N.M. and R.T. designed and executed the H3K27ac ChIP: F.L. and N.M. performed ATACSeq: J.K. and G.R. performed RNASeq: H.R. and A.A. designed and executed the 3D human skin model experiments: G.S. performed the keratinocyte migration assays. N.M., G.S. and A.S. performed all other keratinocyte experiments: C.K. performed integrative analysis across the data sets: S.B. analyzed BRD4 ChIP experiments: C.G.K., F.N. and G.R supervised all RNASeq, ChIPSeq, and ATACSeq analysis: C.Y.L. and T.C.B. performed the core transcriptional circuitry analysis: C.M.O. and H.E. Sheppard built the IRF2 construct: N.M. generated the samples and S.I. performed the IRF2 ChIP. M.F. established the collaboration S.K., F.L., N.M., C.Y.L., C.K. and T.B. wrote and revised the paper.

Competing interests
Most authors (with the exception of C.Y.L., T.B., S.I., C.M.O., H.E.S., S.E., all Baylor, TX) were employees of Novartis during the course of this work, as indicated in the affiliations. C.Y.L. received sponsored travel from Novartis and is a shareholder and inventor of IP licensed to Syros Pharmaceuticals. Some authors own shares in Novartis. The remaining authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-019-12559-x.

Correspondence and requests for materials should be addressed to F.L., C.Y.L. or S.K.

Peer review information Nature Communications thanks Christopher French and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019