

1 **A genome-wide long noncoding RNA CRISPRi screen identifies *PRANCR* as a**
2 **novel regulator of epidermal homeostasis**

3

4 Pengfei Cai^{1*}, Auke BC Otten^{2*}, Binbin Cheng², Mitsuhiro A. Ishii², Wen Zhang¹, Beibei
5 Huang¹, Kun Qu^{1,3**}, Bryan K. Sun^{2**}

6 ¹Division of Molecular Medicine, Hefei National Laboratory for Physical Sciences at
7 Microscale, the CAS Key Laboratory of Innate Immunity and Chronic Disease, Department of
8 Oncology of the First Affiliated Hospital, Division of Life Sciences and Medicine, University of
9 Science and Technology of China, Hefei, China

10 ²Department of Dermatology, University of California–San Diego, La Jolla, California, USA

11 ³CAS Center for Excellence in Molecular Cell Sciences, University of Science and
12 Technology of China, Hefei, China

13

14 *Co-first authors

15 **Corresponding authors: bryansun@ucsd.edu; qkun@ustc.edu.cn

16

17 **Running title:** Control of epidermal homeostasis by lncRNAs

18 **ABSTRACT**

19 Genome-wide association studies indicate that many disease susceptibility regions reside in
20 non-protein coding regions of the genome. Long noncoding RNAs (lncRNAs) are a major
21 component of the noncoding genome, but their biological impacts are not fully understood.
22 Here, we performed a CRISPR interference (CRISPRi) screen on 2,263 epidermis-
23 expressed lncRNAs and identified nine novel candidate lncRNAs regulating keratinocyte
24 proliferation. We further characterized a top hit from the screen, *progenitor renewal-*
25 *associated noncoding RNA (PRANCR)*, using RNA interference-mediated knockdown and
26 phenotypic analysis in organotypic human tissue. *PRANCR* regulates keratinocyte
27 proliferation, cell cycle progression, and clonogenicity. *PRANCR*-deficient epidermis
28 displayed impaired stratification with reduced expression of differentiation genes that are
29 altered in human skin diseases, including keratins 1 and 10, filaggrin, and loricrin.
30 Transcriptome analysis showed that *PRANCR* controls expression of 1,136 genes, with
31 strong enrichment for late cell cycle genes containing a *CHR* promoter element. In addition,
32 *PRANCR* depletion leads to increased levels of both total and nuclear CDKN1A (also known
33 as p21), which is known to govern both keratinocyte proliferation and differentiation.
34 Collectively, these data demonstrate that *PRANCR* is a novel lncRNA regulating epidermal
35 homeostasis and identify other lncRNA candidates that may have roles in this process as
36 well.

37

38 **INTRODUCTION**

39 Whole-exome sequencing has accelerated the discovery into genetic causes of disease.
40 However, the majority of whole-exome studies do not identify a causative mutation (Yang et
41 al. 2013), potentially reflecting the fact that protein-coding regions comprise only ~1-2% of
42 the human genome, and indicating the potential contribution of non-protein coding mutations.
43 Consistent with this, genome-wide association studies (GWAS) indicate that over 93% of

44 disease-linked single nucleotide polymorphisms (SNPs) map to the non-coding genome (Tak
45 and Farnham 2015). These observations therefore suggest a potential underappreciated
46 disease relevance of non-coding elements such as enhancers and non-coding RNAs (Zhang
47 and Lupski 2015).

48 A major component of the noncoding genome are long noncoding RNAs (lncRNAs),
49 transcribed RNA elements greater than 200 nucleotides with no apparent protein-coding
50 potential (Wilusz et al. 2009; Da Sacco et al. 2012). Compared to their protein-coding
51 counterparts, lncRNA transcripts have a lower expression level and demonstrate strong
52 tissue-specific expression patterns, implying potential functions particular to specific
53 biological states (Derrien et al. 2012; Liu et al. 2017). To date, nearly 28,000 human lncRNAs
54 have been catalogued (Hon et al. 2017), but only about 200 have been functionally
55 characterized (Quek et al. 2015). This is in part due to the challenges of studying lncRNAs.
56 They are less sequence-conserved than proteins (Kellis et al. 2014), and some functional
57 lncRNAs are primate or even human-specific (Awan et al. 2017). This limits the application of
58 classical genetic systems, such as murine models, to studying human lncRNAs. In addition,
59 the majority of lncRNAs are functional in only one cell type (Liu et al. 2017). Therefore,
60 understanding of lncRNAs requires examination in the appropriate cell/tissue context. Lastly,
61 it has been a matter of ongoing debate whether noncoding elements are broadly functional at
62 all, or whether many of these elements have minimal biological significance (Doolittle 2013).

63 A recent evaluation of lncRNA expression across species and development indicates that
64 lncRNAs are dynamically expressed and conserved in organs, suggesting that they have
65 evolved specialized functions in human tissues (Sarropoulos et al. 2019). Here, we use
66 human epidermis as a model system for evaluating the function of lncRNAs in tissue
67 homeostasis. In the epidermis, there is a dynamic, ongoing balance between progenitor
68 proliferation and differentiation. Genetic disruptions of proliferation or differentiation disrupt
69 this homeostasis and underlie common diseases such as eczema, psoriasis and keratinocyte
70 cancers, which collectively impact >20% of the population (Lopez-Pajares et al. 2013).

71 Understanding the potential role of lncRNAs in epidermal homeostasis therefore has broad
72 relevance to human health. Using the human epidermis as a model also offers several
73 unique advantages. First, the skin is accessible, allowing collection of primary tissue and
74 cells for study. The *ex vivo* culture conditions for primary human progenitor keratinocytes are
75 well-developed and permit complex genetic manipulations. Finally, epidermal progenitors can
76 be reconstituted into organotypic tissue, allowing studies in a three-dimensional context (Oh
77 et al. 2013). These advantages are particularly useful in studying lncRNAs, whose
78 phenotypes in cultured cells do not always extrapolate to primary tissue or *in vivo* contexts
79 (Bassett et al. 2014). In this report, we perform a CRISPR interference (CRISPRi) screen to
80 systematically identify functional lncRNAs that have roles in human epidermal homeostasis
81 and characterize the tissue and molecular phenotype of a novel lncRNA hit from the screen.

82

83 **RESULTS**

84 **Transcriptome analysis and CRISPRi screen to identify functional lncRNAs**

85 The epidermis consists primarily of organized layers of keratinocytes. Stem/progenitor
86 keratinocytes reside in the innermost layer atop a basement membrane and are capable of
87 self-renewal or differentiation. Differentiating keratinocytes detach from the basement
88 membrane and migrate upwards to form the suprabasal layers, which serve as a structural
89 and functional barrier. At the outermost layers, keratinocytes are enucleated to form a
90 cornified layer, which is eventually sloughed off the surface. To sustain homeostasis, the
91 epidermis dynamically balances cell renewal, differentiation, and cell loss.

92 To identify lncRNAs regulating tissue homeostasis, we first focused on identifying lncRNAs
93 governing epidermal progenitor renewal. As expression level is the most significant predictor
94 of functional lncRNAs (Liu et al. 2017), we combined RNA expression profiling of human
95 epidermis with a CRISPRi screen (Fig. 1A). Using RNA-sequencing data from duplicate
96 biological samples of clinically-normal human epidermis (Sun et al. 2015), we identified 8,634

97 gene transcripts at an RPKM of >1. Classification of expressed transcripts into protein-
98 coding, lncRNAs, microRNAs, and small nucleolar RNAs led to the assignment of 2,263
99 elements as lncRNAs (Fig 1A). The average expression level of all lncRNAs was lower than
100 that of protein-coding genes (Supplemental Fig S1), consistent with previous reports (Derrien
101 et al. 2012; Liu et al. 2017; Tuck et al. 2018).

102 To systematically interrogate the potential roles of these lncRNAs in epidermal progenitor
103 growth, we designed a CRISPRi screen. In a classic version of CRISPRi, a catalytically dead
104 (d)Cas9 protein is fused to a KRAB transcriptional repressor domain, which is guided to its
105 genomic target by a single guide RNA (sgRNA) (Qi et al. 2013). This system has proven
106 useful for loss-of-function targeting of lncRNAs, which are not reliably inactivated by the short
107 indels generated by CRISPR/Cas9 nuclease approaches (Liu et al. 2017).

108 We constructed a custom sgRNA library comprised of 5 independent sgRNAs against each
109 of the 2,263 lncRNAs using Sequence Scan for CRISPR (SSC) (Xu et al. 2015) and included
110 250 non-targeting sgRNA controls. dCas9-expressing keratinocytes were generated by
111 lentiviral transduction of a dCas9-KRAB expression cassette into the clone 103 epidermal
112 keratinocyte cell line (Sun et al. 2015) and selection of a high-expressing clone (see
113 Methods). Into these keratinocytes, the CRISPRi sgRNA library was transduced at a
114 multiplicity of infection of 0.3, selected cells were propagated in culture, and genomic DNA
115 collected at the initial time point and after 28 days of continuous proliferation (Fig 1A). Deep
116 sequencing was used to quantify sgRNA representation at each time point. After
117 normalization and mean-variance modeling, we determined enriched and depleted sgRNAs
118 (Fig 1B). The experiment was performed in technical and biological duplicates, which both
119 showed high reproducibility (R^2 of 0.91 and 0.86 respectively, Supplemental Fig S1).

120

121 **CRISPRi screen identifies novel lncRNAs regulating progenitor replication**

122 We assessed screening results using Model-based Analysis of Genome-wide CRISPR-Cas9
123 Knockout (MAGeCK), a computational analysis tool developed to robustly identify essential
124 gene elements from CRISPR screens (Li et al. 2014). The MAGeCK algorithm uses a mean-

125 variance model to test if sgRNA abundances differ significantly before vs. after the screen
126 with respect to neutral variation of sgRNA abundance, as assessed by the non-targeting
127 sgRNAs. A robust rank aggregation (RRA) algorithm (lncRNA gene ranking) computes p -
128 values, false discovery rates (FDR) and RRA scores (Li et al. 2014) based on screen results.
129 We chose to initially benchmark our screen results against *TINCR*, one of the few well-
130 characterized lncRNAs known to be involved in epidermal homeostasis (Kretz et al. 2013).
131 To do so, we considered lncRNAs scored with a false discovery rate (FDR) value similar or
132 lower than *TINCR* (FDR = 0.07; threshold at FDR < 0.10) as positive hits. We also stipulated
133 that at least three independent sgRNAs against a lncRNA candidate “hit” must change
134 consistently within the screen. Using these relatively strict parameters, we identified nine
135 novel candidate lncRNA hits, all of which were positive regulators of proliferation (red dots in
136 Fig 1C). We excluded two pseudogenes (*MEMO1P1* and *GUSBP4*) because their high
137 sequence homology to protein-coding genes complicates the use of RNA interference and
138 downstream analysis. The remaining candidates were ranked, rendering our candidate
139 lncRNA list (Fig 1D-E). These lncRNA hits had higher expression values than the global
140 lncRNA average (Supplemental Fig S1), consistent with expression levels being an indicator
141 of functional lncRNAs (Liu et al. 2017).

142 To further validate the screen, we aimed to further characterize the phenotype and function
143 of a top candidate. We chose to focus on *RP11-611E13.2* for several reasons. First,
144 ENCODE data showed that its transcriptional start site is enriched with histone 3 lysine 27
145 acetylation (H3K27ac) and histone 3 lysine 4 trimethylation (H3K4me3) in neonatal human
146 epidermal keratinocytes (Fig 2A), a pattern consistent with active genes (Kellis et al. 2014).
147 In addition, our RNA-seq data on human epidermis included subcategorization of basal vs
148 suprabasal transcripts by laser capture microdissection, which allowed us to evaluate counts
149 from basal (progenitor) and suprabasal (differentiated) epidermal layers (Sun et al. 2015).
150 *RP11-611E13.2* is more highly expressed in the basal layers of the epidermis (average
151 RPKM of 5 in basal layers vs 0.02 in suprabasal layers), consistent with its expression where
152 epidermal progenitors are actively dividing (Fig 2B). Based on these observations and the

153 screening results, we termed this lncRNA *progenitor renewal-associated noncoding RNA*
154 (*PRANCR*) and performed more detailed evaluation of its phenotype and function.

155

156 ***PRANCR* is essential for proliferation and clonogenicity of primary keratinocytes**

157 CRISPRi can block genome elements by establishing a repressive chromatin state at the
158 genomic locus, by blocking transcription of a functional RNA, or both. To distinguish between
159 these possibilities, we used short hairpin RNA (shRNA)-mediated knockdown to deplete
160 *PRANCR* RNA without interfering with the genomic locus. Of the six independent, non-
161 overlapping shRNAs (shLNCs), five resulted in *PRANCR* RNA depletion ranging from 30-
162 78% (Supplemental Fig S2A). We employed the two shLNCs that achieved the greatest
163 *PRANCR* knockdown for most downstream experiments (Fig 2C).

164 *PRANCR* is transcribed divergently from the same promoter region as the protein-coding
165 gene *CNOT2* (Fig 2A). Therefore, we performed Western Blotting to evaluate if *CNOT2*
166 protein levels were altered with *PRANCR* shRNA-mediated depletion. We observed no
167 change in *CNOT2* expression using multiple independent shRNAs against *PRANCR* (Fig 2D
168 and Supplemental Fig S2), arguing against the phenotype arising from changing expression
169 of the divergently transcribed gene. We also evaluated if the phenotype identified by the
170 CRISPRi screen could be recapitulated with *PRANCR* RNAi-mediated depletion by
171 assessing proliferation of keratinocyte progenitors. Our results confirmed markedly inhibited
172 keratinocyte proliferation with *PRANCR* knockdown (Fig 2E), a result also observed for all
173 five effective shRNAs (Supplemental Fig S2B). Together, these results support a primary role
174 for *PRANCR* lncRNA in causing the observed phenotype.

175 Next, we evaluated whether *PRANCR* influenced the stem cell potential of epidermal
176 progenitors. In human skin, epidermal keratinocytes can be classified into different
177 populations with varying clonogenic potential (Barrandon and Green 1987). Holoclones, the
178 population with the greatest renewal and proliferative capacity, are critical for long-term
179 epidermal renewal. To assess the impact of *PRANCR* on clonogenic potential, control vs.
180 *PRANCR*-depleted keratinocytes were seeded on fibroblast feeders. *PRANCR* depletion

181 resulted in a significantly reduced number of holoclones, demonstrating an intrinsic loss of
182 clonogenic capacity resulting from loss of *PRANCR* (Fig 2F-G).

183 The robust phenotype of reduced proliferative capacity prompted us to investigate how
184 *PRANCR* depletion impacted cell cycle progression. To quantitate cell cycle state, we
185 performed flow cytometric analysis of propidium iodide (PI)-stained cells. Upon depletion of
186 *PRANCR* (Fig 2H), we observed a decrease in the proportion of cells in the S phase (~8-9%
187 vs 21-22% in control cells) and an increase in the proportion of cells in G₂/M phase (~37-41%
188 vs 21-26%), with minimal changes in the fraction of cells in the G₁/G₀ phase (~44-46% vs 45-
189 52%). Independent biological replicates confirmed these findings (summarized in Fig 2I and
190 Supplemental Fig S3). An increase of cells in G₂/M phase can lead to induction of apoptosis.
191 Therefore, we assessed whether *PRANCR* depletion promoted apoptosis. We quantitated
192 apoptosis induction by flow cytometric analysis of FITC-Annexin-V labeling of
193 phosphatidylserine (PS) exposure (Supplemental Fig S2). We observed no difference in
194 apoptosis induction upon *PRANCR* depletion. Together, these experiments established
195 *PRANCR* as a novel lncRNA essential for epidermal progenitor renewal, proliferation, and
196 clonogenic potential, without effects on cell apoptosis.

197

198 ***PRANCR* is required for formation of organotypic stratified epidermis**

199 Organotypic epidermal tissue displays superior correlation to *in vivo* biological skin states
200 compared to cultured cells (Ridky et al. 2010). We therefore sought to assess the impact of
201 *PRANCR* depletion in organotypic human epidermal tissue. Tissues were generated with
202 *PRANCR*-depleted and control (scrambled) progenitors (Fig 3). *PRANCR*-depleted
203 progenitors displayed disrupted architecture of the outermost epidermal layers (H&E stains,
204 Fig 3A). *PRANCR*-depleted epidermis was notably thinner than controls (Fig 3B). Proper
205 organization of the outermost epidermal layers is essential for skin barrier function and the
206 ability to prevent evaporative water loss. Corroborating the histological findings,
207 immunofluorescence of epidermal differentiation markers keratin 10 (KRT10) and filaggrin
208 (FLG) demonstrated a deficiency in expression of structural and barrier proteins required for

209 functional epidermis (Fig 3A, C-D). We assessed mRNA expression levels of the epidermal
210 differentiation markers keratins 1 and 10, filaggrin, and loricrin, critical genes disrupted in
211 human skin diseases (O'Driscoll et al. 2002; Virtanen et al. 2003; Smith et al. 2006). Their
212 expression was consistently reduced in *PRANCR*-depleted epidermis (Fig 3E). Finally, to
213 assess progenitor replication at the basal epidermal stratum, we performed staining for the
214 replication marker MKI67 (Guillaud et al. 1991), which revealed a marked decrease in
215 MKI67-positive cells in *PRANCR*-depleted epidermis (Fig 3A, F). Collectively, these data
216 indicate that *PRANCR* impacts epidermal tissue homeostasis by regulating both proliferation
217 and differentiation.

218

219 ***PRANCR* regulates expression of cell cycle and MAPK pathway genes *in trans***

220 To investigate how *PRANCR* controls keratinocyte progenitor function, we performed whole
221 transcriptome sequencing on control and *PRANCR*-depleted primary keratinocytes. After
222 differential expression analysis using DESeq (Anders and Huber 2010) and based on
223 thresholds of absolute \log_2 fold change (LFC) of 1 and a p-value of <0.05 , we identified 1,136
224 differentially expressed genes (DEGs) in *PRANCR* knockdown cells (Fig 4A). Most of these
225 genes (927) showed decreased expression. Gene Ontology (GO) analysis using Metascape
226 (Tripathi et al. 2015) revealed enrichment in GO terms related to the cell cycle, mitotic cell
227 phase transition, and DNA replication initiation (Fig 4B). The 209 upregulated genes showed
228 enrichment for several pathways, most notably genes related to MAPK pathway signaling
229 (Fig 4C). MAPK signaling is integral to keratinocyte renewal and differentiation (Scholl et al.
230 2007), raising the possibility that the *PRANCR* depletion phenotype might function through
231 this pathway.

232 LncRNAs can regulate gene expression of adjacent (*cis*) and/or distal (*trans*) genes (Ulitsky
233 and Bartel 2013). To characterize how *PRANCR* impacts global gene transcription, we
234 assessed the genomic location of the DEGs. Our results suggest no disproportionate
235 enrichment for DEGs on Chromosome 12, where *PRANCR* resides (Supplemental Fig S4; p-

236 value for downregulated genes $p = 0.167$ and for upregulated genes $p = 0.051$;
237 hypergeometric test). We also observed no significant change of expression for transcripts
238 immediately upstream and downstream of *PRANCR* (*CNOT2* and *MYRFL*). In addition to
239 assessing *PRANCR*-depleted cells, we evaluated expression of the five *cis*-adjacent genes
240 in *PRANCR*-depleted organotypic epidermis. As observed in the cultured cell context, we
241 also found that the *cis*-adjacent genes displayed no consistent directional changes in
242 organotypic tissue (Supplemental Fig S4). Based on these results, we conclude that
243 *PRANCR* does not display significantly enriched *cis*-regulation, and affects genes in *trans*.

244

245 ***PRANCR* regulates expression of E2F and FOXM1-targeted genes**

246 LncRNAs can affect the expression of distal genes by controlling transcriptional complexes
247 that globally impact gene expression (Long et al. 2017). The analysis tool *Enrichr* (Chen et
248 al. 2013b; Kuleshov et al. 2016) integrates genome-wide ChIP experiments in the ChEA
249 (Lachmann et al. 2010) and ENCODE Project databases (The ENCODE Project Consortium
250 2012) to identify transcription factors involved in the control of a gene set. For *PRANCR*
251 downregulated DEGs, *Enrichr* identified E2F4 and FOXM1 as the most enriched transcription
252 factors (Fig 5A). These transcription factors are prominent components of two distinct, but
253 closely related protein complexes that govern cell cycle gene expression (Engeland 2018).
254 E2F transcription factors have a predominantly nuclear localization (Magae et al. 1996) and
255 are essential to epidermal development (Ivanova et al. 2005). To determine whether
256 *PRANCR* displayed similar localization, subcellular fractionation was performed, which
257 confirmed the enrichment of *PRANCR* transcripts in the nucleus (Fig 5B). Next, we
258 specifically examined the expression of E2F target genes (Bracken et al. 2004) among the
259 genes altered with the *PRANCR* knockdown (Fig 5C). This analysis confirmed that known
260 E2F transcription factor target genes were markedly downregulated upon *PRANCR* depletion
261 (p -value = 0.0074), consistent with a model of *PRANCR* affecting expression of genes
262 targeted by the E2F family of transcription factors.

263

264 ***PRANCR* regulates expression of cell cycle genes containing a *CHR* promoter element**

265 We further explored the role of *PRANCR* in gene expression regulation by applying *HOMER*
266 motif analysis, which identifies regulatory motifs enriched in the promoters of a gene set
267 (Heinz et al. 2010). For the genes downregulated by *PRANCR* knockdown, this analysis
268 identified enrichment of E2F binding sites (Fig 5D). Additionally, this analysis suggested an
269 even stronger role for the Cell cycle genes Homology Region (*CHR*), a DNA element present
270 in promoters of many cell cycle genes (Muller and Engeland 2010) that are bound by E2F4-
271 and/or FOXM1-containing protein complexes (Chen et al. 2013c; Fischer et al. 2014). Gene
272 expression analysis of 148 late cell cycle genes harboring evolutionary-conserved *CHR*
273 elements (Muller et al. 2014) confirmed a highly significant and consistent downregulation of
274 these genes in *PRANCR*-depleted keratinocytes (p-value = 3.2×10^{-15} , Fig 5E).

275 Analysis of RNA-seq data indicated that *E2F4* is the most highly expressed E2F family
276 member in primary keratinocytes (Fig 5F). Recent studies indicate that expression of cell
277 cycle genes with *CHR* sites is repressed indirectly by TP53– via CDKN1A and E2F4–
278 collectively called the TP53-CDKN1A-DREAM-*CHR* pathway (Quaas et al. 2012). Therefore,
279 we tested whether genes of this specific pathway are affected by *PRANCR* knockdown. The
280 gene expression profile of 210 genes in the TP53-CDKN1A-DREAM-*CHR* pathway (Fischer
281 et al. 2016) indicated that expression of these target genes is strongly repressed upon
282 *PRANCR* depletion (p-value = 2.9×10^{-10} ; Fig 5G). Consistent with the observed G₂/M arrest
283 upon *PRANCR* knockdown (Fig 2I), downregulation of genes through this pathway appears
284 to be important for G₂/M cell cycle control (Fischer et al. 2016). Additionally, prominent
285 examples of TP53-DREAM genes involved in G₂/M checkpoint control- *CHEK2*, *CDK1*,
286 *CCNB1*, *CCNB2* and *CDC25C* (Engeland 2018)- were impaired with *PRANCR* knockdown
287 (p-value < 0.05, Fig 5H). As the regulation of expression of *CHR*-containing genes is a critical
288 shared mechanism across different cell types (Muller and Engeland 2010), we also examined
289 whether *PRANCR* functions similarly in another cell type. We depleted *PRANCR* in primary
290 human fibroblasts and observed impaired proliferation as well as reduced expression of G₂/M
291 checkpoint control genes *CCNB1*, *CCNB2*, *CDC25C* and *CDK1* (Supplemental Fig S5).

292 Together, our results support a model in which *PRANCR* regulates expression of late cell
293 cycle genes containing *CHR* sites. This may represent a general mechanism that is
294 functional in multiple cell types and tissues.

295

296 ***PRANCR* alters CDKN1A expression and localization**

297 Finally, to gain insight to how *PRANCR* may interact with the TP53-CDKN1A-DREAM/E2F4-
298 *CHR* pathway, we evaluated if *PRANCR* depletion altered protein levels or subcellular
299 localization of TP53, CDKN1A, or E2F4 (Fig 5I). TP53 and E2F4 expression were relatively
300 unchanged, but CDKN1A expression increased ~4-fold with *PRANCR* depletion (Fig 5I and
301 Supplemental Fig S6). CDKN1A was also more highly represented in nuclear fractions upon
302 *PRANCR* depletion (Fig 5I and Supplemental Fig S6). CDKN1A has a dual role in the
303 epidermis to negatively regulate both proliferation and differentiation (Devgan et al. 2006), a
304 phenotype that mirrors what we observed with *PRANCR* depletion. Together, the
305 transcriptomic evidence and phenotypic similarities suggest an initial model by which
306 *PRANCR* regulates epidermal homeostasis by modulating CDKN1A expression and nuclear
307 localization.

308

309 **DISCUSSION**

310 The noncoding genome has important undiscovered roles in human development and
311 disease (Yang et al. 2013; Tak and Farnham 2015; Zhang and Lupski 2015). However, the
312 biological significance of most noncoding genetic elements is still unknown. Genome-wide
313 screens are a valuable approach to systematically evaluate their potential functions. In this
314 report, we performed a CRISPR interference screen in human keratinocytes to identify
315 lncRNAs controlling epidermal progenitor replication, a fundamental process underlying skin
316 homeostasis. Applying a relatively stringent threshold, we identified nine lncRNAs that
317 regulate progenitor renewal and represent a foundation for follow-up studies to understand
318 involvement of lncRNAs in this process.

319 A more detailed characterization of a top candidate, *PRANCR*, reveals a novel lncRNA that
320 is required for both proliferation and clonogenicity of epidermal progenitors, as well as tissue
321 stratification/differentiation. We found that experimental *PRANCR* depletion leads to
322 upregulation of total and nuclear CDKN1A, and we hypothesize that this may reflect a
323 principal mechanism by which *PRANCR* controls progenitor replication. CDKN1A promotes
324 formation of the E2F4-containing DREAM complex, which binds E2F and *CHR* promoter
325 motifs to repress target gene transcription of cell cycle genes (Quaas et al. 2012; Fischer et
326 al. 2014). By contrast, when CDKN1A levels are low, this protein complex switches to a
327 FOXM1-containing MMB complex that activates late cell cycle genes with a *CHR* motif
328 (Sadasivam et al. 2012; Chen et al. 2013c; Engeland 2018). *PRANCR* may therefore govern
329 keratinocyte progenitor cell cycle through CDKN1A-mediated regulation of both E2F and
330 FOXM1-targeted genes.

331 Further supporting this overarching hypothesis is the observation that *PRANCR* and
332 CDKN1A both share negative regulatory impacts on both epidermal proliferation and
333 differentiation. Generally, proliferation and differentiation are envisioned as opposing fates in
334 a cell. For instance, the epidermal transcription factor ZNF750 forms unique protein
335 complexes that promote differentiation and block proliferation (Boxer et al. 2014). High levels
336 of CDKN1A in the epidermis, however, inhibits both self-renewal and differentiation of
337 keratinocytes (Missero et al. 1996; Topley et al. 1999; Devgan et al. 2006), similar to what is
338 observed with *PRANCR* depletion. While early CDKN1A induction leads to cell cycle arrest,
339 persistent elevated expression suppresses differentiation through activation of the MAPK
340 cascade (Devgan et al. 2006). Consistent with this, RNA-seq of *PRANCR*-depleted cells
341 revealed upregulation of genes regulating MAPK signaling (Fig 4C).

342 These findings naturally raise additional questions: *PRANCR* depletion leads to increased
343 CDKN1A expression, but does expression of *PRANCR* downregulate CDKN1A? This would
344 oppose the function of TP53, which is known to activate CDKN1A, and raises the possibility
345 that *PRANCR* is pro-oncogenic. Both *E2F* and *CHR* motifs have been reported to be central

346 elements in key genes associated with cancer signaling (Paci et al. 2017), which are
347 orchestrated by TP53 (Engeland 2018). In addition, our initial characterization examined
348 expression of *PRANCR* in bulk culture. However, it is possible that *PRANCR* expression
349 changes dynamically during progression through the cell cycle and/or in response to
350 contextual cues. Future single cell and *in vivo*-level experiments will aim to answer these
351 questions, and further elucidate the role of *PRANCR* in the epidermis.

352 Our study demonstrates the value of CRISPRi screens to provide insight into the functional
353 noncoding genome. Studying lncRNAs presents several challenges. Accurate identification
354 of lncRNA transcriptional coordinates is critical for CRISPRi screens, which function optimally
355 by targeting regions close to the transcriptional start site (TSS) (Gilbert et al. 2014). Different
356 annotation systems demonstrate variability in lncRNA annotations, and to best address these
357 differences, we combined RefSeq, UCSC, and ENCODE/GENCODE annotations to identify
358 and map lncRNAs. Since the initiation of our work, the FANTOM5 consortium reported a
359 comprehensive atlas of human lncRNA genes with high-confidence 5' ends using cap
360 analysis gene expression (CAGE) (Hon et al. 2017). These efforts have improved the
361 mapping of lncRNA TSS, which will further improve efficiency in the design and function of
362 future lncRNA CRISPR screens.

363 Disruptions of epidermal homeostasis underlie many skin diseases. *PRANCR* and other
364 epidermal lncRNAs may contribute to the pathogenesis of these conditions by controlling cell
365 proliferation and differentiation, as well as other fundamental biological processes that form
366 the basis of human tissue development and health. The CRISPRi lncRNA screening strategy
367 presented here can also be adapted to interrogate the potential roles of *PRANCR* and other
368 epidermal lncRNAs in processes such as carcinogenesis, cell-cell communication, and
369 response to microbes. These genome scale screening approaches will help elucidate the
370 functions of the vast noncoding genome in human tissue development and disease.

371

372 **METHODS**

373 Primary keratinocyte culture

374 Primary epidermal keratinocytes were isolated from discarded neonatal foreskin from
375 circumcisions, collected upon written informed consent under an Institutional Review Board
376 protocol approved by the University of California, San Diego. Cells were isolated based on
377 the protocol described previously (Aasen and Izpisua Belmonte 2010) and propagated in
378 50:50 mixture (“50:50 media”) of K-SFM and 154 media (Life Technologies) with
379 recommended supplements and 1× Antibiotic-Antimycotic (Thermo Fisher Scientific), at 37°C
380 and 5% CO₂.

381

382 RNA sequencing of human skin biopsies

383 We analyzed RNA sequencing performed previously on clinically normal sun-protected
384 human skin (Sun et al. 2015). Reads were aligned to the hg19 genome assembly using
385 TopHat (Trapnell et al. 2009). At the time of study design, hg19 was chosen as a reference
386 because of its inclusion of more complete ENCODE and Epigenomics Mapping Consortium
387 datasets. The choice of reference sequence does not significantly alter principal conclusions:
388 *PRANCR* is present in both hg19 and hg38. These RNA-seq data consisted of laser micro-
389 dissected epidermis from two unrelated individuals. To obtain a high confidence
390 transcriptome, we integrated the transcript annotation with priority of RefSeq > UCSC >
391 ENCODE/GENCODE databases and provided only one transcript annotation if it was defined
392 in multiple databases. Transcripts belonging to the same gene were merged and reads per
393 kilobase per million (RPKM) were assigned to each gene. Genes with average RPKM<1
394 were excluded. The categorization of each gene as protein-coding, lncRNA, miRNA or
395 snoRNA is based on their annotation in RefSeq, UCSC, and ENCODE/GENCODE
396 databases.

397

398 sgRNA library design

399 The CRISPR library was generated with sgRNAs designed against each of the 2,263
400 epidermally-expressed lncRNAs using the SSC algorithm optimized for CRISPR interference

401 (Xu et al. 2015). The genomic interval from -50 to +450 relative to the transcriptional start site
402 (TSS) was used for each target lncRNA transcript. All potential sgRNAs were sorted by
403 efficiency score and negative scores were discarded. For each lncRNA transcript, the top 5
404 scoring sgRNAs were selected, and if any of the top 5 scoring sgRNAs overlapped by more
405 than 5 nucleotides, the lower scoring sgRNA was replaced with the next highest-scoring
406 sgRNA targeting this lncRNA. Once sgRNA selection was complete, candidate sgRNAs were
407 evaluated for the presence of a 'G' nucleotide in the -20 position to facilitate efficient
408 transcription by the pol III promoter. If the -20 position was not a 'G', then a 'G' was
409 substituted at this position. For negative control sgRNAs, 250 sequences of randomly-
410 generated 19mers were generated and verified for their inability to match human genome
411 sequences. A 'G' nucleotide was then prepended to each 19mer.

412

413 **Construction of the CRISPR-Cas9 sgRNA library**

414 All designed library sgRNA sequences were prepended and appended with linker sequences
415 (5' linker: CTTGTGGAAAGGACGAAACACC; 3' linker:
416 GTTTAAGAGCTATGCTGGAAACAGC) to facilitate polymerase chain reaction amplification
417 and to serve as invariant sequence overhangs for InFusion cloning (Clontech). The
418 oligonucleotides were synthesized (CustomArray) and delivered as a single oligonucleotide
419 pool. The pool was amplified for 15 cycles (Forward primer:
420 ATCTTGTGGAAAGGACGAAACA, Reverse primer: CTGTTTCCAGCATAGCTCTTAAAC)
421 with CloneAmp HiFi PCR premix (Clontech). The entire PCR reaction was resolved on a 2%
422 agarose gel, and the product was retrieved by gel slice isolation. A lentiviral vector pSICO-
423 (F+E) was derived from the vector pSLQ1651 (Addgene plasmid # 51024 (Chen et al.
424 2013a)) by insertion of a 1.9 kb BsmBI stuffer fragment between the U6 promoter and
425 downstream tracrRNA. InFusion cloning was performed to assemble the sgRNA library into
426 BsmBI-digested pSICO-(F+E). Lentivirus was generated by the transfection of lentiviral
427 helper plasmids and the CRISPRi plasmid library into 293T cells, and lentiviral supernatant
428 collected 48 hours afterward. Supernatant was concentrated using Lenti-X Concentrator

429 (Clontech) and frozen at -80°C in replicate aliquots. One aliquot was thawed, and infection
 430 titration performed on keratinocytes to determine appropriate dosing to achieve a multiplicity
 431 of infection of 0.3.

432

433 **CRISPR interference screen**

434 pLEX-KRAB-dCas9-Blast was created by cloning the KRAB-dCas9 open reading frame
 435 (Addgene plasmid # 60954 (Gilbert et al. 2014)) into the pLEX-MCS vector (Thermo Fisher
 436 Scientific) using the BamHI/XhoI restriction sites. The puromycin resistance cassette was
 437 swapped for a blasticidin resistance cassette, yielding pLEX-KRAB-dCas9-Blast. Clone 103
 438 keratinocytes (Sun et al. 2015) were infected with pLEX-KRAB-dCas9-Blast in the presence
 439 of $3\ \mu\text{g/ml}$ of polybrene, selected in $2\ \mu\text{g/ml}$ of blasticidin for 72 hours, then expanded in
 440 limited dilution plating to isolate individual clones. Expanded clones were evaluated by
 441 Western blot to select clones with the highest-expressing KRAB-dCas9. 10^7 cells from a
 442 KRAB-dCas9 expressing keratinocyte line were infected with the titrated CRISPRi library at
 443 MOI 0.3 and selected for 48 hours with puromycin. Starting cell numbers were chosen to
 444 achieve $>300\times$ sgRNA overrepresentation. After selection, 4×10^6 cells were reserved and
 445 snap frozen at the pre-selection timepoint. The remaining cells were distributed onto 150 mm
 446 tissue culture plates at 10^6 cells per plate and grown in 50:50 media. Cells were monitored
 447 visually every day and were split upon reaching $\sim 70\%$ confluence and re-plated at a
 448 minimum of 4×10^6 cells to maintain $>300\times$ library overrepresentation. The screen was taken
 449 to 28 days, which allows enough keratinocytes cell doublings to detect significant changes in
 450 sgRNA abundance, and cells were harvested at the endpoint. The screen was performed in
 451 technical duplicates on two different keratinocyte clones. Genomic DNA was isolated from
 452 cell pellets using a Genomic DNA Isolation Kit (Qiagen). sgRNAs were quantitated by
 453 amplification from genomic DNA using PrimeStar (Clontech) using primers that flanked the
 454 sgRNA sequence (forward primer:
 455 ACACGACGCTCTTCCGATCTTGTGGAAAGGACGAAACACC and reverse primer:
 456 GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCGCTGTTCCAGCATAGCTCTTAA). To

457 increase the complexity of the amplicons and facilitate improved clustering on the Illumina
458 sequencer, we used a mixture of forward primers with staggered-length scramble sequences
459 (ACACGACGCTCTCCGATCTNNTGTGGAAAGGACGAAACACC,
460 ACACGACGCTCTCCGATCTNNNGTGGAAAGGACGAAACACC,
461 ACACGACGCTCTCCGATCTNNNNGTGGAAAGGACGAAACACC). The first round of
462 PCR was performed for 20 cycles from a minimum of 16 µg genomic DNA to assure genomic
463 DNA oversampling and to reduce sgRNA amplification bias. The resulting product was
464 column-purified with a PCR purification kit (Macherey-Nagel) and the entire product
465 introduced into a second round of PCR for 7 cycles to introduce Illumina sequencing primers
466 and unique barcodes for each experiment (forward:
467 AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCT
468 and reverse: CAAGCAGAAGACGGCATACGAGAT<Illumina index 6-bp
469 barcode>GTGACTGGAGTTCAGACGTG). Libraries were quantified using the NEBNext
470 Library Quantitation Kit (New England Biolabs) and mixed in equimolar ratios for sequencing
471 on an Illumina HiSeq 4000 using a read length of 75 bp.

472

473 **Analysis of CRISPRi screen results**

474 MAGeCK (0.5.5) was used to analyze the screening sequencing data (Li et al. 2014). The
475 quality of raw data in FASTQ format was evaluated using FastQC, and the sequencing reads
476 were adapter trimmed using Cutadapt before alignment (Martin 2011). Reads were then
477 mapped to the screening sgRNA library without tolerating any mismatches, and the raw read
478 counts of sgRNAs of all samples were merged into a count matrix, which was automatically
479 performed in the MAGeCK software with the “count” command (sgRNA ranking). The
480 MAGeCK “test” command was then used to identify the negatively and positively selected
481 lncRNAs. Instead of computing the size factor from all sgRNAs (the default normalization
482 method for MAGeCK), the size factor was estimated from the negative control sgRNAs to
483 gain a more realistic estimation. All the other analyses and outputs of MAGeCK were

484 according to the default parameters. During the final step of MAGeCK algorithm, a robust
485 rank aggregation (RRA) computes p-values, false discovery rates (FDRs) and RRA scores
486 (Li et al. 2014) to rank the interrogated lncRNA genes. To select the most promising lncRNA
487 candidates, lncRNAs with statistical parameters comparable or better than the positive
488 control *TINCR* (FDR = 0.07) were filtered (FDR < 0.1 and > 2 sgRNAs changing concordantly
489 with the lncRNA phenotype).

490

491 **RNA interference-mediated gene knockdown**

492 For short hairpin-targeted gene knockdown of *PRANCR*, shRNAs were cloned into the
493 pLKO.1 vector (the RNAi consortium). Lentivirus was generated by transfection of both
494 packaging and transfer plasmids into 293T cells using Lipofectamine 3000 (Life
495 Technologies). Supernatants containing lentivirus were collected 48 hours after transfection
496 and concentrated with Lenti-X Concentrator (Clontech) and stored at -80°C. For knockdown,
497 5.0×10^5 keratinocytes were infected with scrambled control or *PRANCR*-targeting shRNAs in
498 medium containing 3 µg/ml polybrene and incubated overnight. Infected cells were selected
499 in medium supplemented with 1 µg/ml puromycin. The shRNA sequences targeting *PRANCR*
500 are: shLNC1: 5'-CACTTTGAATGACAACGATTT-3' and shLNC2: 5'-
501 TACTTCACTCCTTTAAGTTTC-3'. Scrambled shRNA sequences are: *SCR1*: 5'-
502 CCTAAGGTTAAGTCGCCCTCG-3' and *SCR2*: 5'-GCAAGCTGACCCTGAAGTTCA-3'.

503

504 **Cell proliferation assay**

505 To assess cell proliferation rates, 5,000 cells were plated on a 24-well plate in duplicate for
506 each condition and each time point. Media was changed every 48 hours. At each time point,
507 cell abundance was assessed using AlamarBlue (Thermo Fisher Scientific). At each time
508 point, alamarBlue reagent was added following manufacturer's instructions and fluorescence
509 was measured following 2 hours incubation at 37°C using the SpectraMax id3 microplate
510 reader (Molecular Devices). To compare between conditions, fluorescent signals at the start
511 of the experiment (day 0) was set to 1. Subsequently, relative proliferation was measured

512 relative to the day 0 fluorescence signal, as changes in fluorescence are directly proportional
513 to the cell number.

514

515 **Holoclone assay**

516 3T3 fibroblasts were treated with mitomycin C (15 $\mu\text{g/ml}$) for 2 hours at 37°C. Next, 500,000
517 mitomycin C-treated fibroblasts were seeded onto 6-well plates and incubated overnight. The
518 next day, 300 keratinocytes were seeded onto each well. Clones were propagated for up to
519 15 days with media changed every 3-4 days. At the endpoint, fibroblasts were dislodged by
520 vigorously washing with PBS. Holoclones were fixed in ice-cold methanol: acetone (1:1) for 3
521 minutes, stained with 0.02% crystal violet for 2 minutes and de-stained with >3 washes of
522 PBS. Cells were air-dried and imaged by scanning. Holoclones detection and counting were
523 then performed on these scans using ImageJ software.

524

525 **Cell cycle analysis**

526 Cells were cultured in low-serum medium for 24 hours after puromycin selection. Cell cycle
527 analysis was performed using the Cell Cycle Phase Determination Kit (Cayman Chemical)
528 according to the manufacturer's instructions. In short, 1×10^5 cells were washed and fixed at -
529 20°C overnight. Thereafter, cells were stained with propidium iodide (PI) and incubated for 30
530 minutes in the dark and at least 10,000 cells per condition were measured by flow cytometry
531 using the Guava EasyCyte 8HT (Millipore). Analysis of the resulting FCS2.0 files was
532 performed with FlowJo software and flow cytometric data was fit using the built-in Dean-Jett-
533 Fox univariate model to assess the relative distribution of cells over the different cell cycle
534 phases.

535

536 **Organotypic culture**

537 Air-dried devitalized human dermis was mounted onto 1.7 cm \times 1.7 cm supports and 500,000
538 keratinocytes were seeded onto the basement membrane. Tissue was grown in media
539 (Gangatirkar et al. 2007) at an air-liquid interface over a course of 7 days, with medium

540 changed daily. Half of the final tissue was collected in TRIzol for RNA isolation
541 (Supplemental Methods) and the other half was embedded in O.C.T. media (Sakura) and
542 sectioned on a cryostat at 7 μ m thickness. Sections were visualized with hematoxylin/eosin
543 or immunofluorescence (Supplemental Methods). Epidermal thickness was measured at
544 three fixed sites across the tissue using ImageJ, measured from the basement membrane
545 and the most superficial aspect of the stratum corneum. The percentage of MKI67-positive
546 cells were counted using ImageJ using the Analyze particles feature on both DAPI-positive
547 and MKI67-positive cells. For quantitation of KRT10 and FLG, the total (Hoechst and
548 KRT10/FLG) and KRT10/FLG fluorescent signals were quantified using ImageJ (“Threshold
549 color” and “Measure” features) and the %KRT10/FLG was measured as the ratio of the
550 KRT10/FLG signal (“area”) over the total fluorescent signal.

551

552 ***PRANCR* knockdown RNA-seq analysis**

553 Primary keratinocytes from two independent donors were infected with two scrambled and
554 two *PRANCR*-targeting shRNAs. After 72 hours, following complete negative (non-infected)
555 puromycin selection, total RNA was isolated using the Direct-zol RNA kit (Zymo research),
556 treated with DNase I and mRNA was enriched by oligo(dT) magnetic beads (Invitrogen).
557 Sequencing was performed on the Illumina HiSeq 4000, using 50 bp single-end reads. Raw
558 data qualities were evaluated by FastQC. Sequence reads are mapped to the human
559 reference genome (hg19) using STAR (Dobin et al. 2013). Read counts of each gene were
560 collected into a matrix and the differential expression analysis was performed using DESeq
561 (Anders and Huber 2010). After identifying differentially expressed genes, Gene Ontology
562 (GO) analysis was performed using Metascape (Tripathi et al. 2015).

563

564 **Subcellular RNA fractionation**

565 Measurement of the abundance of nuclear and cytoplasmic mRNA was performed as
566 described before (Wang et al. 2006). In short, cells were lysed, and nuclei were pelleted by
567 centrifugation and the supernatant was collected as cytoplasmic fraction. Next, RT-PCR was

568 performed on both fractions, followed by qPCR quantification of the relative abundance of
569 specific mRNA transcripts in both fractions. *NEAT1* and *MALAT1* mRNAs were used as a
570 positive control for nuclear localization and *ACTB* and *GAPDH* mRNAs as positive controls
571 for cytoplasmic localization. Primer sequences can be found in the Supplemental Methods
572 section.

573

574 **Protein Isolation and Western Blot**

575 Whole cell protein lysates were prepared in RIPA buffer and quantitated with the BCA Assay
576 (Pierce). To separate cultured cells into cytoplasmic and nuclear/cytoskeletal fractions, we
577 used the Cell Fractionation Kit (Cell Signaling Technologies) according to manufacturer's
578 instructions. Proteins were analyzed using Western Blotting (Supplemental Methods).
579 Quantification was performed using Image Studio software (LI-COR Biosciences).

580

581 **DATA ACCESS**

582 The raw and processed sequencing data generated in this study are available at the NCBI
583 Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>) under accession
584 number GSE125400.

585

586 **ACKNOWLEDGEMENTS**

587 *Funding:* This work was supported by the National Key R&D Program of China
588 (2017YFA0102900 to K.Q.), the National Natural Science Foundation of China grants
589 (81788101, 91640113, 91940306 and 31771428 to K.Q.), the Fundamental Research Funds
590 for the Central Universities (to K.Q.), the National Institute of Arthritis and Musculoskeletal
591 and Skin Diseases of the National Institutes of Health (K08AR067853 to B.K.S.), and the
592 Doris Duke Charitable Foundation (Clinical Scientist Development Award to B.K.S.). We
593 thank the USTC supercomputing center and the School of Life Science Bioinformatics Center

594 for providing supercomputing resources for this project. The funding sources had no input on
595 the design of the study and collection, analysis, and interpretation of data and in writing the
596 manuscript.

597 *Author's contributions:* BKS and KQ conceived of the project; PC, AO, KQ, and BKS
598 designed the experiments; PC, AO, BC, MAI, and BKS performed the experiments and data
599 acquisition. PC, AO, WZ, BH, KQ, and BKS analyzed and interpreted the data. AO and BKS
600 wrote the first draft of the manuscript and PC and KQ participated in manuscript revision. All
601 authors have read and approved the final manuscript.

602

603 **Figure Legends**

604

605 **Figure 1.** Discovery of lncRNAs controlling proliferation of epidermal progenitors. A)
606 Schematic of CRISPRi library design strategy and the CRISPRi screen. RNA sequencing
607 identified 2,263 lncRNAs with RPKM > 1. Five sgRNAs were designed for each target
608 lncRNA transcript, together with 250 non-targeting controls, to form the CRISPRi library. B)
609 Scatter plots of sgRNA abundance at day 0 and day 28 of the screen. Non-targeting sgRNAs
610 are shown in blue. C) FDR values of each lncRNA candidate, as calculated by MAGeCK. A
611 discovery threshold for positive hits (red) was defined by lncRNAs with similar or stronger
612 FDR values as a known positive control, *PLAC2/TINCR* (FDR threshold = < 0.1). Gene
613 enrichment represents average log-scale enrichment of sgRNAs changing concordantly with
614 the lncRNA selection. D) The robust ranking aggregation (RRA) scores of top lncRNA screen
615 hits. E) Normalized read counts of sgRNAs of top 10 ranked lncRNA hits comparing post-
616 screen (day 28) vs. pre-screen (day 0) abundance. Center lines represent median values;
617 box limits represent the interquartile range; whiskers each extend 1.5 times the interquartile
618 range; dots represent outliers.

619 **Figure 2.** *PRANCR* is a novel epidermal lncRNA and is essential for keratinocyte
620 proliferation and cell cycle progression. A) Schematic of *PRANCR* locus on Chromosome 12,
621 with UCSC tracks for transcription, H3K27ac, H3K4me3, H3K4me1 from the ENCODE
622 project and conservation tracks from phyloP. B) *PRANCR* expression in basal and
623 suprabasal layers from micro-dissected human epidermis. Bars represent RPKM values with
624 SEM, n = 2. C) *PRANCR* mRNA expression in control (scrambled; SCR) and *PRANCR*-
625 depleted (shLNC) progenitors. Bars represent mean with SEM, n = 4, expression compared
626 with one-way ANOVA. D) Western blot for CNOT2, a protein expressed divergently from the
627 *PRANCR* genomic locus, in SCR and shLNC progenitor cells. E) Proliferation assay of
628 control vs. *PRANCR*-depleted progenitors, measured with a fluorescence-based cell
629 quantification assay. Plotted values represent relative increase at each timepoint relative to

630 day 0. $n = 4$, dots represent mean value with SEM. Comparisons performed by 2-way
631 ANOVA. F) Holoclone assay of control and *PRANCR*-depleted keratinocytes. Representative
632 images are shown. G) Quantification of holoclones from control and *PRANCR*-depleted
633 keratinocytes, bars represent mean with SEM, $n = 12$ each. Differences tested using 1-way
634 ANOVA. H) Flow cytometric analysis of cell cycle using propidium iodide DNA staining,
635 comparing normal and *PRANCR*-depleted human epidermal keratinocytes. Graphs represent
636 FlowJo analysis of the flow cytometric results of $\geq 10,000$ cells using the Dean-Jett-Fox Model
637 for each cell replicate. I) Quantification of cell cycling phases based on data from
638 experiments in three independent keratinocyte lines, represented in H and Supplemental Fig
639 S3. Bars represent mean with SEM, $n = 3$ independent keratinocyte cell lines, with $\geq 10,000$
640 cells per line. Comparisons performed by 1-way ANOVA followed by Dunnett's Multiple
641 Comparison Test. SCR1/2 = Scrambled short hairpin 1 or 2, shLNC1/2 = short hairpin RNA 1
642 or 2 targeting *PRANCR*.

643 **Figure 3.** *PRANCR* is required for proliferation and differentiation in stratified epidermis. A)
644 Hematoxylin & eosin staining (top row); immunofluorescence of the differentiation proteins
645 KRT10 and FLG (middle rows) and immunofluorescence of proliferation marker MKI67
646 (bottom row) in control and *PRANCR*-depleted (shLNC) epidermal tissue. Nuclei are stained
647 in blue (Hoechst 33342). Scale bars = 100 μm . B) Quantitation of epidermal thickness. Each
648 dot represents the average of three measurements per image at fixed positions. Error bars
649 represent mean with SD, $n = 8$ in control, $n = 10$ in *PRANCR* knockdown. Differences
650 evaluated using Student's *t*-test. C, D) KRT10 and FLG quantitation as a percentage of the
651 total fluorescence signal. Dots represent the average intensities measured from different
652 images taken for each tissue. Error bars represent mean with SD, $n = 4$ tissues in control, n
653 = 3 tissues in *PRANCR* knockdown. E) RNA expression in *PRANCR*-depleted epidermis vs
654 control. Bars represent mean with SEM, $n = 4$. F) MKI67 quantitation as a percentage of total
655 cells. Error bars represent mean with SD, $n = 8$ in control, $n = 10$ in *PRANCR* knockdown.

656 Differences evaluated using Student *t*-test. SCR = Scrambled short hairpin, shLNC1/2 =
657 short hairpin RNA 1 or 2 targeting *PRANCR*.

658 **Figure 4.** *PRANCR* regulates expression of cell cycle and MAPK cascade genes *in trans*. A)
659 RNA-seq analysis of *PRANCR*-depleted keratinocytes. RNA was harvested 72 hours after
660 transduction of shRNAs. Hierarchical clustering of differentially expressed genes (>2-fold
661 change) in *PRANCR*-depleted samples (shLNC, *n* = 4, biological replicates of two
662 independent *PRANCR*-targeting shRNAs) vs. control samples (SCR, *n* = 4, biological
663 replicates of two independent scrambled shRNAs). Gene Ontology (GO) terms of genes that
664 are B) downregulated or C) upregulated upon depletion of *PRANCR*.

665

666 **Figure 5.** *PRANCR* regulates expression of cell cycle genes containing the *CHR* promoter
667 element. A) Combined score of transcription factor enrichment of genes downregulated upon
668 *PRANCR* depletion using *Enrichr* gene set enrichment analysis. B) *PRANCR* mRNA
669 expression in cytoplasmic and nuclear cell fractions, measured by qRT-PCR, compared to
670 known cytoplasmic (*ACTB* and *GAPDH*) and nuclear RNAs (*NEAT1* and *MALAT1*), *n* = 4 cell
671 lines, bars represent mean with SEM. C) Normalized expression of E2F target genes. For
672 box plots, dots represent mean expression level (*n* = 4) of each gene, lines represent median
673 values of the gene set, box limits denote interquartile range, and whiskers extend 1.5-times
674 the interquartile range. Differences evaluated with Student's *t*-test. D) Motif enrichment of
675 downregulated DEGs using HOMER. E) Normalized expression of genes containing a *CHR*
676 promoter element. F) Normalized expression of E2F family members in primary keratinocytes
677 (*n*=4, mean with SD). G) Normalized expression of DREAM target genes. H) Normalized
678 expression of G₂/M genes. I) Western blot of TP53, E2F4, and CDKN1A in whole cell lysate,
679 cytoplasmic (Cyt), and nuclear (Nucl) fractions in control (SCR) and *PRANCR*-depleted cells
680 (shLNC1). LMNA/C (nuclear) and GAPDH (cytoplasmic) represent experimental fractionation
681 controls. Lower panel shows quantification of results by densitometry after normalization of

682 total protein levels and represent average of three biological replicates (Supplemental Fig
683 S6). Results evaluated with paired Student's *t*-test.

684

685 REFERENCES

686

687 Aasen T, Izpisua Belmonte JC. 2010. Isolation and cultivation of human keratinocytes from skin or
688 plucked hair for the generation of induced pluripotent stem cells. *Nat Protoc* **5**: 371-382.

689 Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biol* **11**:
690 R106.

691 Awan HM, Shah A, Rashid F, Shan G. 2017. Primate-specific Long Non-coding RNAs and MicroRNAs.
692 *Genomics Proteomics Bioinformatics* **15**: 187-195.

693 Barrandon Y, Green H. 1987. Three clonal types of keratinocyte with different capacities for
694 multiplication. *Proc Natl Acad Sci U S A* **84**: 2302-2306.

695 Bassett AR, Akhtar A, Barlow DP, Bird AP, Brockdorff N, Duboule D, Ephrussi A, Ferguson-Smith AC,
696 Gingeras TR, Haerty W et al. 2014. Considerations when investigating lncRNA function in
697 vivo. *Elife* **3**: e03058.

698 Boxer LD, Barajas B, Tao S, Zhang J, Khavari PA. 2014. ZNF750 interacts with KLF4 and RCOR1,
699 KDM1A, and CTBP1/2 chromatin regulators to repress epidermal progenitor genes and
700 induce differentiation genes. *Genes Dev* **28**: 2013-2026.

701 Bracken AP, Ciro M, Cocito A, Helin K. 2004. E2F target genes: unraveling the biology. *Trends Biochem*
702 *Sci* **29**: 409-417.

703 Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, Park J, Blackburn EH, Weissman JS, Qi
704 LS et al. 2013a. Dynamic imaging of genomic loci in living human cells by an optimized
705 CRISPR/Cas system. *Cell* **155**: 1479-1491.

706 Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A. 2013b. Enrichr:
707 interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*
708 **14**: 128.

709 Chen X, Muller GA, Quaas M, Fischer M, Han N, Stutchbury B, Sharrocks AD, Engeland K. 2013c. The
710 forkhead transcription factor FOXM1 controls cell cycle-dependent gene expression through
711 an atypical chromatin binding mechanism. *Mol Cell Biol* **33**: 227-236.

712 Da Sacco L, Baldassarre A, Masotti A. 2012. Bioinformatics tools and novel challenges in long non-
713 coding RNAs (lncRNAs) functional analysis. *Int J Mol Sci* **13**: 97-114.

714 Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A,
715 Knowles DG et al. 2012. The GENCODE v7 catalog of human long noncoding RNAs: analysis of
716 their gene structure, evolution, and expression. *Genome Res* **22**: 1775-1789.

717 Devgan V, Nguyen BC, Oh H, Dotto GP. 2006. p21WAF1/Cip1 suppresses keratinocyte differentiation
718 independently of the cell cycle through transcriptional up-regulation of the IGF-I gene. *J Biol*
719 *Chem* **281**: 30463-30470.

720 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013.
721 STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15-21.

722 Doolittle WF. 2013. Is junk DNA bunk? A critique of ENCODE. *Proc Natl Acad Sci U S A* **110**: 5294-
723 5300.

724 Engeland K. 2018. Cell cycle arrest through indirect transcriptional repression by p53: I have a
725 DREAM. *Cell Death Differ* **25**: 114-132.

726 Fischer M, Quaas M, Steiner L, Engeland K. 2016. The p53-p21-DREAM-CDE/CHR pathway regulates
727 G2/M cell cycle genes. *Nucleic Acids Res* **44**: 164-174.

- 728 Fischer M, Steiner L, Engeland K. 2014. The transcription factor p53: not a repressor, solely an
729 activator. *Cell Cycle* **13**: 3037-3058.
- 730 Gangatirkar P, Paquet-Fifield S, Li A, Rossi R, Kaur P. 2007. Establishment of 3D organotypic cultures
731 using human neonatal epidermal cells. *Nat Protoc* **2**: 178-186.
- 732 Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C, Panning B,
733 Ploegh HL, Bassik MC et al. 2014. Genome-Scale CRISPR-Mediated Control of Gene
734 Repression and Activation. *Cell* **159**: 647-661.
- 735 Guillaud P, Vermont J, Seigneurin D. 1991. Automatic classification of cells in cell cycle phases based
736 on Ki-67 antigen quantification by fluorescence microscopy. *Cell Prolif* **24**: 481-491.
- 737 Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010.
738 Simple combinations of lineage-determining transcription factors prime cis-regulatory
739 elements required for macrophage and B cell identities. *Mol Cell* **38**: 576-589.
- 740 Hon CC, Ramilowski JA, Harshbarger J, Bertin N, Rackham OJ, Gough J, Denisenko E, Schmeier S,
741 Poulsen TM, Severin J et al. 2017. An atlas of human long non-coding RNAs with accurate 5'
742 ends. *Nature* **543**: 199-204.
- 743 Ivanova IA, D'Souza SJ, Dagnino L. 2005. Signalling in the epidermis: the E2F cell cycle regulatory
744 pathway in epidermal morphogenesis, regeneration and transformation. *Int J Biol Sci* **1**: 87-
745 95.
- 746 Kellis M, Wold B, Snyder MP, Bernstein BE, Kundaje A, Marinov GK, Ward LD, Birney E, Crawford GE,
747 Dekker J et al. 2014. Defining functional DNA elements in the human genome. *Proc Natl Acad
748 Sci U S A* **111**: 6131-6138.
- 749 Kretz M, Sipsashvili Z, Chu C, Webster DE, Zehnder A, Qu K, Lee CS, Flockhart RJ, Groff AF, Chow J et
750 al. 2013. Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature*
751 **493**: 231-235.
- 752 Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik
753 KM, Lachmann A et al. 2016. Enrichr: a comprehensive gene set enrichment analysis web
754 server 2016 update. *Nucleic Acids Res* **44**: W90-97.
- 755 Lachmann A, Xu H, Krishnan J, Berger SI, Mazloom AR, Ma'ayan A. 2010. ChEA: transcription factor
756 regulation inferred from integrating genome-wide CHIP-X experiments. *Bioinformatics* **26**:
757 2438-2444.
- 758 Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, Irizarry RA, Liu JS, Brown M, Liu XS. 2014. MAGeCK
759 enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout
760 screens. *Genome Biol* **15**: 554.
- 761 Liu SJ, Horlbeck MA, Cho SW, Birk HS, Malatesta M, He D, Attenello FJ, Villalta JE, Cho MY, Chen Y et
762 al. 2017. CRISPRi-based genome-scale identification of functional long noncoding RNA loci in
763 human cells. *Science* **355**.
- 764 Long Y, Wang X, Youmans DT, Cech TR. 2017. How do lncRNAs regulate transcription? *Sci Adv* **3**:
765 eaao2110.
- 766 Lopez-Pajares V, Yan K, Zarnegar BJ, Jameson KL, Khavari PA. 2013. Genetic pathways in disorders of
767 epidermal differentiation. *Trends Genet* **29**: 31-40.
- 768 Magae J, Wu CL, Illenye S, Harlow E, Heintz NH. 1996. Nuclear localization of DP and E2F transcription
769 factors by heterodimeric partners and retinoblastoma protein family members. *J Cell Sci* **109**
770 (Pt 7): 1717-1726.
- 771 Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads.
772 *EMBnetjournal* **17**.
- 773 Missero C, Di Cunto F, Kiyokawa H, Koff A, Dotto GP. 1996. The absence of p21Cip1/WAF1 alters
774 keratinocyte growth and differentiation and promotes ras-tumor progression. *Genes Dev* **10**:
775 3065-3075.
- 776 Muller GA, Engeland K. 2010. The central role of CDE/CHR promoter elements in the regulation of cell
777 cycle-dependent gene transcription. *FEBS J* **277**: 877-893.

- 778 Muller GA, Wintsche A, Stangner K, Prohaska SJ, Stadler PF, Engeland K. 2014. The CHR site:
779 definition and genome-wide identification of a cell cycle transcriptional element. *Nucleic*
780 *Acids Res* **42**: 10331-10350.
- 781 O'Driscoll J, Muston GC, McGrath JA, Lam HM, Ashworth J, Christiano AM. 2002. A recurrent
782 mutation in the loricrin gene underlies the ichthyotic variant of Vohwinkel syndrome. *Clin*
783 *Exp Dermatol* **27**: 243-246.
- 784 Oh JW, Hsi TC, Guerrero-Juarez CF, Ramos R, Plikus MV. 2013. Organotypic skin culture. *J Invest*
785 *Dermatol* **133**: 1-4.
- 786 Paci P, Colombo T, Fiscon G, Gurtner A, Pavesi G, Farina L. 2017. SWIM: a computational tool to
787 unveiling crucial nodes in complex biological networks. *Sci Rep* **7**: 44797.
- 788 Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA. 2013. Repurposing CRISPR
789 as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**: 1173-
790 1183.
- 791 Quaas M, Muller GA, Engeland K. 2012. p53 can repress transcription of cell cycle genes through a
792 p21(WAF1/CIP1)-dependent switch from MMB to DREAM protein complex binding at CHR
793 promoter elements. *Cell Cycle* **11**: 4661-4672.
- 794 Quek XC, Thomson DW, Maag JL, Bartonicek N, Signal B, Clark MB, Gloss BS, Dinger ME. 2015.
795 lncRNADB v2.0: expanding the reference database for functional long noncoding RNAs.
796 *Nucleic Acids Res* **43**: D168-173.
- 797 Ridky TW, Chow JM, Wong DJ, Khavari PA. 2010. Invasive three-dimensional organotypic neoplasia
798 from multiple normal human epithelia. *Nat Med* **16**: 1450-1455.
- 799 Sadasivam S, Duan S, DeCaprio JA. 2012. The MuvB complex sequentially recruits B-Myb and FoxM1
800 to promote mitotic gene expression. *Genes Dev* **26**: 474-489.
- 801 Sarropoulos I, Marin R, Cardoso-Moreira M, Kaessmann H. 2019. Developmental dynamics of
802 lncRNAs across mammalian organs and species. *Nature* **571**: 510-514.
- 803 Scholl FA, Dumesic PA, Barragan DI, Harada K, Bissonauth V, Charron J, Khavari PA. 2007. Mek1/2
804 MAPK kinases are essential for Mammalian development, homeostasis, and Raf-induced
805 hyperplasia. *Dev Cell* **12**: 615-629.
- 806 Smith FJ, Irvine AD, Terron-Kwiatkowski A, Sandilands A, Campbell LE, Zhao Y, Liao H, Evans AT,
807 Goudie DR, Lewis-Jones S et al. 2006. Loss-of-function mutations in the gene encoding
808 filaggrin cause ichthyosis vulgaris. *Nat Genet* **38**: 337-342.
- 809 Sun BK, Boxer LD, Ransohoff JD, Sibrashvili Z, Qu K, Lopez-Pajares V, Hollmig ST, Khavari PA. 2015.
810 CALML5 is a ZNF750- and TINCR-induced protein that binds stratifin to regulate epidermal
811 differentiation. *Genes Dev* **29**: 2225-2230.
- 812 Tak YG, Farnham PJ. 2015. Making sense of GWAS: using epigenomics and genome engineering to
813 understand the functional relevance of SNPs in non-coding regions of the human genome.
814 *Epigenetics Chromatin* **8**: 57.
- 815 The ENCODE Project Consortium 2012. An integrated encyclopedia of DNA elements in the human
816 genome. *Nature*, **489**: 57-74.
- 817 Topley GI, Okuyama R, Gonzales JG, Conti C, Dotto GP. 1999. p21(WAF1/Cip1) functions as a
818 suppressor of malignant skin tumor formation and a determinant of keratinocyte stem-cell
819 potential. *Proc Natl Acad Sci U S A* **96**: 9089-9094.
- 820 Trapnell C, Pachter L, Salzberg SL. 2009. TopHat: discovering splice junctions with RNA-Seq.
821 *Bioinformatics* **25**: 1105-1111.
- 822 Tripathi S, Pohl MO, Zhou Y, Rodriguez-Frandsen A, Wang G, Stein DA, Moulton HM, DeJesus P, Che J,
823 Mulder LC et al. 2015. Meta- and Orthogonal Integration of Influenza "OMICs" Data Defines a
824 Role for UBR4 in Virus Budding. *Cell Host Microbe* **18**: 723-735.
- 825 Tuck AC, Natarajan KN, Rice GM, Borawski J, Mohn F, Rankova A, Flemr M, Wenger A, Nutiu R,
826 Teichmann S et al. 2018. Distinctive features of lincRNA gene expression suggest widespread
827 RNA-independent functions. *Life Sci Alliance* **1**: e201800124.
- 828 Ulitsky I, Bartel DP. 2013. lincRNAs: genomics, evolution, and mechanisms. *Cell* **154**: 26-46.

- 829 Virtanen M, Smith SK, Gedde-Dahl T, Jr., Vahlquist A, Bowden PE. 2003. Splice site and deletion
830 mutations in keratin (KRT1 and KRT10) genes: unusual phenotypic alterations in Scandinavian
831 patients with epidermolytic hyperkeratosis. *J Invest Dermatol* **121**: 1013-1020.
- 832 Wang Y, Zhu W, Levy DE. 2006. Nuclear and cytoplasmic mRNA quantification by SYBR green based
833 real-time RT-PCR. *Methods* **39**: 356-362.
- 834 Wilusz JE, Sunwoo H, Spector DL. 2009. Long noncoding RNAs: functional surprises from the RNA
835 world. *Genes Dev* **23**: 1494-1504.
- 836 Xu H, Xiao T, Chen CH, Li W, Meyer CA, Wu Q, Wu D, Cong L, Zhang F, Liu JS et al. 2015. Sequence
837 determinants of improved CRISPR sgRNA design. *Genome Res* **25**: 1147-1157.
- 838 Yang Y, Muzny DM, Reid JG, Bainbridge MN, Willis A, Ward PA, Braxton A, Beuten J, Xia F, Niu Z et al.
839 2013. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J*
840 *Med* **369**: 1502-1511.
- 841 Zhang F, Lupski JR. 2015. Non-coding genetic variants in human disease. *Hum Mol Genet* **24**: R102-
842 110.
- 843

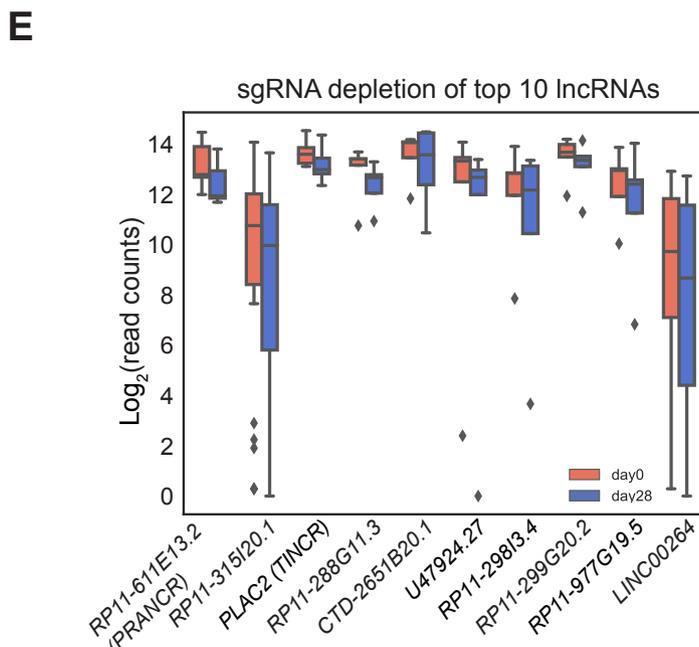
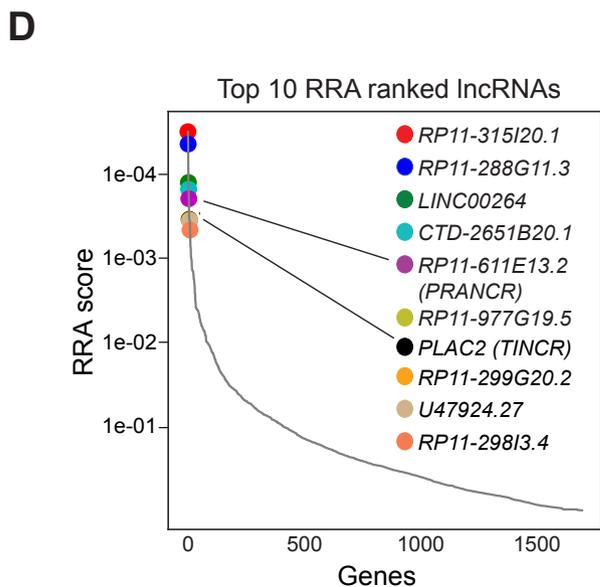
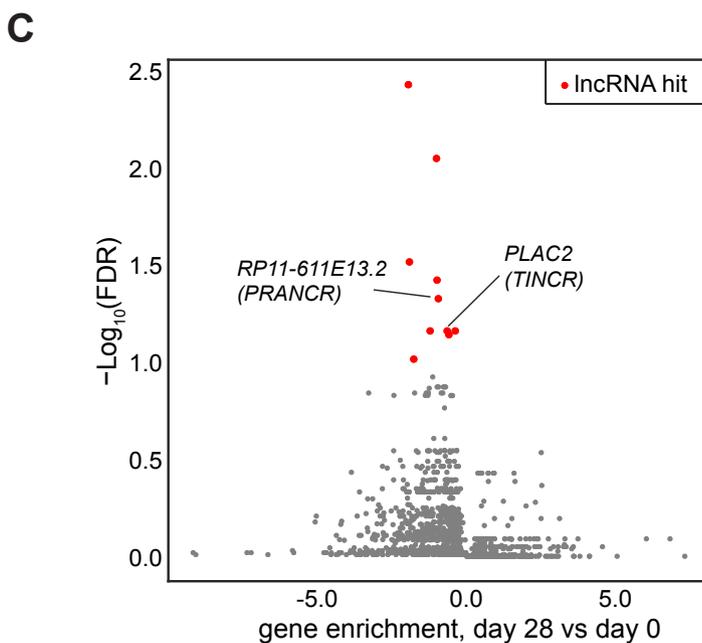
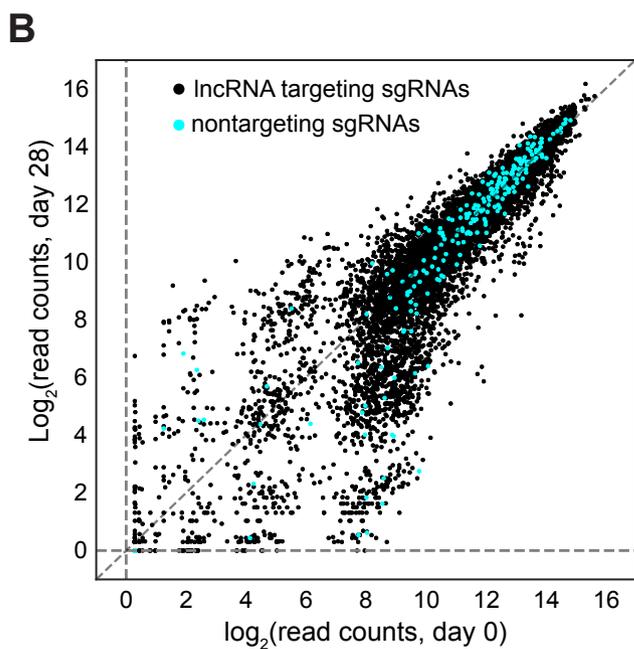
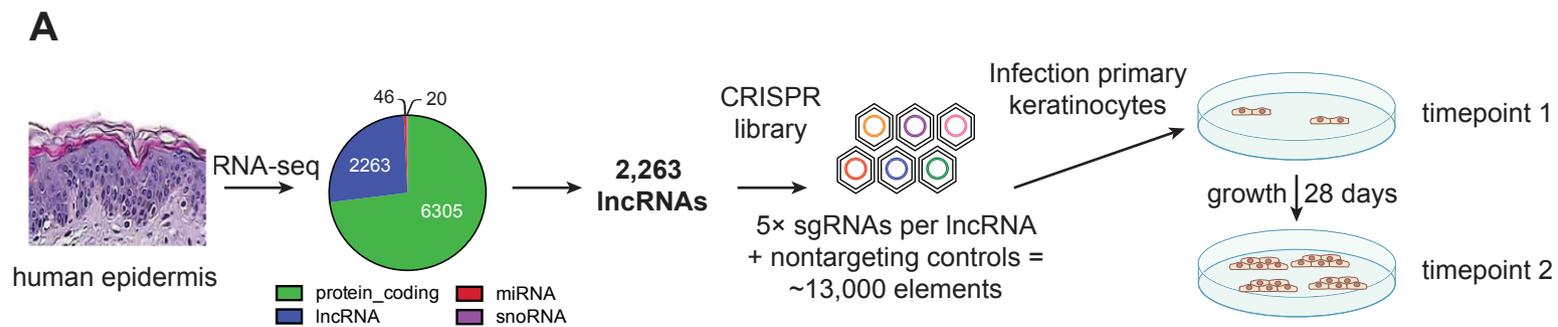


Figure 1

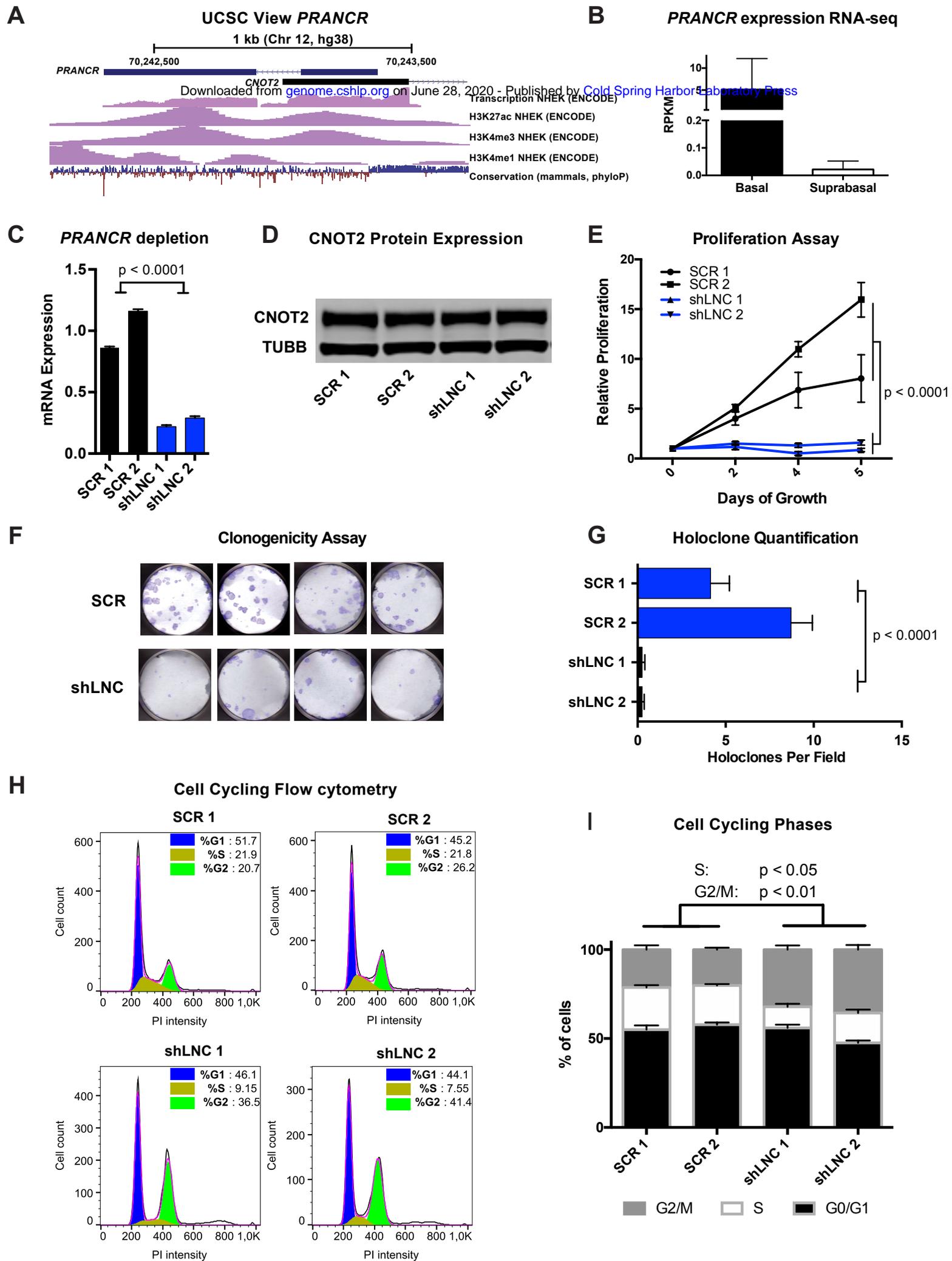
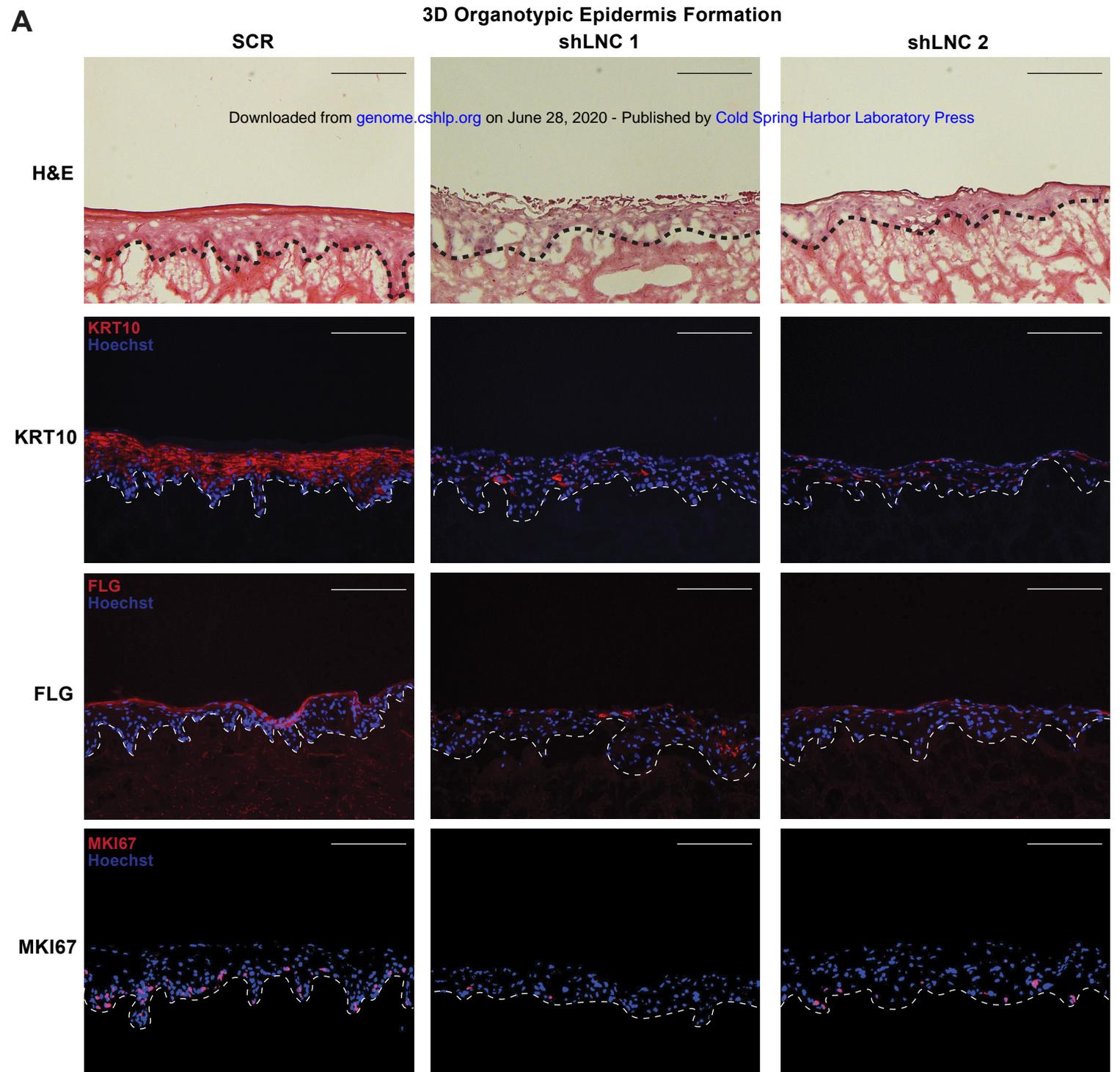


Figure 2



Downloaded from genome.cshlp.org on June 28, 2020 - Published by Cold Spring Harbor Laboratory Press

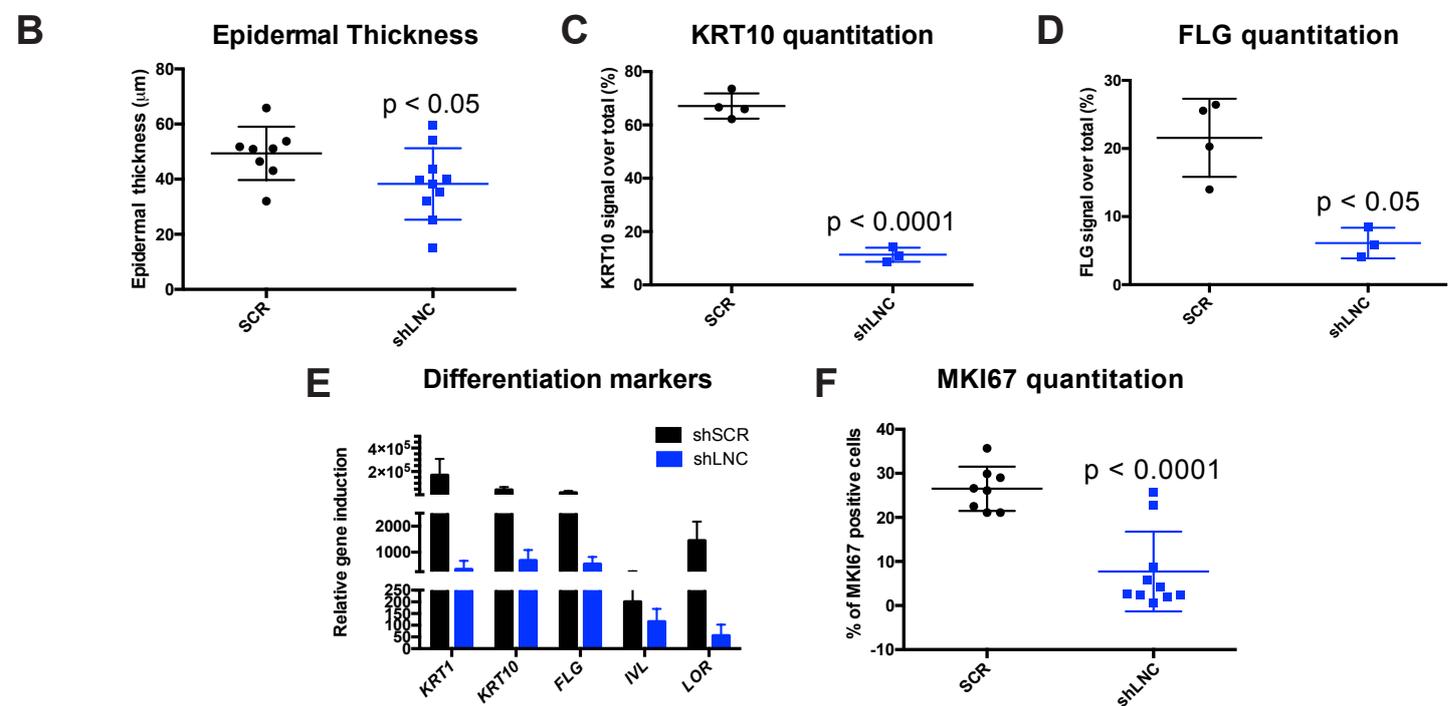


Figure 3

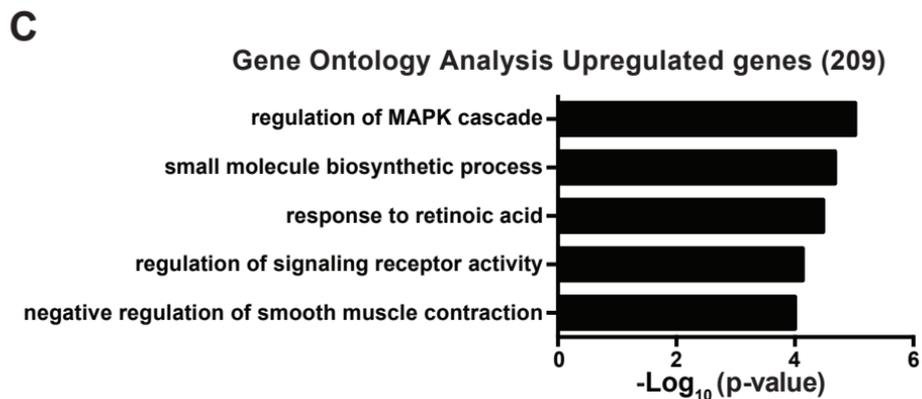
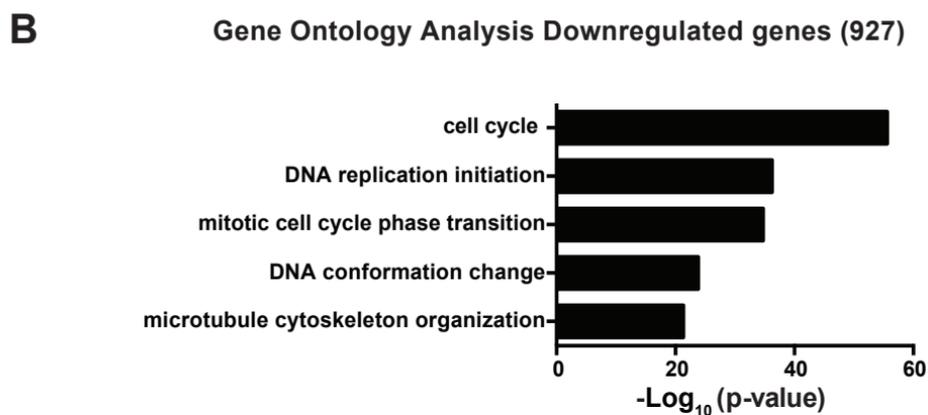
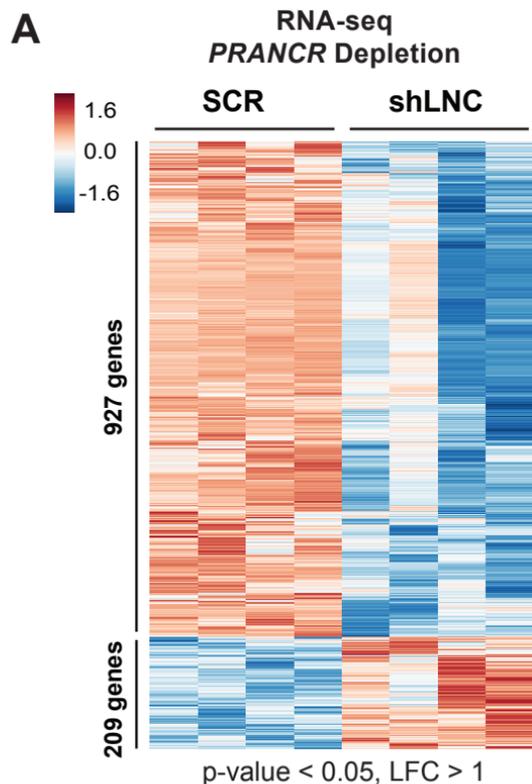


Figure 4

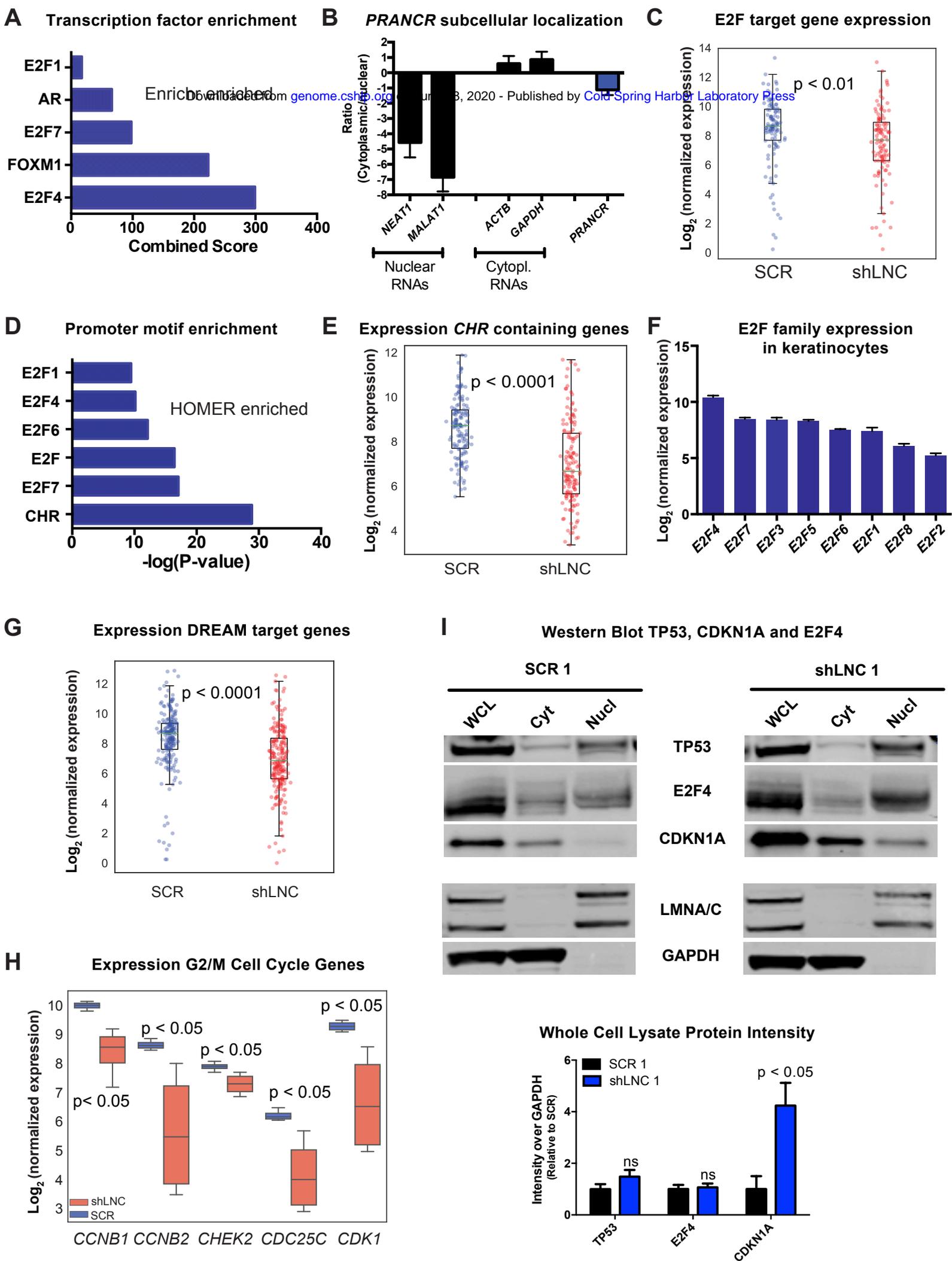


Figure 5



A genome-wide long noncoding RNA CRISPRi screen identifies *PRANCR* as a novel regulator of epidermal homeostasis

Pengfei Cai, Auke BC Otten, Binbin Cheng, et al.

Genome Res. published online December 4, 2019

Access the most recent version at doi:[10.1101/gr.251561.119](https://doi.org/10.1101/gr.251561.119)

Supplemental Material <http://genome.cshlp.org/content/suppl/2019/12/23/gr.251561.119.DC1>

P<P Published online December 4, 2019 in advance of the print journal.

Accepted Manuscript Peer-reviewed and accepted for publication but not copyedited or typeset; accepted manuscript is likely to differ from the final, published version.

Creative Commons License This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genome.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

A banner advertisement for ThruPLEX HV and Takara. The text "ThruPLEX® HV" is in large white font on a dark blue background, with "failproof DNA-seq of FFPE & cfDNA" below it. To the right is the Takara logo, which includes a stylized "T" in a circle and the word "Takara" in blue, with "Contech Wako cellartis" in smaller text below.

To subscribe to *Genome Research* go to:
<http://genome.cshlp.org/subscriptions>
