



Published in final edited form as:

Nat Cell Biol. 2016 January ; 18(1): 111–121. doi:10.1038/ncb3275.

Strand-specific *in vivo* screen of cancer-associated miRNAs unveils a role for miR-21* in SCC progression

Yejing Ge^{1,4}, Liang Zhang^{1,3,4}, Maria Nikolova¹, Boris Reva², and Elaine Fuchs^{1,5}

¹Howard Hughes Medical Institute and Laboratory of Mammalian Cell Biology and Development, Rockefeller University, New York, New York 10065, USA

²Department of Genetics and Genomic Sciences, Mt. Sinai School of Medicine, New York, New York 10029, USA

Abstract

MicroRNAs play diverse roles in both normal and malignant stem cells. Focusing on miRs and/or miR*’s abundant in squamous cell carcinoma (SCC) stem cells, we engineer an efficient, strand-specific expression library, and apply functional genomics screening in mice to identify which of 169 cancer-associated miRs are key drivers in malignant progression. Not previously linked functionally to cancer, miR-21* was the second top hit, surfacing in >12% of tumours. miR-21* also correlates with poor prognosis in human SCCs and enhances tumour progression in xenografts. On deleting the *miR-21* gene and rescuing each strand separately, we document the dual, but independent, oncogenicity of miR-21 and miR-21*. A cohort of predicted miR-21* targets inversely correlate with miR-21* in SCCs. Of particular interest is *Phactr4*, which we show is a miR-21* target in SCCs, acting through the Rb/E2F cell cycle axis. Through *in vivo* physiological miR screens, our findings add an interesting twist to an increasingly important oncomiR locus.

MicroRNAs (miRNAs or miRs) are evolutionally conserved, small non-coding RNAs that regulate gene expression at the post-transcriptional level. They guide an Argonaute-containing multiprotein complex to specific messenger RNAs by binding to partially complementary sequences in the 3′ untranslated regions (3′UTRs), thereby suppressing translation and/or inducing mRNA decay¹. Each miRNA gene gives rise to a stem-loop

Reprints and permissions information is available online at www.nature.com/reprints

⁵Correspondence should be addressed to E.F. (fuchslb@rockefeller.edu).

³Present addresses: SIBS (Institute of Health Sciences)—Changzheng Hospital Joint Center for Translational Research, Institutes for Translational Research (CAS-SMMU), Shanghai 200025, China; Key Laboratory of Stem Cell Biology, Institute of Health Sciences, SIBS, Chinese Academy of Sciences, Shanghai JiaoTong University School of Medicine, Shanghai 200025, China; Collaborative Innovation Center of Systems Biomedicine, Shanghai 200025, China (L.Z.).

⁴These authors contributed equally to this work.

Note: Supplementary Information is available in the online version of the paper

AUTHOR CONTRIBUTIONS

Y.G. and E.F. designed the experiments and wrote the manuscript. L.Z. designed and characterized the SA-miR vector. Y.G. and L.Z. performed the miR sequencing on FACS-sorted cells and made the lentiviral SA-miR constructs (Figs 1a and 2). Y.G. performed all other experiments and analyses (Figs 1b,c and 3–8). B.R. carried out the TCGA data analyses on miR-21 and miR-21* targets. M.N. assisted in CRISPR/CAS of *Phactr4*. All authors provided intellectual input, and vetted and approved the final manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

precursor, which on processing, generates miR-5p and miR-3p strands. Depending on their relative stabilities, they are conventionally termed guide (miR) or passenger (miR^{*}) strands. Although many miR^{*}s are degraded, some are stable, and because miR/miR^{*} harbour distinct seed sequences, each will target a largely non-overlapping mRNA cohort.

Known as key regulators of stem cell (SC) physiology and pathology, miRNAs change patterns markedly on cell fate alteration, including during malignant progression. As a cohort, miRNAs can aid in stratifying cancer subtypes and patient prognosis, rendering them attractive biomarkers². A few of these cancer-associated miRNAs act as functional drivers in tumour progression and/or maintenance^{3–5}. This is also true for squamous cell carcinomas (SCCs), life-threatening and metastatic cancers that occur frequently in stratified epithelia of the head and neck, oesophagus, lung, and skin, where miR-21, miR-203 and miR-125b have been shown to functionally impact tumorigenicity^{6–10}.

The complexity of differentially expressed miRNAs and their targets poses significant hurdles in evaluating not only their cause-versus-consequence roles in cancer progression, but also their relative degrees of potency in exerting their effects. Although individual oncomiRs have been characterized *in vivo*, so far, the functional significance of a cancer-associated miRNA pattern has not been interrogated in an unbiased fashion in a physiological context. Prerequisites to such analyses are first, high-throughput sequence analyses to elucidate the expression dynamics of miRs and miR^{*}s in tumour-initiating cells of the cancer; second, a robust, strand-specific miRNA expression platform compatible with functional genomics; and finally, an *in vivo* system for rapid functional screening of a large pool of relevant miRNAs in a particular cancer. Our current study explores these possibilities and performs a strand-specific *in vivo* screen of cancer-associated miRNAs to unveil key drivers and their oncogenic targets in SCCs.

RESULTS

In vivo miRNA landscapes of stem cells in homeostasis and SCCs

We began by performing miRNA deep sequencing on basal epithelial cells purified by fluorescence-activated cell sorting (FACS) from HRas-induced, pathology-diagnosed malignant SCCs, where basal cells (BCs) are known to be enriched for tumour-initiating potential^{11–13}. Similar analyses were carried out on adult and/or embryonic progenitors of normal epidermis and hair follicles (HFs; Supplementary Fig. 1a–c). Hierarchical clustering based on miRNA expression levels partitioned these populations into three main groups (Fig. 1a), exposing a dynamic miRNA landscape in SCCs versus normal SC-enriched populations.

One hundred and sixty-nine miR/miR^{*}s were abundantly expressed in SCCs. Differential expression analysis for sequence count data¹⁴ (DESeq) identified 97 of these miRs and 11 miR^{*}s that changed by $2\times$ ($P < 0.05$) in SCC-BCs relative to normal adult and/or embryonic counterparts (Fig. 1b and Supplementary Fig. 1d and Supplementary Table 1). A number of miRNAs previously implicated in epithelial cancers, including SCCs, were on our SCC signature^{6–10}. The physiological relevance of the rest of the >100 miR/miR^{*}s remained unexplored. Quantitative PCR (qPCR) and *in situ* hybridizations (ISH) validated and further

documented these patterns (Supplementary Fig. 1e,f). Intriguingly, some miRNAs were expressed in embryonic and normal skin, but declined during tumorigenesis; others were induced at the benign (papilloma) stage or became more prominent in invasive SCCs (Fig. 1c). These patterns underscored the remarkable responsiveness of cancer-associated miRNA genes to various stimuli throughout different stages of development and tumorigenesis.

***In vivo* screen identifies SCC-driving miRNAs**

Previous studies show that gene expression, hence probably miRNA expression, in SCs is highly sensitive to the niche microenvironment¹⁵. Assessing the functional relevance of this myriad of tumour-associated miRNAs thus necessitated an *in vivo* strategy. Compounding this hurdle is the need for a lentiviral miRNA-expression backbone to faithfully and efficiently express the unique miR or miR* strand independently of the other.

Termed small accurate (SA)-miR, our vector used an optimized, artificially engineered precursor backbone to express either side of the unprocessed stem-loop (Fig. 2a). SA-miRs conferred specific expression of the desired strand at ~2× endogenous levels (Fig. 2b,c), with high fidelity in the 5' end of the miRNA processing (Supplementary Fig. 2a,b). Reflective of these refinements, SA-miRs exhibited robust activity (Supplementary Fig. 2c).

With the expression tools and miRNA landscape in hand, we set out to perform a pooled *in vivo* functional screen to identify which of the 169 cancer-associated miRs or miR*s can drive skin SCCs. To this end, we first prepared an SA-miR lentiviral (LV) library of our cohort and 5 scrambled (Scr) controls (Supplementary Table 2). To control for any nonspecific effects from viral infections, we also made a LV-control library composed of only scrambled SA-miRs (SA-Scr).

We next employed our recently devised technology for LV delivery *in utero* into the amniotic sacs of E9.5 embryos¹⁶. This selectively and efficiently transduces single-layered surface epithelium, resulting in clonal expansion and expression in adult skin. Previously, we used this approach to identify protein-coding genes with oncogenic or tumour-suppressor activities^{17,18}. Here, we injected our SA-miR libraries at a multiplicity of infection (MOI) << 1 to ensure that no more than one SA-miR gene was stably integrated into the genome of each cell. To achieve robustness, we used a coverage so that ~100× different E9.5 cells were transduced with the same SA-miR. SA-miR representation was determined through DNA deep sequencing of epidermal progenitors at E12.5 and postnatal day 50 (P50; Fig. 3a).

Characteristic of most proteins with oncogenic potential, the cancer-associated miRNAs on their own did not exhibit signs of tumours by P50 (Fig. 3b). We therefore subjected animals transduced with SA-Scr and SA-miR libraries to chemical carcinogenesis with 7,12-dimethylbenz[a]anthracene (DMBA), which typically induces oncogenic Ras mutations^{10,19,20} and then 12-*O*-tetradecanoylphorbol 13-acetate (TPA), which leads to benign tumours within 2–4 months, some of which can later progress to SCC.

Indeed, for both SA-Scr- and SA-miR-transduced animals, tumours began to appear within this time frame (Fig. 3b). Notably, however, both tumour initiation and growth were markedly accelerated in mice transduced with the SA-miR library compared with the SA-

Scr (Fig. 3b,c). Moreover, whereas SA-Scr-transduced tumours exhibited smooth bordered invaginations characteristic of benign papillomas, SA-miR-transduced tumours exhibited features of malignant SCCs, including ill-defined borders, heightened mitoses and stromal invasion (Fig. 3c and Supplementary Fig. 3a).

Deep sequencing of integrated strand-specific SA-miR genes revealed evidence of clonal dominance in 79 of 107 independent tumours analysed (Supplementary Table 3). Twenty-two of these tumours were enriched for the established oncomiR, miR-21. Unexpectedly, however, was miR-21* as the second top hit, elevated in 13 tumours (Fig. 3d). This was not due to passenger strand contamination because miR-21 and miR-21* were separately encoded by their strand-specific SA-miR vectors (Fig. 2c). Notably, neither SA-miR-21 nor SA-miR-21* showed skewed representation in control cohorts (Fig. 3d). These findings provided compelling evidence that miR-21* harboured oncogenic activity independent of miR-21.

Another intriguing feature of our screen was the independent appearance of tumours that seemed to be fuelled by different members of particular miRNA families, which share a common seed sequence. In this regard, we observed tumour-specific enrichment of miR-10b/99b/100, miR-181a/b/c, miR-200a/141 and miR-200b/200c/429 families (Fig. 3d and Supplementary Fig. 3b). Many of these individual miRNAs have been previously implicated in different types of human cancer^{21–23}. As expected of oncomiRs, their clonal domination was not seen in controls (Fig. 3d).

To confirm that miRNAs surfacing in the screen are indeed induced and/or abundant during tumorigenesis, we performed ISH. Consistently, miR-21 and miR-21* expression was first induced in benign papillomas and further elevated in invasive SCCs; in contrast, miR-200a and miR-200b were detected in normal skin, and remained abundant in tumorigenesis (Fig. 3e). Their prevalent clonal selection during tumorigenesis and their high expression levels made these miRNAs attractive candidates that drive SCC progression.

Validate candidate oncomiRs in SCCs

To directly test whether miRs and miR*s that surfaced in our screen indeed drive SCCs in a physiological setting, we returned to our *in utero* infection procedures. We started by testing one well-established oncomiR (miR-21) along with three candidates, namely miR-21*, miR-200a and miR-200b, which had not been previously implicated in SCCs. This time, we subcloned these SA-miRs into a lentiviral backbone expressing a Cre recombinase (LV-Cre) and infected E9.5 embryos harbouring an inducible oncogenic HRas^{G12V} allele HRas^{lox-wild-type-stop-lox-G12V} (LSL-HRas^{G12V}; ref. 24) as well as a Cre-reporter allele Rosa26^{lox-stop-lox-YFP} (R26-LSL-YFP; Fig. 4a). In this regimen, transduced cells receiving miRNA should switch on YFP and simultaneously activate one HRas^{G12V} allele to initiate papillomas¹⁷.

Recapitulating the *in vivo* pooled screen yet with faster kinetics than the library, each individual SA-miR promoted tumour progression (Fig. 4a). These results underscored the efficacy of our approach. The independent effects of miR-21* were of particular interest, as most *in vivo* studies that have implicated miR-21 in cancer invariably activate or delete the

entire locus^{7,25,26}. A priori, because the seed sequences share no identity and hence the oncogenic potential of a miR* cannot be inferred from a miR, this observation raises the distinct possibility that miR-21* may be a contributing factor to some of the previously reported tumorigenic effects.

Importantly, although miR-21 has been reported to upregulate its expression through an auto-regulatory loop^{27,28}, SA-miR-21 did not directly induce miR-21* expression in our context, and vice versa (Supplementary Fig. 4a). In addition, the miR-21 and miR-21* probes shared little or no homology with each other or any of the other miRNAs on our SCC miR signature list. No background signal was observed in miR-21 knockout skin, thereby validating the specificity of *in situ* probes (Supplementary Fig. 4b). Given its complete absence in normal skin and marked induction in SCCs versus benign papillomas (Fig. 4b), we decided to further investigate the possible role of miR-21* in cancer.

Notably and in contrast to benign papillomas formed in LV-Cre-SA-Scr-transduced LSL-HRas^{G12V} animals, miR-21- and miR-21*-transduced tumours frequently exhibited irregular edges with discontinuous integrin β_4 immunolabelling (Fig. 4c). Other signs of invasiveness included the expansion of the basal marker keratin K5, loss of the differentiation marker K10, and an expanded domain of proliferation as measured by S-phase-specific incorporation of the nucleotide analogue EdU (Fig. 4c,d). In line with these *in vivo* observations, HF-SCs isolated from healthy mice transduced with either miR-21 or miR-21* showed enhanced colony formation efficiency in culture (Supplementary Fig. 4c). Moreover, under suspension conditions that normally halt proliferation and induce apoptosis (anoikis), primary mouse keratinocytes transduced with miR-21 or miR-21* showed protection, with the most significant effect observed from co-expression of both strands (Supplementary Fig. 4d).

To completely uncouple the two miRs, we first disrupted the endogenous miR-21/miR-21* locus in malignant mouse SCC cells using CRISPR/CAS technology²⁹, thereby obliterating expression of both miRs (Fig. 5a,b and Supplementary Fig. 5). As expected, ablating the locus (SCC) significantly crippled tumorigenic activity of parental SCCs (pSCCs) in grafting experiments (Fig. 5c,d), in line with the previously reported oncomiR-21 addiction phenotype in lymphoma²⁶. Notably, re-expressing miR-21* or miR-21 on its own partially rescued tumour initiation and growth as judged by limiting dilution assays and time course analyses, respectively (Fig. 5b–d). Additionally, expressing both miRs together restored full-blown SCC growth (Fig. 5d), arguing against off-target effects (see Methods). Interestingly, in a competition assay, when SCCs \pm SA-Scr, SA-miR-21 or SA-miR-21* was combined at equal ratios with pSCC + SA-Scr and then transplanted, pSCC cells remained the primary contributors to tumour growth (Fig. 5e), underscoring the cell-autonomous roles of miR-21 and miR-21*. Together, these findings provided compelling evidence that miR-21* drives tumorigenesis independent of miR-21, and both strands are essential for sustaining malignant growth.

miR-21* is a physiologically relevant oncomiR in human

Analysis of ISH and The Human Cancer Genome Atlas (TCGA) revealed high miR-21 and miR-21* in human head and neck SCCs (HNSCC) relative to neighbouring non-phenotypic

tissue (Fig. 6a,b). miR-21* was also elevated in other epithelial cancers, haematolymphoid malignancies and sarcomas (Fig. 6c). Although lower than miR-21, miR-21* levels are considerably higher than passenger strands of most established oncomiRs and tumour-abundant miRs (Supplementary Fig. 6a). Consistently, only miR-21*, and to a much lesser extent miR-615*, scored among all of the passenger strand miR*s in our library (Supplementary Table 3).

Interestingly, whereas miR-21 remained invariably high in human SCCs, miR-21* levels intriguingly varied across >2-log values in these cancers, despite their co-transcription and processing. Specifically, HNSCC patients with the highest miR-21* levels, but not with high miR-21, trended towards poorer prognosis (Fig. 6d,e and Supplementary Fig. 6b).

To test whether the elevated expression of miR-21* is physiologically relevant to human cancer as it is for mouse SCC, we applied our lentiviral SA-miR to human SCC (FaDu) cells as well as human immortalized epithelial (HaCaT) cells, a line that behaves more similarly to normal human epidermal keratinocytes. We then xenografted them into immunocompromised (Nude) mice. Interestingly, elevating miR-21* by ~4–5× accelerated tumour growth of FaDu cells over controls and induced the potential of HaCaT cells to undergo *de novo* tumorigenesis (Fig. 6f,g and Supplementary Fig. 6c). Notably, the miR-21* strand on its own was comparable to that of miR-21 in stimulating tumour growth even though both miRNAs required additional oncogenic hits to elicit their tumour-promoting effects. In addition, disrupting the miR-21 locus in human A431 SCCs by CRISPR/CAS suppressed their tumorigenesis in xenografts, suggesting a conserved function of this locus in sustaining malignant growth in human SCCs (Supplementary Fig. 6d–f).

Our focus on the *HRas* mutant background to examine miR-21* oncogenic potential is most relevant because 25% of human HNSCCs and 30% of lung SCC patients show alterations (mutation or amplification) in the RAS/MAPK pathway. It is possible that miR-21* interacts with additional oncogenic pathways, given that HaCaT cells harbour p53 mutations and A431 cells have EGFR amplifications. Although interesting, this goes beyond the scope of our current study.

miR-21* directly targets *Phactr4* and regulates Rb/E2F

To identify miR-21* targets that are responsible for its tumorigenic effects, we perused the TCGA database and found >7,500 mRNAs that are changed by >2× in HNSCCs versus their normal tissue counterparts (Supplementary Table 4). Applying TargetScan and <http://microRNA.org>^{30,31}, we identified 67 SCC-downregulated (>2×) genes as putative targets of miR-21* (Supplementary Table 5). Of these, nine showed significant ($P < 0.05$) inverse correlations with the levels of miR-21* across SCCs (Fig. 7a,b and Supplementary Fig. 7a–c, Supplementary Table 6). Similar analyses revealed six different putative targets of miR-21, including a previously reported miR-21 target in epithelial cancers³² (Supplementary Fig. 7a–c).

Given the seemingly poorer prognosis of TCGA HNSCC patients expressing the highest miR-21* levels, we concentrated our efforts on putative miR-21* targets. *PHACTR4* (phosphatase and actin regulator 4) was particularly intriguing in that it was not only

repressed in HNSCC patients with high miR-21^{*} levels, but also when diminished, significantly associated with poor prognosis among HNSCC patients (Fig. 7a,b). Notably, *PHACTR4* had previously surfaced in an *in vitro* screen for genes that antagonize the proliferation of immortalized mammary epithelial cells, and reduced growth of cancer lines in xenografts³³.

The predicted miR-21^{*}-binding site on the *Phactr4* 3' UTR is conserved between mouse and human (Fig. 7c). Moreover, miR-21^{*} suppressed the activity of a luciferase reporter containing the *Phactr4* 3' UTR as well as the wild type (wt), but not mutant (mut), version of the predicted miR-21^{*}-binding site (MBS; Fig. 7d). Consistent with *Phactr4* as a direct target of miR-21^{*}, the endogenous levels of *Phactr4* were diminished in primary mouse keratinocytes when cells were transduced with SA-miR-21^{*} (Fig. 7e and Supplementary Fig. 8a). This regulation on *Phactr4* was specific to miR-21^{*} and was not observed with miR-21. Finally, if *Phactr4* is a bona fide miR-21^{*} target, its levels should be de-repressed in our CRISPR/CAS-edited isogenic SCC line that specifically lacked miR-21^{*} (SCC+SA-miR-21). This was indeed the case (Fig. 7e). Mechanistically, the ability of miR-21^{*} to target *Phactr4* resulted in the phosphorylation and inactivation of the Rb tumour suppressor, which in turn led to upregulation of E2F/Rb targets and enhanced cell proliferation (Fig. 7f,g). It has been proposed that *PHACTR4* regulates protein phosphatase 1 (PP1) thereby controlling RB phosphorylation³⁴. We return to this point below.

The tumorigenic activity of miR-21^{*} is mediated by *Phactr4*

Finally, to test whether *Phactr4* is a physiological mediator of the tumorigenic activity of miR-21^{*} *in vivo*, we knocked down *Phactr4* in skin progenitors *in utero* with two independent short hairpin RNA (shRNA) hairpins, and then challenged the adult mice to HRAS-induced tumorigenesis. Notably, although neither *Phactr4* shRNA caused tumorigenesis on its own, both promoted HRAS-induced tumorigenesis, phenocopying miR-21^{*} overexpression *in vivo* (Fig. 8a). Remarkably, *Phactr4* knockdown was also sufficient to restore the tumorigenic properties of CRISPR/CAS-mediated miR-21^{*}-deficient cells in allografts (Fig. 8b).

To test whether the effects of *Phactr4* are truly linked to miR-21^{*} and significantly contribute to the activity of miR-21^{*} as an oncogenic driver in SCC, we re-expressed *Phactr4* lacking its 3' UTR and examined the consequences on miR-21^{*}-driven tumorigenesis. As shown in Fig. 8c, this miR-21^{*}-refractive *Phactr4* suppressed SA-miR-21^{*}-driven tumorigenesis in engraftment assays. To rule out confounding problems caused by the complete absence of the 3' UTR of *Phactr4*, we used CRISPR/CAS to delete only the miR-21^{*} target site within the endogenous *Phactr4* 3' UTR. Of significance, in the *Phactr4* mutant SCC cell clones, not only the level of *Phactr4* becomes resistant to miR-21^{*} regulation, but also these SCCs are no longer sensitive to the tumour-promoting effects of miR-21^{*} overexpression (Fig. 8d). Additionally, in line with the regulation of RB by *PHACTR4* via PP1, a mutant version of *PHACTR4* lacking its PP1-interacting domain³⁴, even though expressed at a comparable level to WT, largely lost its suppressive effects on the tumorigenic activity of miR-21^{*} (Fig. 8c and Supplementary Fig. 8b). Together, these pieces of evidence strongly corroborated *Phactr4* as a key and direct mediator of the tumorigenicity of miR-21^{*} in SCCs.

DISCUSSION

We began these studies by profiling the miRNAs that distinguish malignant SCC stem cells from their normal skin counterparts. We then built robust SA-miR tools and executed a functional assay in mice to screen 169 different cancer-associated miRNAs for the ones that contribute to the rapid progression from benign papillomas to malignant SCCs. This illustrates, to our knowledge, the first study that applies functional genomics to simultaneously interrogate a cohort of miRNAs *in vivo* for their relative functional importance.

Of the 169 miRNAs screened, only a few exhibited strong tumorigenic potential in our assay, suggesting that these miRNAs play special roles in malignant progression. Remarkably, 20% of the tumours analysed exhibited a strong bias towards miR-21. Finding miR-21, as well as miR-10b and miR-125b, in this handful of tumour-enhancing miRNAs demonstrated the robustness of our system, as these are well-established oncomiRs of epithelial cancers^{10,21,26}.

What was unexpected was the appearance of miR-21* and miR-200 family members as strand-specific drivers in malignancy. miR-21*, independently of miR-21, was especially potent with 12% of the tumours exhibiting clonal dominance for this miRNA. Moreover, although miR-21 is a well-established oncomiR with well-defined targets, relatively little was known about its passenger strand. Its appearance as a top hit in our screen took on all the more significance in that miR-21* also associated with poor patient prognosis in human SCCs.

By exploiting the TCGA database and applying both CRISPR/CAS and rapid lentiviral-based genetics to skin and head and neck epithelia in mice, we verified the miR-21-independent tumorigenic activity of miR-21* and uncovered *Phactr4* as a key target whose suppression confers much of the oncogenic activity of this hitherto ignored miRNA. *PHACTR4* is one of four members of a family of PP1- and actin-binding proteins. Although it has been implicated in integrin-mediated cell migration³⁴, it has only recently been implicated in cancer, where it surfaced in an *in vitro* RNAi screen for genes that restrain normal cell proliferation³³. Here, we substantiate in SCCs its implicated actions through PP1- and RB-dependent pathways, but most importantly, link its regulation at least in part to miR-21*.

In summary, we have unravelled the physiological relevance of a potent tumour-enhancing miRNA, miR-21*, that was masked within one of the most critical loci implicated in cancer. The physiological link to cancer between this little-studied miRNA and one of its targets, also poorly explored, merits further attention in the future. Overall, these findings underscore the importance of taking an unbiased *in vivo* screening approach to simultaneously and comparatively interrogate the functional significance of a complex signature of miRNAs, regardless of strand identity, that are associated by expression with a particular state, in this case malignant progression.

METHODS

Human HNSCCs and normal tissues

Human HNSCC and normal tissue samples from anonymous patients were obtained as discarded material after surgery and after obtaining informed consent from the patients, at the Memorial Sloan Kettering Institute and according to a protocol following NIH regulations and approved by the Institutional Review Boards of the MSKI and the Rockefeller University.

Mouse strains

Mice used in our experiments were on a C57BL/6 background, including wild type purchased from Harland, *Gt(ROSA)26Sor^{tm1(EYFP)Cos+}* (Jackson Laboratories, donated by A. McMahon, University of Southern California, USA) and *FR-Hras^{G12V}* (HRas^{lox-wild-type-stop-lox-G12V}; ref. 24), *Mir21a^{tm1Yoli/J}* (Jackson Laboratories)⁷. For *in utero* injections, pregnant females were used at E9.5. For the rest of the experiments, treatments on both males and females were started at adult age of P50 (second telogen).

Animal care and use

All animal experiments were performed in the AAALAC-accredited Comparative Bioscience Center at The Rockefeller University. Experiments were in accordance with NIH guidelines for Animal Care and Use, approved and overseen by The Rockefeller University's Institutional Animal Care and Use Committee.

Fluorescence-activated cell sorting (FACS) and analysis

Purification of embryonic progenitor populations was performed using K14–H2BGFP (E12, P0) and Lhx2–EGFP (E17) transgenic mice. Purification of adult populations was performed using K14–H2BGFP mice. Purification of tumour populations was performed using WT background, DMBA/TPA-treated tumour cell transplants (papilloma) or TGFβRII-deficient background, DMBA/TPA-treated tumour cell transplants (SCCs). Cell suspensions were from skins/tumours of transgenic animals expressing the basal progenitor marker *K14–H2BGFP* or the hair placode marker *Lhx2–GFP* (for E17 HF). Cells were first gated against DAPI to exclude dead cells, and then with forward and side scatters to gate for singlets. Lineage-negative gating was as follows: CD31 for endothelial cells, CD45 for immune cells, CD117 for melanocytes, and CD140a for fibroblasts. E12 epidermal (epi) progenitors were further gated as α6⁺K14–H2BGFP⁺, E17 epi progenitors α6⁺Lhx2–GFP⁺, E17 placodes α6⁺Lhx2–GFP⁺, P0 epi progenitors α6⁺K14–H2BGFP⁺Sca1⁺, P0 hair follicle (HF) progenitors α6⁺K14–H2BGFP⁺Sca1⁺, adult P28 epiSCs α6⁺K14–H2BGFP⁺Sca1⁺CD34⁺, HF-SC α6⁺K14–H2BGFP⁺Sca1⁺CD34⁺, outer root sheath (ORS) α6⁺K14–H2BGFP⁺Sca1⁺CD34⁺. Tumour grafts formed from previously characterized aggressive or non-aggressive tumour cells were gated as α6^{Hi}β1^{Hi} for basal stem cells (BCs) and α6^{Lo}β1^{Lo} for differentiated progenies. Whether cultured, or injected directly into recipient mice, these basal SCC cells have been previously shown to have tumour-initiating properties characteristic of cancer stem cells^{11–13}. Our findings here corroborate this behaviour as shown by the tumour-initiating assays in figures throughout the manuscript. For embryonic

cell isolation, the backskin was placed dermal side down in dispase (Sigma) for 2 h at 37 °C; for adult or tumour cell isolation, the backskin (dermis side down) or tumour (chopped) was placed in collagenase (Sigma, 0.25% in HBSS) for 1 h at 37 °C. The dermal fractions were collected by scraping the dermal side using a scalpel. The remaining epidermal side (embryo and adult) or cell mixture (tumour) was then transferred to trypsin (Gibco, 0.25% in PBS) at 37 °C for 10 min. Single-cell suspensions were obtained by scraping the skin (adult) or pipetting (embryo and tumour) gently. The cells were then filtered with 70 µm followed by 40 µm strainers, and pelleted at 300 g 4 °C. For FACS analysis, cells were first stained with Live/Dead Blue (Life Tech, 1:100) and Annexin-Pacific Blue (Life Tech, 1:100), and then fixed, permeabilized, stained with anti-GFP (1:1,000), anti-K5 (1:1,000) or anti-K10 (1:1,000), followed by Click-iT reactions and Alexa Fluor staining. For FACS sorting, cell suspensions were incubated with the appropriate antibodies for 30 min on ice. The following antibodies were used: CD34-eFluor660 (1:100, eBioscience 50-0341-80), α6-PE (1:100, BD Bioscience 551129), β1-Alexa647 (1:1,000, eBioscience 14-0291-81), Sca1-PerCP_Cy5.5 (1:1,000, eBioscience 45-5981-80), CD140a-PE_Cy7 (1:100, eBioscience 14-1401-81), CD31-PE_Cy7 (1:1,000, eBioscience 25-0311-81), CD117-PE_Cy7 (1:1,000, eBioscience 25-1171-82), CD45-APC_eFluor450 (1:1,000, eBioscience 48-0451-80). DAPI (0.2 µg ml⁻¹) was used to exclude dead cells. Sorting was performed on a BD FACSaria II equipped with Diva software (BD Biosciences). Analyses were performed on LSRII FACS analysers and FlowJo.

miRNA sequencing, hierarchical clustering and differential expression analysis

FACS-purified cell populations were lysed in TrizolLS (Invitrogen), and total cellular RNAs were isolated with Direct-zol RNA MiniPrep kit (Zymo Research) and submitted to Weill Cornell Medical College Genomics Resources Core Facility for IlluminaTruSeq Small RNA Sample Preparation and miRNA sequencing (HiSeq2000 Single Read, 51 Cycles). Raw reads are then mapped to miRBase v16 with Python. All 226 mouse-expressed miRNAs, using 100 RPM (reads per million miRNAs mapped) as the cutoff, were then input into Clustering and Treeview to generate an expression heatmap. The DESeq R package was used to generate lists of differentially expressed miRNAs. SCC-SC-expressed miRNAs were compared with either embryonic or adult SC populations. Dispersions (variability) for each miRNA were estimated using a local fit to the data for each sample, and miRNAs with a *P* value <0.05 by the negative binomial test were considered for downstream analysis. Differentially expressed miRNAs with fold change (FC) >2 (upregulated, UP) or <0.5 (downregulated, DOWN) were further filtered to retain conserved miRNAs between mouse and human, resulting in SCC-SC signature miRNAs.

miRNA real-time PCR and *in situ* hybridization

miRNA quantitative PCR was performed with TaqMan MicroRNA Assays (Life Tech) as per the manufacturer's protocol. U6 small RNA is used as a loading reference. Skin squamous cell cancer tissue array (SK802a) was obtained from US Biomax. miRNA *in situ* hybridizations in skin were performed as previously described³⁶. Briefly, frozen sections (mouse normal backskin or tumour) were fixed with 4% PFA in PBS for 20 min, or deparaffinized (human tissue array) with sequential ethanol washes, and subjected to 10 min acetylation followed by 10 min proteinase K (5 µg ml⁻¹) treatment. At 50 °C, sections were

pre-incubated with hybridization buffer for 4 h before incubation with miRNA probes (Exiqon miRCURY LNA Detection probe, 3' end DIG-labelled with Roche labelling kit) at 0.1 μ M overnight, and then washed with 5 \times SSC followed by 3 washes in 0.2 \times SSC for 30 min each. Sections were then equilibrated at room temperature with B1 buffer, blocked for 1 h, incubated with anti-DIG antibody (Roche 11333062910, 1:2,000) at 4 °C O/N, washed with B1 buffer followed by B3 buffer, and developed with BM Purple (Roche). Bright field images were acquired on a Leica Axioskop2 using a 10 \times /0.8 air objective.

***In vivo* pooled screen using SA-miRNA library**

The lentiviral SA-miRNA plasmids were constructed from the pLKO.1 backbone. Lentiviral LV-Cre-SA-miRNA plasmids were constructed using LV-Cre (ref. 16) and SA-miRNA. miRNAs that were more than twofold upregulated in SCC-SCs compared with either embryonic progenitors or adult SCs (SCC-SC UP signature, 75 cloned) or were expressed above the cutoff (100 RPM, 94 cloned) in any of the populations sequenced were cloned into the SA-miRNA vector, along with Scrambled control sequences (5 cloned). One hundred and seventy-four library clones (Supplementary Table 2) were then pooled and maxi-prepped DNA was used for higher-titre lentivirus production as previously described¹⁶. The library was titred to achieve a 15% infection rate measured by FACS analysis. The control library composed of SA-Scramble alone was similarly packaged and titred. At least 10 wild-type C57BL/6 embryos per library were transduced with a theoretical initial coverage = $(1.5 \times 10^5 \text{ surface ectoderm progenitors per embryo}) \times (15\% \text{ infection rate}) \times 10 \text{ embryos/174 library clones} = 129\text{-fold}$. Topical DMBA/TPA treatment was performed as previously described¹⁰. Briefly, P50 mice in second telogen were shaved and treated with 400 nM DMBA in 100 μ l acetone and thereafter, mice were treated with 17 nM TPA in 100 μ l acetone twice weekly for 20 weeks. Genomic DNA was isolated from backskin at E12.5 shortly post infection, at P50 tumour onset (starting DMBA/TPA treatment), or in the tumours formed as well as normal backskin at the time of tumour collection with the DNeasy Blood & Tissue Kit (Qiagen), and was used as a template in a 50 μ l pre-amplification reaction 30 cycles with 40 indexed SA-miRNA primers (Supplementary Table 7) and Phusion High-Fidelity DNA Polymerase (NEB). PCR products were run on a 2% agarose gel, and a clean ~200 bp band was isolated using Agencourt AMPure XP beads (beckmancoulter) and submitted to the Rockefeller University Genomics Resource Center for Illumina sequencing. Illumina reads were trimmed to the 22-nucleotide-long sequence matching the miRNA position using the FASTX-Toolkit and aligned to the SA-miRNA library with BWA (v0.6.2; ref. 37) using a maximum edit distance of 3 as previously described¹⁷. Tumours with dominant clones (defined as more than 75% reads mapped to the same miRNA-X) were scored as miRNA-X Tumour. In total 107 tumours were sequenced, with 79 tumours dominated by one or two miRNAs (Supplementary Table 3).

Tumour-free survival and tumour growth curves

Wild-type or Hras/Yfp animals were transduced at E9.5 with low-titre SA-miRNA library or LV-Cre-SA-miRNA, respectively. Transduced animals were monitored for 12 weeks, assessed every 2–3 days, and scored positive when tumours were larger than 2 mm in diameter. On P60, tumours were measured along their short and long axis using a digital calliper. Transplantation of mouse SCC or human cells transduced with SA-miRNAs into

immunocompromised nude recipients was performed as previously described¹⁰, and animals were monitored every 3 days for a month. Tumour size was measured using a digital calliper, and tumour volume was calculated using the formula (π (length \times width)²)/6. For human HNSCC patient data, level 3 miRNAseq data and clinical data were downloaded via DataPortal at The Cancer Genome Atlas (TCGA: <http://cancergenome.nih.gov>). GraphPad Prism software was used to generate the Kaplan–Meier curves and to calculate the *P* value for tumour-free survival (mouse) or survival (human) by two-tailed log-rank test.

Immunofluorescence microscopy and immunoblot analyses

The following primary antibodies were used: chicken anti-GFP (1:2,000; Abcam ab13970); guinea pig anti-K5 (1:500; E. Fuchs); rabbit anti-K10 (1:1,000; Covance PRB-159P), rat anti-integrin β_4 (eBioscience 346-11A, 1:1,000), rabbit anti-Phactr4 (Abcam ab192882, 1:1,000) and mouse anti- α -tubulin (Cell Signaling no. 2144, 1:10,000). Secondary antibodies were conjugated to Alexa488, 546, 647 (1:1,000, Life Technologies A-11006) or HRP (1:1,000, Life Technologies A10549). EdU was administered through intraperitoneal injection at 100 μ l (5 mg ml⁻¹) per 20 g mouse weight for 4 h before euthanasia. EdU was visualized by Click-iT EdU Alexa Fluor 647 Imaging Assay (Life Tech Molecular Probes). Tissues were processed as previously reported¹⁰, and mounted in ProLong Gold with DAPI (Life Tech). Images were captured on a Zeiss Axioplan2 using a Plan-Apochromat 20 \times /0.8 air objective, and processed using ImageJ and Adobe Photoshop CS5. Immunoblot was performed as previously described³⁸.

RNA sequencing and quantitative PCR

Cells were lysed with TrizolLS (Invitrogen) and total RNA was isolated with the Direct-zol RNA MiniPrep kit (Zymo Research) and submitted to the Genomics Resources Core Facility of the Weill Cornell Medical College for quality control (determined using Agilent 2100 Bioanalyzer, with all samples passing the quality threshold of RNA integrity numbers (RIN) 8), library construction using IlluminaTruSeq Stranded mRNA Sample Prep Kit, and sequencing using Illumina HiSeq2000. Results were analysed as previously described³⁹ using TopHat to initiate mapping and Cufflinks for transcript assembly and expression level estimation with mm9 genome assembly as the reference genome. For RNA quantitative PCR, complementary DNAs were generated from 1 μ g of total RNA using the SuperScript Vilo cDNA synthesis kit (Life Tech), diluted and used as templates for real-time PCR performed with the 7900HT Fast Real-Time PCR System (Applied Biosystems) and gene-specific primers listed in Supplementary Table 8. GAPDH was used as a loading reference.

Cell culture

No cell lines were found in the database of commonly misidentified cell lines maintained by ICLAC and NCBI Biosample. The cell lines were not authenticated. The mouse SCC-SC cell lines isolated from early stage and malignant SCCs were generated previously in the Fuchs' laboratory¹¹. The human HaCaT keratinocytes, A431 SCC and FaDu SCC cell lines were purchased from ATCC. Mouse HF-SCs were isolated and cultured as previously described³⁹. All human cells were cultured in E Medium (Rheinwald and Beckett 1980). Mouse SCC and HF-SC cells were cultured in E intermediate calcium medium (contains 300 μ M calcium); mouse epidermal keratinocytes were cultured in E low calcium medium

(contains 50 μM calcium). For colony formation assay, equal numbers of cells were plated, in triplicate, onto mitomycin C-treated dermal fibroblasts. After 14 days in culture, cells were fixed and stained with 1% (w/v) rhodamine B (Sigma). For suspension cultures, cells were trypsinized and plated into the Costar Clear Flat Bottom Ultra Low Attachment Multiple Well Plates in culture medium. Apoptosis was assayed by the Pacific Blue Annexin V/SYTOX AADvanced Apoptosis Kit (Life tech Molecular Probes). The nucleotide analogue EdU (10 μM) was added to cell culture media 2 h before assaying by the Click-iT EdU Alexa Fluor 647 Flow Cytometry Assay Kit (Life tech Molecular Probes).

CRISPR in SCC cells

sgRNAs against the miR-21 locus were selected (4 each for mouse and human) from the GeCKO library, or 2 sgRNAs designed against putative miR-21* targeting sites within the Phactr4 3'UTR (Supplementary Table 8) and cloned into lentiGuide-Puro vector²⁹. sgRNAs against Scrambled sequences were used as controls and done side-by-side with sgRNAs against target sites throughout the experiments to rule out phenotypic changes due to nonspecific editing. Cas9-Blast vector was first transfected into previously established SCC cells¹¹ and selected with blasticidin (10 $\mu\text{g ml}^{-1}$) to obtain stable SCC-Cas9 cells. Lenti-sgRNA-Puro vector was then transfected into SCC-Cas9 cells and selected with puromycin (3 $\mu\text{g ml}^{-1}$) to obtain stable knockout pools. Surveyor assay and real-time PCR were performed to screen for the most effective sgRNA, which was then used to isolate single colonies to generate isogenic knockout cell lines. Genomic DNAs from single clones were isolated, from which the targeted miR-21 locus was PCR amplified and Sanger sequenced to confirm editing. Three independent clones were analysed and one representative clone was shown. The top three predicted off-target sites were also sequenced to rule out off-target mutations (Supplementary Table 8).

Luciferase assays and transfections

Dual Luciferase reporter assays in cultured HEK293FT cells were conducted using the Dual-Glo Luciferase Assay System following the manufacturer's protocol. The oligonucleotides for cloning wild-type or mutant miRNA-binding sites (MBS) into the 3'UTR of the pmiRGLO (Promega) dual luciferase reporter are listed in Supplementary Table 8. Phactr4 3'UTR constructs were purchased from Origene (NM_001048183.1). Transfection was performed in a 96-well format with Effectene reagent (Qiagen), with a mixture of 5 ng Reporter plasmid + 100 ng SA-miRNA to overexpress miRNA.

Association between a gene's expression level and outcome: TCGA data analyses

To determine the potential association between a gene's expression level and survival in human HNSCC cancer, we used RNAseq tumour profiles and related clinical information available from The Cancer Genome Atlas data portal. The following algorithm was implemented. First, all tumours were sorted by a value of expression level assessed by the RSEM protocol (PMID: 21816040). Second, the sorted tumours were divided into two classes, as follows: the first k tumours with the lowest expression level were put in class 0 and the remaining $N - k$ tumours were put in class 1 (N is used for a number of tumours in a data set; all possible two-class separations were done for $k = 1, \dots, N - 1$). Third, for every class separation a difference in survival between two classes was estimated by a P value

computed using the log-rank test method (PMID 5910392). The Kaplan–Meier curves that illustrate the differences in survival and correspond to the lowest observed *P* values are plotted. For miR-21* association with patient prognosis and phactr4 level, all patients with available miR-21* read data were ranked from high to low according to miR-21* expression. The top 10% of the rank was taken as ‘high’ versus the bottom 10% taken as ‘low’.

Statistics and reproducibility

All experiments were independently repeated in the laboratory. Quantitative data were collected from experiments performed in at least triplicate, and expressed as mean \pm s.d. Differences between groups were assayed using repeated-measure ANOVA tests for tumour growth curves, log-rank (Mantel–Cox) tests for survival curves, and two-tailed Student *t*-tests for the rest of the analysis. Differences were considered to be significant when *P* < 0.05. No statistical method was used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Accession numbers

Data generated during the work have been deposited in NCBI’s Gene Expression Omnibus (Edgar *et al.*, 2002) and are accessible through GEO Series accession number GSE67900.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We especially thank D. Schramek for intellectual input into the design of the allograft assay. We also thank D. Schramek and A. Sandoel for advice regarding data presentation; J. Racelis for technical assistance; A. Aldeguer, L. Polak, J. Levorse, S. Hacker, M. Sribour, D. Oristian and N. Stokes for assistance with mouse handling and experiments; and Z. Shen, N. Oshimori, S. Luo, K. Lay, M. Kadaja, B. Keyes, A. Asare, M. Laurin, R. Adam and A. Kulukian for helpful discussions. We thank the Comparative Bioscience Center (AAALAC accredited) for care of mice in accordance with National Institutes of Health (NIH) guidelines; Rockefeller Genomics Resource Center (C. Zhao, Director) and Weill Cornell Genomics Resource Center (J. Xiang, Director) for sequencing; Flow Cytometry facility (S. Mazel, Director) for FACS sorting. E.F. is an Investigator of the Howard Hughes Medical Institute. Y.G. is the recipient of a Women & Science Postdoctoral Fellowship, and a Department of Defense Breast Cancer Postdoctoral Fellowship.

References

1. Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell*. 2009; 136:215–233. [PubMed: 19167326]
2. Lu J, et al. MicroRNA expression profiles classify human cancers. *Nature*. 2005; 435:834–838. [PubMed: 15944708]
3. Lujambio A, Lowe SW. The microcosmos of cancer. *Nature*. 2012; 482:347–355. [PubMed: 22337054]
4. Di Leva G, Garofalo M, Croce CM. MicroRNAs in cancer. *Annu Rev Pathol*. 2014; 9:287–314. [PubMed: 24079833]
5. Adams BD, Kasinski AL, Slack FJ. Aberrant regulation and function of MicroRNAs in cancer. *Curr Biol*. 2014; 24:R762–R776. [PubMed: 25137592]

6. Darido C, et al. Targeting of the tumor suppressor GRHL3 by a miR-21-dependent proto-oncogenic network results in PTEN loss and tumorigenesis. *Cancer Cell*. 2011; 20:635–648. [PubMed: 22094257]
7. Ma X, et al. Loss of the miR-21 allele elevates the expression of its target genes and reduces tumorigenesis. *Proc Natl Acad Sci USA*. 2011; 108:10144–10149. [PubMed: 21646541]
8. Bhandari A, et al. The Grainyhead transcription factor Grhl3/Get1 suppresses miR-21 expression and tumorigenesis in skin: modulation of the miR-21 target MSH2 by RNA-binding protein DND1. *Oncogene*. 2013; 32:1497–1507. [PubMed: 22614019]
9. Benaich N, et al. Rewiring of an epithelial differentiation factor, miR-203, to inhibit human squamous cell carcinoma metastasis. *Cell Rep*. 2014; 9:104–117. [PubMed: 25284788]
10. Zhang L, Ge Y, Fuchs E. miR-125b can enhance skin tumor initiation and promote malignant progression by repressing differentiation and prolonging cell survival. *Genes Dev*. 2014; 28:2532–2546. [PubMed: 25403182]
11. Schober M, Fuchs E. Tumor-initiating stem cells of squamous cell carcinomas and their control by TGF- β and integrin/focal adhesion kinase (FAK) signaling. *Proc Natl Acad Sci USA*. 2011; 108:10544–10549. [PubMed: 21670270]
12. Lapouge G, et al. Skin squamous cell carcinoma propagating cells increase with tumour progression and invasiveness. *Embo J*. 2012; 31:4563–4575. [PubMed: 23188079]
13. Oshimori N, Oristian D, Fuchs E. TGF- β promotes heterogeneity and drug resistance in squamous cell carcinoma. *Cell*. 2015; 160:963–976. [PubMed: 25723170]
14. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol*. 2010; 11:R106. [PubMed: 20979621]
15. Adam RC, et al. Pioneer factors govern super-enhancer dynamics in stem cell plasticity and lineage choice. *Nature*. 2015; 521:366–370. [PubMed: 25799994]
16. Beronja S, Livshits G, Williams S, Fuchs E. Rapid functional dissection of genetic networks via tissue-specific transduction and RNAi in mouse embryos. *Nat Med*. 2010; 16:821–827. [PubMed: 20526348]
17. Beronja S, et al. RNAi screens in mice identify physiological regulators of oncogenic growth. *Nature*. 2013; 501:185–190. [PubMed: 23945586]
18. Schramek D, et al. Direct *in vivo* RNAi screen unveils myosin IIa as a tumor suppressor of squamous cell carcinomas. *Science*. 2014; 343:309–313. [PubMed: 24436421]
19. Quintanilla M, Brown K, Ramsden M, Balmain A. Carcinogen-specific mutation and amplification of Ha-ras during mouse skin carcinogenesis. *Nature*. 1986; 322:78–80. [PubMed: 3014349]
20. Nassar D, Latil M, Boeckx B, Lambrechts D, Blanpain C. Genomic landscape of carcinogen-induced and genetically induced mouse skin squamous cell carcinoma. *Nat Med*. 2015; 21:946–954. [PubMed: 26168291]
21. Ma L, Teruya-Feldstein J, Weinberg RA. Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. *Nature*. 2007; 449:682–688. [PubMed: 17898713]
22. Gregory PA, et al. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat Cell Biol*. 2008; 10:593–601. [PubMed: 18376396]
23. Taylor MA, Sossey-Alaoui K, Thompson CL, Danielpour D, Schiemann WP. TGF- β upregulates miR-181a expression to promote breast cancer metastasis. *J Clin Invest*. 2013; 123:150–163. [PubMed: 23241956]
24. Chen X, et al. Endogenous expression of Hras(G12V) induces developmental defects and neoplasms with copy number imbalances of the oncogene. *Proc Natl Acad Sci USA*. 2009; 106:7979–7984. [PubMed: 19416908]
25. Hatley ME, et al. Modulation of K-Ras-dependent lung tumorigenesis by microRNA-21. *Cancer Cell*. 2010; 18:282–293. [PubMed: 20832755]
26. Medina PP, Nolde M, Slack FJ. OncomiR addiction in an *in vivo* model of microRNA-21-induced pre-B-cell lymphoma. *Nature*. 2010; 467:86–90. [PubMed: 20693987]
27. Fujita S, et al. miR-21 gene expression triggered by AP-1 is sustained through a double-negative feedback mechanism. *J Mol Biol*. 2008; 378:492–504. [PubMed: 18384814]

28. Talotta F, et al. An autoregulatory loop mediated by miR-21 and PDCD4 controls the AP-1 activity in RAS transformation. *Oncogene*. 2009; 28:73–84. [PubMed: 18850008]
29. Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods*. 2014; 11:783–784. [PubMed: 25075903]
30. John B, et al. Human MicroRNA targets. *PLoS Biol*. 2004; 2:e363. [PubMed: 15502875]
31. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*. 2005; 120:15–20. [PubMed: 15652477]
32. Frankel LB, et al. Programmed cell death 4 (PDCD4) is an important functional target of the microRNA miR-21 in breast cancer cells. *J Biol Chem*. 2008; 283:1026–1033. [PubMed: 17991735]
33. Solimini NL, et al. STOP gene Phactr4 is a tumor suppressor. *Proc Natl Acad Sci USA*. 2013; 110:E407–414. [PubMed: 23319639]
34. Kim TH, Goodman J, Anderson KV, Niswander L. Phactr4 regulates neural tube and optic fissure closure by controlling PP1-, Rb-, and E2F1-regulated cell-cycle progression. *Dev Cell*. 2007; 13:87–102. [PubMed: 17609112]
35. Pink RC, et al. The passenger strand, miR-21-3p, plays a role in mediating cisplatin resistance in ovarian cancer cells. *Gynecol Oncol*. 2015; 137:143–151. [PubMed: 25579119]
36. Yi R, et al. Morphogenesis in skin is governed by discrete sets of differentially expressed microRNAs. *Nat Genet*. 2006; 38:356–362. [PubMed: 16462742]
37. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2009; 25:1754–1760. [PubMed: 19451168]
38. Ge Y, Sun Y, Chen J. IGF-II is regulated by microRNA-125b in skeletal myogenesis. *J Cell Biol*. 2011; 192:69–81. [PubMed: 21200031]
39. Keyes BE, et al. Nfatc1 orchestrates aging in hair follicle stem cells. *Proc Natl Acad Sci USA*. 2013; 110:E4950–E4959. [PubMed: 24282298]

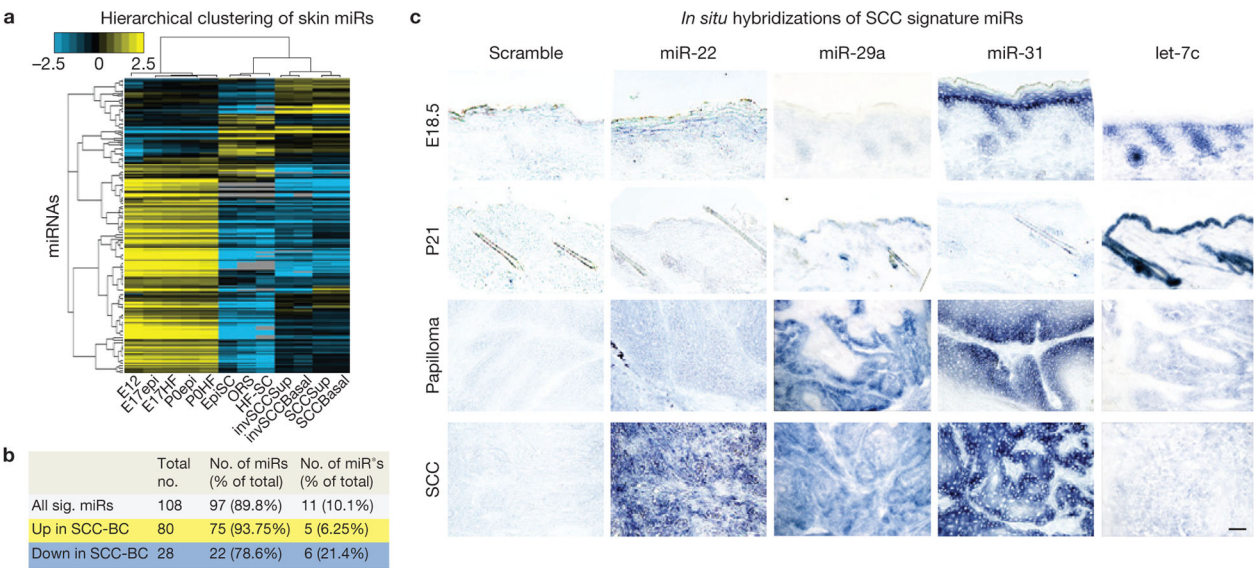
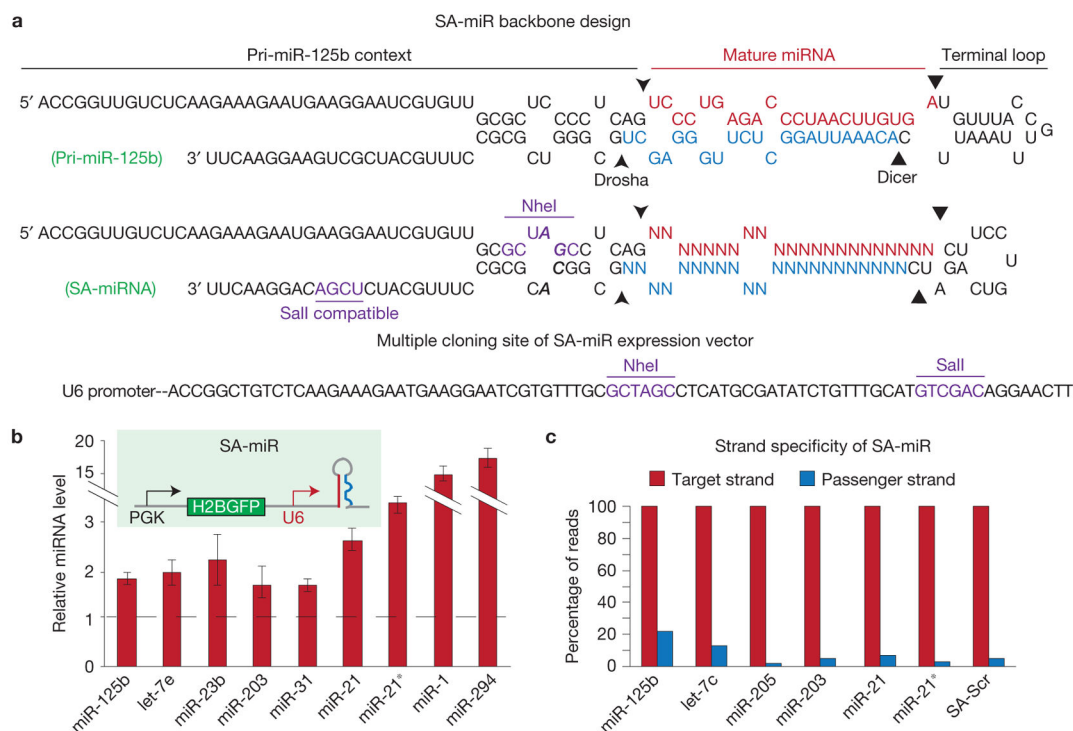
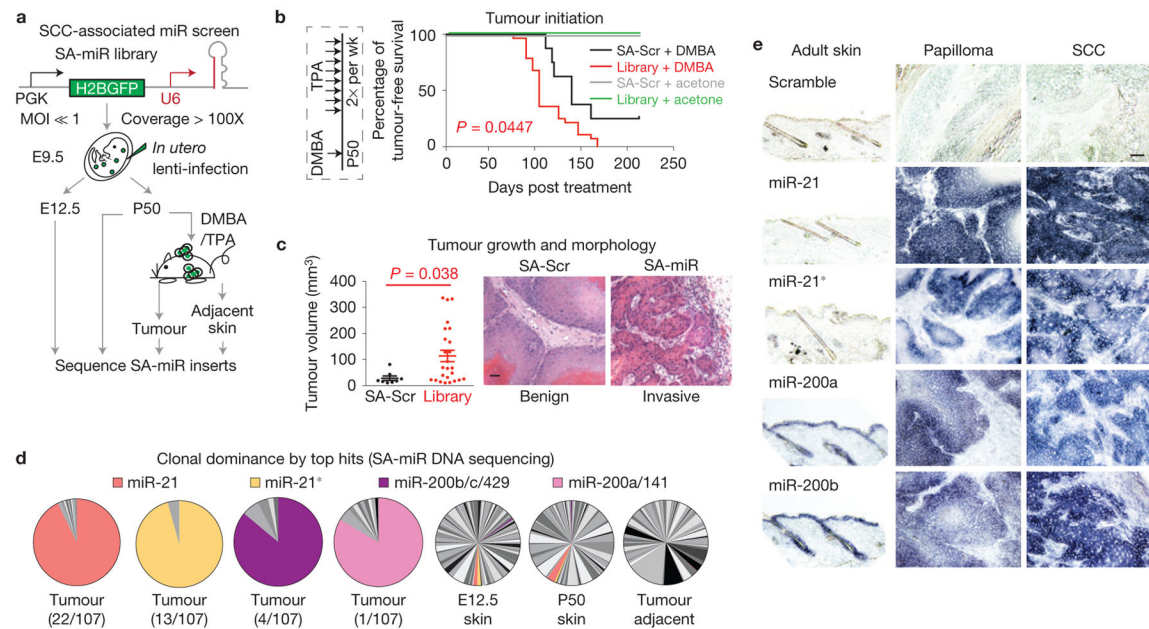


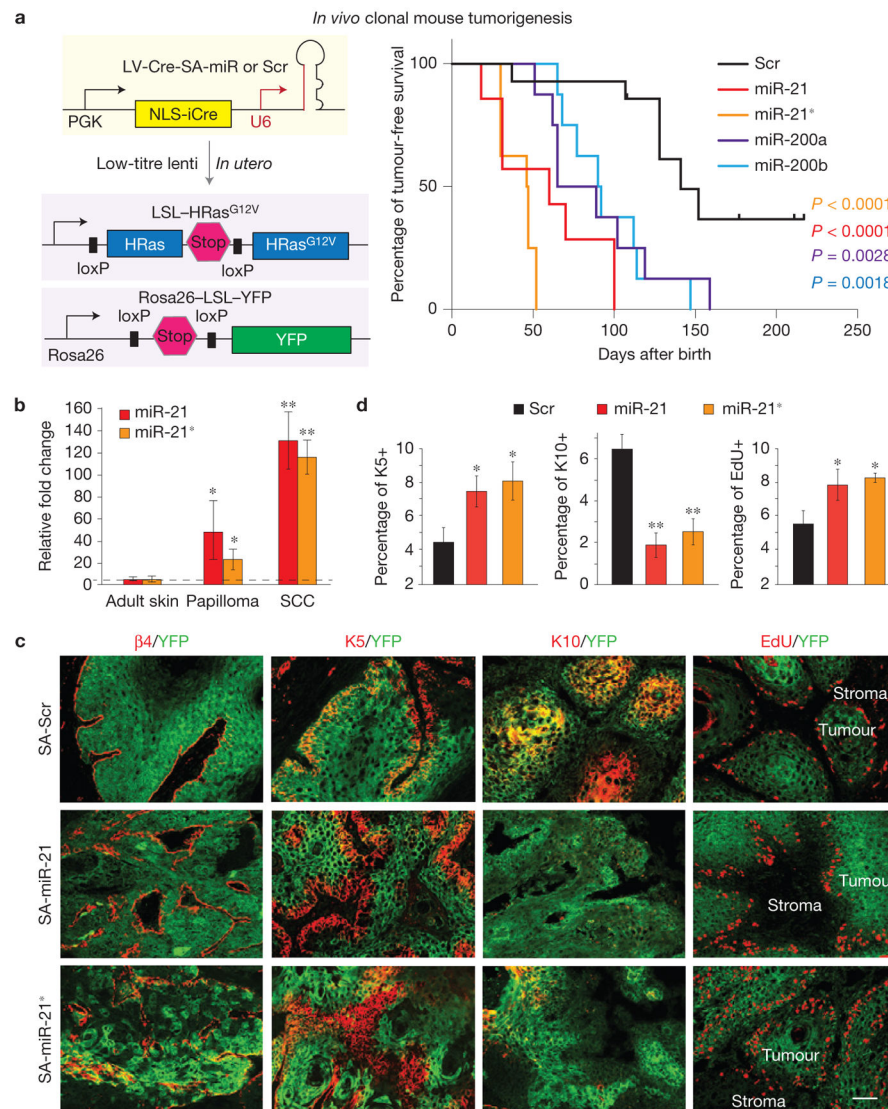
Figure 1. miRNA profiling reveals complex *in vivo* miRNA landscapes in stem cells under homeostasis and tumorigenesis. **(a)** Hierarchical clustering and heatmap showing dynamic miRNA patterns in FACS-purified stem and progenitor cells within the developing epithelium and tumour. Normalized reads per million mapped miRNAs (RPMM) are mean centred and log₂ transformed. miR patterns of epidermal (epi) and hair follicle (HF) progenitors from embryonic E12, E17 and neonates P0 cluster tightly. Adult HF-SCs and their immediate progenies (outer root sheath, ORS) and epidermal stem cells (epiSCs) form a second cluster. Basal progenitors and their differentiated suprabasal (Sup) progenies of early or invasive (inv) stage squamous cell carcinoma (SCC) form a third cluster. **(b)** Summary of the numbers of SCC signature miRs and miR*s defined by DESeq. **(c)** *In situ* hybridizations illustrate miRNA dynamics during malignant transformation. At least three biological replicates were performed; shown are representative images. Scale bar, 50 μ m.

**Figure 2.**

In vivo tool small accurate (SA)-miR for efficient and faithful miRNA expression. **(a)** Pri-miR-125b was used as a template to design the SA-miRNA backbone, with the pre-miR terminal loop shortened and an NheI site added for rapid cloning. Asymmetric thermal stabilities were introduced at the two ends to increase the probability of mature strand selection. The sequence of the multiple cloning site of the SA-miRNA expression vector following the U6 promoter is shown; SA-miRNA also carries a H2BGFP cassette driven by the PGK promoter to be used as an internal marker. **(b)** SA-miRs were transfected into cultured mouse keratinocytes and qPCR was used to quantify their overexpression relative to endogenous levels (black dashed line, SA-scrambled-sequence (SA-Scr)-transduced). In the cases of miR-1 and miR-294, the expression is unusually high owing to the extremely low level of endogenous miRs. At least five biological replicates were performed; shown are mean \pm s.d. $n = 3$ independent experiments. **(c)** Mouse SCC cells were exposed to lentiviral-based SA-miRs harbouring a GFP transgene, sorted for GFP 2 d later, and subjected to deep sequencing. The relative amount of undesired passenger strand as a percentage of targeted guide strand (set as 100%) is plotted to demonstrate preferential guide strand selection with SA-miRNAs. The reads for each miR had been normalized with that from SA-Scr-transduced cells (which reflect endogenous miR levels) and are presented as sole contributions from the SA-miR vector expression. A different Scrambled sequence was used from the SA-Scr shown in the graph.

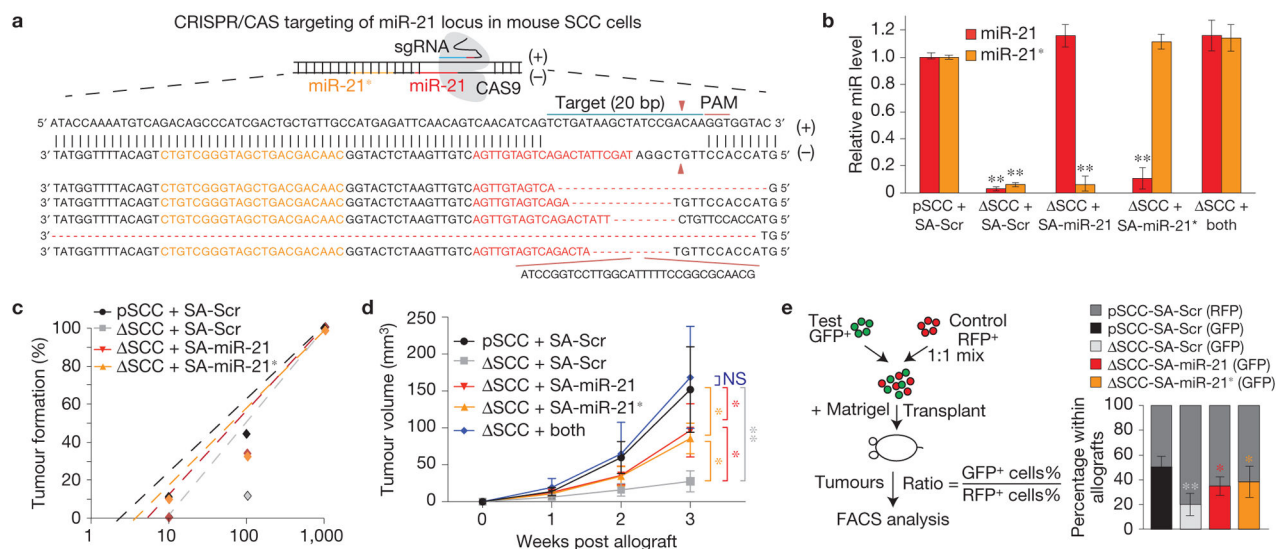
**Figure 3.**

In vivo pooled screen with the SA-miRNA library identifies candidate oncomiRs in skin SCCs. **(a)** Schematic of our SA-miRNA and SA-Scr library screens. For the SA-miR library, at least ten E9.5 embryos were used to ensure >100× coverage. The library was infected at MOI < 1 to minimize double transduction events per cell. Tumours were collected after reaching about 1–1.5 cm diameter. SA-miR representation was determined by Illumina sequencing. **(b,c)** Animals transduced with the SA-miR library exhibited accelerated tumour onset **(b)**, growth rate **(c, left panel)** and invasive morphology **(c, right panel)** compared with controls. Log-rank Mantel–Cox test was used for **b**. $n = 8$ mice for SA-Scr cohort. $n = 25$ mice for SA-miR cohort. Unpaired *t*-test was used for **c** left panel. For histology of SA-Scr and SA-miR tumours **(c, right panel)**, at least five biological replicates were performed; shown are representative images. Scale bar, 50 μ m. **(d)** Raw SA-miR DNA sequencing reads were mapped and the relative percentage of a particular miRNA read versus the total reads was calculated. In parentheses are the number of tumours with clonal dominance of the particular miRNA (defined as more than 75% reads mapped to the same miR), out of 107 total sequenced tumours. Each grey slice represents the rest of the mapped individual miRs not scored as dominant. **(e)** *In situ* hybridization shows that miR-21 and miR-21* are elevated in benign tumours and markedly upregulated in malignant tumours in mouse. miR-200a and miR-200b are expressed in normal skin and abundant in tumours. At least three biological replicates were performed; shown are representative images. Scale bar, 50 μ m.

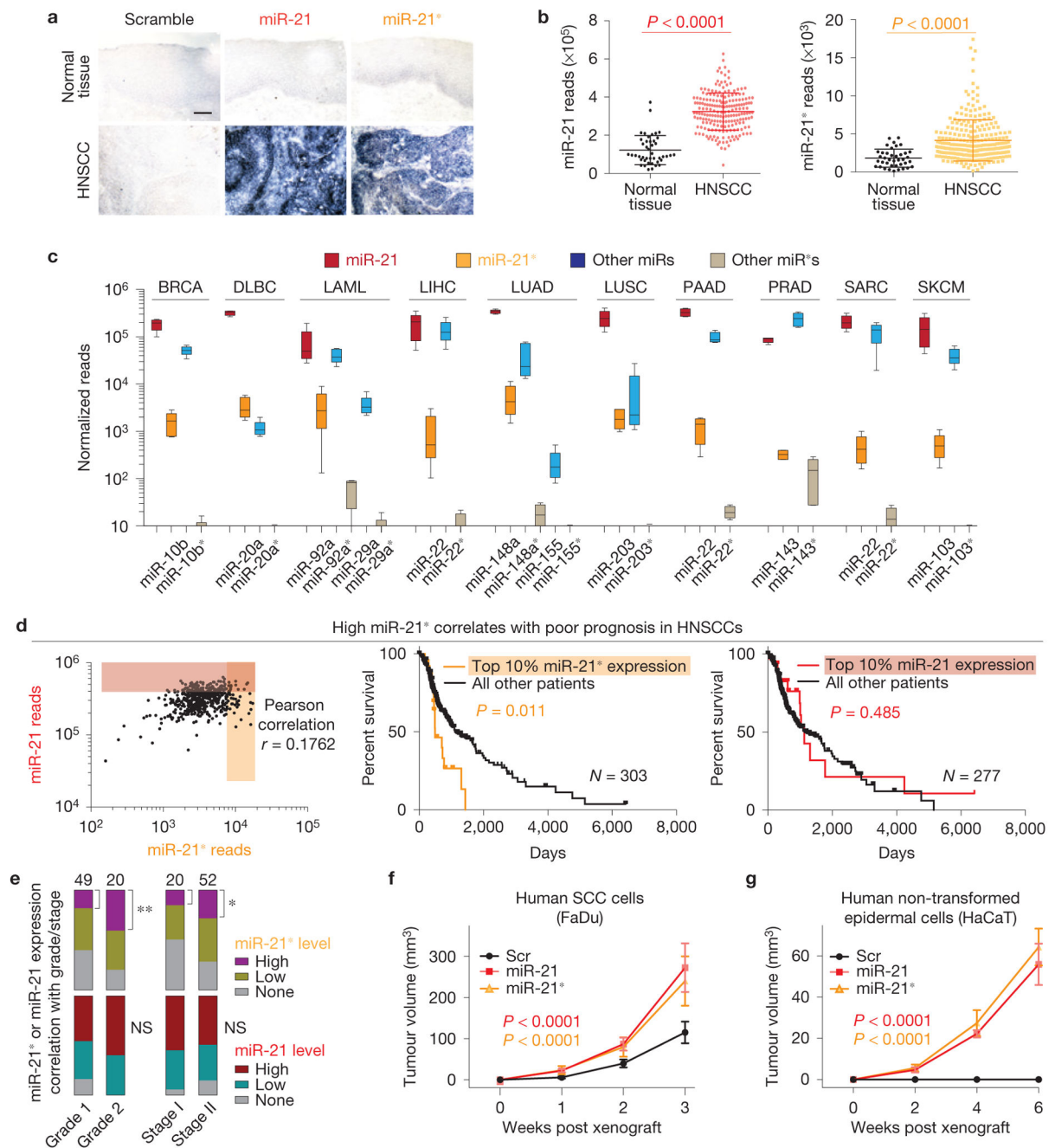
**Figure 4.**

In vivo mouse tumorigenesis validates candidate oncomiRs including miR-21* in driving SCC progression. **(a)** Left panel: schematic of *in vivo* functional test for oncomiRs in mice (see text). Right panel: tumour-free survival curves from Cre-induced HRas^{G12V}/YFP animals from LV-Cre-SA-miR-21*, miR-21, miR-200a and miR-200b. $n = 6$ mice for Scr cohort, $n = 7$ mice for miR-21 cohort, $n = 9$ mice for miR-21* cohort, $n = 8$ mice for miR-200a, $n = 8$ mice for miR-200b cohort. Log-rank Mantel–Cox test was used. **(b)** qPCR measurement of miR-21 or miR-21* expression levels, expressed as fold change in benign papilloma and SCCs relative to normal adult skin. **(c)** Immunofluorescence microscopy of skin sections of loxP-stop-loxP HRas^{G12V}/YFP animals transduced at E9.5 with LV-Cre-SA-miR/Scr viruses and then induced to form tumours in adults. Epifluorescence signals and/or secondary antibodies are colour-coded as indicated. EdU signals were detected using the azide-647 Click-iT kit. Note that Cre induction of SA-Scr-transduced mice resulted in early skin lesions with smooth tumour edges, typical of benign papilloma. Cre induction of either

miR-21- or miR-21^{*}-overexpressing skin resulted in accelerated progression to SCC, with irregular tumour edges, expansion of proliferative basal layers and loss of differentiation layers. Scale bar, 50 μ m. **(d)** Quantification of K5, K10 and EdU in tumour cells FACS-purified from the animals from **c**. For **b** and **d**, at least three biological replicates were performed. Paired *t*-test was used, shown are mean \pm s.d. *n* = 3 independent experiments. **P* < 0.05, ***P* < 0.01.

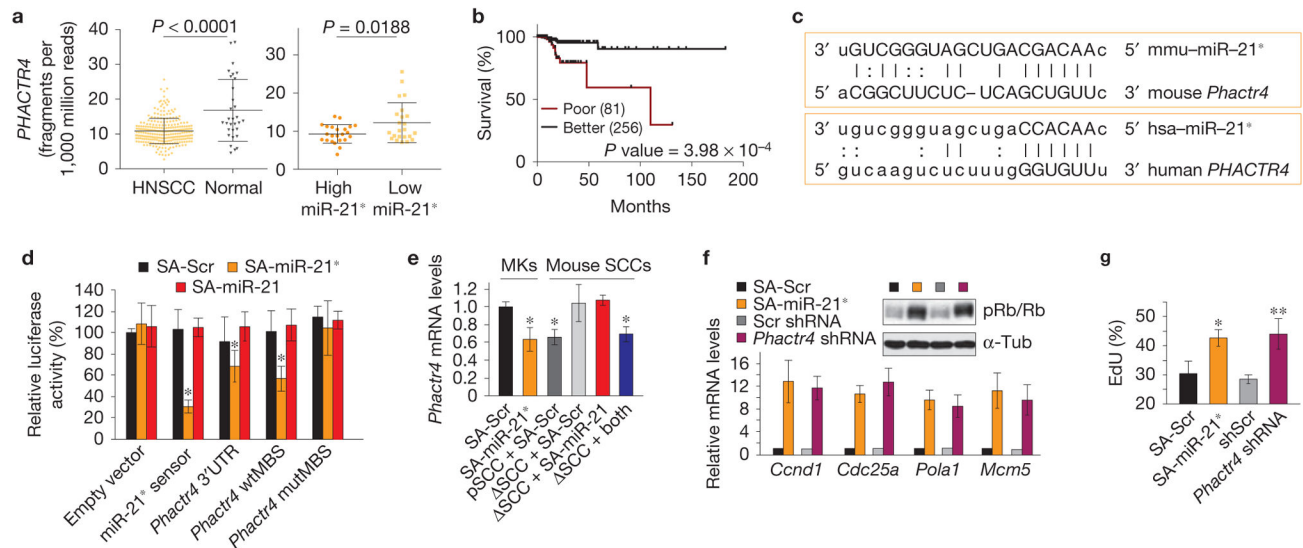
**Figure 5.**

miR-21* is required to sustain tumour growth independent of miR-21. **(a)** Sequences of single clones isolated from CRISPR/CAS-edited murine SCC63B tumour-initiating cells¹¹. In each case, the genome-editing events disrupted the miR-21 locus. **(b)** qPCR of miR-21 locus-edited SCC cells (ΔSCC) shows efficient loss of both miR-21 and miR-21*. Strand-specific miR replacements using lentiviral SA-miR vectors shows efficient and selective restoration of the appropriate strand following lentiviral SA-miR transduction. **(c)** Limiting dilution assay showing the dampened tumour initiation activity from CRISPR-mediated miR-21-locus ΔSCC cells, and partially restored by re-expressing either miR-21 or miR-21*. **(d)** ΔSCC cells, generated by CRISPR/CAS-mediated deletion of the miR-21 locus (three independent clones were analysed and shown here is a representative clone) in previously established SCC cells¹¹, were transduced with SA-miRs to generate isogenic ΔSCC lines that express miR-21, miR-21*, neither or both. Engraftment of these lines revealed that relative to Scr (control) and the parental SCC (pSCC, generated with sgRNA-Scr alongside sgRNA-miR), each strand possessed tumour-forming ability. Without them, tumorigenesis was severely compromised; together, they restored tumorigenicity to full capacity. Two-way ANOVA with repeated measurement was performed. Shown are mean ± s.d. $n = 3$ independent experiments. Log-rank Mantel–Cox test was used. * $P < 0.05$, ** $P < 0.01$. NS, not significant. **(e)** Allograft competition assay. Test (GFP marked) and control (RFP marked) cells are mixed at equal ratios, combined with Matrigel and then transplanted into the backs of Nude mice. Emerging tumours (>1 cm) are then subjected to FACS analysis to gauge efficiency of contribution of test and control cells to the tumours. In this assay, a red and green control population behave similarly but test cells that have lost the miR-21 locus or with one strand restored are outcompeted or partially rescued, respectively. For **b,e**, paired t -tests were used to compare edited cells with parental cells. At least three biological replicates were performed; shown are mean ± s.d. $n = 3$ independent experiments. * $P < 0.05$, ** $P < 0.01$.

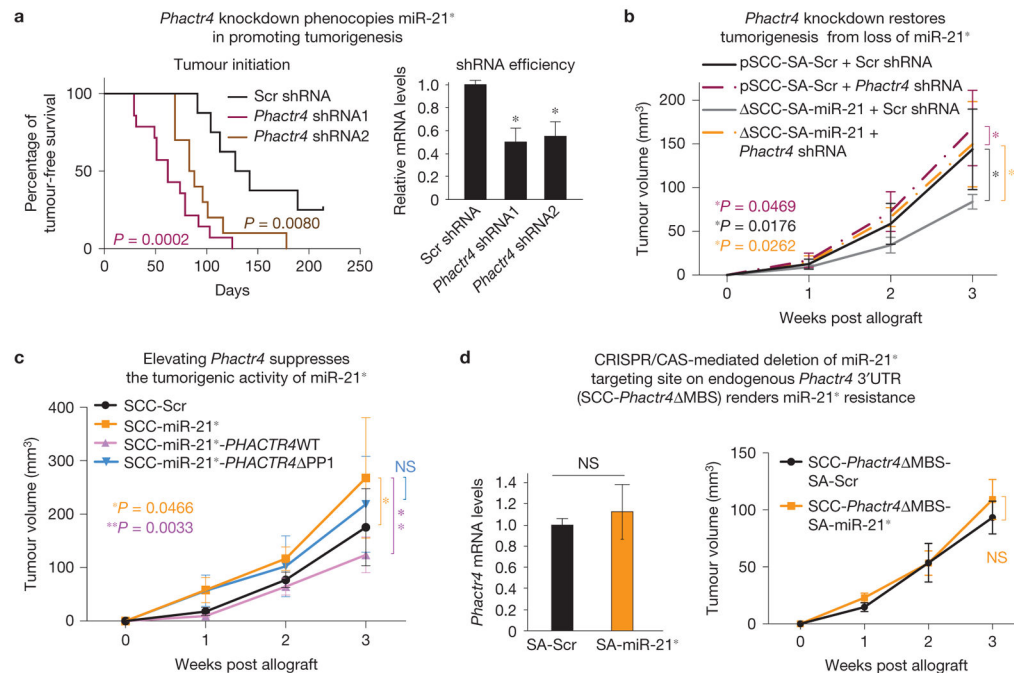
**Figure 6.**

miR-21* is physiologically relevant in human SCCs. **(a,b)** miR-21 and miR-21* are both elevated in human head and neck SCCs (HNSCCs) relative to unaffected tissue. Data are based on *in situ* hybridizations from HNSCC biopsies **(a)** and on TCGA miRNA sequencing data **(b)**. For **a**, at least three biological replicates were performed; shown are representative images. Scale bar, 100 μm . For **b**, unpaired *t*-test was used, shown are mean \pm s.d. $n = 31$ for normal tissues and $n = 263$ for tumours. **(c)** miRNA sequencing data from the TCGA database show that miR-21* is expressed in many types of cancer³⁵ including breast invasive

carcinoma (BRCA), diffuse large B-cell lymphoma (DLBC), acute myeloid leukaemia (LAML), liver hepatocellular carcinoma (LIHC), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), pancreatic ductal adenocarcinoma (PAAD), prostate adenocarcinoma (PRAD), sarcoma (SARC) and cutaneous melanoma (SKCM). Illustrated on the same graph are expression levels of many tumour-associated-miRs whose corresponding miR*s were not similarly stabilized. An exception was miR-143* in PRAD. Box plot maxima, minima and centre are defined at highest, lowest reads and reads at medium. $n = 3$ tumours. **(d)** miR-21* (orange) is considerably more variable than miR-21 (red) in human SCCs (left panel). However, whereas miR-21 is uniformly high, patients with the highest miR-21* expression show a trend towards decreased survival (right and middle panels). Pearson r correlation test was performed. $n = 303$ and 277 for miR-21* and miR-21, respectively. **(e)** Tissue array *in situ* hybridization data were divided into high, low or no expression groups, and correlated with patient tumour grades and stages. Strong miR-21* (but not miR-21) expression was associated with more advanced tumours. $n = 49$, 20 total for grade tumours, and $n = 20$, 52 total for stage tumours. Binomial test was used. ** $P = 0.0098$ for miR-21* grades, * $P = 0.0100$ for miR-21* stages. NS, not significant. **(f,g)** SA-miR-21* and SA-miR-21 promote xenograft growth from human SCC FaDu cells **(f)** and from non-transformed human epithelial HaCaT cells **(g)**. At least five biological replicates were performed; shown are mean \pm s.d. $n = 3$ independent experiments.

**Figure 7.**

miR-21* directly targets *Phactr4* and regulates tumorigenesis through the Rb/E2F axis. **(a)** TCGA inverse correlation between miR-21* and its putative target *PHACTR4* in HNSCCs. *PHACTR4* expression is diminished in HNSCCs compared with normal tissues (left) and is inversely correlated to miR-21* levels in these tumours (right). Unpaired *t*-test was performed. Shown are mean \pm s.d. $n = 31$ for normal tissues and $n = 263$ for tumours. **(b)** *PHACTR4* expression levels inversely correlate with prognosis in HNSCCs. Log-rank Mantel-Cox test was used. **(c)** Sequence alignment of predicted miR-21*-binding site within the *Phactr4* 3'UTR that is conserved between human and mouse. **(d)** Luciferase reporter assays in mouse keratinocytes show that when the miR-21*-binding site of the *Phactr4* 3'UTR is mutated, its potency in repressing luciferase expression in the presence of miR-21* is abrogated, indicating that miR-21* directly targets *Phactr4*. Note that miR-21 has no effect in this assay. **(e)** *Phactr4* expression is reduced by miR-21* in primary keratinocytes (SA-miR-21* compared with SA-Scr) and in parental SCC cells (pSCC+SA-Scr compared with SA-Scr) and it is de-repressed in miR-21*-defective SCC cells (SCC+SA-Scr or SCC+SA-miR-21 compared with pSCC+SA-Scr). **(f)** Western blot from SA-miR-21* overexpression or *Phactr4* knockdown cells show elevated pRb level under serum starvation, corresponding to induced transcription of E2F/Rb targets involved in cell cycle progression, measured with qPCR. At least three biological replicates were performed; shown are representative images. An unprocessed original scan of the blot is shown in Supplementary Fig. 9. **(g)** Increased EdU labelling and proliferation from SA-miR-21* overexpression or *Phactr4* knockdown in keratinocytes under resuspension growth. For **d–g**, at least three biological replicates were performed. Shown are mean \pm s.d. $n = 3$ independent experiments. Paired *t*-test was performed. * $P < 0.05$, ** $P < 0.01$. NS, not significant.

**Figure 8.**

Phactr4 is a main mediator of oncomiR-21* activity *in vivo*. (a) *Phactr4* knockdown by 2 shRNAs *in vivo* phenocopies miR-21* overexpression in promoting tumorigenesis. *Phactr4* shRNA hairpins were subcloned into LV-Cre virus and used to transduce LSL-HRas^{G12V} mice using our *in utero* procedure. $n = 8$ mice for Scr cohort; $n = 14$ mice for shRNA1 cohort, $n = 10$ mice for shRNA2 cohort. Paired *t*-test was performed for qPCR measurement. Log-rank Mantel-Cox test was used for survival analysis. (b) Loss of miR-21* inhibited tumour formation (SCC-SA-miR-21 + Scr shRNA, grey line), whereas *Phactr4* knockdown fully restored its tumour formation efficiency (SCC-SA-miR-21 + *Phactr4* shRNA, orange line) and further enhanced pSCC tumorigenicity (pSCC + *Phactr4* shRNA, purple line). Two-way ANOVA with repeated measurement was performed. (c) Elevating *Phactr4* suppresses the tumorigenic activity of miR-21*. The wild type but not the mutant version lacking the PP1-binding domain (*Phactr4* PP1) of *Phactr4* overexpression reversed the enhanced tumorigenesis mediated by SA-miR-21*. Two-way ANOVA with repeated measurement was performed. (d) SCC-*Phactr4* MBS cells, generated by CRISPR/CAS-mediated deletion of the predicted miR-21* targeting site on the endogenous *Phactr4* 3'UTR, were subjected to SA-Scr or SA-miR-21* expression, followed by *Phactr4* level measurement using qPCR, or allograft assays. *Phactr4* MBS SCCs cells were no longer sensitive to miR-21* regulation. Paired *t*-test was performed for qPCR. Two-way ANOVA with repeated measurement was performed for tumour growth. For a–d, at least three biological replicates were performed; shown are mean \pm s.d. $n = 3$ independent experiments. * $P < 0.05$, ** $P < 0.01$. NS, not significant.