MicroRNA-203 represses selection and expansion of oncogenic Hras transformed tumor initiating cells

Kent Riemondy¹, Xiao-jing Wang², Enrique C. Torchia³, Dennis R. Roop³ and Rui Yi¹,⁴

¹Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, Boulder, CO 80309
²Department of Pathology, University of Colorado Denver Anschutz Medical Campus, Aurora, CO 80045
³Department of Dermatology and Charles C. Gates Center for Regenerative Medicine and Stem Cell Biology, University of Colorado Denver Anschutz Medical Campus, Aurora, CO 80045
⁴Correspondence:
Rui Yi
Department of Molecular, Cellular, and Developmental Biology
University of Colorado, Boulder
Boulder, CO USA 80303
Tel: 303-735-4886
Email: yir@colorado.edu

Competing Interests Statement
The authors declare no competing interests.
Abstract

In many mouse models of skin cancer, only a few tumors typically form even though many cells competent for tumorigenesis receive the same oncogenic stimuli. These observations suggest an active selection process for tumor-initiating cells. Here we use quantitative mRNA- and miR-Seq to determine the impact of Hras\textsuperscript{G12V} on the transcriptome of keratinocytes. We discover that microRNA-203 is downregulated by Hras\textsuperscript{G12V}. Using a knockout mouse model, we demonstrate that loss of microRNA-203 promotes selection and expansion of tumor-initiating cells. Conversely, restoration of microRNA-203 using an inducible model potently inhibits proliferation of these cells. We comprehensively identify microRNA-203 targets required for Hras-initiated tumorigenesis. These targets include critical regulators of the Ras pathway and essential genes required for cell division. This study establishes a role for the loss of microRNA-203 in promoting selection and expansion of Hras mutated cells and identifies a mechanism through which microRNA-203 antagonizes Hras-mediated tumorigenesis.
Introduction

Recent efforts in comprehensively sequencing human cancer genomes have confirmed ~140 protein-coding genes that, when mutated, can drive tumorigenesis [1]. When genome sequencing data were utilized to construct the history of cancer cells in breast cancer, it was revealed that a considerable amount of “molecular time” exists between the common ancestors that harbor the great majority of driver mutations and the phenotypically identified cancer cells that compose the bulk of the tumor [2]. In support of these observations, lineage tracing experiments conducted in genetically engineered mouse models revealed only a few clones give rise to tumors whereas a vast majority of mutated cells are unable to sustain tumorigenesis [3,4]. These results suggest that even after the acquisition of key driver mutations in the nascent cancer cells, these cells must still undergo continuous evolution and likely clonal selection before developing into clinically apparent tumors. To begin to understand the molecular basis underlying such selection, we examined papilloma formation driven by oncogenic Hras in the skin, a well characterized model where Hras has been shown to initiate the formation of tumors that clonally evolve [3,5,6]. Oncogenic Ras mutations are some of the most frequently detected driver mutations in human cancer. Among the three Ras genes (H-, K- and N-ras), Hras is commonly mutated in tumors originated from stratified epithelial tissues including squamous cell carcinoma in the skin, head and neck cancer as well as bladder cancer [7–9]. Experimental and genomic sequencing studies have revealed that the vast majority of Ras mutations are missense, point mutations at amino acid residues glycine 12 (G12), glycine 13 (G13) or glutamine 61 (Q61) [8]. Structural and biochemical studies have further confirmed that all of these mutations generally interfere with the GTP binding pocket and compromise the GTPase activity of Ras proteins. In turn, these mutations lead to uncontrolled activation of downstream effectors
including Raf/MEK/ERK and PI(3)K pathways, resulting in sustained cell survival and proliferation observed in human cancers. Because of the prominent role of Ras mutations in human cancer, extensive efforts have been devoted to uncover and subsequently target downstream pathways that are regulated by Ras mutations. However, the immediate impact of Ras mutations on the transcriptome, in particular with regards to microRNAs (miRNAs) remains unclear.

miRNAs are a class of small, noncoding RNA species that are involved in virtually all biological processes examined in mammals including mouse and human. These regulatory RNA molecules function by repressing the protein producing ability of mRNA targets through destabilization of mRNAs and inhibition of translation [10]. miRNAs typically target a large number of mRNAs in a dosage- and cell context-dependent manner [11]. As prominent proto-oncogenes, Ras mutations have long been recognized to interact with the miRNA pathway. Indeed, Hras, Kras and Nras all harbor multiple binding sites for the let-7 miRNA, a founding member of miRNAs, in their 3’UTRs [12]. Additionally, impaired miRNA biogenesis in the form of Dicer1 disruption has been shown to be a tumor suppressing mechanism for the development of Kras induced lung cancer in a mouse model [13]. A number of individual miRNAs were also found to function as modifiers for Ras induced tumorigenesis that include miR-21, -29 and miR-17~92 as tumor promoting miRNAs and miR-34, -15/16, and miR-143/145 as tumor suppressing miRNAs [14–16]. Collectively, these seminal studies demonstrate unequivocally that the miRNA pathway and individual miRNAs play important roles in Ras-induced tumorigenesis. However, it is unclear how Ras mutations, usually the tumor-initiating drivers, directly alter the landscape of miRNA expression during tumorigenesis. Importantly, it is also unknown whether the changes in miRNA expression play a role in the selection of
oncogenic Ras transformed cells during tumor initiation. Finally, the lack of a comprehensive
survey of high-confidence miRNA targets that may play a role downstream of Ras mutations
hinders our mechanistic understanding and limits the potential to develop miRNA based
therapeutics.

In this study, we utilized our recently improved quantitative miR-Seq techniques to
examine the impact of an oncogenic Hras mutation (HrasG12V) on both mRNA and miRNA
expression. We discovered that miR-203, the most highly expressed miRNA in the skin [17,18],
is downregulated by HrasG12V. Using both knockout (KO) and inducible models, we provide
evidence for an important role of miR-203 in restricting expansion of oncogenic Hras
transformed cells in vitro and in vivo. We comprehensively surveyed skin-specific targets of
miR-203 and identified a number of novel targets that have important implications for Hras-
mediated tumorigenesis. Our results suggest that miR-203 plays a tumor-suppressing role in
inhibiting selection and expansion of tumor-initiating cells early in tumor development.

Results:

HrasG12V profoundly deregulates the mRNA and miRNA transcriptome in the skin

Oncogenic mutation of the Hras gene is one of the initiating drivers in the development
of benign papillomas and malignant squamous cell carcinomas in murine skin chemical
carcinogenesis. However, the molecular consequences defining the cellular changes that
accompany expansion of oncogenic Hras-transformed keratinocytes to initiate papillomas
remain elusive. We first investigated the consequences of HrasG12V activation on the mRNA
transcriptome using a modified form of PolyA+ RNA-Seq, known as 3P-Seq, or 3seq (Figure
1A-C). Compared to traditional RNA-Seq, 3Seq allows both quantification of mRNA transcripts
and detection of changes in alternative 3’ UTR formation [19]. To examine the immediate impact of *Hras*<sup>G12V</sup> on primary skin cells, we used primary keratinocytes isolated from newborn skin and performed 3Seq after *Hras*<sup>G12V</sup> transduction. We did not observe widespread shortening or alternative formation of 3’UTRs, which are often ascribed to oncogenic transformation when comparing tumor cell lines to normal cells (data not shown). This is similar to our previous observation that alternative 3’UTR usage is infrequent within the skin lineages [19]. Over 1,100 transcripts were differently expressed (two-fold change and FDR <0.05) in keratinocytes expressing *Hras*<sup>G12V</sup>, compared to the control (Figure 1C and Figure 1-source data). Gene ontology functional analysis revealed profound deregulation in three core processes by *Hras*<sup>G12V</sup>: activation of cellular migration, upregulation of pro-angiogenic pathways, and suppression of the terminal differentiation program (Figure 1D). All of these three processes are identified as hallmarks of human cancer [20]. The observed widespread changes in the transcriptome also endorse the driver role of *Hras*<sup>G12V</sup> in skin tumorigenesis. Importantly, transcripts upregulated by *Hras*<sup>G12V</sup> in our primary keratinocytes strongly and significantly overlapped with the putative cancer stem cell signatures obtained from murine squamous cell carcinoma (SCC) models [21]. In addition, transcripts upregulated by *Hras*<sup>G12V</sup> significantly overlapped with transcripts known to be targets of the *c-Fos* transcription factor in a genetic model of SCC [22]. Furthermore, known core components of the *Hras* signaling pathway were also among the differentially detected genes [23] (Figure 1E). These transcriptome data indicate that we have captured the initiating changes induced by oncogenic *Hras* in the keratinocytes.

To define the impact of the oncogenic *Hras* on the landscape of miRNA, we applied our recently developed, quantitative miRNA-Seq [24] to *Hras*<sup>G12V</sup> transformed keratinocytes. Overall, we detected 15 differentially expressed miRNAs upon *Hras*<sup>G12V</sup> expression (FDR <0.05, two-
fold change) (Figure 1F-H). Two key patterns emerged in these profiles. First, the epithelial tissue-specific miRNAs, miR-203 and miR-205, which represent the most abundant miRNAs expressed in murine skin and primary keratinocytes were strongly suppressed by Hras<sup>G12V</sup>. Additionally, members of the abundantly expressed, miR-23/24/27 miRNA cluster were also downregulated by Hras<sup>G12V</sup>. Secondly, miR-21 was strongly induced becoming the most highly expressed miRNA, consistent with its direct activation by oncogenic Ras reported in other systems [25]. The upregulation of miR-21 is also consistent with its well-appreciated oncogenic function in skin cancer [26]. miR-146 was also induced by Hras<sup>G12V</sup>. However this miRNA is expressed 2-orders of magnitude lower than miR-21, suggesting that its upregulation may have limited contribution to Hras initiated tumorigenesis at this early stage.

We further measured mature miR-21, miR-203, and miR-205 RNAs by qPCR. In support of the quantitative performance of our miR-Seq, the differential expression of all three miRNAs measured by qPCR was nearly identical to the quantification by our miR-Seq (Figure 2A-B). To initially probe the mechanism through which HRAS<sup>G12V</sup> suppresses miR-203 expression, we examined the level of miR-203 primary transcripts. We previously characterized the transcribed genomic region of miR-203 including the promoter region and transcription start site [17]. Because the primary transcript of miR-203 harbors a polyadenylation [Poly(A)] signal and generates a Poly(A) tail, we directly quantified the abundance of the primary transcripts by counting the 3’end reads of the primary miRNA obtained by 3seq (Figure 2C). This result was further confirmed by qPCR measurement specific to the pri-miR-203 (Figure 2D). The degree of downregulation for both mature and primary miR-203 transcripts was similar as judged by these two independent assays. Collectively, we conclude that the repression of miR-203 by Hras<sup>G12V</sup> is
most likely mediated by suppressing the production of primary miR-203 transcripts at an early stage of oncogenic cellular transformation.

**miR-203 silencing is an early event in mouse and human SCC**

Our data revealed the early silencing of miR-203 by oncogenic Hras in an in vitro model. We further investigated miR-203 expression with in situ hybridization during skin tumorigenesis in vivo. In classic chemical carcinogenesis models initiated by DMBA/TPA treatment of mouse skin, Hras is preferentially mutated, which compromises the GTPase activity and results in constitutive Hras activation. Mutating Hras at Q61 leads to papilloma development and infrequent malignant transformation to SCCs [27]. We first examined miR-203 expression in benign papillomas. Consistent with our in vitro results, miR-203 expression was absent from epithelial compartments adjacent to tumor stroma, the region where putative cancer stem cells reside. Although we also observed moderately expressed miR-203 in tumor regions with evidence of cellular differentiation (keratin pearls and large differentiated morphology), the levels of miR-203 was considerably lower in these regions compared to the suprabasal cells of adjacent normal skin, where miR-203 is normally expressed (Figure 2E). Overall, miR-203 expression levels were strongly reduced in tumor tissues compared to adjacent epidermal regions. In an independent mouse model of SCC, we also found that miR-203 expression was gradually lost in a Kras$^{G12D}/$Smad4 cKO model where skin tumors progressed to invasive SCC through serial passages of tumors [28] (Figure 2F). Taken together, these results demonstrate that miR-203 is significantly downregulated at both early and late stages of papilloma and SCC formation in mouse models.
To evaluate the relevance of the loss of miR-203 in human skin cancer, we examined 9 tumor samples obtained from patients with the early, middle and late stages of skin cancer. In these human skin SCC samples, miR-203 was already downregulated at the onset of tumorigenesis and progressively lost during the course of tumor progression, similar to the pattern observed in the mouse models. In the poorly differentiated SCC samples, the miR-203 signal was completely absent, yet readily detectable in surrounding hyperplastic or normal epidermal regions (Figure 2G-O). In addition, we mined publicly available datasets from The Cancer Genome Atlas (TCGA) and detected a significant reduction in miR-203 and an elevation of miR-21 in head and neck SCC samples compared to patient matched normal tissues (Figure 2P). Altogether, these results corroborate our observations in mouse models and validate the strong correlation between the loss of miR-203 and the development of skin cancer at multiple stages of tumorigenesis. These expression analyses suggest that the loss of miR-203 coincides with the tumor-initiating events and miR-203 might function as a tumor suppressing mechanism in skin cancer.

Genetic deletion of miR-203 impacts early epidermal development in mouse

We then generated a conditional KO mouse model to assess the function of miR-203 in murine skin development and carcinogenesis in vivo. We have characterized the genomic locus of miR-203 and determined that miR-203 is located in an intergenic region, 3.3 kbp downstream from the Asp gene and 15.3 kbp upstream of the Kif26a gene (Figure 3-figure supplement 1). Two loxP sites were inserted to flank the miR-203 hairpin (Figure 3A). miR-203 was deleted by first mating miR-203flo/flo mice with Actb-Flpe mice to remove the Neo cassette, followed by breeding with El2a-Cre or Krt14-Cre mice, resulting in complete miR-203 loss from all tissues or
only from skin tissues, respectively. In both cases, ablation of miR-203 was confirmed by qPCR on isolated epidermis and in situ hybridization (Figure 3B-C). We previously demonstrated that expression of miR-203 was largely restricted to stratified epithelial tissues. Within the skin, the differentiated skin lineages express miR-203 ~10-fold higher than the stem/progenitor lineages. Consistent with this observation, there are no discernible differences between constitutive miR-203 null (miR-203\(^{-/-}\)) mice and skin-specific cKO (Krt14-Cre/miR-203\(^{fl/fl}\)) mice. Both strains were born at the expected Mendelian ratio. They also showed no signs of gross developmental defects as adults (Figure 3D). For all subsequent studies we used miR-203\(^{-/-}\), which was maintained in the C57BL/6 background.

miR-203 is highly expressed primarily in stratified epithelial tissues, such as the epidermis, tongue, esophagus, and cervix, yet poorly expressed or absent in other tissues such as the small intestine, bladder, lung, kidney, liver, or brain (Figure 3-figure supplement 2). Because miR-203 begins to be expressed by E13 when the epidermis begins to stratify, we examined the proliferation rate and thickness of embryonic skin from E16 to P4. At E16, we observed a mild increase in cell proliferation in the KO, as measured by BrdU incorporation (Figure 3E-F). Although the difference in BrdU incorporation did not achieve statistical significance (p=0.07), the thickness of the KO epidermis was significantly increased, compared to WT littermates (p=0.03). Interestingly, the increase in epidermal thickness was most prominent at early stages (E16 and E17) and waned as skin development progressed (P4) (Figure 3G). At P4, miR-203\(^{-/-}\) mice displayed normal histological development of the epidermis and hair follicles and the difference in proliferation and epidermal thickness between KO and WT became indistinguishable (Figure 3G). In addition, we found no evidence of perturbed differentiation in miR-203\(^{-/-}\) mice based upon analysis of early and late epidermal differentiation markers, Keratin
1 and Loricrin, for the spinous and granular layers, respectively (Figure 3H-I). Together, these results provide evidence that miR-203 limits cell division when the proliferation rate is high during early embryonic skin development but not at later stages when the proliferation rate wanes.

We noted that the impact of miR-203 loss was correlated with the rate of cell proliferation. To further test whether miR-203 functions to restrict the expansion of highly proliferative cells, we investigated the roles for miR-203 in regulation of primary and established keratinocyte cell lines, respectively. We observed ~2-fold higher colony forming capacity of miR-203−/− keratinocytes, compared to the WT controls (Figure 3J). To further confirm the ability of miR-203 to suppress cell proliferation cell-autonomously, we generated a miR-203fl/fl keratinocyte cell line. By treating these cells with an adenoviral vector to express Cre (Ad-Cre), we determined that within 48 h of Ad-Cre exposure miR-203 was completely depleted (<1% remaining as measured by qPCR). We again observed a similar, ~2-fold higher colony forming capacity by the Ad-Cre treated miR-203fl/fl cells, compared to the Ad-GFP treated control cells (Figure 3K). In both cases, although we detected some slightly larger clones with more cells formed by the miR-203 null keratinocytes, the biggest differences were the significantly increased number of clones formed by the miR-203 null cells. These results suggest that miR-203 limits the clonogenicity of normal keratinocytes. When miR-203 is deleted, more cells are likely to become clonogenic.

Loss of miR-203 enhances chemical carcinogenesis in the skin

We determined that the loss of miR-203 is an early event in the initiation and development of both mouse and human skin SCCs. Furthermore, genetic deletion of miR-203 confers ~2-fold higher colony forming ability on primary keratinocytes. Because oncogenic Hras
is a potent driver for tumorigenesis in the skin and our miR-seq data revealed a rapid and strong repression of miR-203 by Hras\(^{G12V}\), these observations prompted us to investigate whether the loss of miR-203 plays a role in skin carcinogenesis. We subjected WT and miR-203 null mice to two-stage chemical DMBA/TPA carcinogenesis (Figure 4A). The chemical carcinogenesis experiments were terminated at week 21 when tumor burden had reached a maximum. Because our mice were generated in the C57BL/6 background and these mice are known to be resistant to two-stage chemical carcinogenesis [27], all tumors generated in our mice were papillomas with no evidence for malignant conversion to squamous cell carcinomas during the time frame of our study. Nevertheless, we examined the temporal and numeric characteristics of tumor formation in our mice for the role of miR-203 in tumor initiation. During the course of carcinogenesis, we observed a slightly earlier tumor formation pattern on the backskin of miR-203 null animals (Figure 4B). Furthermore, miR-203 null animals developed ~2.5-fold more tumors, when compared to WT control animals (Figure 4B). Measurement of tumor sizes at week 17 showed no statistically significant difference in tumor sizes between genotypes although miR-203 null animals were more susceptible to tumor formation (Figure 4C). These results suggest that loss of miR-203 increases the number of tumor-initiating cells but does not significantly alter the type of tumors e.g. conferring tumors with more aggressive phenotypes.

Hematoxylin and Eosin staining revealed that the papillomas produced had similar morphology, displaying exophytic lesions with evidence for squamous differentiation (Figure 4D-E). Assessment of proliferation (Ki67), differentiation markers (Krt1, Lor) revealed similar dynamics between miR-203\(^{+/+}\) and miR-203\(^{-/-}\) tumors (Figure 4F). To further probe the mechanistic differences between miR-203 WT and KO tumors, we assessed the mutation spectra of these tumors and found that they possess the canonical Hras\(^{Q61L}\) mutation at similar
frequencies, 80% and 85.7% for WT and KO tumors, respectively (Figure 4-figure supplement 1). Taken together, these results provide direct evidence for miR-203’s tumor suppressing roles at the stage of tumor initiation in the classic DMBA/TPA tumorigenic model.

**miR-203 represses clonal selection of oncogenic Hras transformed cells in vitro**

To further probe miR-203’s role in clonal selection, we infected miR-203 WT and null primary keratinocytes with pBabe-Hras\(^{G12V}\). At the passage 1, the loss of miR-203 led to ~40% increase in colony forming capacity immediately after the initial plating (Figure 5A). When we passaged the Hras\(^{G12V}\) transduced cells, we began to observe reduced numbers of colonies that were formed by both the WT and null cells (Figure 5A). This was likely due to oncogenic stress caused by Hras\(^{G12V}\) induction. Strikingly, in contrast to the WT cells that generally formed smaller colonies in subsequent passages, the miR-203 null cells generated more and bigger clones compared to WT control cells at each passage (Figure 5A). Collectively, the serial passage assay supported the enhanced selection of tumor-initiating cells in the absence of miR-203 upon oncogenic Hras induction.

To further characterize the ability of miR-203 to suppress the growth of oncogenic Hras transformed cells, we used our previously established Krt14-rtTA/pTre2-miR-203 inducible model [17]. After infecting the inducible keratinocytes with either the pBabe vector control or pBabe-Hras\(^{G12V}\), we treated the cells with 5ug/ml doxycycline to induce ~4-7-fold increase in miR-203 expression, a physiologically relevant level of miR-203 typically observed during epidermal differentiation (Figure 5B). Introduction of miR-203 resulted in suppression of keratinocyte proliferation and colony formation ability (Figure 5C-D), as noted previously [17,18,29]. Furthermore, whereas oncogenic Hras enhanced S-phase entry, short-term (~24 h)
induction of *miR-203* completely abolished the gain of S-phase entry (Figure 5D). Over a longer term, continuous induction of *miR-203* severely compromised the colony forming capacity of the transduced cells (Figure 5C). We did not detect any evidence for enhanced apoptosis caused by *miR-203*, as measured by the absence of sub-G1 keratinocytes. Taken together, our results suggest that the loss of *miR-203* is critical for the initial selection and expansion of primary keratinocytes harboring the oncogenic *Hras* mutation and the gain of *miR-203* can effectively suppress the growth of these cells.

**Comprehensive identification of *miR-203* targets reveals regulation of the Ras signaling pathway**

Our data so far have suggested a role of *miR-203* in suppressing the expansion of tumor-initiating cells driven by oncogenic *Hras* mutations. To decode the underlying mechanism, we carried out comprehensive analyses to identify *miR-203* targets in the skin. Recent studies demonstrated that miRNAs’ impact on gene expression could be largely captured by measuring the changes of mRNA levels upon manipulation of miRNA expression [30,31]. In parallel, the High-Throughput Sequencing of RNA isolated by CrossLinking ImmunoPrecipitation (HITS-CLIP) approach directly crosslinks miRISC to their targets and identifies miRNA targets through physical interaction [32]. Therefore, we took a combinatorial approach integrating multiple datasets obtained from our KO and inducible mouse models with Ago2 HITS-CLIP analysis.

We applied several different profiling techniques including ribosome profiling, RNA-Seq (3Seq) and microarray methods to determine upregulated genes in *miR-203* KO samples and downregulated genes in *miR-203* induced samples. For ribosome profiling and 3Seq, we used primary keratinocytes, which are amenable to effective cycloheximide treatment and therefore
also allowed us to investigate miR-203’s impact on translation efficiency (Figure 6A). For microarray analysis, we used freshly isolated epidermis obtained from miR-203 WT and KO animals at P4.5. To interrogate the impact of miR-203 gain-of-function, we used the Krt14-rtTA/pTre2-miR-203 inducible mouse model that allows us to control the duration and dosage of miR-203 expression [17]. We applied microarray profiling to determine downregulated genes in FACS-purified neonatal epidermal and hair follicle progenitors with a short-term (24 hr) induction of miR-203 (Figure 6A). Altogether, we accrued 20 genome-wide expression datasets from miR-203 WT, KO and inducible samples. This comprehensive collection of functional genomic data allowed us to characterize the action of miR-203 on the transcriptome.

We first analyzed the miR-203 overexpression samples. GO-analysis demonstrated that genes involved in regulation of mitotic cell cycle, DNA synthesis and metabolism were prominently downregulated upon miR-203 induction (Figure 6-source data 1). miR-203 upregulation in both epidermal and hair follicle progenitors resulted in strong suppression of transcripts harboring perfect seed matches to miR-203 in their 3’UTRs (Figure 6-figure supplement 1). Among them, transcripts containing the 8-mer matches were strongly suppressed, compared to transcripts without a miR-203 seed match (p-value = 1.5x10^{-23} for epidermis and 2.6x10^{-25} for hair follicle, respectively) (Figure 6 - figure supplement 1A). In contrast to miR-203 overexpression, deletion of miR-203 did not perturb transcript levels genome-wide as strongly as observed under the induced conditions, and offered little insight into miR-203’s function based upon GO-analysis (Figure 6-source data 1). However, despite the relatively mild changes in the transcript levels in the KO skin and primary keratinocytes, transcripts that contain 8-mer 3’UTR matches were still more significantly upregulated than transcripts without any miR-203 seed matches (p <0.05) (Figure 6-figure supplement-1B). Lastly, a ranked sum analysis that ranked
genes based on expression changes correlated with the levels of miR-203 showed that genes with the 8-mer or 7-mer seed ranked significantly higher than those transcripts without seed matches (Figure 6-figure supplement-1C).

By demanding all potential targets be upregulated in all the miR-203 KO samples and downregulated in all the miR-203 induced samples, we identified 294 transcripts as our candidates for miR-203 targets (Figure 6B). Among the top 20 most differentially expressed transcripts when miR-203 was deleted or induced, 15 of them contain at least one 7-mer or 8-mer match in their 3’UTRs (Figure 6C and Figure 6-source data 2). In support of this approach, an unbiased de novo motif search using the 3’UTRs of these 294 transcripts revealed that the most enriched motif (ACAUUUCA, p=1x10^{-13}) perfectly matched to nucleotide position 2-9 of miR-203, the 8-mer seed sequences of miR-203 (Figure 6D). Additional statistical analysis of the number of genes containing perfect 3’UTR matches to the 7-mer or 8-mer seed among these candidates revealed significant enrichment over the background distribution among mRNAs expressed in the skin (Figure 6D). Therefore, we selected genes containing 7-mer or 8-mer seed matched 3’UTR sites as miR-203 target candidates and obtained a collection of 100 potential miR-203 targets (Figure 6B).

We then applied HITS-CLIP to interrogate the direct interaction between miR-203 and its mRNA targets. We generated four Ago2 HITS-CLIP libraries from primary WT keratinocytes, which abundantly express miR-203 (Figure 1G). Ago2-RNA complexes were isolated from a region extending from approximately 110kd – 130kd as expected (Figure 6-figure supplement 2A). Sequencing libraries from Ago2 HITS-CLIP samples were generated and analyzed using previously described methods [32,33]. The HITS-CLIP faithfully captured miRNA species expressed in keratinocytes, including miR-203 (Figure 6-figure supplement 2B and E). Overall
our HITS-CLIP reads and clusters alignment were similar to previous published results with a significant portion aligned to 3’UTRs (Figure 6-figure supplement 2D). To further validate our approach, we analyzed the positional enrichment of miRNA seed sequences within the 3’UTR HITS-CLIP clusters and detected strong enrichment over dinucleotide shuffled cluster sequences or randomly distributed 3’UTR regions (Figure 6-figure supplement 3A). Additionally, de novo motif searching for 8mer motifs identified the seed sequences of miRNAs that are highly expressed in keratinocytes (Figure 6-figure supplement 3B). The motif corresponding to miR-203 detected by the HITS-CLIP was also ACAUUUCA (p=1x10^{-19}), identical to the motif detected by our transcriptome analysis (Figure 6E and Figure 6-figure supplement 3B). A total of 113 mRNAs were detected to have miR-203 seed containing Ago2 binding sites. We next examined the ranked sum expression of these targets and determined that transcripts with Ago2 HITS-CLIP miR-203 binding sites were ranked significantly higher compared to non-targeted mRNAs, indicating that many of these targets are functional (Figure 6-figure supplement 4). Finally, we did not detect any evidence for global regulation at the translation level for miR-203 targets based on our analysis of translational efficiency changes upon miR-203 deletion (Figure 6-figure supplement 5). Together, these HITS-CLIP data independently validated that miR-203 uses its seed sequences for mRNA targeting and predicted 113 miR-203 targets based on Ago2 binding.

By combining the targets detected by our differential expression datasets and by the HITS-CLIP, we identified a list of high-confidence targets for miR-203 (Figure 6F). Importantly, we found a number of key regulators in the Ras signaling pathway and important genes involved in regulation of cell division including Hbegf, Ccnd1, Snai2, Met, and Pola1 in this list (Figure 6F). This collection of miR-203 targets suggests that miR-203 targets multiple pathways, including several components of the Ras signaling pathway, to suppress cell proliferation.
miR-203 directly targets Hbegf and Pola1

Our genome-wide analyses identified a number of novel targets of miR-203. Given our findings that the loss of miR-203 promotes the selection and expansion of oncogenic Hras transformed cells both in vivo and in vitro, we were interested in understanding the underlying mechanism. Overall, miR-203 targets identified in the meta-analysis were enriched in the upregulated transcripts in Hras\textsuperscript{G12V} transformed keratinocytes, compared to non-targeted transcripts, consistent with the downregulation of miR-203 in these cells (Figure 7-figure supplement 1A ). Among these targets, we were intrigued by the observation that multiple regulators of the Ras signaling pathway and critical factors for DNA replication and cell cycle progression were among our high confidence targets and additionally were among the most upregulated miR-203 targets in Hras\textsuperscript{G12V} transformed keratinocytes (Figure 7-figure supplement 1B ). To begin to validate the high-confidence set of miR-203 targets, we selected Pola1 and Hbegf for further study. Pola1 is the catalytic subunit of the DNA-POL-alpha holoenzyme, which is required in initiation of DNA replication during S-phase [34]. Hbegf is an Egf-like ligand that activates MAPK signaling through activation of EGF-receptors, Erbb1 and Erbb4. In keratinocytes, Hbegf is mitogenic and promotes keratinocyte migration [35]. In an epithelial cancer cell line, Hbegf acts as an oncogene promoting cell proliferation [36]. Pola1 and Hbegf contain 3’ UTR miR-203 target sites (9-mer and 8-mer respectively) that are targeted by miR-203, validated by luciferase assay (Figure 7A). Furthermore, in miR-203 null epidermis, both Pola1 and Hbegf mRNAs were elevated (Figure 7B). In addition to mRNA levels, Pola1 protein levels were also elevated in the absence of miR-203. It was further elevated in the presence of Hras\textsuperscript{G12V} and repressed by miR-203 induction (Figure 7C-D). We were unable to measure the protein level
of Hbegf due to poor antibody performance. Additionally, we observed that the expression of
Ccnd1, an essential cell cycle regulator that is often induced or amplified by oncogenic Ras
[9,37], showed a strong negative correlation to miR-203 (Figure 7C-D). This suggested that loss
of miR-203 increases the levels of Ccnd1 and contributes to the observed upregulation of this
critical gene.

To assess the functional consequences of Hbegf and Polal suppression on keratinocyte
proliferation, we knocked down these targets using three independent shRNAs. Suppression of
Hbegf and Polal strongly suppressed the growth potential of keratinocytes (Figure 7E).
Additionally, we knocked down Hbegf and Polal in established miR-203+/+ and miR-203/−
keratinocytes transduced with HrasG12V and similarly observed potent suppression of
keratinocyte growth potential, demonstrating that these targets are also required in HrasG12V
transformed keratinocytes (Figure 7-figure supplement 2 ). Taken together these results validate
Hbegf and Polal as direct targets of miR-203. We hereby propose a model in which miR-203
restricts selection and expansion of Hras mutated cells by repressing multiple targets, a subset of
which are involved in the Ras signaling pathway (Figure 7F).

Discussion
Identification and experimental validation of driver gene mutations have been
instrumental for our understanding of cancer biology [1]. Although it is clear that not all cells
harboring driver gene mutations develop into tumors, the mechanism of clonal selection within
cells that have acquired driver gene mutations remains poorly understood. In this study, we used
an oncogenic Hras induced tumor model to determine the molecular consequences of this
oncogenic driver mutation and to study the selection process. We quantitatively measured the
impact of oncogenic $Hras^{G12V}$ on the landscape of mRNA and miRNA expression. Whereas our
mRNA-Seq data confirmed the profound ability of $Hras^{G12V}$ in promoting cellular transformation,
our miR-Seq data revealed new insights into the dynamic changes in highly expressed miRNAs
caused by $Hras^{G12V}$ (Figure 1). In particular, the downregulation of $miR-203$, one of the most
highly expressed miRNAs in the skin, likely occurs at the transcriptional level (Figure 2C-D). By
examining tumor samples obtained from both mouse and human at multiple stages, we showed
that the downregulation of $miR-203$ is associated with tumor initiation and progression (Figure
2E-P). To determine the role of $miR-203$ in this process, we generated a KO mouse model. Initial
insights into the function of $miR-203$ came from the analysis of $miR-203$ null epidermis during
skin development. Whereas we and others have shown the potent inhibition of epidermal
proliferation by $miR-203$ with gain-of-function approaches [17,18,29,38], it is unknown how the
loss of $miR-203$ affects the skin and mouse development in general. By analyzing the KO mouse,
we revealed an early, albeit mild, hyperproliferative phenotype in the embryonic but not
postnatal skin (Figure 3E-I). However, when we subjected $miR-203$ WT and null cells obtained
from postnatal skin for their colony forming ability, we observed an increased ability of the $miR-
203$ null cells to give rise to colony forming cells (Figure 3J-K). This observation was indicative
of a role for $miR-203$ in restricting the expansion of highly proliferative cells. Of note, we did not
observe any defects in the formation of the differentiated layers of the skin in the $miR-203$ KO
mice. This suggests that the primary function of $miR-203$ is to restrict cell proliferation but not to
promote epidermal differentiation per se. When we examined the role of $miR-203$ during
DMBA/TPA mediated chemical carcinogenesis, we observed strong increase in the number of
papillomas formed in the KO skin (Figure 4A-C). In vitro colony formation assays with serial
passages further illustrated the enhanced ability for selecting tumor-initiating cells with $miR-203$
KO cells when subjected to \( Hras^{G12V} \) transformation (Figure 5A). However, due to tumor resistance of our C57BL/6 strain, we were unable to determine if the loss of \( miR-203 \) also promotes malignant transition. Of note, a recent study showed interesting results that restoration of \( miR-203 \) suppresses human SCC metastasis [29]. Collectively, our results provide experimental evidence for an important role of \( miR-203 \) in suppressing clonal selection and expansion of \( Hras \) mutated nascent cancer cells.

A major challenge in understanding miRNA functions is to identify their high-confidence targets globally. To this end, we employed two independent approaches: transcriptome analysis with our KO and inducible models and direct miRNA target capture by \( Ago2 \) HITS-CLIP (Figure 6). Importantly, with our expansive dataset, we confirmed that \( miR-203 \) utilizes the 5’ seed sequences to target perfectly matched sequences located on the 3’UTRs of its target genes. Consistent with recent genome-wide studies for the impact of miRNA on mRNA levels and translation efficiency in mammals [30,31], we found no evidence for global changes of translational efficiency in the absence of \( miR-203 \), one of the most highly expressed miRNAs in the skin. However, we also noted that \( Trp63 \), a well-established \( miR-203 \) target, showed little change at the mRNA levels under all conditions (not shown). Our analysis with differential gene expression did not identify \( Trp63 \) as a \( miR-203 \) target. In contrast, the recognition of the 3’UTR site of \( Trp63 \) by \( miR-203 \) was robustly detected by the HITS-CLIP (Figure 6E). Thus, it is likely that our identification of \( miR-203 \) targets based on the transcriptome analysis alone was conservative. More studies will be required to further integrate direct miRNA target capture with transcriptome analysis for miRNA target identification. Nevertheless, we still observed strong enrichment of \( miR-203 \) targets in genes required for cell proliferation. These high-confidence targets illustrate the broad targeting of cell proliferation by \( miR-203 \) and, suggest mechanisms by
which loss of miR-203 can facilitate the expansion of oncogenic Hras transformed cells. These
new insights point to the potential of utilizing miR-203 to simultaneously target multiple
effectors required for cell division in human cancers. We also note that miR-203 is known to
antagonize human papillomaviruses (HPV) [39]. Our identification of Pola1, a critical cellular
gene required for HPV genome replication, as a target of miR-203 may provide a potential
mechanism for miR-203-mediated antagonism against HPV infection. Finally, the relatively mild
requirement for miR-203 in slowly proliferating cells suggests that a reduction or increase in the
level of miR-203 may be well tolerated by most normal cells. This unexpected discovery could
be further explored to determine miR-203’s therapeutic potential in treating certain types of
epithelial cancers.

Materials and Methods

3’seq library construction:

3’seq libraries were constructed using methods described previously [19]. Briefly, 500 ng
of total RNA isolated from primary keratinocytes was poly-A purified (Dynabead® Invitrogen)
and chemically fragmented by treatment 95°C for 8 minutes in Fragmentation Buffer (Ambion).
Fragmented RNAs were then oligoDT primed with a P7T20V oligo and reverse transcribed
using SuperScript® III (Invitrogen)(see Supplementary File 2 for Oligo Sequences). Following
ethanol precipitation, ligation competent cDNAs were generated via second-strand synthesis
using RNase H and DNA Pol I enzymes, end-repaired using a mix of T4 DNA polymerase,
Klenow DNA polymerase and T4 Polynucleotide Kinase in the presence of dNTPs (NEB), and A-Tailed using Klenow 3’ to 5’ exo- in the presence of dATP (NEB). Ligation was then performed using the P5 Adaptor with T4 DNA Ligase for 1 hour (NEB). cDNA inserts 80-100 nt in length were isolated from 8% PAGE gels and subject to PCR amplification using RP1 and RT-Primer oligos for 12-16 cycles (Phusion® NEB). PCR products were isolated from PAGE gels and subject to 6 cycles of secondary PCR to introduce unique indices for multiplexed DNA sequencing on a HiSeq 2000 (Illumina).

Bioinformatics processing of 3seq data was performed as described previously, including adaptor trimming, read alignments, peak calling, peak filtering to remove internal polyA priming events, and transcript quantification with minor modifications [19]. Following read alignment to the mm10 genome, alignments from every library were grouped together for peak calling and 3’end filtering to create a master set of high confidence 3’ end peaks. From this 3’end dataset, reads counts for each library were obtained (See Supplementary File 1 for mapping statistics). To obtain transcript read counts, reads from all peaks that passed our internal priming filter and were 3’UTR localized were summed to obtain an overall transcript count for each transcript. Normalized transcript counts were calculated as Reads per Million reads mapped (RPM). To determine differential transcript expression, low-abundance transcripts with less than 10 reads in two libraries were first excluded then the remaining transcripts were analyzed with EdgeR with classical analysis parameters.

Small RNA-Seq:

Small RNA-Sequencing libraries were prepared using minor modifications of a previously published protocol that reduces RNA-ligation biases and enables accurate miRNA quantification [24]. Briefly, two micrograms of total RNA, was first ligated to a pre-adenylated 3’
linker (10 uM) using truncated T4 RNA Ligase 2 (20 Units NEB) in the presence of PEG-8000 (10% w/v) RNase-Out™ (20 Units Invitrogen) for 4 hours at room temperature. Ligation reactions were then resolved on 7 M Urea-15% PAGE gel to isolate the miRNA-3’ adaptor hybrids. Following overnight elution from the acrylamide gel in HSCB buffer (400 mM NaCl, 25 mM Tris-HCl pH 7.5, 0.1% SDS), ligation products were ethanol precipitated, resuspended in 5’ ligation reaction mixture containing a 5’ linker (10 uM), 1 mM ATP, Peg-8000 (20% w/v), 1x T4 Rna Ligase buffer (NEB), denatured for 5 minutes at 70°C, after which RNase-Out™ (20 units) and T4 RNA Ligase 1 (10 units) were added. Following ligation for 37 °C for 2 hours, cDNA was generated via reverse transcription using Superscript® III (Invitrogen) in the presence of a 3’adaptor specific primer. cDNA products were then subjected to 10-14 PCR cycles and resolved on 8% native PAGE gels. Libraries were then eluted as described above, precipitated, and submitted for sequencing on a HiSeq 2000 (Illumina).

Small-RNA sequencing reads were analyzed following previous described methods with the following modifications [24]. Adaptor Sequences were trimmed from reads using CutAdapt using default parameters. Reads were then further trimmed to remove the randomized adaptor dinucleotides on the 5’ and 3’ end. Read were next aligned to a database of mouse miRNA sequences (miRbase) using blastn with the following settings (blastn -word_size 11 -outfmt 6 -strand plus). Blast alignments were then parsed with custom python scripts to extract and count the best miRNA alignment for each read with a minimum read alignment of 18 nucleotides. miRNA counts for each library were then filtered to keep only miRNAs with a minimum count of 50 reads in two libraries, and analyzed for differential expression using EdgeR, with classical analysis parameters.

**Affymetrix microarray analysis:**
For the *miR-203* epidermal loss-of-function microarray analysis, RNA was isolated from total epidermal samples from two-pairs of *miR-203*+/+ and *miR-203*−/− animals at p4. For the doxycycline-inducible *miR-203* over-expression microarray analysis, RNA was isolated from two pairs of doxycycline induced or uninduced *Krt14-rtTA/ pTre-miR-203* animals using FACS sorting for *Krt14-H2B-GFP*+ cells as described previously [17]. The microarray analysis of *miR-203* overexpression in basal epidermis was previously published [17]. Subsequently total RNAs (500 ng) were processed and hybridized to the Mouse Genome 430 2.0 array (Affymetrix, USA) following the manufacturer’s instruction at the MCDB microarray facility. Microarray image files were processed using the R Bioconductor suite and Mas5 normalization. Probesets were then filtered to include only those probes with Present or Absent calls in at least two arrays. Probesets were then collapsed using the probeset with the maximum averaged probeset intensity to represent each GeneID. Log2 fold changes were then computed using the *limma* Bioconductor package.

**Ribosome Profiling:**

Ribosome profiling was performed on primary keratinocyte lysates using the ARTSeq™ Ribosome profiling kit (Epicentre). Briefly, lysates from a 10 cm dish of primary keratinocytes were isolated in the presence of cycloheximide (Sigma Aldrich, 50 μg/ml) and subject to limited RNase digestion (10 units) for 45 min at room temperature. RNase digestion was terminated by addition of 15 μl of Super-RNase IN™ (Ambion) followed by ribosome isolation using illustra™ MicroSpin™ S-400 HR Columns (GE Healthcare). Following RNA extraction and precipitation, rRNA was depleted using the Ribo-Zero Gold™ kit (Epicentre), with the remaining RNA then fractionated through 18% PAGE gels. RNA species 28-32 nt were isolated for adaptor ligation,
reverse transcription, circularization, and PCR amplification following the manufacturer’s
protocol. PAGE gel isolated PCR products were then sequenced on a HiSeq 2000 (Illumina).

Raw reads were first trimmed to removed 3’ adaptors using cutAdapt with default
parameters. Read were aligned to mm10 rRNA, tRNA, and ncRNA (Ensemble annotation)
databases using Bowtie (default settings) to exclude reads aligning to abundant rRNA, tRNA,
and ncRNA sequences. Unaligned reads were then aligned via Tophat using default settings,
with a supplied .gtf annotation file containing combined Refseq and Ensembl gene annotations
(iGenomes Illumina downloaded 9/4/2013). Uniquely aligned read counts were quantified across
each CDS using HTSeq Count (settings: -s yes -m union -t CDS) using the above-mentioned
GTF annotation database. Transcripts with low reads counts were excluded by only keeping
transcripts with at least 50 reads in at least two libraries. Filtered transcript reads count data was
then analyzed for differential expression using EdgeR with classical analysis parameters. To
calculate translation efficiency for each transcript, Reads Per Million Mapped (RPM) values
from 3Seq were divided by Reads per Million Mapped to coding-sequence for the Ribo-Seq. The
change in translation efficiency was then computed as the ratio of translation efficiency in the
miR-203⁻/⁻ and miR-203⁺/+ libraries.

Ago2 HITS-CLIP:

Ago2 hits-clip was performed as described with minor modifications [32]. 15 cm² dishes
of primary keratinocytes were irradiated twice at 200 mJ/cm with 254 nm UVC light. Following
irradiation cell lysates were harvested by scraping and stored at -80°C. Following thawing,
lysates were further lysed by trituration 3 times through pre-chilled 25 and 30 gauge needles.
Lysates were then treated with 10 μl Turbo™ DNase (Promega) and 5 μl RNase-Out™
(Invitrogen) per ml of lysate. Limited RNase digestion was performed using 10 μl per ml of
lysate of a 1:20 dilution of an RnaseA/T1 mix (Sigma/Ambion 1x mix = 3.33 μl RnaseA (2 μg/μl) with 6.66 μl RnaseT1 (1U/μl)). Crosslinked Ago2 was recovered via immunoprecipitation for 2 hours at 4°C with 3 μg of a monoclonal anti-mouse Ago2 antibody (Wako Chemicals USA clone 2D4) complexed with Protein-G Dynabeads™ (Invitrogen). Immunoprecipitates were washed twice with High-Salt Buffer and PNK buffer, then end-labeled with 25 μCi ³²P γ-ATP using PNK 3’ phosphatase minus (NEB) for 5 minutes at 37°C. After washing the beads as listed above, 5’ adaptor ligation was performed for two hours at room temperate using T4 RNA Ligase 1 with 10 uM 5’ RNA Linker, 20% PEG-8000 (w/v final), 1 mM ATP, and RNaseOut™ (Invitrogen). Beads were again washed twice with PNK buffer then resuspended in a phosphatase reaction with 5 μL FastAP (ThermoFisher) with RNaseOUT™. Following washing twice with PNK buffer, Protein-RNA complexes were eluted from the beads using 1x Novex® Loading buffer supplemented with 50 mM DTT at 70°C for 10 minutes. Protein-RNA complexes were then resolved on an 8% Bis-Tris Gel and transferred to nitrocellulose. Membranes were exposed to a phosphor screen for 1-2 hours to obtain an autoradiograph. Subsequently, RNA-protein complexes migrating in the 110-130 kD range were excised. RNA was recovered from the nitrocellulose using Proteinase K treatment followed by phenol-chloroform extraction and ethanol precipitation. Isolated RNA was ligated to a 3’linker using the same ligation reaction conditions as for 5’ ligation, and ligated RNA species were fractionated away from adaptor-adaptor products on 10% UREA Page gels. The RNA was eluted from the PAGE gel with HSCB buffer overnight at 4°C, then ethanol precipitated and resuspended for reverse transcription with SuperScript™ III (Invitrogen). cDNA products were then subjected to PCR amplification for 20-24 cycles and fractionated on an 8% native PAGE
gel. PCR products representing cDNA inserts of 20-50 nts were recovered and subject to sequencing on a HiSeq 2000 (Illumina)

HITS-CLIP reads were analyzed as follows. First, reads were processed to identify miRNA alignments using the same pipeline that we use for Small-RNA-Seq listed above. Reads not mapping to miRNAs were next processed as follows. Reads were trimmed with Cutadapt to remove adaptor sequences using default settings. To avoid PCR duplicates from biasing the analysis, duplicate reads were then collapsed to a single read using Fastx_collapser (default settings). The 5’ and 3’ adaptor sequences contain randomized dinucleotides on their 3’ and 5’ ends respectively, which were next trimmed from the reads. The reads were then aligned to the mm10 genome assembly using NovoAlign requiring a minimum alignment length of 25 nucleotides (settings –s 1 –t 85 –l 25) (Novocraft). All unique alignments from each library were then pooled to identify Ago2 HITS-CLIP clusters. Clusters were defined as two read alignments that overlap by a minimum of one nucleotide. Clusters were next annotated to gene features in a hierarchical manner in which clusters were annotated to protein coding RefSeq 3’ UTRs (with 5kbp extension allowed), RefSeq CDS regions, RefSeq 5’ UTRs, Ensemble ncRNA regions, and RefSeq intron regions. Clusters not found in these regions were annotated as intergenic clusters.

For predicting miRNA target sites, 3’UTR cluster sequences were searched for 6mer seed-sequence matches for miRNA species that accounted for 90% of miRNAs expressed in epidermis based on small-RNA sequencing. From this dataset 117 binding sites in 113 mRNAs were predicted be miR-203 target sites.

Meta-analysis of miR-203 targets:

The miR-203 overexpression datasets used in the meta-analysis included previously published microarrays from sorted H2B-GFP+ epidermal cells with transient miR-203 induction
(GSE45121), and microarrays from sorted H2B-GFPhi+ hair follicle progenitor cells with transient miR-203 induction described in this manuscript. The miR-203 knockout datasets used in the meta-analysis included the microarrays from p4.5 total epidermal samples from miR-203+/+ and miR-203−/− animals, ribosome profiling of miR-203+/+ and miR-203−/− primary keratinocytes, and 3seq of miR-203+/+ and miR-203−/− primary cultures. We also performed 3seq on miR-203−/− cells transformed with HrasG12V, however there was no enrichment for miR-203 seed matches in the upregulated transcripts based on CDF analysis, consistent with the low levels of miR-203 in HrasG12V transformed keratinocytes, and therefore this dataset was excluded from the expression meta-analysis (Data not shown). Gene Symbols were used to compare across microarray, 3Seq and Ribo-Seq datasets. Log2 fold changes were used to assess differential gene expression in each dataset. In total 6,407 genes were detectable in all the datasets, from which 294 satisfied the criteria of being upregulated in all the miR-203 knockout datasets and downregulated in the miR-203 overexpression datasets. Negative control datasets were constructed to analyze the enrichment of genes containing miR-203 seed matches with the following criteria, randomly selected genes from the list of detected transcripts in the meta-analysis, or genes that have a positive correlation to miR-203 expression (downregulated in miR-203 knockout datasets and upregulated in miR-203 overexpression datasets).

In order to compare gain-of-function and loss-of-function datasets in aggregate, a ranked correlation to miR-203 metric was calculated. Transcripts were assigned a rank in each knockout dataset with the most upregulated gene given a rank of 1. Transcripts were next assigned a rank in each overexpression dataset with the most downregulated gene given a rank of 1. The ranked values from each of the 5 datasets were then summed and ranked with the transcript most
upregulated upon miR-203 ablation and downregulated upon overexpression being assigned a value of 1.

**Hierarchical Clustering, GO-Term Analysis, GSEA, 3’UTR motif-searching:**

Normalized transcript abundances (expressed as Reads Per Million) were used for mean-centered unsupervised hierarchical clustering using Cluster 3.0 software. For 3seq data, only transcripts with at least 2-fold change were selected and for miRNA data, only miRNAs with at least an expression level of 1000 RPM and at least a 2-fold change were selected for visualization in JavaTree View. GO-term enrichment analysis was performed by DAVID using Gene-IDs as input, with analysis being performed using GO Biological processes datasets. GSEA analysis was performed using ranked expression values for the displayed dataset, and genesets selected from the referenced publications. De novo motif searching was performed using the HOMER package to search for 7 or 8mer motifs in RefSeq 3’UTR sequences for selected gene-sets. For genes with multiple 3’UTR isoforms, the longest 3’UTR was selected for motif searching. miRNA seed-sequence searches in Ago2-HITS-CLIP clusters were performed with a Python script using regular expressions.

**Analysis of The Cancer Genome Atlas (TCGA) Data:**

Head and Neck SCC miRNA-Seq data was obtained from the TCGA https://tcga-data.nci.nih.gov/tcga/ (Download date 10/23/2014). Patient matched normal solid tissue and tumor miRNA quantification records were identified using custom R scripts (Regular Expression query for tumor and solid tissue normal samples respectively "TCGA-[0-9A-Z]{2}+[0-9A-Z]{4}+[0-9A-Z]{4}+0", "TCGA-[0-9A-Z]{2}+[0-9A-Z]{4}+[0-9A-Z]{4}+11"). The normalized reads_per_million quantification for miR-203 and miR-21 was then plotted to determine the relative expression in normal and tumor tissue samples.
Mouse strains and generation of the conditional miR-203 knockout mouse:

A gene targeting vector was constructed that contained 11 kbp homologous region surrounding the miR-203 locus (Figure 3-figure supplement 1). LoxP were inserted flanking the pre-miR-203 sequences, with a neomycin selection cassette flanked by Frt sites. The construct was electroporated into Cy2.4 ES cells (B6(Cg)-Tyr<c2J> genetic background). Positive clones were identified by Southern Blot analysis using a probe complementary to the 3’ end of the targeted homologous region. ES cells were injected into blastocysts and chimeras were screened based on white/black coat color selection. Upon obtaining germline transmission, the neo cassette was excised by breeding the F1 progeny to an Actb-Flpe line maintained on a C57BL/6J background (obtained from Jackson Labs). miR-203floxed animals were then bred to a EIIa:Cre line maintained on a C57Bl/6 background (obtained from Jackson Laboratory) or a Krt14:Cre line maintained on a mixed background (obtained from E. Fuchs Laboratory), to obtain germline or conditional ablation of miR-203. The EIIa:Cre transgene was removed from the germline miR-203 deleted line by backcrossing to a C57Bl/6 line and subsequently maintained on a C57Bl/6 background. pTre2-miR-203/Krt14-rtTA mice were generated as described previously and maintained on an FVB background [17]. Mice were bred and housed according to the guidelines of IACUC at a pathogen-free facility at the University of Colorado (Boulder, CO, USA).

Immunofluorescence, EdU detection, and miRNA in situ hybridization:

Frozen cryostat sections (8 μM) were fixed in 4% Paraformaldehyde for 10 minutes at room temperature, wash thrice with PBS, and blocked for 10 minutes using Gelatin Block (0.1% Triton X-100, 2% gelatin, 2.5% normal goat serum, 2.5% normal donkey serum, and 1% BSA in PBS). Primary antibodies, diluted in gelatin block, were then incubated overnight (See
Supplementary File 2 for antibody references). Following three washes with PBS, sections were incubated with appropriate Alexa-Fluor secondary antibodies (1:2,000) for 1-2 hours at room temperature. Following three washes with PBS, sections were stained with Hoescht Dye, and mounted in Anti-fade solution. miRNA in-situ hybridization for miR-203 was performed on frozen sections as described previously [18]. EdU detection was performed following manufacturer’s instructions, with the following parameters. P4 animals were IP injected with 50 μg/g EdU 4 hours prior to tissue harvest. Following EdU detection, the sections were blocked and probed with antibodies as described above. BrdU detection was performed as previously described, with the following parameters. Pregnant female mice were IP injected with 50 μg/g BrdU for 2 hours prior to embryo harvest in OCT compound. miR-203 in situ hybridization on FFPE mouse and human tumor samples was performed with the following modifications, after deparaffinization in xylenes, the tissue was treated with proteinase K (20 μg/ml) for an extended period of 20 min at an elevated temperature (37°C). Microscopy images were obtained using a Leica DM5500B microscope with either a Leica camera (brightfield) or Hamamatsu C10600-10B camera (fluorescence) and processed with the Leica image analysis suite, MetaMorph (MDS Analytical Technologies) and Fiji software. BrdU or EdU image quantifications were performed by counting the number of Krt5+/EdU or BrdU positive cells in randomly chosen microscopy fields. The length of the basement membrane was used to represent the length of the epidermis analyzed and was determined by tracing the basement membrane and calculating line length using Fiji software. Epidermal thickness was assessed by tracing a line tangential to the basement membrane and extending to the beginning of the stratum corneum and calculating the line length.

qPCR and western blotting:
qPCR was performed using the Qiagen miR-script RT system and a BioRad CFX-384 machine. Fold-changes were computed using the ΔΔCt formula normalized to sno25 and Hprt values. In all qPCR figures error bars denote standard errors of the normalized mean. Western blotting was performed using 20-40 μg of protein lysate run on 8-12% SDS-PAGE gels. Proteins were transferred to PVDF for detection of Pola1, β-tubulin (Tubb5), or Ccnd1. Primary antibodies were incubated in 5% BSA overnight and detected using HRP-conjugated secondary antibodies and Amersam ECL-Plus reagents (1:10,000). See Supplementary File 2 for antibody descriptions and dilutions. X-ray films were scanned and processed with Fiji software to calculate relative protein abundance.

**Primary Keratinocyte harvesting and cell culture, viral infections, Edu Flow Cytometry, and shRNA knockdown:**

Primary keratinocytes were isolated from neonatal mice using previously described methods with the following modifications [40]. Isolated skin was incubated on a solution of Dispase overnight at 4°C to dissociate the epidermis from the dermis. The following day epidermal sheets were incubated in 37°C Trypsin for 10 minutes to isolate keratinocytes. Primary keratinocytes were then plated in 6 cm or 10 cm dishes with E-Low media supplemented with 0.2 mM Calcium Chloride for the first 24 hours then switched to E-Low media with 0.05 mM Ca++. Lentiviral particles were produced by transient transfection of pLKO-shRNA constructs, PsPax.2, and pVSVG. 24 hours post transfection the media was changed to E-Low calcium. Retroviral particles were produced by transient transfection of pBabe-vector-puro, pBabe-HrasG12V-puro, pBabe-vector-neo, or pBabe-HrasG12V-neo, with PCL-Eco and pAdvantage packaging plasmids. Viral supernatant was harvested every 12 hours for up to 96 hours, pooled and filtered with 0.45 μM filter. Ad-eGFP or Ad-CREeGFP adenoviruses were
obtained from the Iowa Gene Transfer Core and used at MOI of 50. Retroviral and Lentiviral infections were performed 3-4 days after plating primary keratinocytes. Keratinocytes were selected with 2 μg/ml puromycin for 48 hrs or 50 μg/ml neomycin for 7 days, at which time non-infected cell cultures were non-viable. Spontaneously immortalized miR-203+/+, miR-203−/−, and miR-203fl/fl mouse keratinocyte cell lines were also generated via serial passage on mitomycin-C treated NIH-3T3 feeder cell culture layer and utilized for assays as noted in the text. Flow cytometry was performed as previously described, with minor modifications [17]. Cell cultures derived from pTre2-miR-203/Krt14-rtTA animals were treated with 5μg/ml doxycycline for 24 hours to induce miR-203 expression, pulsed with EdU (10 μM) for 30 minutes, harvested and analyzed according to the Click-iT® EdU Plus instructions on an BD Cyan™ flow cytometer (Invitrogen). For colony formation assays, 2,000 cells were split into individual wells of 6-well plates, cultured for 10-14 days, fixed with 4% PFA, and stained with 0.2% Crystal Violet in 70% Ethanol. For induction of miR-203, 24 hours after plating, cells were supplied with fresh media containing doxycycline at 5 μg/ml. Sigma-Aldrich™ TRC lentiviral shRNAs against Hbegf and Pola1 were obtained from the Functional Genomic Facility (University of Colorado at Boulder, sequences listed in Supplementary File 2).

3’UTR Luciferase Assays:

3’UTR reporter constructs were generated by PCR amplification of 3’UTRs from cDNA or gDNA and subcloning of the fragments into pGL3-Control (Promega) (primer sequences listed in Supplementary File 2). 2 ng renilla luciferase control, 20 ng pGL3-3’UTR reporter, and 380 ng of Krt14 empty vector, or Krt14-miR-203 were transiently cotransfected into keratinocytes in each well of a 12 well plate using Mirus Bio LT1 reagent (MirusBio). 24 hours later cell lysates were collected and renilla and firefly luciferase activity were measured using
Dual-Glo Luciferase Assay system (Promega) as described previously [18]. Data are represented as the ratio of firefly to renilla RFU values, normalized to PgI3-Control values. Error bars represent propagated standard deviations.

DMBA/TPA carcinogenesis:

DMBA/TPA carcinogenesis was performed as described previously [27]. The backskin of 7-9 week old miR-203+/+ and miR-203−/− mice was shaved. 48 hours later the backskin was painted with a single dose of 25 μg of DMBA in 200 μl of Acetone. Two weeks following DMBA treatment the mice began receiving bi-weekly treatments of 4 μg of TPA in Acetone. The number of palpable tumors of at least 1 mm in diameter, persisting for at least two weeks, was recorded weekly. Tumor diameters were measured using a digital caliper. Following 21 weeks of TPA treatment, mice were euthanized and tumors were collected for \(^{Hras^{Q61L}}\) genotyping, OCT embedding, and paraffin embedding.

\(^{Hras^{Q61L}}}\) genotyping of DMBA/TPA tumors:

Tumor DNA was isolated by incubating the tissue in a DNA Lysis Buffer (400 mM NaCl, 0.1% SDS, 1 mM EDTA, 1 μg/ml Proteinase K) at 55° C for 4 hours. Lysates were then vortexed and lightly centrifuged to liberate DNA from the partially-digest tumor tissue. The supernatant was then removed and subject to Phenol-Chloroform Extraction, followed by Isopropanol precipitation. Isolated DNA pellets were then resuspended in TE buffer and quantified by UV spectrophotometry (10 mM Tris pH 8.0, 1 mM EDTA). The \(^{Hras}\) gene was PCR amplified using primers that flank exon 2. Following amplification, the PCR reactions were digested with 5 units of XbaI restriction enzyme at 37° C. The reaction products were then resolved and visualized on a 3% agarose gel. DNA isolated from the tails of animals in the DMBA/TPA experiment was treated in parallel as a negative control for detection of the \(^{Hras^{Q61L}}}\) mutation.
Statistical Analysis:

Statistical analysis was performed using either R or Microsoft Excel. Statistical methods employed are indicated in the figure legends. Unpaired two-sided student t-tests were used to assess statistical significance unless indicated otherwise in the figure legends. For comparisons with multiple categories, ANOVA was used with Tukey’s HSD post-hoc test. Non-parametric Whitney-Mann U-Tests were used to assess significance for the tumor multiplicity measurements [27]. The hypergeometric test was used to assess the enrichment of gene lists in genome-wide studies. The Kolmogorov–Smirnov test was used to assess differences in cumulative distributions functions.

Data access:

All sequencing and microarray data are deposited in the Gene Expression Omnibus (GSE66056).

Author contributions

R.Y. conceived and supervised the study. K.R. performed the experiments. X.J.W., E.C.T., and D.R.R. provided critical advice for tumorigenesis experiments and important reagents for analysis of miR-203 expression patterns in mouse and human tumor samples. K.R. and R.Y. designed the experiments and wrote the paper.

Acknowledgements

We thank C. Yang, J. Gao and D. Feng for the generation of the miR-203 KO mouse. We thank members of the Yi laboratory for their critical discussions. We also thank P. Muhlrad for critical reading of the manuscript.
References


doi:10.1016/j.ccr.2011.10.014

doi:10.1038/nprot.2009.120


Figure Legends

Figure 1. Genome-wide profiling of the oncogenic Hras\textsuperscript{G12V} transformed miRNA and mRNA transcriptome in primary keratinocytes. (A) Schematic of experimental approach to identify deregulated mRNA and miRNA networks driven by oncogenic Hras\textsuperscript{G12V} using small-RNA Seq and 3Seq. The 3seq library preparation allows quantitative definition of poly-A+ RNA 3’ends and expression levels. (B) 3Seq reproducibly detects mRNA expression levels over 4 orders of magnitude. Pearson correlation coefficient displayed (C) Unsupervised hierarchical clustering of log-transformed mean-centered mRNA expression levels for all transcripts deregulated 2-fold by oncogenic Hras\textsuperscript{G12V} (n=2 libraries per condition) (D) Gene Ontology analysis of transcripts up and downregulated by Hras\textsuperscript{G12V} (2-fold change FDR <0.05) indicates enrichment for migratory and angiogenic processes, and suppression of keratinocyte differentiation. (E) GSEA analysis of selected gene-sets relevant to skin carcinogenesis. (F) Unsupervised hierarchical clustering of log-transformed mean-centered miRNA expression levels for all transcripts deregulated 2-fold by oncogenic Hras\textsuperscript{G12V} (n=2 libraries per condition) (G-H) Abundant miRNAs such as miR-203, miR-205, and miR-21 are strongly deregulated by oncogenic Ras.
Figure 2. miR-203 is strongly suppressed in mouse and human SCCs. (A-B) qPCR and small-RNA-Seq independently validate downregulation of miR-203 and upregulation of miR-21 driven by oncogenic Hras\textsuperscript{G12V} (n=3 biological rep. qPCR, n=2 small-RNA-Seq, mean +/- SEM displayed, * = p-value <0.05, Students T-Test two-sided). (C) Gene track and quantification the 3’end of the miR-203 primary transcript based on 3Seq. (D) miR-203 primary transcript detection by qPCR (n=3 biological replicates, Mean +/- SEM displayed, * = p-value < 0.05, ns = non-significant, Students T-Test two-sided). (E) miR-203 is downregulated in DMBA/TPA produced papillomas compared to normal adjacent tissue. Epi=epidermis, Der=dermis, T=tumor and S=stroma. The black lines denote the epidermal/dermal and tumor/stroma boundary (F) miR-203 is downregulated in malignant SCCs derived from Kras\textsuperscript{G12D}/Smad4\textsuperscript{cKO} and passaged in immunocompromised mice. (G-O) Reduced miR-203 expression is correlated with increasing malignancy in human skin SCC cancers. Panels G,J,M were taken from regions with more histologically normal regions to demonstrate successful miR-203 hybridization. (P) miRNA-Seq quantification from patient matched normal and tumor tissue obtained from the TCGA consortium data (bar indicates mean value, Students T-Test two-sided). Scale bar = 50 μm.

Figure 3. Loss of miR-203 modestly impairs embryonic epidermal development. (A) Schematic of miR-203 conditional allele generation and knockout strategy. (B) Validation of miR-203 ablation by qPCR from isolated epidermal samples (n=3 biological replicates, * = p-value <0.05, ns = non-significant, Student T-Test two-sided). (C) Validation of miR-203 ablation within the epidermis by in situ hybridization (Scale bar = 50 μm). (D) miR-203 knockout mice are visibly indistinguishable from wild type counterparts. (E-G) miR-203 ablation results in mild epidermal hyperplasia during embryonic development. (n=3 E16, n=4 E17, and n=3 p4 animals,
p-value provided in figure, Student T-test one-sided). (H) Representative Hematotoxylin and Eosin image from p4.5 animals, demonstrating restored normal skin morphology in neonatal animals. (Scale bars = 50 μm for inset and 100 μm for main images) (I) Epidermal differentiation is not compromised by loss of miR-203. (Scale bars = 50 μm) (J) m\textit{i}R-203\textsuperscript{-/-} Primary keratinocytes are more clonogenic than wild type counterparts (representative results from 3 experiments, * = p-value <0.05, student t-test two-sided). (K) Conditional ablation of miR-203 from passaged m\textit{i}R-203\textsuperscript{fl/fl} keratinocytes results in higher clonogenicity (representative results from n=3 independent experiments, mean +/- standard deviation displayed, * = p-value <0.05, student T-test, two-sided).

**Figure 4. Loss of miR-203 sensitizes mice to DMBA/TPA skin carcinogenesis.** (A) Representative images of tumors that were formed in the skin of WT and miR-203 null mice treated with DMBA/TPA. (B) miR-203\textsuperscript{-/-} mice have a larger tumor burden than miR-203\textsuperscript{+/+} counterparts (n=6 and 7 miR-203\textsuperscript{+/+} and miR-203\textsuperscript{-/-} animals respectively, mean +/- SEM displayed, £ = p-value <0.05, Whitney-Mann U-test one-sided) (C) miR-203\textsuperscript{-/-} tumor size distribution is similar to wild type animals (ns = non-significant, student T-test two sided, median displayed as bar). (D-E) miR-203\textsuperscript{+/+} and miR-203\textsuperscript{-/-} papillomas display similar morphologies and histology. (F) Proliferation and differentiation dynamics are similar between miR-203\textsuperscript{+/+} and miR-203\textsuperscript{-/-} tumors. (Scale bars = 50 μm).

**Figure 5. miR-203 antagonizes Hras\textsuperscript{G12V} driven keratinocyte proliferation.** (A) Hras\textsuperscript{G12V} transduced miR-203\textsuperscript{-/-} primary cultures are more colonogenic upon serial passage than wild-type controls (representative of n=2 independent experiments, mean +/- standard deviation displayed.
** = p-value <0.05, , ns = non-significant, Student T-Test two-sided) (B) qPCR of miR-203 induction upon addition of doxycycline in vector and \( Hras^{G12V} \) transduced cells. (mean +/- SEM displayed, n=3 biological replicates) (C) Restoration of miR-203 using a doxycycline inducible transgene, results in suppression of colony formation ability in \( Hras^{G12V} \) transduced and control keratinocytes. miR-203 was induced with doxycycline (5 μg/ml) 24 hours after plating (representative of n=3 independent experiments, mean +/- standard deviation displayed, * = p-value = <0.05, , ns = non-significant). (D) miR-203 restoration suppresses \( Hras^{G12V} \) driven S-Phase entry. miR-203 was induced for 24 hours prior to harvesting for flow cytometry. (n=3, mean +/- standard deviation displayed, * = p-value = <0.05).

**Figure 6. Comprehensive identification of miR-203 targets using genome-wide expression analyses and Ago2 HITS-CLIP.** (A) Schematic of genome-wide expression profiling datasets used in meta-analysis to identify bona-fide miR-203 targets. (B) Genes upregulated in all three miR-203 loss-of-function datasets (786 genes, no fold-change or p-value cut-off) were compared to genes downregulated in both miR-203 gain-of-function datasets (1704 genes, no fold-change or p-value cut-off) to identify a subset of genes with a strong inverse correlation to miR-203 expression (294 genes), of which 100 genes contained miR-203 7mer or 8mer seed sequence matches in their 3’UTRs. (C) Table demonstrating top 20 genes identified in meta-analysis ranked by negative-correlation to miR-203 expression. Genes colored in red contain 3’ UTR miR-203 7 or 8mer seed matches. (D) De novo motif searching identified an 8mer miR-203 seed motif, complementary to the miR-203 seed sequence enriched in the 3’UTR of candidate miR-203 target genes identified in the meta-analysis (294 genes). Table demonstrating enrichment for 7 or 8mer seed matches in the 3’UTR of candidate miR-203 target genes (294) over the
background seed distribution in primary keratinocytes, which is not seen for randomly selected
294 genes expressed in primary keratinocytes or a negative control gene set of genes upregulated
in miR-203 gain-of-function and downregulated miR-203 loss-of-function (353 genes). (E) Schematic of Ago2 HITS-CLIP and the identified miR-203 seed motif. (F) Diagram of genes detected by expression meta-analysis and Ago2-HITS-CLIP. Table of 21 high confidence miR-203 targets identified through expression meta-analysis and that have Ago2-HITS-CLIP 3’UTR peaks with miR-203 seed matches.

Figure 7. Hbegf and Pola1 are direct miR-203 target genes critical for keratinocyte proliferation. (A) 3’UTR luciferase reporter assays demonstrate that miR-203 directly targets Trp63 (positive control), Pola1, and Hbegf in keratinocytes (representative of n=3 independent experiments, mean +/- propagated standard deviation displayed, * = p-value < 0.05, , ns = non-significant, student T-Test two-sided) (B) Hbegf and Pola1 are upregulated in miR-203-/- isolated epidermis (p4) (n=8 and n=10, miR-203+/+ and miR-203-/- animals respectively, mean +/- SEM displayed, * = p-value < 0.05). (C-D) Western blots from lysates with miR-203 overexpression (48 hours) or miR-203 ablation. (E) shRNA knockdown of Hbegf or Pola1 impairs keratinocyte colony formation ability (representative of n=3 independent experiments, * = p-value <0.05, mean +/- standard deviation displayed). (F) Model for the mechanism of miR-203 in restricting Hras-initiated tumorigenesis.

Supplementary Figure Legends

Figure 3-figure supplement 1. Generation of a miR-203 conditional knockout mouse. (A) RNA-Seq and 3Seq detection of the miR-203 primary transcript (B) Schematic of the miR-203
conditional allele. H3 = HindIII (C) Southern Blot confirmation of founder miR-203 conditional knockout mice (D) Small-RNA-seq confirmation of miR-203 loss. Blue lines indicate two-fold change (E) qPCR quantification of miR-203 and miR-205, demonstrating that the miR-203 floxed allele does not alter microRNA levels. Error bars represent S.E.M from reactions performed from n=2 mice in technical triplicate, ns = non-significant.

**Figure 3-figure supplement 2. miR-203 expression in diverse mouse tissues.**

qPCR detection of mature miR-203 in various adult mouse organ tissues. Error bars represent S.E.M from reactions performed in technical duplicate.

**Figure 4-figure supplement 1. The Hras$^{Q61L}$ mutation is common in both miR-203$^{+/+}$ and miR-203$^{-/-}$ tumors.** (A) Representative end-point PCR followed by XbaI digestion. The Hras$^{Q61L}$ allele produces 120bp and 87bp product, whereas wild-type produces a 207bp product (B) Table with quantification of genotyping results. non-significant p > 0.05, Chi-Squared Test.

**Figure 6-figure supplement 1. Transcripts containing 3’UTR miR-203 seed matches are regulated by miR-203.** (A) Genes containing miR-203 seed matches are more likely to be downregulated upon miR-203 overexpression, or (B) upregulated upon miR-203 ablation (Panel E) (p-value <0.05 for comparison of 8mer match to no-match, K-S Test, one-sided). (C) Transcripts are ranked based upon aggregate fold-changes consistent with miR-203 regulation to produce a ranked expression correlation metric. Transcripts with 7 or 8mer miR-203 seed matches are more likely to be regulated by miR-203 modulation. (p-value <0.05) (See methods)
Figure 6-figure supplement 2. Ago2-HITS-CLIP in primary keratinocytes. (A) Example autoradiogram of isolated Ago2-RNA complexes, red box indicates region excised for sequencing (B) Proportion of miRNAs detected by Ago2-HITS-CLIP (C) Number of 3’UTR Ago2 peak containing seed sequences from miRNA families accounting for 90% of all miRNAs in p4 epidermis. (D) Genome-wide distribution of Ago2 peaks and reads (E) Comparison of miRNAs detected by Ago2-HITS-CLIP and small-RNA-Seq.

Figure 6-figure supplement 3. Ago2 HITS-CLIP 3’UTR peaks are enriched in keratinocyte miRNA seed matches, including miR-203. (A) Position of miRNA seed matches for miRNAs highly expressed in total epidermal samples. The peak summit represents nucleotide position 0 (C) De novo motif searching identifies most enriched 8mer motifs in 3’UTR peaks.

Figure 6-figure supplement 4. Predicted miR-203 targets based on HITS-CLIP are regulated by miR-203. (A) Cumulative distributions of miR-203 targets or not targeted transcripts based on HITS-CLIP in miR-203 overexpression datasets. (B) Cumulative distributions of miR-203 targets or not targeted transcripts based on HITS-CLIP in miR-203 knockout datasets (C) Ranked analysis of miR-203 targets detected by HITS-CLIP, based on 6,7,8mer seed only, or through meta-analysis (p value displayed on plots).

Figure 6-figure supplement 5. miR-203 targets do not display translation efficiency changes upon miR-203 ablation. (A) Comparison of translation efficiency for miR-203 targets (red) and non-targeted transcripts (blue) identified by expression meta-analysis. (B) Quantification of the change in translation efficiency in miR-203 KO samples for miR-203 targets identified by
expression meta-analysis. (C) Comparison of translation efficiency for miR-203 targets identified by Ago2-HITS-CLIP (red) and non-targeted transcripts (blue). (D) Quantification of the change in translation efficiency in miR-203 KO samples for miR-203 targets based on Ago2-HITS-CLIP.

Figure 7-figure supplement 1. A subset of miR-203 targets are upregulated by Hras<sup>G12V</sup>. (A) miR-203 target genes identified through meta-analysis containing miR-203 seed matches, are more likely to be upregulated upon Hras<sup>G12V</sup> expression in primary keratinocytes than non-targeted transcripts (p-value = <0.05, K-S Test one-sided). (B) Top 20 miR-203 targets, based on expression meta-analysis, ranked by upregulation by Hras<sup>G12V</sup>, genes shown in red are also identified by Ago2-HITS-CLIP.

Figure 7-figure supplement 2. Hbegf and Pola1 are required for keratinocyte growth potential in Hras<sup>G12V</sup> transformed miR-203<sup>+/+</sup> and miR-203<sup>-/-</sup> cultures. (A) Colony formation assays of established miR-203<sup>+/+</sup> cultures stably infected with pbabe-Hras<sup>G12V</sup>-neo, followed by Pola1 and Hbegf knockdown. (B) Colony formation assays of established miR-203<sup>+/+</sup> cultures stably infected with pbabe-Hras<sup>G12V</sup>-neo, followed by Pola1 and Hbegf knockdown. (p-value determined by ANOVA with Tukey HSD correction, n=3 wells per experiment).

Supplementary Files


Supplementary File 2. Primers, Antibodies, and shRNAs used in this study.
Source Data Files

Figure 1-source data. Log2 Fold Changes for transcripts up or down regulated 2-fold with FDR <0.05 in Hras<sup>G12V</sup> transformed keratinocytes.

Figure 6-source data 1. GO-analysis of selected miR-203 datasets.

Figure 6-source data 2. Putative miR-203 targets detected in the Expression Meta-analysis.
**Figure 1**

**A**
Isolate primary mouse keratinocytes

Infect with Hras^{G12V} or control virus and select for resistance to puromycin

Harvest for small-RNA Seq and PolyA-Selected 3-Seq

**B**
- **3seq**
  - PolyA select RNA and Fragment
  - OligoT primed cDNA Synthesis
  - Second Strand Synthesis and End Repair
  - Size Select and PCR Enrich
  - 3seq Tags

**C**
**3seq**
- Vector
- Hras^{G12V}

1818 genes

**D**
**Upregulated Genes (687)**
- biological adhesion
- cell adhesion
- vasculature development
- cell migration
- blood vessel morphogenesis
- response to wounding
- blood vessel development
- cell motion
- extracellular structure organization

**Downregulated Genes (494)**
- epidermis development
- epidermal cell differentiation
- ectoderm development
- keratinocyte differentiation
- epithelial cell differentiation
- epithelium development
- keratinization
- monosaccharide metabolic process

**E**
- **Hras Pathway**
- **SCC Cancer Stem Cell Signature**
- **c-Fos Targets in Ras Model of Skin Carcinogenesis**

**F**
**miRNA-Seq**
- Vector
- Hras^{G12V}

50 miRNAs

**G**
Log2 Normalized Hras^{G12V} RPM

**H**
<table>
<thead>
<tr>
<th>miRNA</th>
<th>Fold Change</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>mmu-miR-99a-5p</td>
<td>-9.83</td>
<td>3.60 x 10^{-6}</td>
</tr>
<tr>
<td>mmu-miR-335-3p</td>
<td>-7.86</td>
<td>3.00 x 10^{-6}</td>
</tr>
<tr>
<td>mmu-miR-203-5p</td>
<td>-5.87</td>
<td>2.95 x 10^{-4}</td>
</tr>
<tr>
<td>mmu-miR-203-3p</td>
<td>-5.22</td>
<td>3.71 x 10^{-3}</td>
</tr>
<tr>
<td>mmu-miR-335-5p</td>
<td>-4.70</td>
<td>1.16 x 10^{-3}</td>
</tr>
<tr>
<td>mmu-miR-24-3p</td>
<td>-4.20</td>
<td>1.70 x 10^{-2}</td>
</tr>
<tr>
<td>mmu-miR-23b-3p</td>
<td>-3.62</td>
<td>1.37 x 10^{-2}</td>
</tr>
<tr>
<td>mmu-miR-24-2-5p</td>
<td>-2.89</td>
<td>8.31 x 10^{-3}</td>
</tr>
<tr>
<td>mmu-miR-205-3p</td>
<td>-2.83</td>
<td>1.70 x 10^{-2}</td>
</tr>
<tr>
<td>mmu-miR-7a-5p</td>
<td>2.68</td>
<td>1.70 x 10^{-2}</td>
</tr>
<tr>
<td>mmu-miR-322-5p</td>
<td>3.31</td>
<td>1.55 x 10^{-2}</td>
</tr>
<tr>
<td>mmu-miR-21a-5p</td>
<td>3.72</td>
<td>3.73 x 10^{-4}</td>
</tr>
<tr>
<td>mmu-miR-29b-3p</td>
<td>3.98</td>
<td>6.40 x 10^{-4}</td>
</tr>
<tr>
<td>mmu-miR-29a-3p</td>
<td>5.68</td>
<td>6.12 x 10^{-4}</td>
</tr>
<tr>
<td>mmu-miR-146a-5p</td>
<td>13.71</td>
<td>3.00 x 10^{-6}</td>
</tr>
</tbody>
</table>
**Figure 2**

### A

![Graph showing miR-203 in situ hybridization](image)

**B**

<table>
<thead>
<tr>
<th>miRNA</th>
<th>Mature miRNA Fold Change (small-RNA Seq)</th>
<th>Mature miRNA Fold Change (qPCR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-203</td>
<td>-5.22</td>
<td>-8.46</td>
</tr>
<tr>
<td>miR-205</td>
<td>-2.83</td>
<td>-3.06</td>
</tr>
<tr>
<td>miR-21</td>
<td>3.72</td>
<td>4.04</td>
</tr>
</tbody>
</table>

### C

**miR-203 primary transcript (3Seq)**

![Graph showing miR-203 3Seq expression](image)

### D

**Primary miRNA Fold Change**

<table>
<thead>
<tr>
<th>miRNA</th>
<th>miR-203 (3Seq)</th>
<th>miR-203 (qPCR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-203</td>
<td>-2.53</td>
<td>-2.72</td>
</tr>
</tbody>
</table>

### E

- **Normal Skin**
- **DMBA/TPA Papilloma**

### F

**SCC Passage 1**

**SCC Passage 2**

### G

**Human SCC**

- Well Differentiated (4/4)
- Moderately Differentiated (2/2)
- Poorly Differentiated (3/3)

### H

**miR-203 in situ hybridization**

### I

### J

### K

### L

### M

### N

### O

### P

**TCGA miRNA-Seq**

**Head and Neck SCC**

- **miR-203**
  - p-value = 0.0288
  - n = 43

- **miR-21**
  - p-value = 1.07e-17
  - n = 43
Figure 3

A) Schematic diagram showing the miR-203 expression and regulation.

B) Bar graph showing the relative miRNA levels of miR-203 in different genotypes.

C) Images of skin sections showing miR-203 expression in WT and KO mice.

D) Images of WT and KO mice at different developmental stages.

E) Immunofluorescence images of BrdU/Krt5 and EdU/Krt5 staining in WT and KO mice.

F) Bar graph showing the number of BrdU or EdU+ cells/mm in WT and KO mice.

G) Bar graph showing the epidermal thickness in WT and KO mice.

H) Images of skin sections showing Krt1/Krt5, Lor/Krt5, and Trp63/Krt5 expression in WT and KO mice.

I) Images of skin sections showing Krt1/Krt5, Lor/Krt5, and Trp63/Krt5 expression in WT and KO mice.

J) Bar graph showing the number of colonies in different genotypes.

K) Bar graph showing the relative miR-203 levels in different genotypes.

Legend:
- WT: Wild Type
- KO: Knock Out
- E16: Embryonic Day 16
- E17: Embryonic Day 17
- P4: Postnatal Day 4
- P4.5: Postnatal Day 4.5
- miR-203+/+: miR-203 homozygous wild type
- miR-203+/−: miR-203 heterozygous
- miR-203−/−: miR-203 homozygous knockout
- Ad-CRE: Adenovirus Cre recombinase
- Ad-GFP: Adenovirus GFP
- LoxP: Lox P site
- Frt: FRT site
- Cq: Cycle threshold
- μm: Micrometers
- p-values: p<0.05
- ns: Not significant
**Figure 5**

**A**

<table>
<thead>
<tr>
<th>Passage:</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-203&lt;sup&gt;+/+&lt;/sup&gt; Hras&lt;sup&gt;G12V&lt;/sup&gt;</td>
<td>![Image]</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>miR-203&lt;sup&gt;−/−&lt;/sup&gt; Hras&lt;sup&gt;G12V&lt;/sup&gt;</td>
<td>![Image]</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
</tbody>
</table>

**B**

![Graph showing mRNA levels](image)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Vector</th>
<th>Hras&lt;sup&gt;G12V&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>-Dox</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
</tr>
<tr>
<td>+Dox</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
</tr>
</tbody>
</table>

**C**

![Graph showing colony numbers](image)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Vector</th>
<th>Hras&lt;sup&gt;G12V&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>-Dox</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
</tr>
<tr>
<td>+Dox</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
</tr>
</tbody>
</table>

**D**

![Graph showing EdU+ cells](image)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Vector</th>
<th>Hras&lt;sup&gt;G12V&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>-Dox</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
</tr>
<tr>
<td>+Dox</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
</tr>
</tbody>
</table>
### miR-203 Knockout (KO) vs. miR-203 Overexpression

**A** Microarray analysis of purified epidermal progenitors (+DOX) and hair follicle progenitors (+DOX) showed differential expression patterns in miR-203 KO compared to miR-203 overexpression. The analysis revealed 1,410 genes upregulated and 294 genes downregulated in the KO condition.

**B** Genes containing miR-203 7 or 8mer 3'UTR sites were identified, contrasting with 1,410 genes not containing these sites.

**C** A Venn diagram illustrates the overlap of genes detected by Ago2-HITS-CLIP and expression meta-analysis.

**D** De Novo Motif analysis revealed a significant enrichment of a specific motif, with a P-value of $1 \times 10^{-13}$.

**E** Ago2 HITS-CLIP workflow: 7meG UV-Crosslink, Lyse RNase, Ago2 IP, 32P Radiolabel, Sequence cDNA Align and Identify Ago2 binding sites.

**F** Targets detected by expression meta-analysis and Ago2-HITS-CLIP overlapped, indicating a consistent pattern of gene expression.

---

**Additional Data**

- **Genes upregulated in miR-203 KO**
  - Trp63α
  - V0p1
  - Pola1
  - Tap1
  - Tax1bp3
  - Cav1
  - Gbe1
  - Ndufb2
  - Lnx2
  - Pgam5
  - Snhpx2b
  - Yars2
  - Ctps2
  - Ppap2c
  - Prmt6
  - Dhx36
  - Tprkb

- **Genes downregulated in miR-203 overexpression**
  - S100a9
  - Cb1
  - Ccnd1
  - Csl3
  - Snx18
  - Trpo1
  - Znmi4

- **Log2 Fold Change**
  - 2.00
  - 1.50
  - 1.00
  - 0.50
  - 0.00
  - -0.50
  - -1.00
  - -1.50
  - -2.00

**miR-203 Targets (294)**

- **Non-miR-203 Targets (353)**
- **Random Targets (294)**

---

**Additional Information**

- **Information content**
  - 0.0
  - 1.0
  - 2.0

- **Sequence**
  - CAAAUUUC

- **P-value**
  - $1 \times 10^{-13}$

---

**Figure 6**
Figure 7

A. Normalized Luciferase Activity

B. Rel. mRNA Levels (ΔΔCq)

C. Western Blot Analysis

D. miR-203

E. Colony Formation Assay

F. Schematic Representation

**miR-203**