

Signaling amplitude molds the *Ras* mutation tropism of urethane

Siqi Li¹ and Christopher M. Counter^{1*}

¹ Pharmacology and Cancer Biology, Duke University, Durham, North Carolina, USA

* corresponding author, email: count004@mc.duke.edu

Abstract

RAS genes are commonly mutated in cancers yet despite many possible mutations, cancers have a ‘tropism’ towards a specific subset. As driver mutations, these patterns ostensibly originate from normal cells. High oncogenic RAS activity causes oncogenic stress and different oncogenic mutations can impart different levels of activity. Here we show that changing rare codons to common in the murine *Kras* gene to increase translation shifts tumors induced by the carcinogen urethane from arising from canonical Q₆₁ to biochemically less active G₁₂ *Kras* driver mutations, despite the carcinogen still being biased towards generating Q₆₁ mutations. Loss of p53 to blunt oncogenic stress partially reversed this effect, restoring Q₆₁ mutations. Finally, transcriptional analysis revealed similar signaling amongst tumors driven by different mutations and *Kras* alleles. These findings suggest that the RAS mutation tropism of urethane is largely product of selection in normal cells for mutations promoting proliferation without causing oncogenic stress.

Impact statement

The bias towards specific *Kras* driver mutations during urethane carcinogenesis appears to arise predominantly from the selection of a narrow window of oncogenic signaling in normal cells.

Introduction

The RAS genes are mutated in a fifth or more of human cancers (Prior et al., 2020), which is well established to be tumorigenic (Pylayeva-Gupta et al., 2011). There are three individual RAS genes in humans, namely *KRAS*, *NRAS*, and *HRAS*, three primary sites mutated in human cancers, namely G₁₂, G₁₃, and Q₆₁, and six possible amino acid substitutions at each site arising from a point mutation. As such, there are 54 possible oncogenic mutations, not including rare non-canonical mutations (Hobbs et al., 2016). Mapping these mutations to human cancers reveals a distinct pattern or ‘tropism’ in which specific mutations are characteristics of individual cancer types (Li et al., 2018). The same can be said for other oncogenes, for example, the most frequent EGFR point mutation in glioblastoma is G₅₉₈V/A but L₈₆₁Q in lung adenocarcinoma; the most frequent IDH1 mutation in low-grade glioma is R₁₃₂H but R₁₃₂C in melanoma, and so on (Chang et al., 2016). While these mutational biases are well described, the mechanism responsible is not.

One hint to the mechanism underlying RAS mutation tropism is that oncogenic RAS mutations are thought to occur early. Focusing on lung cancer, in humans oncogenic *KRAS* mutations have been detected in premalignant lesions (Kanda et al., 2012) as well as in multiple regions within the same tumor (Zhang et al., 2014a), indicative an early origin (Wistuba and Gazdar, 2006). In mice, oncogenic *Kras* mutations are capable of initiating pulmonary tumors in carcinogen (McCreery and Balmain, 2017) and genetically engineered (Kwon and Berns, 2013) lung cancer models. As early driver mutations, it follows that RAS mutation tropism is a reflection of the normal cells in which the mutation first occurred. While by its very classification oncogenic RAS can induce proliferation (Pylayeva-Gupta et al., 2011), high oncogenic signaling through the MAPK effector pathway of RAS can paradoxically induce a stress response in normal cells mediated by the tumor suppressors p16 and p53, which leads to the growth arrest termed senescence (Munoz-Espin and Serrano, 2014). Indeed, hyperactivation of MAPK signaling via the combination of *Kras*^{G12V}, *Braf*^{D631A}, and loss of the remaining wildtype *Braf* allele activated p53 and impedes lung tumorigenesis, an effect rescued by pharmacologic inhibition of the MAPK pathway (Nieto et al., 2017).

Accumulating evidence suggests that different oncogenic mutations may impart biochemical differences to the RAS proteins (Munoz-Maldonado et al., 2019; Smith et al., 2013). Relevant to

this study, a G₁₂D mutation sterically inhibits the catalytic cleft (Parker et al., 2018) while Q₆₁R replaces the catalytic amino acid (Buhrman et al., 2010). A direct comparison of the G₁₂D mutation to the Q₆₁R mutation in Nras revealed the former has lower GTP loading and tumorigenic potential in the skin (Burd et al., 2014) and hematopoietic system (Kong et al., 2016). Indeed, a panel of G_{12/13} Kras mutants introduced by Cas9-mediated gene editing in the lung revealed widely different tumorigenic potentials between different mutants (Winters et al., 2017). Taken together, we hypothesize that the sensitivity of a normal cell to oncogenic signaling amplitude may dictate the type of mutation selected, or to put it another way, selection underpins RAS mutation tropism.

One challenge to testing this hypothesis is trying to backtrack to catch a single, ostensibly random mutagenic event in one gene from one cell, decades before manifesting as cancer in humans. However, in mice the moment of tumor initiation can be precisely defined as the point of carcinogen exposure, with the added benefit that carcinogens model the spontaneous nature of human cancers. Carcinogens can also model RAS mutation tropism. Case in point, the carcinogen urethane found in the fermented foods and alcoholic products (Gowd et al., 2018) primarily induces pulmonary tumors with a very specific Kras^{Q61L} or Kras^{Q61R} driver mutation, depending on the strain (Dwyer-Nield et al., 2010). These oncogenic mutations are also a strong match to the mutation signature of this carcinogen. Urethane-induced mutations conform to the A→T/G consensus sequence derived from comprehensive whole-exome sequencing urethane-induced tumors (Westcott et al., 2015), and the even more restricted CAN→CT/GN sequence determined shortly after urethane exposure by different ultra-sensitive sequencing approaches (Li et al., 2020; Valentine et al., 2020), which at the position CA₁₈₂A gives rise to a CT/GA mutation encoding Q₆₁L/R oncogenic mutations. We thus capitalized on the extreme mutational tropism of urethane for Kras^{Q61L/R}-mutant pulmonary tumors to elucidate the effect of a normal cell response on the selection of initiating mutations.

To examine the effect of signaling amplitude on the selection of initiating oncogenic mutations, we genetically manipulated the level of endogenous Kras (to increase oncogenic stress) or the expression of the aforementioned tumor suppressor p53 (to inhibit the cellular response to oncogenic stress) in mice exposed to urethane. In regards to the first genetic change,

we compared the native *Kras*^{nat} allele that is naturally enriched in rare codons and correspondingly poorly translated (Lampson et al., 2013) to the *Kras*^{ex3op} allele, in which 27 rare codons in exon 3 (which is not the site of oncogenic mutations) were converted to common, leading to roughly twice as much Kras protein in the lungs or derived cells of mice (Pershing et al., 2015). In regards to the second genetic change, we evaluated retaining or conditionally inactivating the *Trp53* gene in the lung, which has been shown to suppress oncogenic stress due to high mutant Kras expression in this tissue in vivo (Feldser et al., 2010; Junttila et al., 2010). Using this approach, we show here that the canonical Q₆₁L/R mutations are selected against in the more highly expressed *Kras*^{ex3op} allele in urethane-induced tumors, even though the carcinogen favors this mutation. Instead, biochemically less active G₁₂ mutations are detected, which upon the loss of p53, partially shifts back to Q₆₁L/R mutations. p53 loss also promoted the expansion of the normally rare G₁₂ mutants in the native *Kras* allele, which was accompanied by an imbalance between the mutant and wildtype transcripts suggestive of higher expression. Taken together, these data support a very narrow window of signaling conducive to tumor initiation, with p53 loss correcting this by suppressing oncogenic stress in more active mutations (Q₆₁L/R in *Kras*^{ex3op}) and promoting or permitting allelic imbalance of less active mutations (G₁₂ mutations in *Kras*^{nat}). In agreement, the various combinations of mutations and *Kras* alleles that were tumorigenic induced similar levels of oncogenic signaling, as measured by the transcriptional responses of downstream RAS target genes. In sum, we provide evidence that by modulating endogenous genes in vivo in a model of extreme RAS mutation tropism, selection for a narrow signaling amplitude appears to be the main driving force in the type of Kras driver mutations capable of initiating tumorigenesis.

Results

Higher Kras expression accelerates tumorigenesis in the absence of p53. To determine the effect of genetically manipulating oncogene-induced stress on the RAS mutation tropism of urethane, we crossed *Sftpc*^{CreER/CreER}; *Trp53*^{flox/flox} mice into a background with two native *Kras* alleles (*Kras*^{nat/nat}) or with one (*Kras*^{nat/ex3op}) or two (*Kras*^{ex3op/ex3op}) copies of the aforementioned

Kras^{ex3op} allele in which rare codons were altered to common in exon 3 (Pershing et al., 2015). The *Sftpc^{CreER/CreER};Trp53^{flox/flox}* genotype was chosen as injection of tamoxifen into such mice leads recombination and inactivation of the endogenous *Trp53^{flox}* alleles in the type II alveolar cells in the lung (Xu et al., 2012), which is reported to suppress oncogene-induced senescence/apoptosis induced by oncogenic Kras in this organ (Feldser et al., 2010; Junttila et al., 2010). The *Kras^{ex3op}* allele was chosen as a way to increase Kras protein expression of the native gene while leaving the rest of the locus almost entirely intact (Pershing et al., 2015). Cohorts of 21-30 mice from each of these three genotypes were injected with tamoxifen to inactivate the *Trp53* gene in the lung, followed by exposure to urethane via a single intraperitoneal injection to induce *Kras* mutations. One year later these mice were humanely euthanized, the number and size of tumors determined at necropsy (Figure 1A and Supplementary file 1), and the tumors removed and recombination of the *Trp53^{flox}* alleles confirmed by PCR (Figure 1-figure supplement 1A and Supplementary file 1). As a first step, we simply examined the effect of p53 loss on tumorigenesis when Kras expression was altered. In sharp contrast to the previous findings that the *Kras^{ex3op}* allele reduced urethane carcinogenesis (Pershing et al., 2015), the loss of p53 instead increased tumor burden in mice with at least one *Kras^{ex3op}* allele (Figure 1B), similar to what was observed in a *Cdkn2a* null background (Pershing et al., 2015). This appeared to be a product of more tumors (Figure 1C), with a trend towards larger tumors (Figure 1D) that reached statistical significance when the data were censored for large ($\geq 100 \text{ mm}^3$) tumors (Figure 1E). There was also a trend when the genotypes were subdivided into one or two *Kras^{ex3op}* alleles compared to the *Kras^{nat/nat}* background, although no difference was observed between *Kras^{nat/ex3op}* versus *Kras^{ex3op/ex3op}* genotypes (Figure 1-figure supplement 1B-E). This suggests that in the absence of p53, the *Kras^{ex3op}* allele promotes both the initiation and progression of urethane-induced lung tumors, consistent with p53 suppressing oncogene toxicity to allow oncogenic mutations in the *Kras^{ex3op}* allele to exert a more potent signal to drive tumorigenesis.

Loss of p53 reprograms the extreme RAS mutation tropism of urethane. To specifically address the effect of genetically inactivating the *Trp53* gene on the type of oncogenic mutations arising in the native *Kras^{nat}* versus the codon-optimized *Kras^{ex3op}* allele in tumors induced by urethane,

we compared the *Kras* mutation status in tumors from the above *Sftpc^{CreER/CreER};Trp53^{flox/flox}*, *Kras^{nat/ex3op}* mice injected with tamoxifen (termed *Trp53^{-/-}*), as the heterozygous status of the *Kras^{nat/ex3op}* background allows for the most direct comparison, to tumors from a parallel control cohort not injected with tamoxifen (termed *Trp53^{+/+}*) prior to urethane exposure (**Figure 2A** and **Supplementary file 1**). As mentioned above, the *Trp53^{flox}* alleles the were confirmed by PCR to be recombined in tumors from the *Trp53^{-/-}* background. The same analysis was performed on tumors from the *Trp53^{+/+}* cohort, which identified one tumor having a significant degree *Trp53^{flox}* recombination, which was excluded from the analysis of comparing *Trp53^{+/+}* versus *Trp53^{-/-}* mice (**Figure 2-figure supplement 1A**). Consistent with the role of p53 as a tumor suppressor during lung tumor progression (Feldser et al., 2010; Junttila et al., 2010), loss of p53 tracked with larger, but not more tumors (**Figure 2B** and **Figure 2-figure supplement 1B,C**). To examine the effect of *p53* deficiency on the RAS mutation tropism of urethane, we sequenced *Kras* derived from mRNA isolated from these lung tumors and screened for mutations at the three main hotspots of G₁₂, G₁₃, and Q₆₁. In complete agreement with the previous observation that the increased protein expression of the endogenous *Kras^{ex3op}* allele shifts the RAS mutation tropism of urethane from the canonical Q₆₁ to G₁₂ oncogenic mutations (Pershing et al., 2015), tumors with an oncogenic mutation in the *Kras^{ex3op}* allele from control *Trp53^{+/+}* mice similarly had G₁₂ oncogenic mutations in this allele (**Figure 2C**). Having established that urethane behaves identically as previously reported in this regard, we turned our attention to the types of mutations recovered in the *Kras^{ex3op}* allele from the urethane-induced tumors of the *Trp53^{-/-}* mice. Sequencing revealed that 40% hotspot oncogenic mutations in the *Kras^{ex3op}* allele were now detected at Q₆₁ (**Figure 2C**), or to put it another way, the loss of p53 shifted the mutations in the *Kras^{ex3op}* allele back to more of the canonical Q₆₁ mutations of urethane.

The mutation signature of urethane is not affected by the *Kras^{ex3op}* allele. To determine if the shift in oncogenic mutations from Q₆₁ in the *Kras^{nat}* allele to G₁₂ in the *Kras^{ex3op}*, and then back again in the *Trp53^{-/-}* background, resides at the level of the locus or with the amount of encoded protein, we tested whether urethane induces different mutations in the *Kras^{ex3op}* versus *Kras^{nat}* allele. We thus turned to the ultra-sensitive Maximum Depth Sequencing (MDS) assay (Jee et al.,

2016), which we adapted for the mammalian genome to detect urethane-induced mutations within days of carcinogen exposure (Li et al., 2020). Since only a short region of genomic DNA can be sequenced by this approach, it is not possible to track oncogenic mutations in exon 1 or 2 and also determine the identity of the *Kras* allele (*native* versus *ex3op*) based on the codon usage in exon 3. Thus, we compared mutations arising in the *Kras*^{nat/nat} versus *Kras*^{ex3op/ex3op} genotype. To ensure potent mutagenesis for detection purposes, these two strains were crossed into the pure 129 background, which is sensitive to urethane (Malkinson and Beer, 1983; Shimkin and Stoner, 1975), and injected three times instead of just once as above with either the vehicle PBS or urethane. Seven days later, before overt cell selection (Li et al., 2020), the mice were humanely euthanized and their lungs removed and subjected to MDS sequencing to determine both the mutation signature and the type of *Kras* driver mutations induced by urethane (Figure 3A and Supplementary file 2). Mutation frequencies based on MDS sequencing of *Kras* exon 1 and exon 2 were averaged for A→T transversions, log₁₀ transformed, and displayed in a heatmap format. This revealed a trend towards A→T transversions within the context of a 5' C in both the native and ex3op alleles of *Kras* specifically in urethane-exposed mice (Figure 3B), consistent with previously identified bias for urethane (Li et al., 2020; Valentine et al., 2020; Westcott et al., 2015). Moreover, the frequency of these CA→CT mutations, which give rise to Q₆₁L (CA₁₈₂A→CTA), are significantly higher than GG→GA mutations, which give rise to G₁₂D (GG₃₅T→GAT) and G₁₃D (GG₃₈C→GAC), in both *Kras*^{nat/nat} and *Kras*^{ex3op/ex3op} mice (Figure 3C). Similar results were found upon repeating the experiment with a single injection of urethane in the less sensitive 129/B6 mixed strain background using *Sftpc*^{CreER/CreER}; *Trp53*^{flox/flox}; *Kras*^{nat/nat} versus *Sftpc*^{CreER/CreER}; *Trp53*^{flox/flox}; *Kras*^{ex3op/ex3op} mice treated or not with tamoxifen (Figure 3-figure supplement 1A). The only exception was that far fewer CA→CT mutations were detected in general, and perhaps as a consequence, the difference between the frequency of CA→CT and GG→GA mutations was no longer significant, although GG→GA mutations were detected in both the urethane and PBS cohorts, suggesting false positives, although other interpretations are possible (Figure 3-figure supplement 1B,C). Thus, with the above proviso, urethane mutagenesis does not appear to be changed at the *Kras*^{ex3op} locus, at least within the detection limit of the MDS assay. These and the above findings support that the observed bias towards G_{12/13}-driver mutation in the *Kras*^{ex3op}

allele in urethane-induced tumors is a product of negative selection against the biochemically more active Q₆₁ oncogenic mutation, rather than a change in the mutational spectrum of urethane. In other words, oncogenic stress molds the type of RAS mutations conducive to initiate tumorigenesis irrespective of the mutational signature.

p53 loss tracks with an mRNA allelic imbalance of less active oncogenic mutations. Not only was there was a shift in the oncogenic mutations in the *Kras*^{ex3op} allele upon the loss of p53, as noted above, but surprisingly also in the *Kras*^{nat} allele. In more detail, we found that 60% hotspot mutations in the *Kras*^{nat} allele of urethane-induced tumors from the *Sftpc*^{CreER/CreER};*Trp53*^{flox/flox}; *Kras*^{nat/ex3op} mice in which the *Trp53* gene was recombined in the lung occurred at codon G₁₂, and in one case also G₁₃ (**Figure 2C**). We also note that the percentage of tumors that have *Kras* hotspot mutations is higher in p53^{-/-} mice (**Figure 4-figure supplement 1A**). In addition, even though oncogenic mutations are generally more frequent in the *Kras*^{ex3op} allele, the percentage of tumors with mutations in the native allele is higher in p53^{-/-} mice (**Figure 4-figure supplement 1B**). One interpretation of these findings is that the absence of p53 enhances the ability of G_{12/13} mutations to be productive. Given the above tight relationship between *Kras* expression and mutation type, we explored a possible relationship between p53 loss and higher expression of *Kras* alleles with a G_{12/13} mutation. To this end, we calculated the ratio of mutant to wildtype *Kras* mRNA based on the number of cDNA sequencing reads matched to the mutant or wildtype allele from the above analysis. This revealed a clear demarcation in *Kras* mRNA levels between those alleles with a Q₆₁ versus a G_{12/13} oncogenic mutation. In detail, a waterfall plot revealed the mutant:wildtype ratio was higher in G_{12/13}-mutant *Kras* alleles, with an average ratio of ~2 copies for the mutant *Kras* transcript to each copy of wildtype *Kras* counterpart. Conversely, the mutant:wildtype ratio was lower in Q₆₁-mutant *Kras* alleles, with an average ratio of ~0.8 copies for the mutant *Kras* transcript to each copy of wildtype *Kras* counterpart (**Figure 4A** and **Figure 4-figure supplement 1C,D**). The allelic imbalance in the G_{12/13}-mutant tumors appears to be important for tumorigenesis. Namely, cross referencing the mutant:wildtype *Kras* ratio to the mutation type and size of tumors revealed that G_{12/13}-mutant tumors with a mutant:wildtype ratio ≥ 1.5 are larger, reaching the same size of Q₆₁-mutant tumors (**Figure 4B**). This suggests

that to achieve optimal signaling amplitude, *Kras* with less active G_{12/13} oncogenic mutations undergo a selection for higher expression to compensate for the lower signaling, while if anything, the reverse was seen for more active Q₆₁ oncogenic mutations.

Allelic mRNA imbalance of G_{12/13}-mutant *Kras* alleles tracks with Ras signaling. To test the possibility that there is a selection for a specific signaling amplitude, we measured the degree of oncogenic RAS signaling through the MAPK pathway, the very pathway promoting proliferation (Drosten and Barbacid, 2020; Hymowitz and Malek, 2018; Ryan et al., 2015) or senescence (Munoz-Espin and Serrano, 2014), by assaying the transcriptional level by qRT-PCR of the three well-described downstream target genes *Dusp6* (Buffet et al., 2017; Kidger and Keyse, 2016; Unni et al., 2018; Zhang et al., 2010), *Egr1* (Esnault et al., 2017; McMahon and Monroe, 1995; Swarbrick et al., 2008), and *Fosl1* (Chung et al., 2017; Esnault et al., 2017; Gillies et al., 2017; Vallejo et al., 2017a; Vallejo et al., 2017b). Binning the relative expression of these three genes into *Kras* G_{12/13}-mutant tumors with a high (> 1.5) versus low (\leq 1.5) mutant:wildtype ratio revealed that *Kras* G_{12/13}-mutant tumors with a high ratio expressed of these three genes more strongly than those with a low ratio. Furthermore, when we compared the relative expression of these three genes in *Kras* Q₆₁-mutant tumors, *Dusp6* and *Egr1* were expressed higher compared to the *Kras* G_{12/13}-mutant tumors with a low ratio (**Figure 4C-E**). Plotting the signaling intensity against the mutant:wildtype cDNA ratio of individual tumors showed that signaling intensity is general correlated with the allelic ratio for G_{12/13} mutations (**Figure 4-figure supplement 1E-G**). As such, the increase in expression of the G_{12/13}-mutant *Kras* alleles, as measured by a high mutant:wildtype ratio, manifests as an increase in Ras signaling, matching that achieved by the more potent Q₆₁-mutated *Kras* alleles. To assess the effect of p53 on this relationship, we compared *Dusp6*, *Egr1*, and *Fosl1* expression between tumors with and without p53. This revealed higher expression of all three genes in the latter tumors, consistent with the loss of p53 permitting higher Ras signaling (**Figure 4F-H**). p53 loss also tracked with higher mutant:wildtype allelic ratio in tumors with *Kras* G_{12/13}-mutations in the *Kras*^{ex3op} allele (**Figure 4-figure supplement 1H**). Finally, we compared *Dusp6*, *Egr1*, and *Fosl1* expression in tumors of different sizes, which revealed higher Ras signaling in larger tumors, regardless of how this

achieved (**Figure 4-figure supplement 1I-K**). Collectively, these findings suggest that the absence of p53 expands the spectrum of driver *Kras* mutations to include less active mutations by permitting or even fostering an increase in their expression; once again favoring a model whereby an optimal oncogenic amplitude is selected to initiate tumorigenesis.

Discussion

Here we show that p53 loss reprograms RAS mutation tropism. On one hand, loss of this tumor suppressor shifts the canonical Q₆₁ oncogenic mutations normally detected the native *Kras*^{nat} allele of urethane-induced tumors towards the usually rare G_{12/13} mutations in conjunction with an increase in the mutant:wildtype ratio. On the other hand, the loss of p53 shifts the prevalence of G_{12/13} mutations in the codon optimized and more highly translated *Kras*^{ex3op} allele towards Q₆₁ mutations. The one common theme in both these shifts is oncogenic amplitude, which argues that the degree of oncogenic signaling dictates the type of mutation conducive to initiate tumorigenesis (**Figure 5**). In agreement, oncogenic signaling, as measured by the expression of known RAS downstream target genes, was similar between *Kras*^{ex3op} with Q₆₁ mutations and *Kras*^{nat} with G_{12/13} mutations exhibiting an increase in the mutant:wildtype ratio, suggesting that expression difference render Q₆₁ and G_{12/13} mutations fungible. While there is no question that the level of oncogenic signaling can influence tumorigenesis, what is new here is that the sensitivity of normal cells to oncogenic signaling underlies the type of oncogenic mutations selected, which is relevant to the RAS mutation tropism of human cancers. We also note that neither the loss of p53 nor the change in codon usage in *Kras* altered the mutation signature of urethane, arguing that a selection for an optimal oncogenic signal supersedes the mutational bias of the urethane. Such a finding speaks to the somewhat complexing discordance often observed between the mutation signature of tumors and the corresponding driver mutation in human cancers (Dietlein et al., 2020; Temko et al., 2018), even in cancers in which the mutational signature can be ascribed to a specific mutagenic process (Buisson et al., 2019).

There are four suppositions to the above model (**Figure 5**). First, Q₆₁ oncogenic mutations in *Kras* activate the MAPK pathway more potently than G_{12/13} mutations. In agreement, a Q₆₁R

mutation in *Nras* and/or *Kras* have a much (1,000 times) slower intrinsic hydrolysis rate, higher GTP-loading, and more robust activation of the MAPK pathway, and when activated in the skin or hematopoietic system in vivo, are more tumorigenic than the corresponding G₁₂D mutation (Burd et al., 2014; Kong et al., 2016; Pershing et al., 2015). We also find that the expression of downstream Ras target genes is higher in tumors with a Q₆₁ mutation compared to G_{12/13} mutation in the absence of an allelic imbalance. Second, increasing *Kras* level increases signaling. In agreement, the *Kras*^{ex3op} allele has been documented to express more protein in the lung (Pershing et al., 2015) and increase ERK phosphorylation in hematopoietic stem cells (Sasine et al., 2018). Altering rare codons to common in ectopic or endogenous human *KRAS* also leads to higher translation, *KRAS* protein expression, and/or MAPK activation (Ali et al., 2017; Fu et al., 2018; Lampson et al., 2013; Pershing et al., 2015; Peterson et al., 2020). Nevertheless, it remains possible that rare codons may be differentially translated (Quax et al., 2015) or higher *Kras* protein expression reduces MAPK signaling through negative feedback (Shin et al., 2009) specifically in the cell-of-origin of these pulmonary tumors. Third, high oncogenic signaling is a negative selection, and fourth, this selection is p53-dependent; two points well supported in multiple model systems (Munoz-Espin and Serrano, 2014). As such, it is reasonable to propose that oncogenic amplitude is a driving force in the selection of type of oncogenic *Kras* mutations initiating tumorigenesis (Figure 5).

We note that urethane has a bias towards CAN→CT/GN mutations that is absent in both strand orientations at codons encoding G₁₂ and G₁₃, and the mutagenic signature of urethane was the same in the *Kras*^{nat} and *Kras*^{ex3op} alleles. This suggests that G_{12/13} oncogenic mutations are extremely rare, with urethane either inducing these mutations directly at a low frequency, or inducing a cooperating mutation or non-mutational event that stimulates a rare, pre-existing *Kras*-mutant cell to expand. In either scenario, the very fact that tumors arise with G_{12/13} oncogenic mutations argues for potent selection of an extremely rare oncogenic mutation over the far more common mutations favored by the carcinogen. Furthermore, if these mutations are indeed pre-existing, such a finding elevates the importance of promotional events (secondary mutations, inflammation, etc.) in the process of tumor initiation (Balmain, 2020). We also note an imbalance in the mRNA ratio of the mutant:wildtype *Kras* alleles in some urethane-induced

tumors. Amplification of oncogenic *Kras* alleles has been previously documented in various mouse cancer backgrounds (Chung et al., 2017; Mueller et al., 2018; Westcott et al., 2015). What is new, however, is that this tracked with the type of oncogenic mutation, and even more interesting, in opposite directions. Namely we observed a bias towards a high mutant:wildtype mRNA ratio in G_{12/13}-mutant *Kras* alleles in the absence of p53. Few tumors show amplification of the G_{12/13} mutant allele based on genomic copy number analysis (**Figure 4-figure supplement 2**), pointing towards transcriptional or post-transcriptional mechanisms. One interpretation of these findings is that loss of p53 allows expansion of tumors with G_{12/13}-mutant *Kras* alleles that are more highly expressed, an observation borne out by signaling analysis. Increasing MAPK signaling by paradoxical activation with a BRAF inhibitor expands the number of cell types that are tumorigenic upon activation of the *LSL-Kras*^{G12D} allele in the lung (Cicchini et al., 2017), hence it is quite possible that a similar phenomenon is at play here as well. Such higher expression could be a product of natural variation of gene expression, a mutation caused by urethane, or p53 loss itself. We also observed a bias towards a low mutant:wildtype mRNA ratio in Q₆₁-mutant *Kras* alleles. Such a finding is consistent with this mutation inducing a strong oncogenic signal, and hence a selection against increased expression. Interestingly, this ratio was low irrespective of the p53 status. Perhaps there is an upper limit to oncogenic signaling, even in the absence of p53, that is tumor-promoting. In support, we note that G_{12/13}-mutant tumors with a high mutant:wildtype mRNA ratio had the same Ras signaling output as Q₆₁-mutant tumors, at least as assessed by the expression level of genes downstream of Ras. Indeed, RAS-mutant human tumors, which often have disruption of p53 and other tumor suppressors, typically do not have BRAF or EGFR mutations and vice versa outside of drug resistance settings (Sanchez-Vega et al., 2018). Moreover, co-expressing oncogenic RAS with oncogenic BRAF or EGFR can be growth suppressive (Cisowski et al., 2016; Unni et al., 2018).

In conclusion, multiple independent lines of investigation support a model whereby the sensitivity of a normal cell to a narrow range of oncogenic signaling dictates the nature of mutation conducive to tumor initiation. Too low, a mutation fails to induce proliferation, as supported by the finding that less active G₁₂ mutants are rarely recovered in the native *Kras* allele unless accompanied by an allelic imbalance permitted or fostered by the absence of p53. Too

high, a mutation induces arrest, as supported by the finding that more active Q₆₁ mutants are not recovered from the more highly expressed *Kras*^{ex3op} allele unless oncogene-induced senescence is suppressed by the loss of p53 (**Figure 5**). Moreover, this selection pressure far exceeds the mutational bias of the mutagen, as supported by the observations that G₁₂ mutations fail to match the urethane mutational signature, yet nevertheless arise at a high frequency in urethane-induced tumors in a *Kras*^{ex3op} background. Such findings speak directly to the extreme mutational bias of urethane-induced tumors, and perhaps more broadly, provides mechanistic insight into RAS mutation tropism in human cancers.

Materials and Methods

Key Resource Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
Tamoxifen	Sigma	Cat# T5648
Corn oil	Sigma	Cat# C8267
Urethane	Sigma	Cat# U2500
RLT buffer	Qiagen	Cat# 79216
β -mercaptoethanol	Thermo Fisher Scientific	Cat# 21985023
proteinase K	New England Biolabs	Cat# P8107S
Trizol LS	Thermo Fisher Scientific	Cat# 10296010
dNTP	New England Biolabs	Cat# N0447S
Agarose	EMD Millipore	Cat# 2120-OP
Tris	EMD Millipore	Cat# 9210-OP
EDTA	VWR	Cat# 97061-406
Sodium dodecyl sulfate	Sigma	Cat# L4509
Phenol	Sigma	Cat# P1037
Chloroform	Macron Fine Chemicals	Cat# 4440-04
Ethanol	VWR	Cat# 89125-190
RNase A	Sigma	Cat# R4642
EcoRV	New England Biolabs	Cat# R3195
EcoRI	New England Biolabs	Cat# R3101
Xmnl	New England Biolabs	Cat# R0194
Exonuclease I	New England Biolabs	Cat# M0293
10X exonuclease I buffer	New England Biolabs	Cat# B0293S
Sodium Chloride	EMD Millipore	Cat# SX0420
Potassium chloride	VWR	Cat# BDH0258
Potassium phosphate monobasic	Sigma	Cat# 795488
Sodium phosphate dibasic	Sigma	Cat# RDD022
Critical Commercial Assays		
iScript cDNA synthesis kit	Bio-Rad	Cat# 1708890
Platinum Taq Polymerase	Thermo Fisher Scientific	Cat# 10966083
QIAquick PCR Purification Kit	Qiagen	Cat# 28104
Q5 Hot Start High-Fidelity DNA Polymerase	New England Biolabs	Cat# M0493
Agencourt AMPure XP	Beckman Coulter	Cat# A63880
iTaq Universal SYBR Green Supermix	Bio-Rad	Cat# 1725120
PrimeTime Gene Expression Master Mix	Integrated DNA Technologies	Cat# 1055770
Deposited Data		
Tumor sequencing	This study	SRA: PRJNA663179
MDS	This study	SRA: PRJNA663179
Experimental Models: Organisms/Strains		

Mouse: 129 - <i>Kras</i> ^{ex3op}	Persing et al., 2015 ; MMRRC	Stock 050601-UNC
Mouse: B6.129 - <i>Trp53</i> ^{flox/flox}	The Jackson Laboratory	Stock 008462
Mouse: B6.129 - <i>Sftpc</i> ^{CreER/CreER}	Xu et al., 2012 , gift from Mark Onaitis	N/A
Mouse: B6.129 - <i>Sftpc</i> ^{CreER/CreER} ; <i>Trp53</i> ^{flox/flox} ; <i>Kras</i> ^{ex3op}	This study	N/A
Oligonucleotides		
<i>Kras</i> ^{ex3op} genotyping F - TGGTAGGGTAGAACTAGGATTC	This study	N/A
<i>Kras</i> ^{ex3op} genotyping R - GAGTACACAGAGAGACCATTTC AAC	This study	N/A
<i>Trp53</i> genotyping F - CACAAAAACAGGTTAAACCCA	This study	N/A
<i>Trp53</i> genotyping WT R - AGCACATAGGAGGCAGAGAC	This study	N/A
<i>Trp53</i> genotyping Del R - GAAGACAGAAAAGGGGAGGG	This study	N/A
<i>Sftpc</i> genotyping F - GCTTCACAGGGTCGGTAG	This study	N/A
<i>Sftpc</i> genotyping R - GAGGCACCGCTCCGCGAG	This study	N/A
<i>Sftpc</i> genotyping CreER R - CAACTCACAACGTGGCACTG	This study	N/A
Tumor sequencing primers	Supplementary file 3	N/A
qPCR primers	Supplementary file 3	N/A
MDS assay primers	Supplementary file 3	N/A
Software and Algorithms		
fastq-join	https://usegalaxy.org/	Galaxy Version 1.1.2-806.1
Filter by Quality	https://usegalaxy.org/	Galaxy Version 1.0.0
Trim	https://usegalaxy.org/	Galaxy Version 0.0.1
Filter sequences by length	https://usegalaxy.org/	Galaxy Version 1.1
Group	https://usegalaxy.org/	Galaxy Version 2.1.4
Barcode Splitter	https://usegalaxy.org/	Galaxy Version 1.0.0
PEAR pair-end read merger	Zhang et al., 2014b	https://cme.hits.org/exelixis/web/software/pear/
Morpheus	https://software.broadinstitute.org/morpheus	N/A
Prism 6	GraphPad	N/A

362

363

364

Mice

Kras^{ex3op} mice were previously described (Pershing et al., 2015), *Trp53^{flox/flox}* mice (Marino et al., 2000) were purchased from The Jackson Laboratory (Jax # 008462), and *Sftpc^{CreER/CreER}* mice (Xu et al., 2012) were kindly provided by Mark Onaitis (University of California at San Diego). *Sftpc^{CreER/CreER};Trp53^{flox/flox};Kras^{ex3op}* mice were from a mixed 129 x C57BL/6 background. *Kras^{ex3op}* mice used for mutagenesis studies were from a pure 129 background. Mice were genotyped using the following primers:

Kras native and *ex3op* alleles:

Kras F: 5'-TGGTAGGGTAGAACTAGGATTC-3'

Kras R: 5'-GAGTACACAGAGAGACCATTTC AAC-3'

Products: 504 bp (*Kras^{ex3op}*) or 614 bp (*Kras^{nat}*)

Sftpc^{CreER} allele:

Sftpc F: 5'-GCTTCACAGGGTCGGTAG-3'

Sftpc R: 5'-GAGGCACCGCTCCGCGAG-3'

Spc-CreER R: 5'-CAACTCACAACGTGGCACTG-3'

Products: 550 bp (*Sftpc*) or 500 bp (*Sftpc^{CreER}*)

Trp53 allele:

Trp53 F: 5'-CACAAAAAACAGGTAAACCCA-3'

Trp53 WT R: 5'-AGCACATAGGAGGCAGAGAC-3'

Trp53 Del R: 5'-GAAGACAGAAAAGGGGAGGG-3'

Products: 288 bp (wildtype *Trp53*), 370 bp (unrecombined *Trp53^{fl}*), and 612 bp (recombined *Trp53^{fl}*).

All animal experiments were approved by Duke IACUC.

Tumor experiments

A 20 mg/ml tamoxifen (Sigma T5648) solution was made by dissolving tamoxifen in corn oil. Six- to eight-week-old mice of the indicated genotypes were injected intraperitoneally with 0.25 mg/g body weight tamoxifen every other day for four doses. Urethane (Sigma U2500) was dissolved in PBS and injected intraperitoneally at dose of 1 mg/g body weight 1 week after the last injection of tamoxifen. ~12 months after urethane injection, all mice were humanely euthanized, after which lung tumors were counted, measured, and microdissected for RNA and DNA extraction. The 12 month time point was chosen based

on similar previous studies (Dwyer-Nield et al., 2010; Gurley et al., 2015; Miller et al., 2003; You et al., 1989) but also to take into account a potentially longer tumor latency in the mixed 129 x C57BL/6 background (Malkinson and Beer, 1983; Shimkin and Stoner, 1975). Tumor volume was calculated as $\frac{1}{2}$ (length \times width²). Tumor burden was calculated as the sum of tumor volumes.

Mutagenesis experiments

Six- to eight-week-old *Kras^{nat/nat}* or *Kras^{ex3op/ex3op}* mice were intraperitoneally injected daily for three days with either urethane or PBS as above. These mice were humanely euthanized 1 week later and the lungs collected for the extraction of genomic DNA. Alternatively, six- to eight-week-old *Sftpc^{CreER/CreER};Trp53^{flox/flox};Kras^{nat/nat}* or *Sftpc^{CreER/CreER};Trp53^{flox/flox};Kras^{ex3op/ex3op}* mice were or were not injected intraperitoneally with 0.25 mg/g body weight tamoxifen every other day for four doses. One week after the last injection of tamoxifen, the same mice were intraperitoneally injected with one dose of either urethane or PBS. The mice were humanely euthanized 1 month later and the lungs collected for the extraction of genomic DNA.

RNA and DNA extraction

Tumors or normal lung tissues were lysed in RLT buffer (Qiagen 79216) with 1% β -mercaptoethanol (Thermo Fisher Scientific 21985023) and 5 units/ml proteinase K (New England Biolabs P8107S) at 55°C for 30 minutes. RNA and DNA were then extracted from the lysate using Trizol LS (Thermo Fisher Scientific 10296010) following the manufacturer's instructions. RNA was subsequently converted to cDNA using iScript cDNA synthesis kit (Bio-Rad 1708890) following the manufacturer's instructions.

PCR and sequence analysis of tumors

Kras exon 1 to 3 was amplified from the cDNA of tumor or normal lung tissue using nested PCR. Primers are listed in Supplementary file 3. PCR reactions were as follows:

PCR1: 1 μ l cDNA, 0.5 μ l of 10 μ M forward primer and 0.5 μ l reverse primer, 2 μ l of 2.5 mM dNTP (New England Biolabs No447S), 0.75 μ l of 50 mM MgCl₂, 2 μ l of 5X buffer, and 0.1 μ l Platinum Taq Polymerase (Thermo Fisher Scientific 10966083) in a total

volume of 25 μ l. PCR cycles were as follows: one cycle at 94°C for 2 min and 25 cycles at 94°C for 30 seconds, 56°C for 30 seconds, 72°C for 30 seconds.

PCR2: 2.5 μ l of PCR1 reaction, 0.5 μ l of 10 μ M forward primer and 0.5 μ l reverse primer, 2 μ l of 2.5 mM dNTP, 0.75 μ l of 50 mM MgCl₂, 2 μ l of 5X buffer, and 0.1 μ l Platinum Taq Polymerase in a total volume of 25 μ l. PCR cycles were as follows: one cycle at 94°C for 2 min and 20 cycles at 94°C for 30 seconds, 58°C for 30 seconds, 72°C for 35 seconds. 10 μ l product for PCR2 was then analyzed with gel electrophoresis to check PCR efficiency. For samples with no product, *Kras* exon 1 and 2 was amplified from the DNA of the same tumor separately using nested PCR. Primers were listed in Supplementary file 3. PCR reactions were as follows:

PCR1: 1 μ l cDNA, 0.5 μ l of 10 μ M forward primer and 0.5 μ l reverse primer, 2 μ l of 2.5 mM dNTP, 0.75 μ l of 50 mM MgCl₂, 2 μ l of 5X buffer, and 0.1 μ l Platinum Taq Polymerase in a total volume of 25 μ l. PCR cycles were as follows: one cycle at 94°C for 2 min and 25 cycles at 94°C for 30 seconds, 53°C for 30 seconds, 72°C for 15 seconds.

PCR2: The reactions were comprised of 5 μ l PCR1, 0.5 μ l of 10 μ M forward primer and 0.5 μ l reverse primer, 2 μ l of 2.5 mM dNTP, 0.75 μ l of 50 mM MgCl₂, 2 μ l of 5X buffer, and 0.1 μ l Platinum Taq Polymerase in a total volume of 25 μ l. PCR cycles were as follows: one cycle at 94°C for 2 min and 20 cycles at 94°C for 30 seconds, 57°C for 30 seconds, 72°C for 18 seconds.

Products from PCR2 were pooled and purified with Ampure XP beads according to the manufacturer's protocol (Beckman Coulter A63880). The library was sequenced using MiSeq v2 Nano 250 PE at Duke Center for Genomic and Computational Biology. All primers were synthesized by Integrated DNA Technologies.

Sequencing data analysis

Raw sequencing data were uploaded to usegalaxy.org (Jalili et al., 2020). For analysis of amplicon from cDNA and exon 2 from genomic DNA, paired-end reads were joined with fastq-join tool, with the maximum percentage difference between matching segments set at 20% and the minimum length of matching segments set at 10. For analysis of amplicon from exon 1 from genomic DNA, only read 1 was used. The joined-read for cDNA and exon 2, or read 1 for exon 1, were then processed with Filter by Quality Tool, requiring 75% of bases have quality equal to or higher than the cut-off value of 20. For the analysis of

amplicon from cDNA, the forward and reverse index was then extracted from the filtered reads through the Trim Tool. Bases encoding codons 12, 13, 61, as well as nucleotide 96 (where a SNP exists that differentiates 129 from B6 strains) and 8 nucleotides in the middle of exon 3 (where the sequence differs between native and ex3op versions of *Kras*) were appropriate were also extracted through the Trim Tool. Collapsing by the index (representing individual samples) and counting the reads for each variant of the extracted region-of-interest were then performed by the Group Tool. Mutation and allele information were then assigned to each extracted region-of-interest by comparing the sequences of the extracted region-of-interest with reference sequences for all possible missense mutations in codon 12, 13, 61, SNP at nucleotide 96, and either native or codon-optimized exon 3 using the Excel program. The fraction of each variant of the extracted region was calculated by dividing the counts for that variant by the total number of counts per sample. For analysis of amplicon from genomic DNA, quality filtered-reads were split into separate files using the Barcode Splitter Tool and 5' index. Each of the files were then trim from 3' end to expose the 3' index. Trimmed files were further split into separate files corresponding to different samples using the Barcode Splitter Tool and 3' index. For each sample, bases covering codons 12 and 13 in exon 1 and codon 61 in exon 2 were extracted through the Trim Tool. The mutation was assigned to extracted bases by comparing them against reference sequences for all possible missense mutations in codon 12, 13, and 61 using the Excel program. The fraction of each variant was calculated by dividing the counts for that variant by the total number of counts per sample. Samples with less than 30 total reads or variants with fraction less than 8% were excluded from analysis.

qPCR analysis of *Trp53* recombination

Quantitative PCR (qPCR) reactions were performed using iTaq Universal SYBR Green Supermix (Bio-Rad 1725120) and CFX384 touch real-time PCR detection system (Bio-Rad). Reaction were comprised of 200 ng gDNA, 5 µl Supermix, 0.5 µl forward primer, 0.5 µl reverse primer in a total volume of 10 µl. PCR conditions were one cycle at 95°C for 3 minutes, 40 cycles at 95°C for 10 seconds, 55°C for 30 seconds, and a melt curve cycle (65-95°C at 0.5°C increments at 5 seconds/step). Primers were designed to detect

unrecombined *Trp53^{flox}* allele (p53 WT), recombined *Trp53^{flox}* allele (p53 Del) and the reference gene *Tf1c*. Primer sequences were:

Trp53 WT and Del F: 5'-ATCCTTTATTCTGTTCGATAAGCTTG-3'

Trp53 WT R: 5'-AGGACTACACAGAGAAACCCT-3'

Trp53 Del R: 5'-GCTATTGTAGCTAGAACTAGTGGAT-3'

Tf1c F: 5'-ACCAAATGGTTCGTACAGCA-3'

Tf1c R: 5'-ATGACAGTAGTTTGCTGTTATACATC-3'

The relative levels of *Trp53* WT and *Trp53* Del genomic DNA were calculated using ΔC_t method in comparison to *Tf1c*. The fraction of *Trp53* Del was then calculated by dividing the relative level of *Trp53* Del by the sum of the relative level of *Trp53* WT and *Trp53* Del.

qPCR analysis of *Kras* copy number

qPCR reactions were performed using PrimeTime Gene Expression Master Mix (Integrated DNA Technologies 1055770) and CFX384 touch real-time PCR detection system (Bio-Rad). The reaction were comprised of 200 ng gDNA, 5 μ l master mix, and 0.5 μ l forward primer, 0.5 μ l reverse primer, 0.25 μ l probe for *Kras^{nat}* allele, *Kras^{ex3op}* allele, and reference gene *Tert* in a total volume of 10 μ l. The conditions were one cycle at 95°C for 3 minutes, 40 cycles at 95°C for 10 seconds, 57°C for 30 seconds, and a melt curve (65-95°C at 0.5°C increments at 5 sec/step). Primers sequences were:

Kras^{nat} F: 5'-GGAATAAGTGTGATTTGCCTTCT-3'

Kras^{nat} R: 5'-ACCTGTCTTGTCTTTGCTGA-3'

Kras^{ex3op} F: 5'-AAGTGCGACCTCCCTAGC-3'

Kras^{ex3op} R: 5'-CTGTCTTGTCTTGGCGCT-3'

Tert F: 5'-CCTGACCATCTGGTGACAC-3';

Tert R: 5'-GTGCCTTCTCAGAGAACACA-3'.

Probe sequences were:

Kras^{nat}: 5'-/5Cy5/AACAGTAGA/TAO/CACGAAACAGGCTCAGGA/3IAbRQSp/-3'

Kras^{ex3op}: 5'-/5HEX/AACCGTGGA/ZEN/CACCAAGCAGGCC/3IABkFQ/-3'

Tert: 5'-/56-FAM/TGGAACCAA/ZEN/ACATACATGCAGGTGCAG/3IABkFQ/-3'.

The relative levels of *Kras^{nat}* and *Kras^{ex3op}* alleles were calculated using ΔC_t method in comparison to *Tert*. The copy number of *Kras^{nat}* and *Kras^{ex3op}* allele were then determined by comparing the relative DNA level in tumor samples to normal lung tissues

from *Kras^{nat/nat}* or *Kras^{ex30/ex30p}* mice, which were used as reference for two copies of *Kras^{nat}* or *Kras^{ex30p}* allele, respectively.

qPCR analysis of MAPK signaling

qPCR reactions were performed using iTaq Universal SYBR Green Supermix (Bio-Rad 1725120) and CFX384 touch real-time PCR detection system (Bio-Rad). The reaction were comprised of 1 µl gDNA, 5 µl Supermix, 0.5 µl forward primer, 0.5 µl reverse primer in a total volume of 10 µl. The conditions were one cycle at 95°C for 3 minutes, 40 cycles at 95°C for 10 seconds, 58°C for 30 seconds, and a melt curve (65-95°C at 0.5°C increments at 5 sec/step). The primers sequences were:

Dusp6 F: 5'-ACTTGGACGTGTTGGAAGAGT-3'

Dusp6 R: 5'-GCCTCGGGCTTCATCTATGAA-3'

Egr1 F: 5'-CCTGACCACAGAGTCCTTTTCT-3'

Egr1 R: 5'-AGGCCACTGACTAGGCTGA-3'

Fosl1 F: 5'-CAGGAGTCATACGAGCCCTAG-3'

Fosl1 R: 5'-GCCTGCAGGAAGTCTGTCAG-3'

Actin F: 5'-CGTGAAAAGATGACCCAGATCATGT-3'

Actin R: 5'-CGTGAGGGAGAGCATAGCC-3'.

Gene expression values were calculated using the comparative Ct (-ΔΔCt) method (Livak and Schmittgen, 2001), using actin as internal control.

Isolation of genomic DNA for maximum depth sequencing assay

Lung tissues were cut into fine pieces and resuspended in 500 µl lysis buffer (100 mM NaCl, 10 mM Tris pH 7.6, 25 mM EDTA pH 8.0, and 0.5% SDS in H₂O, supplemented with 20 µg.ml⁻¹ RNase A (Sigma R4642). Samples were incubated at 37°C for 1 hour. 2.5 µl of 800 U.ml⁻¹ proteinase K (New England Biolabs P8107S) was then added to each sample, then the samples were vortexed and incubated at 55°C overnight. Genomic DNA was isolated by phenol/chloroform extraction followed by ethanol precipitation using standard procedures.

Maximum depth sequencing (MDS)

This method was adapted from published protocols (Jee et al., 2016; Li et al., 2020). In detail, 20-50 µg of genomic DNA was incubated with EcoRV (New England Biolabs R3195), EcoRI (New England Biolabs R3101), and XmnI (New England Biolabs R0194) for the analysis of the non-transcribed strand of *Kras* exon 1 and exon 2. Reaction conditions were 5 units of each of the indicated restriction enzymes and per 1µg DNA per 20 µl reaction. Digested genomic DNA was column purified using QIAquick PCR Purification Kit following the manufacturer's protocol (Qiagen 28104) and resuspended in ddH₂O (35 µl H₂O per 10 µg DNA). The barcode and adaptor were added to the target DNA by incubating purified DNA with the appropriate barcode primers (see below) for one cycle of PCR. PCR reactions were comprised of 10 µg DNA, 2.5 µl of 10 µM barcode primer (see below), 4 µl of 2.5 mM dNTP, 10 µl of 5X buffer, and 0.5 µl Q5® Hot Start High-Fidelity DNA Polymerase (New England Biolabs M0493) in a total volume of 50 µl. The number of PCR reactions was scaled according to the amount of DNA. PCR conditions were 98°C for 1 minute, 60°C for 15 seconds, and 72°C for 1 minute using the barcoding primer for exon 2, followed by the addition of the barcoding primer for exon 1 to the same reaction, 98°C for 1 minute, 68°C for 15 seconds, and 72°C for 1 minute. 1 µl of 20,000 U.ml⁻¹ exonuclease I (New England Biolabs M0293) and 5 µl of 10X exonuclease I buffer (New England Biolabs B0293S) was then added to each 50 µl reaction to remove unused barcoded primers and incubated at 37°C for 1 hour and then 80°C for 20 minutes. Processed DNA were column-purified using QIAquick PCR Purification Kit as above and resuspended in ddH₂O (35 µl H₂O per column). The concentration of purified product was measured with SimpliNano spectrophotometer (GE Healthcare Life Sciences). Samples were linear amplified with forward adaptor primer (see below). PCR reactions were comprised of 1.5 µg DNA, 2.5 µl of 10 µM forward-adaptor primer, 4 µl of 2.5 mM dNTP, 10 µl of 5X buffer, and 0.5 µl Q5® Hot Start High-Fidelity DNA Polymerase in a total volume of 50 µl. The number of PCR reactions was scaled according to the amount of DNA. PCR conditions were as follows: 12 cycles of 98°C for 15 seconds, 70°C for 15 seconds, and 72°C for 10 seconds. 2.5 µl of 10 µM exon-specific reverse primers (see below) and 2.5 µl of 10 µM reverse-adaptor primer (see below) were then added to each 50 µl reaction. The mixtures were then subjected exponential amplification. PCR conditions were as follows: 4 cycles of 98°C for

15 seconds, 62°C for 15 seconds, 72°C for 10 seconds, 20 cycles of 98°C for 15 seconds, 70°C for 15 seconds, and 72°C for 10 seconds. The final library was size selected and purified with Ampure XP beads according to the manufacturer's protocol (Beckman Coulter A63880). Sequencing was performed using NovaSeq 6000 S Prime 150bp PE at Duke Center for Genomic and Computational Biology.

Primers for MDS

Barcode primer: [Forward adaptor][Index][Barcode][Primer]

Where

[Forward adaptor] =

5'-TACGGCGACCAACGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-3'

[Index] = variable length of known sequences from 0 to 7 nucleotides (Supplementary file 3)

[Barcode] = NNNNNNNNNNNNNNNN

Kras exon 1 [Primer] = 5'-ATCTTTTTTCAAAGCGGCTGGCT-3'

Kras exon 2 [Primer] = 5'-TCTTCAAATGATTTAGTATTATTTATGGC-3'

Forward-adaptor primer: 5'-AATGATACGGCGACCAACGAGAT-3'

Exon-specific reverse primer: [Reverse adaptor][Index][Primer]

Where

[Reverse adaptor] =

5'-CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'

[Index] = variable length of known sequences from 0 to 7 nucleotides (Supplementary file 3)

Kras exon 1 [Primer] = 5'-TATTATTTTTTATTGTAAGGCCTGCTGA-3'

Kras exon 2 [Primer] = 5'-GACTCCTACAGGAAACAAGT-3'

Reverse-adaptor primer: 5'-CAAGCAGAAGACGGCATACGAGA-3'

All primers were synthesized by Integrated DNA Technologies.

Analysis of MDS data

Raw data were uploaded Galaxy Cloudman (Jalili et al., 2020). Read 1 and read 2 were joined via PEAR pair-end read merger (Zhang et al., 2014b). The reads were then filtered

by quality by requiring 90% of bases in the sequence to have a quality core ≥ 20 . Filtered reads were split into different files based on assigned sample indexes and variation in sequence lengths using the Barcode Splitter tool and Filter sequences by length tool. The reads were trimmed down to the barcode and the target exon. Trimmed reads were grouped by barcode. Barcode families containing ≥ 2 reads and have $\geq 90\%$ reads being identical are selected. Sequences from selected barcode families were compared against annotated reference mutant sequences containing all possible single nucleotide substitutions in the exon of interest and the mutation in the reference mutant sequence was assigned to the matched barcode family. The frequency of the corresponding mutation was calculated by dividing the counts of the families containing the mutation by the total number of families.

Generation of heatmaps

All heatmaps were generated using Morpheus (<https://software.broadinstitute.org/morpheus>). The mutation frequencies used in heatmaps in **Figure 3** and **Figure 3-figure supplement 1** were corrected by the addition of 1×10^{-5} (the detection limit at a barcode recovery of 1×10^5), \log_{10} transformed and plotted.

Statistics

The number of independent experiments and the statistical analysis used are indicated in the legends of each figure. Data are represented as mean \pm SEM. p values were determined by Dunn's multiple comparison test following Kruskal-Wallis test, Holm-Sidak multiple comparisons test following one-way ANOVA, two-tailed Mann-Whitney U test, or two-sided Fisher's exact test. For correlation analysis, Rho and p values were derived from the Spearman correlation test. All statistical test were performed using GraphPad Prism 6.

Data Availability

All raw sequencing data has been deposited to NCBI Sequence Read Archive (SRA) under accession number PRJNA663179.

Acknowledgements

We thank the head of and/or members of the laboratories of Drs. James Alvarez, Christopher Counter, David MacAlpine, and Nikoleta Tsvetanova (Duke University) for thoughtful discussions. This work was supported by the National Cancer Institute (R01CA94184 and P01CA203657 to CMC).

Author Contributions

SL performed while CMC oversaw all experiments, both authors contributed to the writing of the manuscript.

Declaration of Interests

The authors declare no competing interests.

References

- Ali, M., Kaltenbrun, E., Anderson, G.R., Stephens, S.J., Arena, S., Bardelli, A., Counter, C.M., and Wood, K.C. (2017). Codon bias imposes a targetable limitation on KRAS-driven therapeutic resistance. *Nat Commun* 8, 15617.
- Balmain, A. (2020). The critical roles of somatic mutations and environmental tumor-promoting agents in cancer risk. *Nat Genet* 52, 1139-1143.
- Buffet, C., Hecale-Perlemonne, K., Bricaire, L., Dumont, F., Baudry, C., Tissier, F., Bertherat, J., Cochand-Priollet, B., Raffin-Sanson, M.L., Cormier, F., *et al.* (2017). DUSP5 and DUSP6, two ERK specific phosphatases, are markers of a higher MAPK signaling activation in BRAF mutated thyroid cancers. *PLoS One* 12, e0184861.
- Buhrman, G., Holzapfel, G., Fetics, S., and Mattos, C. (2010). Allosteric modulation of Ras positions Q61 for a direct role in catalysis. *Proc Natl Acad Sci USA* 107, 4931-4936.
- Buisson, R., Langenbucher, A., Bowen, D., Kwan, E.E., Benes, C.H., Zou, L., and Lawrence, M.S. (2019). Passenger hotspot mutations in cancer driven by APOBEC3A and mesoscale genomic features. *Science* 364, eaaw2872.
- Burd, C.E., Liu, W., Huynh, M.V., Waqas, M.A., Gillahan, J.E., Clark, K.S., Fu, K., Martin, B.L., Jeck, W.R., Souroullas, G.P., *et al.* (2014). Mutation-specific RAS oncogenicity explains NRAS codon 61 selection in melanoma. *Cancer Discov* 4, 1418-1429.
- Chang, M.T., Asthana, S., Gao, S.P., Lee, B.H., Chapman, J.S., Kandoth, C., Gao, J., Socci, N.D., Solit, D.B., Olshen, A.B., *et al.* (2016). Identifying recurrent mutations in cancer reveals widespread lineage diversity and mutational specificity. *Nat Biotechnol* 34, 155-163.
- Chung, W.J., Daemen, A., Cheng, J.H., Long, J.E., Cooper, J.E., Wang, B.E., Tran, C., Singh, M., Gnad, F., Modrusan, Z., *et al.* (2017). Kras mutant genetically engineered mouse models of human cancers are genomically heterogeneous. *Proc Natl Acad Sci USA* 114, E10947-10955.
- Cicchini, M., Buza, E.L., Sagal, K.M., Gudiel, A.A., Durham, A.C., and Feldser, D.M. (2017). Context-dependent effects of amplified MAPK signaling during lung adenocarcinoma initiation and progression. *Cell Rep* 18, 1958-1969.
- Cisowski, J., Sayin, V.I., Liu, M., Karlsson, C., and Bergo, M.O. (2016). Oncogene-induced senescence underlies the mutual exclusive nature of oncogenic KRAS and BRAF. *Oncogene* 35, 1328-1333.
- Dietlein, F., Weghorn, D., Taylor-Weiner, A., Richters, A., Reardon, B., Liu, D., Lander, E.S., Van Allen, E.M., and Sunyaev, S.R. (2020). Identification of cancer driver genes based on nucleotide context. *Nat Genet* 52, 208-218.
- Drosten, M., and Barbacid, M. (2020). Targeting the MAPK pathway in KRAS-driven tumors. *Cancer Cell* 37, 543-550.
- Dwyer-Nield, L.D., McQuillan, J., Hill-Baskin, A., Radcliffe, R.A., You, M., Nadeau, J.H., and Malkinson, A.M. (2010). Epistatic interactions govern chemically-induced lung tumor susceptibility and Kras mutation site in murine C57BL/6J-ChrA/J chromosome substitution strains. *Int J Cancer* 126, 125-132.

700 Esnault, C., Gualdrini, F., Horswell, S., Kelly, G., Stewart, A., East, P., Matthews, N., and Treisman, R. (2017).
701 ERK-induced activation of TCF family of SRF cofactors initiates a cChromatin modification cascade
702 associated with transcription. *Mol Cell* 65, 1081-1095.

703 Feldser, D.M., Kostova, K.K., Winslow, M.M., Taylor, S.E., Cashman, C., Whittaker, C.A., Sanchez-Rivera,
704 F.J., Resnick, R., Bronson, R., Hemann, M.T., *et al.* (2010). Stage-specific sensitivity to p53 restoration
705 during lung cancer progression. *Nature* 468, 572-575.

706 Fu, J.J., Dang, Y.K., Counter, C., and Liu, Y. (2018). Codon usage regulates human KRAS expression at both
707 transcriptional and translational levels. *J Biol Chem* 293, 17929-17940.

708 Gillies, T.E., Pargett, M., Minguet, M., Davies, A.E., and Albeck, J.G. (2017). Linear integration of ERK
709 activity predominates over persistence detection in Fra-1 regulation. *Cell Syst* 5, 549-563.

710 Gowd, V., Su, H., Karlovsky, P., and Chen, W. (2018). Ethyl carbamate: An emerging food and
711 environmental toxicant. *Food Chem* 248, 312-321.

712 Gurley, K.E., Moser, R.D., and Kemp, C.J. (2015). Induction of lung tumors in mice with urethane. *Cold*
713 *Spring Harb Protoc* 2015, pdb.prot077446.

714 Hobbs, G.A., Der, C.J., and Rossman, K.L. (2016). RAS isoforms and mutations in cancer at a glance. *J Cell*
715 *Sci* 129, 1287-1292.

716 Hymowitz, S.G., and Malek, S. (2018). Targeting the MAPK pathway in RAS mutant cancers. *Cold Spring*
717 *Harb Perspect Med* 8, a031492.

718 Jalili, V., Afgan, E., Gu, Q., Clements, D., Blankenberg, D., Goecks, J., Taylor, J., and Nekrutenko, A. (2020).
719 The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update.
720 *Nucleic Acids Res* 48, W395-W402.

721 Jee, J., Rasouly, A., Shamovsky, I., Akivis, Y., Steinman, S.R., Mishra, B., and Nudler, E. (2016). Rates and
722 mechanisms of bacterial mutagenesis from maximum-depth sequencing. *Nature* 534, 693-696.

723 Junttila, M.R., Karnezis, A.N., Garcia, D., Madriles, F., Kortlever, R.M., Rostker, F., Brown Swigart, L., Pham,
724 D.M., Seo, Y., Evan, G.I., *et al.* (2010). Selective activation of p53-mediated tumour suppression in high-
725 grade tumours. *Nature* 468, 567-571.

726 Kanda, M., Matthaei, H., Wu, J., Hong, S.M., Yu, J., Borges, M., Hruban, R.H., Maitra, A., Kinzler, K.,
727 Vogelstein, B., *et al.* (2012). Presence of somatic mutations in most early-stage pancreatic intraepithelial
728 neoplasia. *Gastroenterology* 142, 730-733.

729 Kidger, A.M., and Keyse, S.M. (2016). The regulation of oncogenic Ras/ERK signalling by dual-specificity
730 mitogen activated protein kinase phosphatases (MKPs). *Semin Cell Dev Biol* 50, 125-132.

731 Kong, G., Chang, Y.I., You, X., Ranheim, E.A., Zhou, Y., Burd, C.E., and Zhang, J. (2016). The ability of
732 endogenous Nras oncogenes to initiate leukemia is codon-dependent. *Leukemia* 30, 1935-1938.

733 Kwon, M.C., and Berns, A. (2013). Mouse models for lung cancer. *Mol Oncol* 7, 165-177.

734 Lampson, B.L., Pershing, N.L., Prinz, J.A., Lacsina, J.R., Marzluff, W.F., Nicchitta, C.V., MacAlpine, D.M., and
735 Counter, C.M. (2013). Rare codons regulate KRas oncogenesis. *Curr Biol* 23, 70-75.

736 Li, S., Balmain, A., and Counter, C.M. (2018). A model for RAS mutation patterns in cancers: finding the
737 sweet spot. *Nat Rev Cancer* 18, 767-777.

738 Li, S., MacAlpine, D.M., and Counter, C.M. (2020). Capturing the primordial Kras mutation initiating
739 urethane carcinogenesis. *Nat Commun* 11, 1800.

740 Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time
741 quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* 25, 402-408.

742 Malkinson, A.M., and Beer, D.S. (1983). Major effect on susceptibility to urethan-induced pulmonary
743 adenoma by a single gene in BALB/cBy mice. *J Natl Cancer Inst* 70, 931-936.

744 Marino, S., Vooijs, M., van Der Gulden, H., Jonkers, J., and Berns, A. (2000). Induction of medulloblastomas
745 in p53-null mutant mice by somatic inactivation of Rb in the external granular layer cells of the cerebellum.
746 *Genes Dev* 14, 994-1004.

747 McCreery, M.Q., and Balmain, A. (2017). Chemical carcinogenesis models of cancer: Back to the future.
748 *Ann Rev Cancer Biol* 1, 295-312.

749 McMahon, S.B., and Monroe, J.G. (1995). Activation of the p21ras pathway couples antigen receptor
750 stimulation to induction of the primary response gene *egr-1* in B lymphocytes. *J Exp Med* 181, 417-422.

751 Miller, Y.E., Dwyer-Nield, L.D., Keith, R.L., Le, M., Franklin, W.A., and Malkinson, A.M. (2003). Induction of
752 a high incidence of lung tumors in C57BL/6 mice with multiple ethyl carbamate injections. *Cancer Lett* 198,
753 139-144.

754 Mueller, S., Engleitner, T., Maresch, R., Zukowska, M., Lange, S., Kaltenbacher, T., Konukiewicz, B., Ollinger,
755 R., Zwiebel, M., Strong, A., *et al.* (2018). Evolutionary routes and KRAS dosage define pancreatic cancer
756 phenotypes. *Nature* 554, 62-68.

757 Munoz-Espin, D., and Serrano, M. (2014). Cellular senescence: from physiology to pathology. *Nat Rev Mol*
758 *Cell Biol* 15, 482-496.

759 Munoz-Maldonado, C., Zimmer, Y., and Medova, M. (2019). A comparative analysis of individual RAS
760 mutations in cancer biology. *Front Oncol* 9, 1088.

761 Nieto, P., Ambrogio, C., Esteban-Burgos, L., Gomez-Lopez, G., Blasco, M.T., Yao, Z., Marais, R., Rosen, N.,
762 Chiarle, R., Pisano, D.G., *et al.* (2017). A Braf kinase-inactive mutant induces lung adenocarcinoma. *Nature*
763 548, 239-243.

764 Parker, J.A., Volmar, A.Y., Pavlopoulos, S., and Mattos, C. (2018). K-Ras populates conformational states
765 differently from its isoform H-Ras and oncogenic mutant K-RasG12D. *Structure* 26, 810-820 e814.

766 Pershing, N.L., Lampson, B.L., Belsky, J.A., Kaltenbrun, E., MacAlpine, D.M., and Counter, C.M. (2015). Rare
767 codons capacitate Kras-driven de novo tumorigenesis. *J Clin Invest* 125, 222-233.

768 Peterson, J., Li, S., Kaltenbrun, E., Erdogan, O., and Counter, C.M. (2020). Expression of transgenes
769 enriched in rare codons is enhanced by the MAPK pathway. *Sci Rep* 10, 22166.

770 Prior, I.A., Hood, F.E., and Hartley, J.L. (2020). The frequency of Ras mutations in cancer. *Cancer Res* 80,
771 2969-2974.

772 Pylayeva-Gupta, Y., Grabocka, E., and Bar-Sagi, D. (2011). RAS oncogenes: weaving a tumorigenic web.
773 *Nat Rev Cancer* *11*, 761-774.

774 Quax, T.E., Claassens, N.J., Soll, D., and van der Oost, J. (2015). Codon bias as a means to fine-tune gene
775 expression. *Mol Cell* *59*, 149-161.

776 Ryan, M.B., Der, C.J., Wang-Gillam, A., and Cox, A.D. (2015). Targeting RAS-mutant cancers: is ERK the key?
777 *Trends Cancer* *1*, 183-198.

778 Sanchez-Vega, F., Mina, M., Armenia, J., Chatila, W.K., Luna, A., La, K.C., Dimitriadou, S., Liu, D.L., Kantheti,
779 H.S., Saghafeina, S., *et al.* (2018). Oncogenic signaling pathways in the Cancer Genome Atlas. *Cell* *173*, 321-
780 337 e310.

781 Sasine, J.P., Himburg, H.A., Termini, C.M., Roos, M., Tran, E., Zhao, L., Kan, J., Li, M., Zhang, Y., de Barros,
782 S.C., *et al.* (2018). Wild-type Kras expands and exhausts hematopoietic stem cells. *JCI Insight* *3*, e98197.

783 Shimkin, M.B., and Stoner, G.D. (1975). Lung tumors in mice: application to carcinogenesis bioassay. *Adv*
784 *Cancer Res* *21*, 1-58.

785 Shin, S.Y., Rath, O., Choo, S.M., Fee, F., McFerran, B., Kolch, W., and Cho, K.H. (2009). Positive- and
786 negative-feedback regulations coordinate the dynamic behavior of the Ras-Raf-MEK-ERK signal
787 transduction pathway. *J Cell Sci* *122*, 425-435.

788 Smith, M.J., Neel, B.G., and Ikura, M. (2013). NMR-based functional profiling of RASopathies and
789 oncogenic RAS mutations. *Proc Natl Acad Sci USA* *110*, 4574-4579.

790 Swarbrick, A., Roy, E., Allen, T., and Bishop, J.M. (2008). Id1 cooperates with oncogenic Ras to induce
791 metastatic mammary carcinoma by subversion of the cellular senescence response. *Proc Natl Acad Sci*
792 *USA* *105*, 5402-5407.

793 Temko, D., Tomlinson, I.P.M., Severini, S., Schuster-Bockler, B., and Graham, T.A. (2018). The effects of
794 mutational processes and selection on driver mutations across cancer types. *Nat Commun* *9*, 1857.

795 Unni, A.M., Harbourne, B., Oh, M.H., Wild, S., Ferrarone, J.R., Lockwood, W.W., and Varmus, H. (2018).
796 Hyperactivation of ERK by multiple mechanisms is toxic to RTK-RAS mutation-driven lung adenocarcinoma
797 cells. *Elife* *7*, e33718.

798 Valentine, C.C., 3rd, Young, R.R., Fielden, M.R., Kulkarni, R., Williams, L.N., Li, T., Minocherhomji, S., and
799 Salk, J.J. (2020). Direct quantification of in vivo mutagenesis and carcinogenesis using duplex sequencing.
800 *Proc Natl Acad Sci U S A* *117*, 33414-33425.

801 Vallejo, A., Perurena, N., Guruceaga, E., Mazur, P.K., Martinez-Canarias, S., Zandueta, C., Valencia, K.,
802 Arricibita, A., Gwinn, D., Sayles, L.C., *et al.* (2017a). An integrative approach unveils FOSL1 as an oncogene
803 vulnerability in KRAS-driven lung and pancreatic cancer. *Nat Commun* *8*, 14294.

804 Vallejo, A., Valencia, K., and Vicent, S. (2017b). All for one and FOSL1 for all: FOSL1 at the crossroads of
805 lung and pancreatic cancer driven by mutant KRAS. *Mol Cell Oncol* *4*, e1314239.

806 Westcott, P.M., Halliwill, K.D., To, M.D., Rashid, M., Rust, A.G., Keane, T.M., Delrosario, R., Jen, K.Y., Gurley,
807 K.E., Kemp, C.J., *et al.* (2015). The mutational landscapes of genetic and chemical models of Kras-driven
808 lung cancer. *Nature* *517*, 489-492.

809 Winters, I.P., Chiou, S.H., Paulk, N.K., McFarland, C.D., Lalgudi, P.V., Ma, R.K., Lisowski, L., Connolly, A.J.,
810 Petrov, D.A., Kay, M.A., *et al.* (2017). Multiplexed in vivo homology-directed repair and tumor barcoding
811 enables parallel quantification of Kras variant oncogenicity. *Nat Commun* 8, 2053.

812 Wistuba, II, and Gazdar, A.F. (2006). Lung cancer preneoplasia. *Annu Rev Pathol* 1, 331-348.

813 Xu, X., Rock, J.R., Lu, Y., Futtner, C., Schwab, B., Guinney, J., Hogan, B.L., and Onaitis, M.W. (2012).
814 Evidence for type II cells as cells of origin of K-Ras-induced distal lung adenocarcinoma. *Proc Natl Acad Sci*
815 *USA* 109, 4910-4915.

816 You, M., Candrian, U., Maronpot, R.R., Stoner, G.D., and Anderson, M.W. (1989). Activation of the Ki-ras
817 protooncogene in spontaneously occurring and chemically induced lung tumors of the strain A mouse.
818 *Proc Natl Acad Sci USA* 86, 3070-3074.

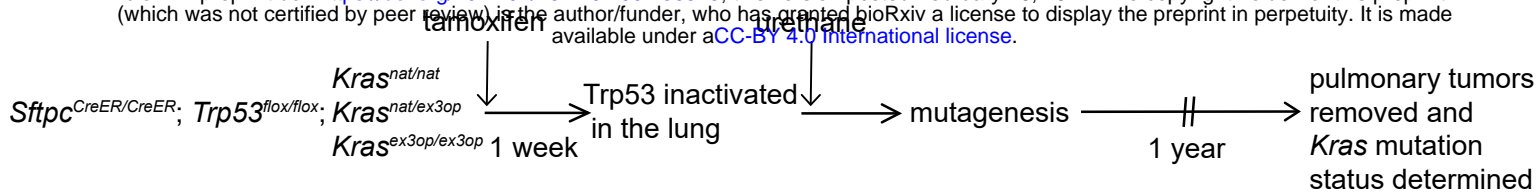
819 Zhang, J., Fujimoto, J., Zhang, J., Wedge, D.C., Song, X., Zhang, J., Seth, S., Chow, C.W., Cao, Y., Gumbs, C.,
820 *et al.* (2014a). Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion
821 sequencing. *Science* 346, 256-259.

822 Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014b). PEAR: a fast and accurate Illumina Paired-End
823 reAd mergeR. *Bioinformatics* 30, 614-620.

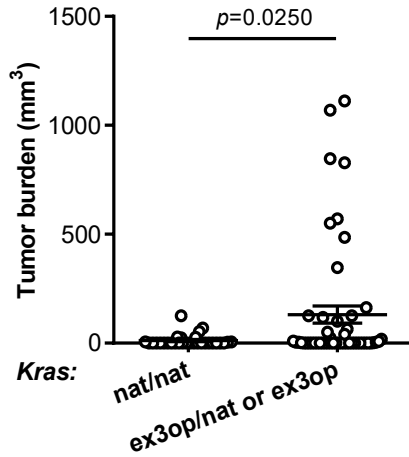
824 Zhang, Z., Kobayashi, S., Borczuk, A.C., Leidner, R.S., Laframboise, T., Levine, A.D., and Halmos, B. (2010).
825 Dual specificity phosphatase 6 (DUSP6) is an ETS-regulated negative feedback mediator of oncogenic ERK
826 signaling in lung cancer cells. *Carcinogenesis* 31, 577-586.

A

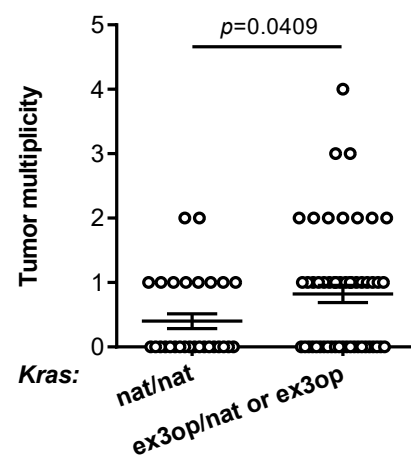
bioRxiv preprint doi: <https://doi.org/10.1101/2021.02.09.430515>; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



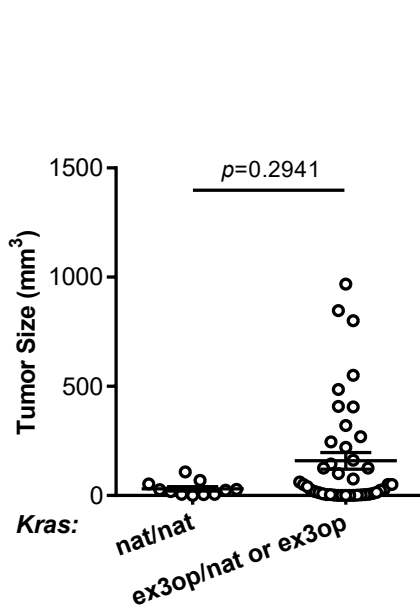
B



C



D



E

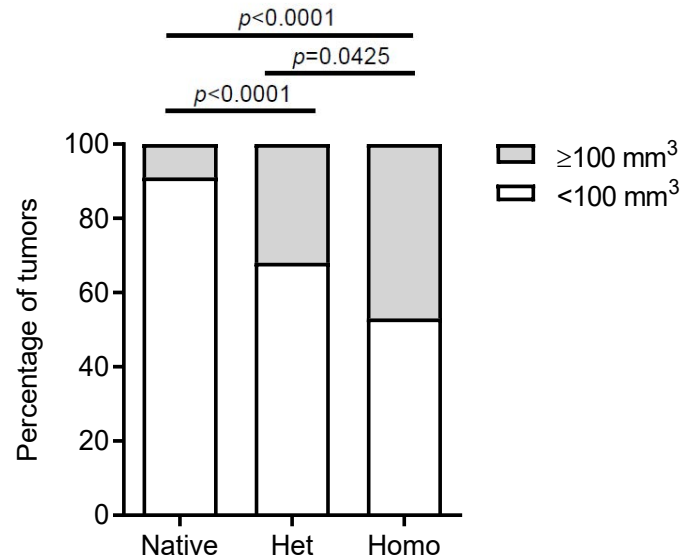


Figure 1. Loss of p53 converts the *Kras*^{ex3op} allele from suppressing to enhancing urethane carcinogenesis.

(A) Experimental design to evaluate the effect of inactivating p53 specifically in the lung on urethane carcinogenesis upon increase *Kras* expression.

(B-D) Mean ± SEM of urethane-induced tumor (B) burden, (C) multiplicity, and (D) size in tamoxifen-treated *Sftpc*^{CreER}/*CreER*; *Trp53*^{f/f} mice in a homozygous native ((B and C) n= 30 mice, (D) n=11 tumors) and heterozygous or homozygous ((B and C) n= 51 mice, (D) n=42 tumors) ex3op *Kras* background. Mann-Whitney test.

(E) % of tumors ≥ (grey bar) or < (white bar) 100 mm³ in tamoxifen-treated *Sftpc*^{CreER/CreER}; *Trp53*^{f/f} mice in *Kras*^{nat/nat} (n=11 tumors), *Kras*^{ex3op/nat} (n=25 tumors), or *Kras*^{ex3op/ex3op} (n=17 tumors) background after urethane exposure. Two-sided Fisher's exact test.

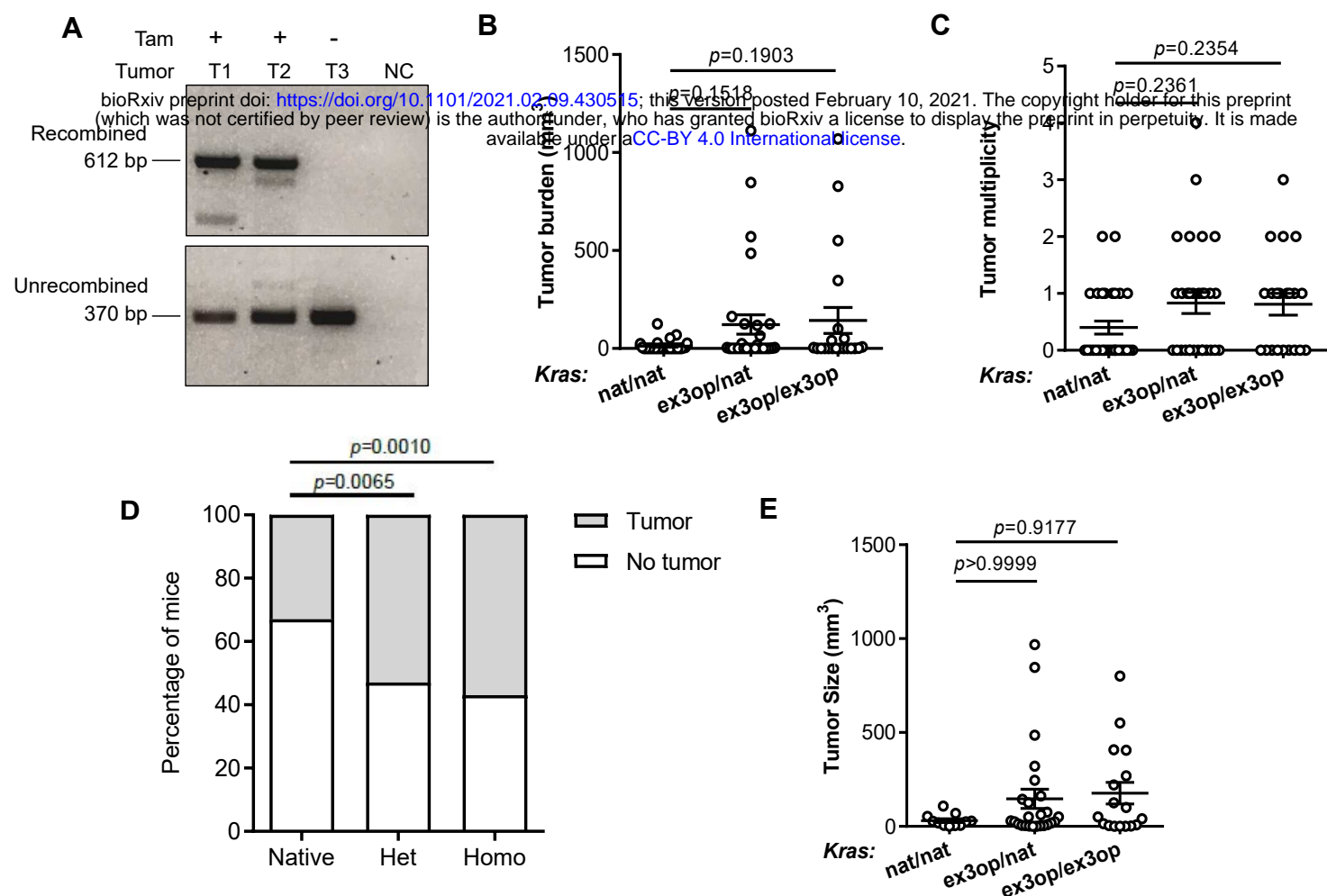


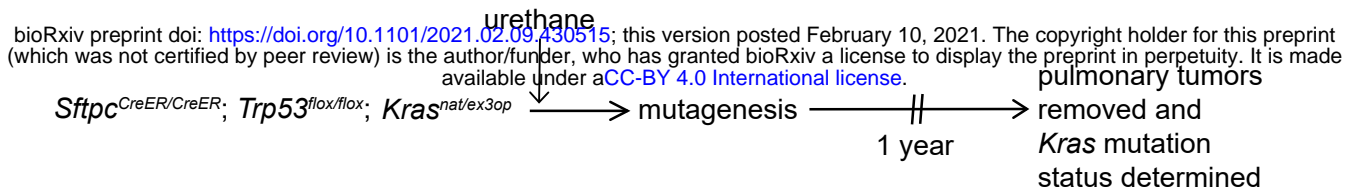
Figure 1-figure supplement 1. The effect of *Kras*^{ex3op} allele on urethane-mediated lung tumorigenesis in the absence of p53.

(A) PCR analysis of the status of the *Trp53*^{fl/ox} allele in DNA isolated from lung tumors in *Sftpc*^{CreER/CreER};*Trp53*^{fl/fl};*Kras*^{ex3op/nat} (T1,T3) or *Sftpc*^{CreER/CreER};*Trp53*^{fl/fl};*Kras*^{ex3op/ex3op} (T2) mice treated (+) or not treated (-) with tamoxifen. NC: no DNA control.

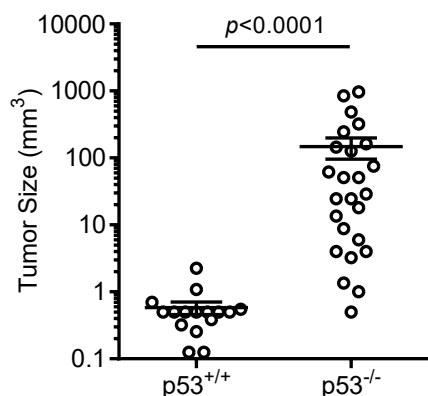
(B, C, and E) Mean ± SEM of urethane-induced tumor (B) burden, (C) multiplicity, and (E) size in tamoxifen-treated *Sftpc*^{CreER/CreER};*Trp53*^{fl/fl} mice in a homozygous native ((B and C) n=30 mice, (E) n=11 tumors), heterozygous ((B and C) n=30 mice, (E) n=25 tumors) and homozygous ((B and C) n=21 mice, (E) n=17 tumors) ex3op *Kras* background. Dunn's multiple comparison test following Kruskal-Wallis test.

(D) % of mice with (grey bar) or without (white bar) a tumor in tamoxifen-treated *Sftpc*^{CreER/CreER};*Trp53*^{fl/fl} mice in *Kras*^{nat/nat} (n=30 mice), *Kras*^{ex3op/nat} (n=30 mice), or *Kras*^{ex3op/ex3op} (n=21 mice) background after urethane exposure. Two-sided Fisher's exact test.

A



B



C

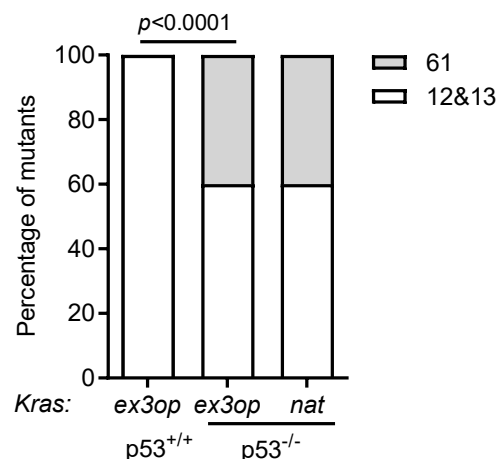


Figure 2. Loss of p53 reprograms the RAS mutation tropism of urethane.

(A) Experimental design to obtain urethane-induced lung tumors from p53^{+/+} mice.

(B) Mean ± SEM of urethane-induced tumor size in *Sftpc*^{CreER/CreER}; *Trp53*^{fl/fl}; *Kras*^{ex3op/nat} mice not treated (p53^{+/+}, n=16 tumors) or treated with tamoxifen (p53^{-/-}, n=25 tumors). Mann-Whitney test.

(C) % of urethane-induced tumors with an oncogenic mutation at codon G12/13 (white bar) versus Q61 (grey bar) in the *Kras*^{nat} versus *Kras*^{ex3op} allele in *Sftpc*^{CreER/CreER}; *Trp53*^{fl/fl}; *Kras*^{ex3op/nat} mice not treated (p53^{+/+}) or treated with tamoxifen (p53^{-/-}) where indicated. n=4 tumors ex3op p53^{+/+}, 5 tumors nat p53^{-/-} and 10 tumors ex3op p53^{-/-}. Two-sided Fisher's exact test.

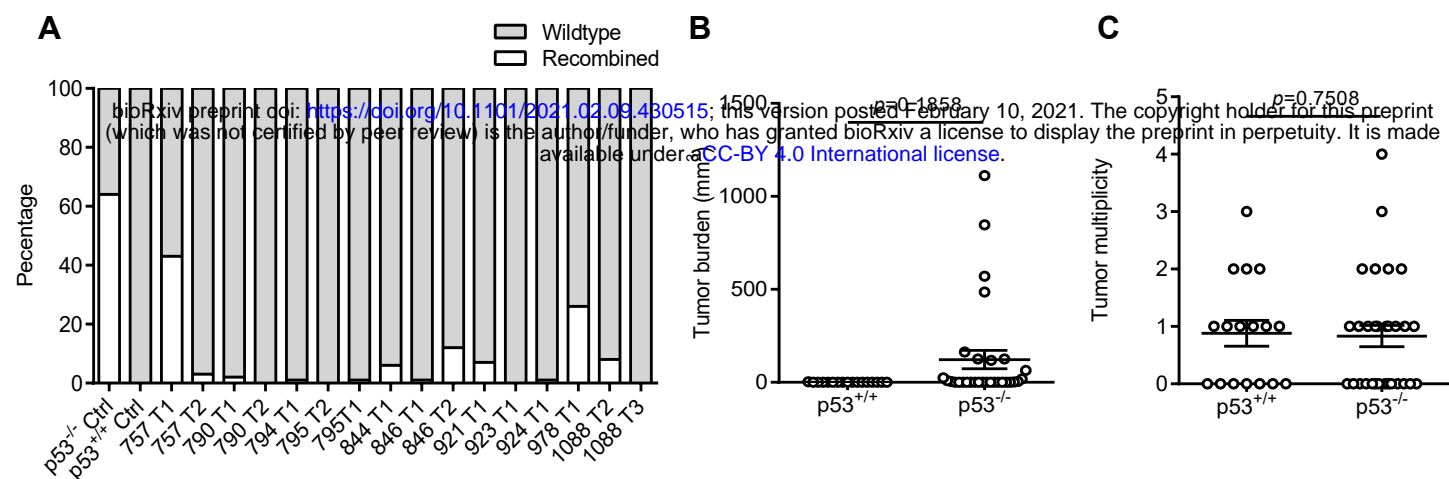
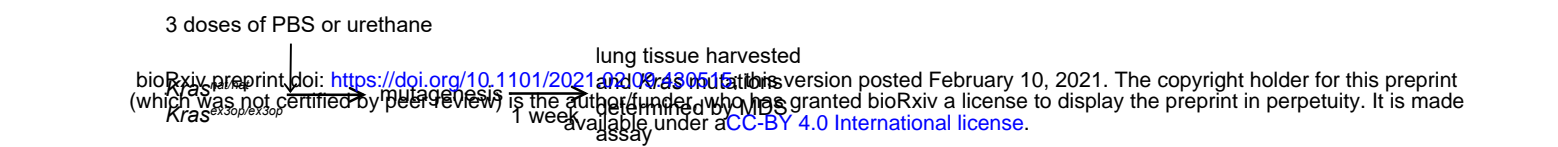


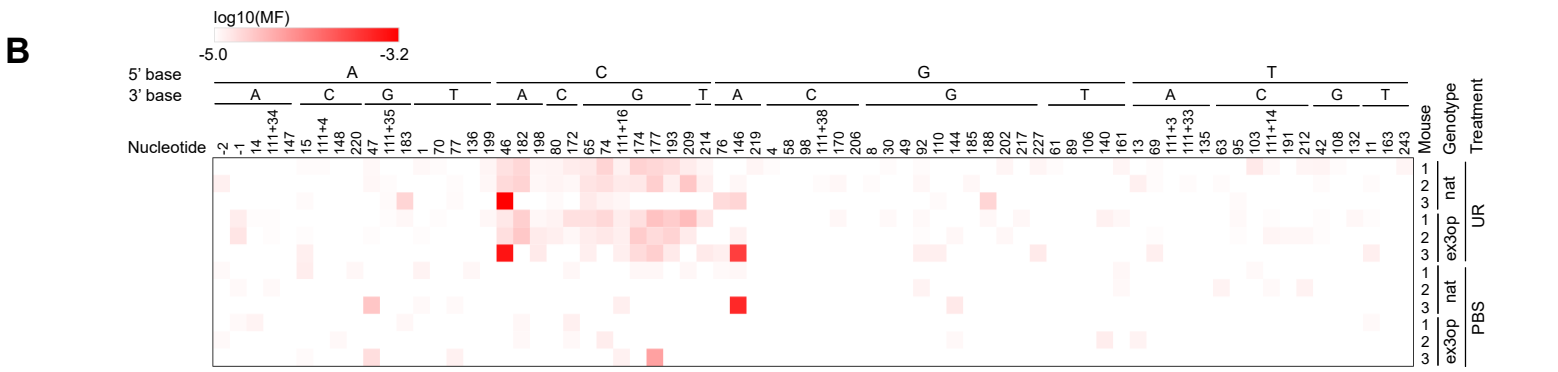
Figure 2-figure supplement 1. The effect of p53 loss on tumor burden and multiplicity.

(A) Proportion of recombined p53 allele and wildtype allele in tumors from mice not treated with tamoxifen by qPCR. (B,C) Mean \pm SEM tumor (B) burden and (C) multiplicity of *Sftpc*^{CreER/CreER}; *Trp53*^{fl/fl}; *Kras*^{ex3op/nat} mice not treated (p53^{+/+}, n=17 mice) or treated (p53^{-/-}, n=30 mice) with tamoxifen. Mann-Whitney test.

A



B



C

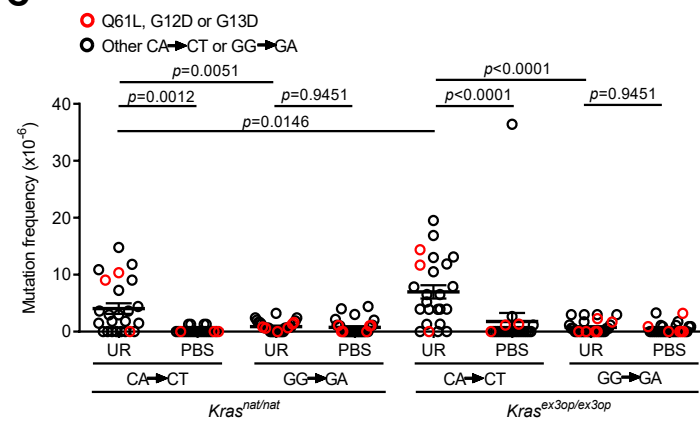


Figure 3. The mutation signature of urethane is not affected by the *Kras^{ex3op}* allele.

(A) Experimental design to identify mutations induced by urethane in mouse lung in a *Kras^{nat}* versus *Kras^{ex3op}* background.

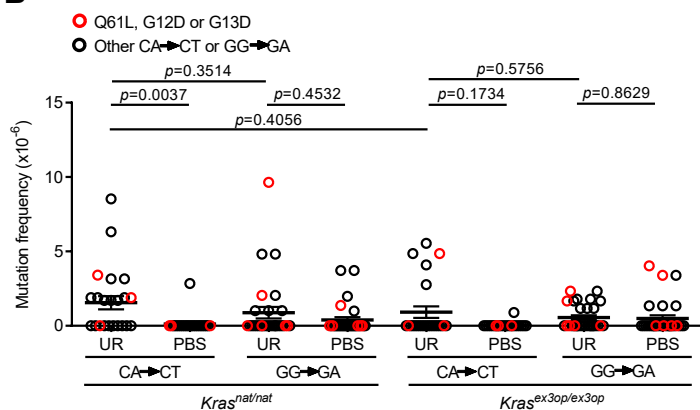
(B) Heatmap of the log transformed mutation frequency (MF) of A→T transversions determined by MDS sequencing the exon 1 and exon 2 of *Kras* in from the lungs of mice exposed to urethane (UR) in a *Kras^{nat/nat}* (nat) (n=3 mice) versus *Kras^{ex3op/ex3op}* (ex3op) (n=3 mice) background. Nucleotide number as well as the 5' and 3' base of the substituted A are shown at the top; “-” indicates nucleotides upstream of ATG start codon in 5'UTR; “111+” indicates nucleotides in the intron downstream of exon 1.

(C) Mean ± SEM mutation frequency of all CA→CT mutations in *Kras* exon 2, with Q61L mutation highlighted in red, as well as all GG→GA mutations in *Kras* exon 1, with G12D and G13D mutations highlighted in red, derived from the aforementioned MDS sequencing of *Kras* exon 1 and 2 from the lungs of *Kras^{nat/nat}* versus *Kras^{ex3op/ex3op}* mice treated with either urethane or PBS (n=3 mice each). Holm-Sidak multiple comparisons test following one-way ANOVA.

A



B



C

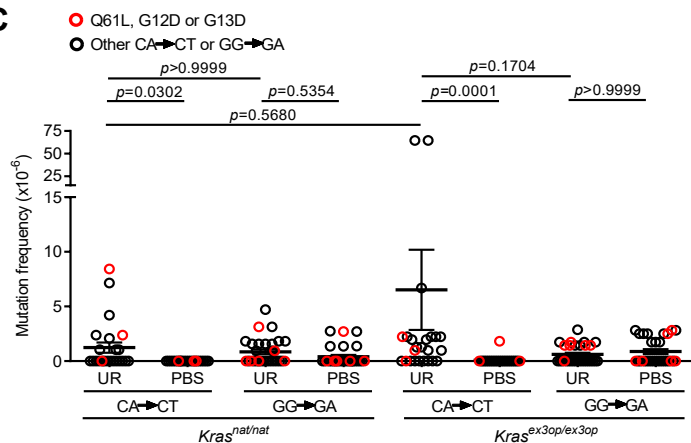


Figure 3-figure supplement 1. Mutagenesis profile of *Sftpc*^{CreER/CreER};*Trp53*^{fl/fl};*Kras*^{nat/nat} and *Sftpc*^{CreER/CreER};*Trp53*^{fl/fl};*Kras*^{ex3op/ex3op} mice.

(A) Heatmap of the log transformed mutation frequency (MF) determined by MDS sequencing the exon 1 and exon 2 of *Kras* in from the lungs of mice treated (+) or not treated with tamoxifen (-), exposed to urethane (UR) or PBS, in a *Sftpc*^{CreER/CreER};*Trp53*^{fl/fl};*Kras*^{nat/nat} (nat) or *Sftpc*^{CreER/CreER};*Trp53*^{fl/fl};*Kras*^{ex3op/ex3op} (ex3op) background (n=3 mice) for each A>T transversions (nucleotide number as well as the 5' and 3' base of the substituted A are shown at the top, "-" indicates nucleotides upstream of ATG start codon in 5'UTR, "111+" indicates nucleotides in the intron downstream of exon 1).

(B,C) Mean \pm SEM mutation frequency of all CA>CT mutations in *Kras* exon 2, including Q61L mutation highlighted in red, as well as all GG>GA mutations in *Kras* exon 1, including G12D and G13D mutations highlighted in red, derived from MDS sequencing of *Kras* exon 1 and 2 from the lungs of *Sftpc*^{CreER/CreER};*Trp53*^{fl/fl};*Kras*^{nat/nat} versus *Sftpc*^{CreER/CreER};*Trp53*^{fl/fl};*Kras*^{ex3op/ex3op} mice (B) not treated or (C) treated with tamoxifen and exposed to either urethane or PBS (n=3 mice). (B) Holm-Sidak multiple comparisons test following one-way ANOVA. (C) Dunn's multiple comparison test following Kruskal-Wallis test.

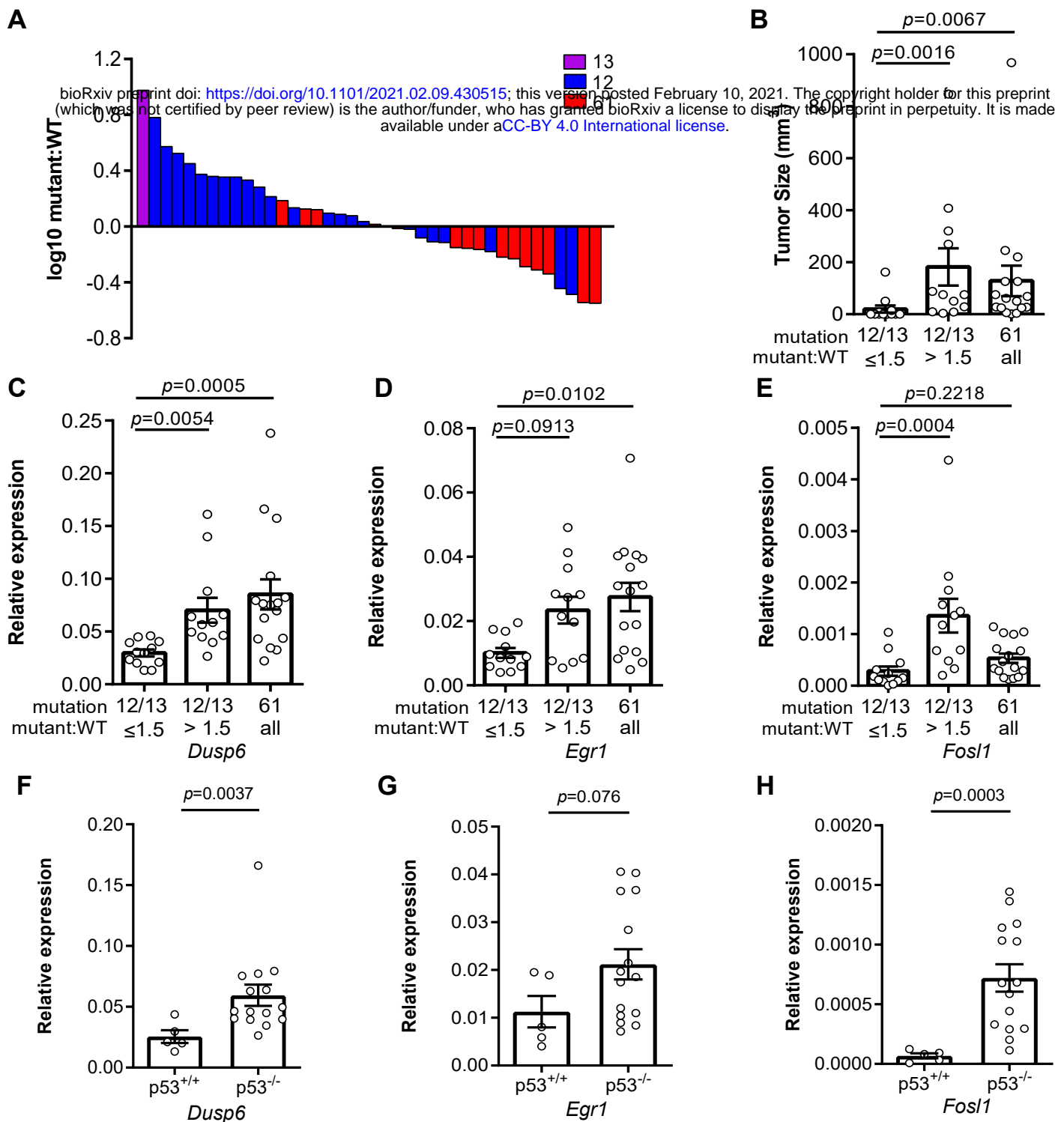


Figure 4. Loss of p53 promotes higher expression of weaker oncogenic mutations.

(A) Log10-transformed ratio of mutant to wildtype *Kras* mRNA determined by RT-qPCR in all *Kras* hotspot-mutant tumors (n=40) derived from figures 1 and 2.

(B) Mean \pm SEM size of tumors with a G12/13 oncogenic *Kras* mutation with a high (>1.5 , n=12 tumors) versus low (≤ 1.5 , n=12 tumors) mutant:WT ratio versus tumors with a Q61 oncogenic *Kras* mutation (n=16 tumors). Dunn's multiple comparison test following Kruskal-Wallis test.

(C-H) Mean \pm SEM levels of the indicated mRNAs normalized to β -actin (Relative expression) in **(C-E)** tumors with a G12/13 oncogenic *Kras* mutation with a high (>1.5 , n=12 tumors) versus low (≤ 1.5 , n=12 tumors) mutant:WT ratio versus tumors with a Q61 oncogenic *Kras* mutation (n=16 tumors) or **(F-H)** tumors from *Sftpc*^{CreER/CreER}; *Trp53*^{fl/fl}; *Kras*^{ex3ap/nat} mice not treated (*p53*^{+/+}, n=5 tumors) or treated with tamoxifen (*p53*^{-/-}, n=15 tumors) partitioned by *p53* mutation status. **(C-E)** Dunn's multiple comparison test following Kruskal-Wallis test. **(F-H)** Mann-Whitney test.

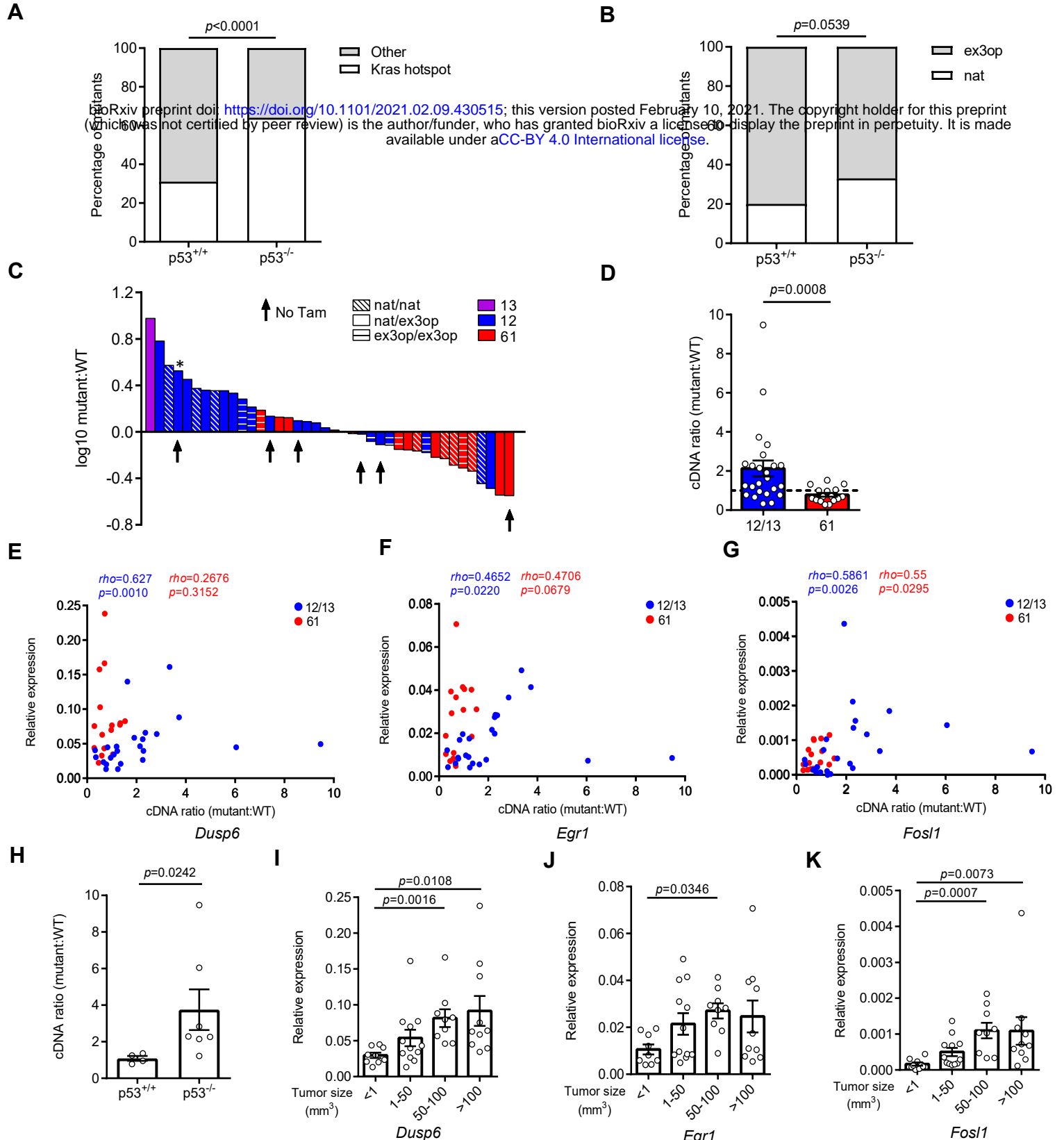


Figure 4-figure supplement 1. Allelic imbalance and MAPK signaling in Kras hotspot-mutant tumors.

(A) % of tumors with an oncogenic mutation at Kras hotspot (white bar) versus other tumors (grey bar) from *Sftpc*^{CreER/CreER};*Trp53*^{fl/fl};*Kras*^{ex3op/nat} mice not treated ($p53^{+/+}$, n=16 tumors) or treated ($p53^{-/-}$, n=25 tumors) with tamoxifen. Two-sided Fisher's exact test.

(B) % of tumors with Kras hotspot mutations occurring in the native (white bar) or ex3op (grey bar) allele from *Sftpc*^{CreER/CreER};*Trp53*^{fl/fl};*Kras*^{ex3op/nat} mice not treated ($p53^{+/+}$, n=5 tumors) or treated ($p53^{-/-}$, n=15 tumors) with tamoxifen. Two-sided Fisher's exact test.

(C) Log10-transformed ratio of mutant to wildtype *Kras* mRNA determined by RT-qPCR in all Kras hotspot-mutant tumors (n=40 tumors) derived from Figures 1 and 2. Asterisk indicates tumor 757T1 with a p53 deficiency.

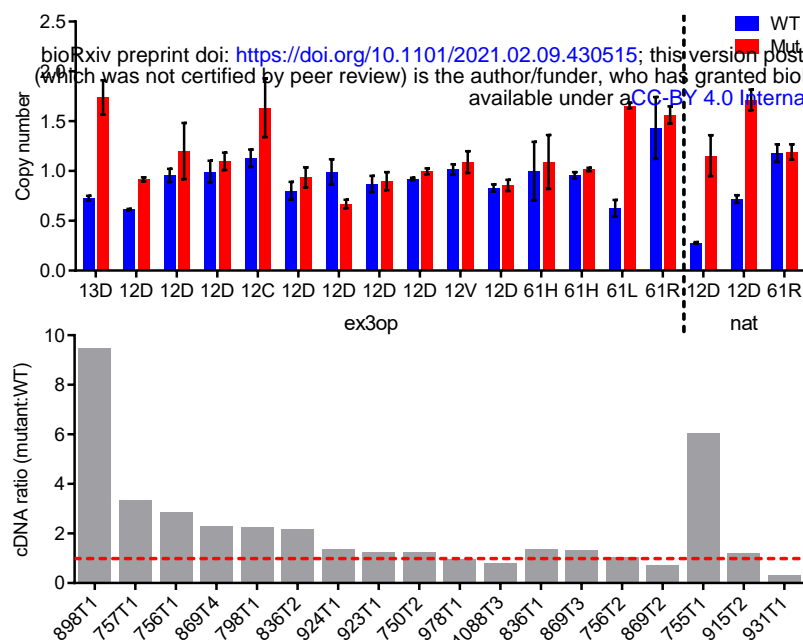
(D) Mean \pm SEM ratio of mutant to wildtype *Kras* mRNA in tumors with G12/13 (n=24 tumors) and Q61 (n=16 tumors) mutations.

(E-G) Correlation between the levels of the indicated mRNAs normalized to β -actin (Relative expression) and the ratio of mutant to wildtype *Kras* mRNA. Rho and p values are from Spearman correlation analysis. Mann-Whitney test.

(H) Mean \pm SEM ratio of G12/13 mutant *Kras*^{ex3op} to wildtype *Kras*^{nat} mRNA in tumors from *Sftpc*^{CreER/CreER};*Trp53*^{fl/fl};*Kras*^{ex3op/nat} mice not treated ($p53^{+/+}$, n=4 tumors) or treated with tamoxifen ($p53^{-/-}$, n=7 tumors). Mann-Whitney test.

(I-K) Mean \pm SEM levels of the indicated mRNAs normalized to β -actin (Relative expression) for tumors that are <1 (n=9 tumors), 1-50 (n=12 tumors), 50-100 (n=9 tumors), and >100 (n=10 tumors) mm³. Dunn's multiple comparison test following Kruskal-Wallis test.

A



B

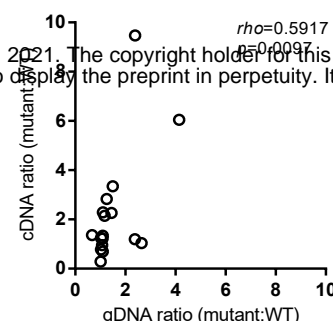


Figure 4-figure supplement 2. The imbalance at mRNA level could not be fully attributed to the imbalance of DNA copy number.

(A) Top: Copy number of the mutant or wildtype allele in tumors from *Sftpc^{CreER/CreER};Trp53^{fl/fl};Kras^{ex3op/nat}* mice estimated by qPCR copy number assay (Tert as reference gene). Mutant type and the allele with mutation are indicated. Bottom: the ratio of mutant to wildtype *Kras* mRNA and ID of the tumors listed in the top graph. Data shown are mean \pm SEM of two technical replicates.

(B) Correlation between the mRNA ratio and genomic DNA ratio of mutant to wildtype *Kras* allele. Rho and p values are from Spearman correlation analysis.

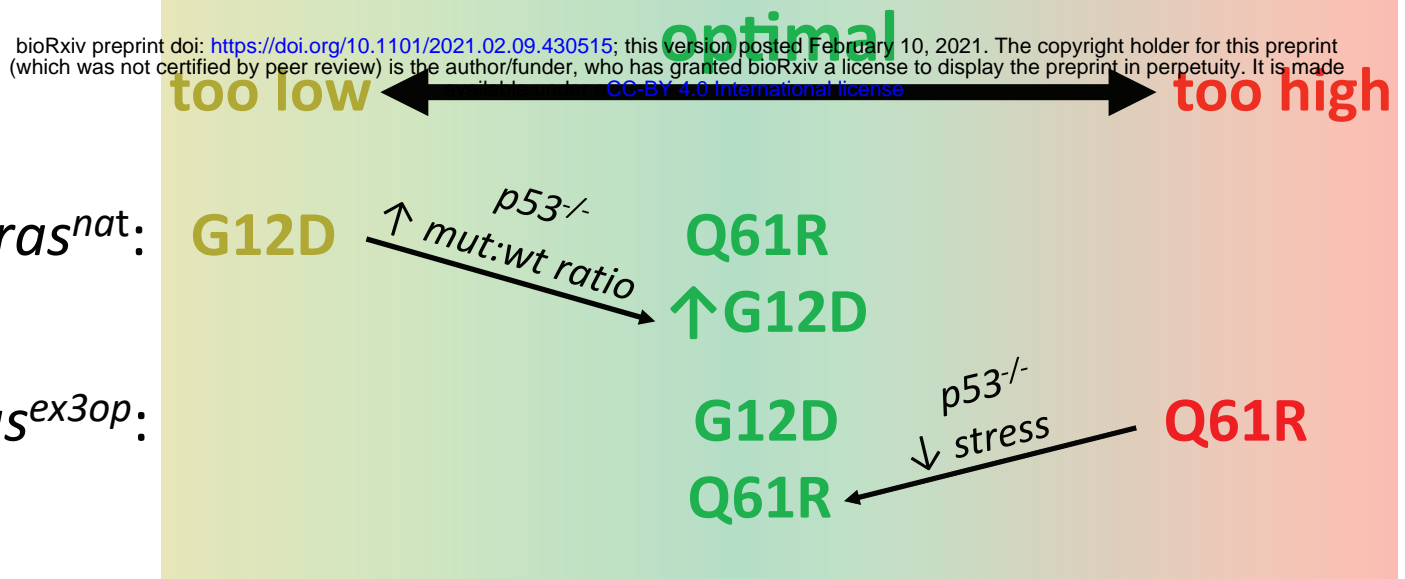


Figure 5. Optimal signaling is required for effective tumor initiation.

Signaling from a G12D mutation in the native (nat) *Kras* allele and from a Q61R mutation in the codon optimized (ex3op) *Kras* allele are outside of the window of optimal signaling amplitude achieve by *Kras^{nat}*(G12D) and *Kras^{ex3op}*(Q61R). Loss of p53 alleviates the selection against oncogenic stress and allows the recovery of a Q61R mutation in *Kras^{ex3op}* allele or a G12D mutation in the *Kras^{nat}* allele with elevated mutant:wildtype (mut:wt) mRNA allelic ratio.