# Characterizing the genetic interactions of *RAS* alleles Variability of RAS mutations: mutational patterns, biology, and interactors Characterizing the origin, distribution and genomic interactions of RAS alleles across cancer types

## Abstract

Mutations in one of the three oncogenic *RAS* genes (*KRAS*, *NRAS*, and *HRAS*) are initiating events in many tumor types. The mutations primarily occur at 4 hotspots surrounding the nucleotide binding pocket, and though the mutant alleles are naïvely similar, the biology of each is distinct and clinically relevant. However, an explanation for the different frequency of alleles in *RAS*-driven cancers remains elusive and theirbroader effects on the genomic landscape during oncogenesis have yet to be explored. To this end, 13,581 samples were collated from 5 tumor types with the highest frequency of mutation in *KRAS* or *NRAS*, colorectal adenocarcinoma (COADREAD), lung adenocarcinoma (LUAD), multiple myeloma (MM), pancreatic adenocarcinoma (PAAD), and skin cutaneous melanoma (SKCM). Investigating the origin of mutation revealed that all *RAS* mutations were produced by base changes that were infrequent in most cancers and over 90% of them were attributable to aging except for LUAD, MM, and SKCM where exogenous mutagens played significant role. Though the origin of these mutations was explained, the relative frequency of the alleles within each cancer could only be predicted for *NRAS* Q61 in SKCM and *KRAS* G12 in LUAD. To characterize the genetics of *KRAS*-mutant tumors in an allele-specific fashion, an investigation of co-mutated genes revealed that the *RAS* alleles have similar patterns of mutual exclusivity within a cancer, though few shared co-mutational partners, reinforcing the hypothesis that each allele had its own unique path to achieve an oncogenic endpoint. Finally, broadening the scope of investigation to co-mutating signaling networks uncovered allele-specific enrichment of mutations in key pathways. The general trend was that the *RAS* alleles of a cancer all had a high frequency of mutation in regions of known oncogenes and tumor suppressors, but each tended to have additional allele-specific modules of co-mutation. This study contributes evidence for *KRAS* allele-specific effects on the genetic landscape of a tumor, potentially resulting in distinct behavior of the cancer.

## Introduction

*KRAS* and *NRAS* are two prominent members of the *RAS* family of highly homologous, small GTPases, ubiquitously expressed in humans [﻿3304147, ﻿22589270]. They operate as GTP-regulated signaling hubs, relaying extracellular signals to key intracellular functions such as growth, proliferation, metabolism, and motility. When bound to GTP, Ras proteins activate their downstream effectors via protein-protein interactions until they are deactivated by hydrolyzing the GTP to GDP with assistance from a GTPase-activating protein (GAP). The release of the GDP, facilitated by a guanine nucleotide exchange factor (GEF), allows Ras to again bind a GTP molecule, returning to its active state [﻿3304147, ﻿28630043]. Mutations of *KRAS* and *NRAS* that increase the time spent in the GTP-bound state are frequently found in cancer, though are highly enriched in colorectal adenocarcinoma (COADREAD), lung adenocarcinoma (LUAD), multiple myeloma (MM), pancreatic adenocarcinoma (PAAD), and skin cutaneous melanoma (SKCM) (**Figure 1A**).

Along with cancer specificity, the particular mutations commonly found in *RAS* are different in many fundamental aspects. Most mutations of *KRAS* and *NRAS* are found at codons 12, 13, and 61, however, the overwhelming majority of mutations in COADREAD, LUAD, and PAAD are at *KRAS* G12, while *NRAS* Q61 mutations are the most common in SKCM, and MM features a large portion of *KRAS* Q61 mutations (**Figure 1B and C**). Further, *KRAS* A146 mutations are almost exclusive to COADREAD, with a few instances in MM. Finally, *RAS* alleles can be highly tissue specific, such as *KRAS* G12R in PAAD or *KRAS* G13D in COADREAD and MM (**Figure 1C**). To confound the strict specificity, many amino acid substitutions at the three hotspots, including all at G12 save for proline, are transformative *in vitro*, though only a handful of substitutions are found in patients [3304147]. In addition, using an adeno-associated virus/CRISPR-Cas9 approach to induce mutations in *KRAS* at G12 and G13 in the lungs of mice revealed surprisingly high rates of G12R and G13R tumors in lung, while the G12C allele, the most common in human LUAD patients, was found to drive tumor progression about half as frequently as G12D [29233960]. The same analysis in the pancreas revealed specificity for the location of G12R and G12D alleles within the organ, but not for G12V [29233960]. Importantly, we are unaware of any such patterns in the specificity of the common *NRAS* alleles. These examples of striking specificity of *KRAS* mutations argue that each allele has unique properties beyond its ability to hyperactivate the protein.

The alleles differ in their fundamental biochemistry. While mutations at all three hotspots decrease intrinsic or GAP-mediated GTP hydrolysis, G13 and Q61, but not G12 mutations increase GDP-exchange rates [26037647, 23487764]. In addition, the different amino acid substitutions induce varying levels Ras activation [26037647]. The classification of *RAS* alleles as decreasing GTP-hydrolysis or increasing GDP exchange can extend to less common alleles such as *KRAS* A146T which increases the rate of nucleotide exchange over wild-type *KRAS* with little change to the rate of GTP hydrolysis [16969076, 20570890, Poulin 2018]. Interestingly, the degree of *KRAS* activation does not correlate to the prevalence of the mutation *in vivo* [﻿26037647], suggesting additional oncogenic properties of *KRAS* are possessed by the more common alleles to promote their selection.

The common substitutions alter effector binding in different ways [﻿29444439], thus the dysregulation of downstream signaling is *RAS* allele-specific. For example, Kras G12D and G12V exhibit reduced binding affinity for Raf at the head of the Mek/Erk signaling pathway, though still induce hyperactivity due to their low GTPase rates [26037647]. Additionally, *KRAS* G12D connotes anchorage-independent growth and protection against apoptosis through increased levels of activated APC [11118062]. One proposed mechanism for how the oncogenic mutations near the GTP-binding site of *RAS* perturb effector binding is by altering which of the two GTP-bound conformations (known as switch I and II) is energetically preferred [15589837, 24441586]. Recently, using molecular dynamics simulations, another model has been proposed claiming that each *KRAS* allele effects the conformation of the effector binding region through a network of hydrophobic residues [30199525]. Whatever the mechanism, it is clear that each allele has substantially variable consequences on downstream signaling of *KRAS*, manifesting in specific behaviors of the tumor.

These behaviors are apparent in comparisons of the clinical results by *RAS* allele. For instance, the *KRAS* G12D allele is associated with reduced overall survival in advanced PAAD when separately compared to patients with wild type *KRAS*, *KRAS* G12V, or *KRAS* G12R (the other two most common *KRAS* variants found in PAAD) [27010960]. Another remarkable case of allele-specific clinical outcomes is how, in contrast to the established belief that *KRAS*-mutant tumors are unresponsive to anti-EGFR therapy, *KRAS* G13D tumors are sensitive to cetuximab [20978259]. Kinetics modeling has suggested that reduced binding of Kras G13D to the GAP NF1 increases the regulation of the wild-type allele, making it more susceptible to anti-EGFR therapies [McFall, 2018, *bioRxiv*]. The clinical relevance of the specific *KRAS* allele warrants investigation into the specific attributes of each.

Taken together, it is clear that each *RAS* mutation uniquely dysregulates its signaling network. However, though *RAS* mutations are often an early event in oncogenesis [﻿25301631 (lung), 22226782 (pancreas), ﻿30420765 (review)], what is not well understood is the impact of these mutations on the evolution of the tumor’s mutational landscape. We hypothesize that the individual characteristics of each *KRAS* allele helps to determine supplemental mutations to the tumor’s genome to drive oncogenesis and progression. Therefore, the impact of endogenous and exogenous mutational processes was measured and used to predict the likelihood of obtaining each *RAS* mutation. In addition, the impact of the *RAS* allele on the mutational landscape of the cancers was estimated by identifying patterns of co-mutation and mutual exclusivity. Finally, the effects of the differential signaling processes were investigated by identifying allele-specific synthetic lethal interactions and mutational hotspots in distal signaling modules. Overall, this study aims to highlight the possibility that the different properties of the *RAS* alleles determine specific supplementary perturbations of the cancer’s genetics, ultimately contributing to the tumor’s behavior and the patient’s clinical outcome.

## *KRAS* allele frequency in *KRAS*-driven cancers

*KRAS* and *NRAS* share over 90% sequence identity in the catalytic domain (though they differ considerably in the hypervariable region responsible for interacting with the cell membrane) and the same mutational hotspots [﻿3304147, 28958387]. Of these hotspots, codon 12 mutations accounted for over the 70% of all ­mutations followed by codons 61 (10%), 13 (5%) and 146 (less than 5%, almost exclusively found in COADREAD). Of the *KRAS* or *NRAS*-driven cancers, COADREAD, LUAD, MM, PAAD, and SKCM, there was variability in in the frequency of the *RAS* alleles. PAAD and SKCM were the least variable: only *KRAS* was mutated in PAAD with almost all mutations occurring at G12 (G12D/V/R), and SKCM harbored primarily *NRAS* Q61 mutations (Q61R/K/L and rarely H). Alternatively, MM was the most variable; both *KRAS* and *NRAS* were commonlymutated at both codons 13 and 61.

Most *RAS* mutations were caused by single nucleotide variants (**Figure 1B)**. An exception was the exceedingly rare, though transformative *in vitro* [﻿3304147, Hahn’s saturated mutagenesis of KRAS?], G12F allele (< 1%) caused by the dinucleotide substitution c.34\_35GG>TT. Glycine at codon 12 can be transformed to 6 different amino acids (C, D, S, R, A, and V) through single nucleotide changes in the first two cytosines. Glutamine at codon 61 can mutate to only 4 other amino acids (H, K, L, and R). All 96 amino acid substitutions can be represented by the combination of the 6 possible pyrimidine to purine base substitutions (C>T, C>A, C>G, T>A, T>C, T>G) and all the possible 3’ and 5’ flanking bases. Comparing the spectrum of *RAS* mutations to the genome-wide mutational spectrum, the *RAS* mutations were caused by relatively rare events (**Figure 2A**). For example, although *KRAS* G12D is created by a C>T transition, the specific trinucleotide context within which this mutation arises (A[C>T]C) is not the canonical CpG context that is particularly common in tumors of the colon and pancreas (**Figure 1B and 2A**). Driver mutations are known to follow a different pattern than the average overall mutation landscape of tumor samples [http://dx.doi.org/10.1101/485292].

## Mutational processes partially elucidate *RAS* mutation origins

To elucidate the origin of these mutations, the likelihood that specific *RAS* alleles were caused by active mutational processes in each tumor sample was calculated exploiting mutational signatures analysis [﻿23945592] (**Supplementary Figure 1 and 2**, see methods for further detail). Distinguishing between “clock-like” Signatures 1 and 5 (mutations believed to accumulate with age [﻿26551669]) from all the other signatures, *RAS* mutations in COADREAD, MM, and PAAD could be attributed to aging in about 70% to 100% of the cases (**Figure 2B**, second column). Signature 9, mutations introduced by polymerase η linked to activation-induced deaminase (AID) [citation needed for pol eta statement, 20887897], played a large role in some MM samples on specific *RAS* alleles, namely at *KRAS* codon 61 (**Figure 2B, C**). Specifically, there was an enrichment of Signature 9 in *KRAS* Q61H MM samples when compared to WT and all other alleles (**Figure 2C**). In SKCM, Signature 7 and 38 (signatures of UV-light exposure) were responsible for the remaining 50% of *NRAS* Q61 mutant samples (**Figure 2B**, fifth row). As all of the RAS mutated samples showed levels of Signature 7 and 38 comparable to WT (**Figure 2C**) and none of the *NRAS* Q61 mutations have a particularly higher chance to be caused by UV-light exposure (**Figure 2B**), it is unlikely that any of the NRAS Q61 alleles are selectively mutated by UV exposure. LUAD was the only cancer were *KRAS*-mutated samples have different levels of a signature according to the specific allele. Samples with *KRAS* G12A/C/V and G13C mutations were primarily burdened with mutations attributable to tobacco smoke (Signature 4 [27811275]) while *KRAS* G12D samples were split into about 25% Signature 4 and about 75% Signature 1 and 5 (**Figure 2B** second row and **Figure 2C**).

In COADREAD, Signature 18, likely caused by erroneous base excision repair of damage caused by reactive oxygen species [28551381, 28127763], was tended to be elevated in samples with *KRAS* G12 mutations, particularly in G12C, when compared to levels in *KRAS* WT samples. Signature 18, although not present at high levels in any of the COADREAD samples (**Supplementary Figure 1**), likely provided a high burden of C>A mutations that could not be accounted for by Signature 1 (**Supplementary Figure 2**), the most prevalent mutational signature in COADREAD. A connection between Signature 18 and KRAS G12C has been already speculated in the literature [28127763] and the rarity of the signature itself could explain why is relatively less present than KRAS G12D/Vmutation in COADREAD.

## Modeling the prevalence of *RAS* mutations

With the exception of Signature 9 in MM, the vast majority of mutational signatures contributing to the mutation of *RAS* were active prior to tumorigenesis and, thus, could be proposed to be causal of the initiating oncogenic mutation. However, the reason for obtaining one *RAS* mutation over another given the same mutagenic background was not apparent. Thus, we tested to what extent the frequency of each *RAS* allele in a cancer could be predicted from the genome-wide mutations in a sample by calculating the fraction of the sample’s mutations that were identical to the driving *RAS* mutation. For example, if in a given sample the number of mutations of the type A[C>T]C (*KRAS* G12D) was twice the number of C[C>A]A (*KRAS* G12C), the chance of randomly generating a G12D mutation was twice that of G12C. These probabilities were averaged across samples with the same *RAS* allele and compared to the observed prevalence of each *RAS* allele. We restricted the possible alleles to those observed in at least 5 samples of the same tumor type, so that there were enough alleles to model G12 in COADREAD, LUAD, MM, and PAAD and Q61 in MM and SKCM. Regression lines were shown for those tumors with at least 3 possible alleles observed per position (**Figure 2D** and methods). A well-defined linear relationship was be established for G12 mutations in LUAD (adjusted = 0.98, F-test p-value = 0.0026) and Q61 in SKCM (adjusted = 0.99, F-test p-value = 0.04), demonstrating that the prevalence of each allele could be determined by genome-wide mutational processes in tumors with a strong presence of exogenous mutagens (smoke and UV-light). In MM, a linear relationship for *KRAS* G12 and *KRAS* Q61 was pronounced (adjusted = 0.88 and 0.98, F-test p-value = 0.038 and 0.011, respectively), suggesting that *KRAS* G12D and *KRAS* Q61H alleles were the most frequently observed event because they had more chances to randomly occur. Alternatively, in MM, models of allele frequency on *NRAS* were highly unstable. COADREAD and PAAD did not show a defined linear relationship across all alleles, though a few trends were noticeable. In particular, *KRAS* G12D and G12V mutations in both tumors were highly prevalent and predictable from the genome-wide mutational events like in the other cancers analyzed. Alleles below the regression lines (such as *KRAS* G12R in PAAD and *KRAS* G12A in COADREAD) had a low predicted chance to occur but high observed prevalence. These results suggest that despite the low-likelihood of occurrence, the properties of these alleles provide a strong selective advantage, causing them to fix easily in the population. On the other hand, *KRAS* G12S in COADREAD and *KRAS* G12C in PAAD, could be considered “less-fit” alleles for these cancers: despite a relatively high probability of occurrence, their presence was rare. To note, *KRAS* G13 mutations were predicted to occur very frequently but were relatively rare in tumor samples. This too was likely due to a lower oncogenic “fitness” of G13 mutations than G12 mutations. Models that allow every possible allele to have the same chance of occurring (both across tumor types and across positions) were unable to predict the population’s allele distribution, showing in particular a clear difference in the properties of each of the position (12, 13, 61, 146) and gene (*KRAS* and *NRAS*). This is not surprising since mutations like G12C and G13C are caused by the same exact type of SNV (C[C>A]A) but in tumors like LUAD, G12C is ~10 times more frequent (14% and 1.3% of LUAD samples, respectively).

## Co-mutation and mutually exclusive genetic interactions

While the previous section investigated the origin and relative prevalence of *RAS* alleles, it did not describe the characteristics of each mutation in relation to the rest of the genome. To identify patterns of co-occurent and mutually exclusive mutations, a one-tail Fisher’s exact test (a right-tail to detect co-mutation, a left-tail for mutual exclusivity) was used to compare the association between the over 20,000 human genes and each *RAS* allele. Given the large number of tests (on the order of 105 tests), a standard p-value correction with no *a priori* hypothesis resulted in no significant associations. To increase sensitivity, the test was only applied to driver genes mutated in over 1% of the samples and expressed either in the normal tissue or in the cancer (see **methods**). With these restrictions, recurrent patterns of mutual exclusivity and co-occurrence were identified. In general, *RAS* mutations were mutually exclusive with the mutations of other gain-of-function oncogenes; for example *KRAS* and *EGFR*  and *KRAS* and *ERBB2* in LUAD (**Figure 3A**) and *KRAS* and *BRAF* in SKCM, PAAD, and COADREAD (**Figure 3B** and **Supplementary Figure 5**). *TP53,* which show behaviors of typical tumor suppressors and gain-of-functions oncogenes [26024390], was found mutually exclusive with several alleles in both LUAD and COADREAD. In MM, *KRAS* and *NRAS* are mutually exclusive but not with *BRAF* (mutated in ~7% of the samples), in agreement with reports of multiclonal tumors with a mutation in both RAS and *BRAF* genes in different cells [29209042]. The most prominent patterns of allele-specific co-mutation were found with known tumor suppressors with little overlap between different alleles and across cancers. In COADREAD, all RAS alleles co-mutate with *APC*, but specific alleles showed unique interactions like *KRAS* G12D and *SMAD2* and *KRAS* G12S and *ADAMTS9* (**Figure 3A**)*.* Similarly in LUAD, *STK11, KEAP1* and *ATM* co-mutate with multiple *KRAS* alleles but *KRAS* G12D significantly co-mutates with *GNAS* (which are known to cooperate in Intraductal papillary mucinous neoplasm [26257060])*,* *KRAS* G13C with *ROR2* and *KRAS* G12C with *POLE* (**Figure 3B**). These observations let us speculate that all RAS alleles act on the same pathway with the possible help of a specific tumor suppressor but tend to not “overcharge” the pathway itself with a secondary gain-of-function event. An exception to this rule can be seen in two different examples. In PAAD, *RAS* mutations were so frequent (mutated in ~90% of the samples) that there were no clear patterns of mutual exclusivity and little co-mutation (**Figure 3C**). In PAAD, *KRAS* requires the cooperation of tumor suppressors like *TP53,* CDKN2Aand *SMAD4,* but there is no apparent allele-specific determination [29229669, 24388967]. In COADREAD *KRAS* can cooperate with *PIK3CA* (mutated in about 20% of the samples). This relationship has been investigated in the literature and the coexistence is possible because, while there is no selective pressure to mutate *EGFR*, *BRAF* and *RAS* at the same time (they all contribute to the same MAPK signaling), *PIK3CA* is also involved in a different downstream signal (PI3K pathway) [20619739].

Previously, *PIK3CA* alterations were divided in exon 9 (mostly E545K and E542K) and exon 20 (mostly H1047R) mutations, showing how the co-mutation pattern is observable on exon 9 but not on 20 [20619739, 30361395, 23785428]. Since the sequencing panels collated for COADREAD all target *PIK3CA*, *KRAS*, and *BRAF* (over 4,500 samples), we have enough statistical power to run the same kind of co-occurrence and mutual exclusivity analysis allele-wise. *KRAS* G12D, G12V and G13D all significantly co-mutated with *PIK3CA* E545K and *KRAS* G12D also co-mutated with *PIK3CA* Q546K (**Figure 3C**). When aggregated by position, G12, G13, and A146 all co-occurred with exon 9 mutations but not exon 20 (**Supplementary Figure 5**). This observation is consistent with the literature (*KRAS* cooperating with *PIK3CA*) but refines the relationship between the two genes. For example, *KRAS* Q61 mutations do not tend to co-mutate with *PIK3CA* (only 1 case contained both *PIK3CA* E545 and *KRAS* Q61 mutations in a total of 92 samples with a *KRAS* Q61 mutation) while A146, an allele with reduced activity compared to other oncogenic mutations [30952657], co-mutates in a similar fashion as the more frequent G12 and G13. The allele-wise analysis is instead in agreement with the gene-wise analysis on *BRAF*, where essentially all *KRAS* alleles tend to be mutually exclusive with *BRAF* V600E. *BRAF*, on the other hand, is able to significantly co-mutate with *PIK3CA* H1047K (exon 20) and *PIK3CA* Q546K (exon 9), but not with the more frequent E545K (that elicits *KRAS* more often). This contradicts previous reports on the inability of *BRAF* to cooperate with *PIK3CA* [30361395] and delineates a clearer interplay among these genes.

## Allele-specific synthetic lethality

The distinct characteristics of *KRAS* alleles and the consequential differences in signaling properties open up the possibility of allele-specific genetic dependencies. Termed “synthetic lethality” [16110319, 16572121, 25354106], the oncogenic state induced by an activating *KRAS* allele can make the cancerous cells more dependent on specific pathways or cellular functions, thus creating new potential drug targets out of normally non-essential genes. The Cancer Dependency Map (DepMap) project has worked to elucidate these dependencies by systematically measuring the effect of individually knocking-out 17,634 coding genes in over 500 cancer-derived cell lines [28753430, 29083409]. While the distinct dependencies of *KRAS*-mutant tumors have been investigated [28753431, 19847166], the individual alleles have been largely ignored.

The DepMap project screened 27 COADREAD, 76 LUAD, 16 MM, 25 PAAD, and 34 SKCM derived cell lines, of which 13 COADREAD, 26 LUAD, and 23 PAAD had a *KRAS* mutation, 8 SKCM had a *NRAS* mutation, and 9 MM had either a *KRAS* or *NRAS* mutation (only including mutations in hotspot codons; **Supplementary Figure JHC1A**). Only alleles found in at least three cell lines of a cancer were used in the following analysis, which restricted the analysis to the most common *KRAS* alleles of COADREAD, LUAD, and PAAD cell lines. The metric for genetic dependency was a measure of the depletion of the CRISPR sgRNA for the given target gene [29083409]; thus a negative score indicated a greater dependency. Confirming the importance of the *KRAS* oncogenic mutation [19477428, 22341439], most of the *KRAS*-mutant cell lines used were strongly dependent on *KRAS*(**Supplementary Figure JHC1D**). A gene was deemed to be synthetic lethal if it had a dependency score within or less than one standard deviation of genes described as pan-essential (were uniformly lethal across cancer types) [29083409]. Of these synthetic lethal genes, they were determined to be allele-specific using an ANOVA test across the *KRAS* mutants (*p*-value < 0.1). To understand the functions of the allele-specific synthetic lethal genes, WebGestalt was used to identify enriched pathways and cellular functions [28472511]. Unsurprisingly, across cancer types, the most commonly enriched cellular functions were related to cellular respiration, transcription, translation, and replication. However, there were still distinct genetic interactions between some target gene and *KRAS* alleles. Below, a few instances of allele-specific synthetic lethality were described, though more were made available in the Supplementary Information (**Supplementary Figures JHC2-4** for COADREAD, LUAD, and PAAD, respectively).

In COADREAD cell lines, *KRAS* G12V had many allele-specific genetic interactions, primarily with subunits of mitochondrial ribosomal and NADH dehydrogenase. In contrast, *KRAS* G13D had the lowest overall allele-specific dependency on mitochondrial genes (**Figure JHC1A**). Of note, *POLRMT*, mitochondrial DNA-directed RNA polymerase, was also weakly mutually exclusive with *KRAS* G12V (*p*-value = 0.07), and *FASTKD5*, mitochondrial FAST kinase domain-containing protein 5, was weakly mutually exclusive with *KRAS* G12D (*p*-value = 0.0930). Concordantly, these two genes appear to be more critical for the survival of cell lines of the same *KRAS* genotype (highlighted nodes in Figure JHC1A). While synthetic lethality with mitochondrial genes was expected, the specificity of the genetic dependencies may reflect distal effects of the *KRAS* allele’s signaling or specific metabolic requirements.

LUAD cell lines displayed strikingly clear separation of dependency on genes implicated in cell cycle regulation (**Figure JHC1B**). Two distinct clusters of genes were identified, one demonstrating strong association with *KRAS* G12C and the other with *KRAS* G12V, suggesting the presence of allele-specific dependencies for the regulation of the cell cycle. *NUMA1*, nuclear mitotic apparatus protein 1, was mutated in a weakly mutually exclusive manner with *KRAS* G12C in human LUAD samples (*p*-value = 0.0905) and consistently had a lower dependency score with this allele compared to *KRAS* G12V. Though further experimentation is required to confirm any hypotheses, these results suggest that there are unique properties of cell cycle regulation dependent upon the *KRAS* alleles of LUAD tumors.

Interestingly, the distinction between synthetic lethal interactions of the *KRAS* alleles was less clear in PAAD cell lines. This concurred the trend of low co-mutation and mutual exclusive genetic interactions found in tumor samples (**Figure 3C**). Of the more distinct differences was the greater dependency of *KRAS* G12R on mitochondrial ribosomal subunits than *KRAS* G12D or G12V (**Figure JHC1C**). Additionally, cell lines with *KRAS* G12V mutations were consistently less dependent on mitochondrial genes than the other two alleles. These data provide further evidence to suggest that the metabolic requirements of PAAD tumors rely, in part, on the specific *KRAS* allele.

Of particular interest were genes implicated in canonical oncogenic signaling pathways. Thus, allele-specific genetic dependencies were found for genes involved in apoptosis, cell cycle, Notch/Wnt signaling, p53 signaling, and Ras signaling [27899662] using the same tests as above. In COADREAD cell lines (**Figure JHC2A**), *KRAS* G12D/V and G13D cell lines had differing levels of dependence on the cell cycle regulators *CCND1* and *CDK2*: whereas *KRAS* G12D cell lines were strongly dependent on both genes, *KRAS* G12V cell lines were consistently more reliant upon *CCND1* than *CDK2*. In addition, KRAS G12D cell lines showed stronger dependence on *MDM2*, a gene coding for a E3-ubiquitin ligase important in the negative regulation of p53 mediated cell cycle arrest and apoptosis [9450543, 9153395, 9153396]. This could suggest a greater reliance on the down-regulation of p53 for *KRAS* G12D than for *KRAS* G12V or G13D in COADREAD. In LUAD, compared to cell lines with KRAS G12V mutations, those with *KRAS* G12C were consistently more dependent upon *PTK2*, which codes for focal adhesion kinase (FAK) and is critical for the regulation of cell adhesion, apoptosis, and the cell cycle [21118706, 28213315]. Alternatively, KRAS G12V cell lines had a greater dependence upon *SOCS3*, a negative regulator of cytokine signaling through the JAK/STAT pathway. This was noteworthy as JAK/STAT signaling activates the MEK/ERK pathway through *KRAS*, suggesting a greater role for this mechanism in KRAS *G12V* mutant LUAD tumors. Lastly, as mentioned previously, the distinctions between the genetic dependencies of the *KRAS* alleles in PAAD cell lines was less clear than in other cancer types. However, compared to cell lines with either *KRAS* G12D or G12V mutations, there tended to be a reduced reliance of *KRAS* G12R mutants on *GRB2*, a critical signaling adapter between growth receptors and Kras. Though intriguing patterns of *KRAS* allele-specific genetic interactions have already emerged, more will likely be discovered as the DepMap project continues to screen additional cell lines.

## Co-mutation of modules of the protein-protein interaction network

Because proteins rarely act alone, a cellular process can be perturbed through the mutation of a number of various proteins. Thus, after inspecting *KRAS* allele-specific genetic interactions with individual genes, measuring mutation rates of proteins that interact with each other can further highlight modules of the cellular signaling network that demonstrate allele-specific disruptions. Therefore, HotNet2 [25501392] was employed to identify modules of the protein-protein interaction (PPI) network [﻿22846459] that tended to be mutated in samples with *RAS* mutations. Briefly, HotNet2 selects regions of the PPI that exhibited a high rate of mutation by treating mutational frequency as heat, allowing it to partially diffuse across the network, and identifies regions that are “hotter” than expected. Using this method, mutated signaling modules were recognized for each *RAS* allele across the five cancers, though only alleles with at least 20 samples demonstrated sufficient stability and reproducibility. Overall, while samples with different *KRAS* alleles generally all had high mutational frequency of known oncogenes and tumor suppressors, they also had allele-specific regions of high mutation.

COADREAD samples with *KRAS* G12C/D/V or G13D mutations all displayed high mutations rates in *APC* and *TP53* signaling modules, though each had additional components exhibiting uniquely high rates of mutation (**Supplementary Fig JHC2** and **Fig JHC2**). *KRAS* G12D samples had an exceptionally high proportion of samples with mutations in a portion of the complement pathway, an integral component of the innate immune system (*p*-value = 0.00175, **Fig JHC2A**) [24161035]. This could point to a disruption of the immune surveillance of the tumor microenvironment [20720586, 28248200]. Samples harboring a *KRAS* G12C allele also had a higher rate of mutation in genes involved in a module of PTEN signaling (*p*-value = 0.000819, **Fig JHC2B**). This may suggest an allele-specific role for this signaling pathway of *PTEN*, a prominent tumor suppressor important to COADREAD due to its antagonistic relationship with *PIK3CA* [18767981, 18794886]. Additionally, samples with *KRAS* G12C mutations had a higher frequency of mutation in a module of proteins important for regulating Wnt and cell cycle, including *ROBO2*, *RPS6KA5*, and *MDFIC* (*p*-value = 0.00100, **Fig JHC2C**). With respect to the importance of Wnt signaling to COADREAD, the frequent mutations in this signaling module in *KRAS* G12C samples may point to an additional mode of Wnt hyperactivation unique to this *KRAS* allele. Thus, a network-level analysis of co-mutation in *KRAS*-mutant COADREAD samples identified allele-specific disruption of signaling modules that were implicated in key oncogenic pathways. Though *TP53* tends to be mutually exclusive with most of *KRAS* alleles in COADREAD (63% of the samples harbors a *TP53* but only 56% of them are *KRAS* mutants), there were differential rates of mutation of its nearby interactors. The most striking example was the almost doubled rate of mutation in *KRAS* G12C samples in the cell adhesion and cell cycle regulator molecules, *NRCAM*, *MAGI2/3*, *DSCAML1*, *ACVR2A* and *INHBA* (*p*-value = 0.0102, **Fig JHC2D**). This fit the trend of mutations in cell membrane-associated adhesion and growth signaling proteins in *KRAS* G12C COADREAD tumor samples as noted previously.

*KRAS* G12A/C/D/V LUAD samples demonstrated both remarkable similarity and diversity of co-mutating signaling modules (**Supplementary Fig JHC3** and **Fig JHC3**). For instance, *KRAS* G12V samples were significantly enriched in mutations of genes involved in two cell-surface to cell-cycle regulatory mechanisms. The first included the interactions of cadherin-13 and hyaluronan mediated motility receptor connecting to the interaction of Jagged-1 and Notch-2 (*p*-value = 0.000949, **Fig JHC3A**). The second incorporated the signaling of the adhesion DLGAP1-Shank1/2 complex to Timeless, a regulator integral to the timing and stability of the DNA replication (*p*-value = 0.000367, **Fig JHC3B**). The higher rate of mutation in these pathways in tumors with *KRAS* G12V mutations may hint at a specific disruption of the cell-cycle that creates a selective advantage in the context of this *KRAS* allele. Another example of allele-specific co-mutation in LUAD is the enrichment of mutations in a signaling module containing proepiregulin and Robo-1, molecules important in the regulation of growth and cell motility (*p*-value = 0.00736, **Fig JHC3C**). In *KRAS* G12A LUAD samples, there was an enrichment of mutations in the complex of chromatin modifiers and transcription initiation factors and co-activators (p-value = 0.0000490, **Fig JHC3D**). The disruption of this signaling subnetwork may misregulate the expression of genes that complement the unique properties of *KRAS* G12A signaling, thus providing a selective advantage.

The tumor suppressor *STK11* was mutated in 15% of LUAD tumors total and 20% in *KRAS*-mutant LUAD tumors, though there was variation in the rate of mutation of its signaling partners in samples with different *KRAS* alleles. LUAD samples with *KRAS* G12A mutations had an enrichment of mutations in *FAIM3*, *PLD3*, *PHF23*, and *TBC1D3*, all direct interactors with STK11 (*p*-value = 0.00331, **Fig JHC3E**). Alternatively, *KRAS* G12C samples were enriched in mutations of *ARG2*, *D2HGDH*, *ERLEC1*, and *FAIM2* (*p*-value = 0.000000592, **Fig JHC3F**). Finally, *KRAS* G12C/D samples showed enrichment of mutations in the tumor suppressor *SERPINB5* and its interactors *RAB9B*, *RABGGTA*, and *CHML* (*KRAS* G12C *p*-value = 0.0000117, *KRAS* G12D *p*-value = 0.0418, **Fig JHC3G**). Though confirming any causal associations is beyond the scope of the present work, these findings suggested that specific disruptions to *STK11* signaling complement the different *KRAS* alleles due to their specific signaling properties.

## Discussion

The quick brown fox jumps over the lazy dog