Title: Tissue-specific oncogenic activity of K-Ras^{A146T}

Authors:

Emily J. Poulin[1,2], Asim K. Bera[3*], Jia Lu[3*], Yi-Jang Lin[1,2], Samantha Dale Strasser[1,4,5], Joao A. Paulo[6], Tannie Q. Huang[7], Carolina Morales[7], Wei Yan[3], Joshua Cook[1,2], Jonathan A. Nowak[8], Douglas K. Brubaker[1,2,4], Brian A. Joughin[9], Christian W. Johnson[1,2], Rebecca A. DeStefanis[1,2], Phaedra C. Ghazi[1,2], Sudershan Gondi[3], Thomas E. Wales[10], Roxana E. Iacob[10], Lana Bogdanova[7], Jessica J. Gierut[1,2], Yina Li[1,2], John R. Engen[10], Pedro A. Perez-Mancera[11], Benjamin S. Braun[7], Steven P. Gygi[6], Douglas A. Lauffenburger[4], Kenneth D. Westover[3], Kevin M. Haigis[1,2,12]

Affiliations:

1. Cancer Research Institute, Beth Israel Deaconess Medical Center, Boston, MA, USA.
2. Department of Medicine, Harvard Medical School, Boston, MA, USA.
3. Departments of Biochemistry and Radiation Oncology, The University of Texas Southwestern Medical Center at Dallas, Dallas, TX, USA.
4. Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA.
5. Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, USA.
6. Department of Cell Biology, Harvard Medical School, Boston, MA, USA.
7. Department of Pediatrics and Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, CA, USA.
8. Department of Pathology, Brigham and Women's Hospital, Boston, MA, USA.
9. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA, USA.
10. Department of Chemistry and Chemical Biology, Northeastern University, Boston, MA, USA.
11. Department of Molecular and Clinical Cancer Medicine, Institute of Translational Medicine, University of Liverpool, Liverpool, UK.
12. Harvard Digestive Disease Center, Harvard Medical School, Boston, MA, USA.

*Equally contributing authors

Running title: Allele- and tissue-specific K-Ras function

Corresponding authors:

Kevin M. Haigis, Beth Israel Deaconess Medical Center, 3 Blackfan Circle, CLS409, Boston, MA, 02115. Phone: 617-735-2056; Email: khaigis@bidmc.harvard.edu.

Kenneth D. Westover, UT Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75390. Phone: 214-645-0323; Email: Kenneth.Westover@UTSouthwestern.edu

Disclosure of Potential Conflicts of Interest: The authors have no conflicts to declare.

ABSTRACT

KRAS is the most frequently mutated oncogene. The incidence of specific *KRAS* alleles varies between cancers from different sites, but it is unclear whether allelic selection results from biological selection for specific mutant K-Ras proteins. We used a cross-disciplinary approach to compare K-Ras^{G12D}, a common mutant form, and K-Ras^{A146T}, a mutant that occurs only in selected cancers. Biochemical and structural studies demonstrated that K-Ras^{A146T} exhibits a marked extension of switch 1 away from the protein body and nucleotide binding site, which activates K-Ras by promoting a high rate of intrinsic and GEF-induced nucleotide exchange. Using mice genetically engineered to express either allele, we found that K-Ras^{G12D} and K-Ras^{A146T} exhibit distinct tissue-specific effects on homeostasis that mirror mutational frequencies in human cancers. These tissue-specific phenotypes result from allele-specific signaling properties, demonstrating that context-dependent variations in signaling downstream of different K-Ras mutants drive the *KRAS* mutational pattern seen in cancer.

SIGNIFICANCE: Although epidemiological and clinical studies have suggested allele-specific behaviors for *KRAS*, experimental evidence for allele-specific biological properties is limited. We combined structural biology, mass spectrometry, and mouse modeling to demonstrate that the selection for specific K-Ras mutants in human cancers from different tissues is due to their distinct signaling properties.

INTRODUCTION

Effective implementation of precision medicine requires a detailed understanding of the genetic, biochemical, biological, and clinical/epidemiological properties of cancer genes and their corresponding proteins. *KRAS* is the most frequently mutated oncogene in cancer. Its oncoprotein products, K-Ras4A and K-Ras4B, play an active role in tumor pathogenesis by engaging downstream signal transduction cascades to promote proliferation and survival. Efforts to target K-Ras and its effector signaling pathways have largely failed, making oncogenic K-Ras a major barrier to precision medicine for cancer. As a small GTPase, K-Ras activity is modulated by its nucleotide binding state. While wild-type (WT) K-Ras cycles between the GDP-bound (inactive) state and the GTP-bound (active) state with the help of guanine nucleotide exchange factors (GEFs) and GTPase activating proteins (GAPs), somatic mutations found in cancer promote the GTP-bound state (1).

K-Ras can be activated by missense mutation at any of a number of residues. Most cancer-associated mutations occur at glycine 12 and these impair GAP-induced hydrolysis (2). Studies of genetically engineered mouse models have revealed that expression of K-Ras^{G12D}, the most common allele in human cancers, is sufficient to disrupt homeostasis in a variety of tissues (3-7). Some K-Ras mutations are restricted to certain tumor types. For example, alanine 146 (A146) is the fourth most common site of K-Ras mutation across all cancer types, but, A146 mutations are found almost exclusively in cancers of the intestinal tract and blood (8). Relative to other mutant forms of K-Ras, little is known about the biochemical and biological properties of K-Ras mutated at A146. A146V, the first mutant allele characterized, was originally identified through random mutagenesis almost three decades ago and was determined to alter H-Ras GTP binding affinity and to increase the rate of intrinsic nucleotide exchange (9). It was not until 2006 that A146 mutations (A146V and A146T) were identified as cancer-associated *KRAS* alleles (10,11). Although *KRAS* alleles are not typically differentiated in clinical practice, epidemiological studies suggest that cancers expressing different mutant forms of K-Ras exhibit distinct clinical behaviors (1,12). For example, in colorectal cancer (CRC) codon 146 mutations are associated with better overall survival relative to codon 12 mutations, however, like codon 12 mutants, they promote resistance to anti-Epidermal Growth Factor Receptor (EGFR) therapy (11,13,14). Whether the biological properties of K-Ras proteins activated through different biochemical mechanisms are similar or distinct, and whether distinct clinical behaviors derive from distinct biological properties of *KRAS* alleles, is unresolved. In this study, we undertook a comprehensive, multifaceted approach to systematically analyze the similarities and differences between mutant forms of K-Ras in order to understand mechanisms driving allelic selection in cancer.

RESULTS

A146T promotes nucleotide exchange

Missense mutations activate Ras proteins by altering their abilities to bind and/or hydrolyze GTP. To determine how mutation of codon 146 impacts the core biochemical properties of K-Ras, we first measured GDP dissociation +/- SOS1, a Ras GEF (15). We found that K-Ras4B^{A146T} had a ~12-fold higher rate of intrinsic GDP dissociation than K-Ras4B^{WT}, which was further increased by addition of SOS1 (Fig. 1A,B). We noted that SOS1 stimulated exchange more effectively for K-Ras4B^{A146T} than it did for K-Ras4B^{WT} or K-Ras4B^{G12D} and microscale thermophoresis demonstrated a roughly 100-fold increase in affinity of SOS1 for K-Ras4B^{A146T} relative to WT protein (Fig. 1C). Next, we measured the rate of intrinsic and GAP-stimulated GTP hydrolysis. Intrinsic rates were similar between K-Ras4B^{A146T} and K-Ras4B^{G12D}, but reduced relative to K-

Ras4B^{WT} (Fig. 1D), while GAP-stimulated GTP hydrolysis was significantly impaired for K-Ras^{G12D}, but only mildly impaired for K-Ras4B^{A146T} (Fig. 1E). We also measured the affinity of K-Ras4B for its downstream effector RAF by using an AlphaScreen-based assay (16). All proteins exhibited similar binding affinity, indicating that in its GTP-bound form, K-Ras4B^{A146T} assumes a more typical closed conformation that is able to engage RBD (Supplementary Fig. S1). These data are consistent with a model in which K-Ras4B^{A146T} activation derives primarily from an increase in intrinsic and GEF-induced nucleotide exchange, not GAP insensitivity.

Structural basis of rapid nucleotide exchange

To understand the molecular basis for the increased basal and GEF-stimulated exchange rates of K-Ras^{A146T}, we solved a crystal structure of K-Ras4B^{A146T} in complex with GDP (Supplementary Table 1). The structure is notable for a pronounced extension of switch 1 and β2 away from the body of the protein and nucleotide binding pocket (Fig. 2A, Supplementary Fig. S2). In most K-Ras crystal structures, switch 1 participates in nucleotide binding, forming interactions with the base, sugar, and phosphates(17-19). Nevertheless, in the K-Ras4B^{A146T} structure, the C-terminal region of the α1 helix undergoes a large conformational change, forming a new anti-parallel β interaction with what was originally β2 (now $\beta3_{new}$ in K-Ras4B^{A146T}) and the usual β2-β3 interaction is disrupted (Fig. 2A, Supplementary Fig. S2). In this position, switch 1 no longer participates in GDP binding. The active site is notable for loss of the magnesium ion (Fig. 2B), leading to alterations in switch 2 (Fig. 2A). Switch 1 destabilization appears to result from a new interaction with N7 of the purine base and sidechain oxygen of the mutated residue at Thr146 which draws the guanine ring closer to residue 146 and slightly alters the pitch of the guanine base. In conjunction, the loop containing Thr146 is shifted towards α1, bringing Lys147 closer to the base and enabling a new interaction between N3 of the base and the Lys147 side chain (Fig. 2B). Lys147 in this new position would be expected to clash with Phe28 in the closed switch 1 position. This clash and/or alterations in the nucleotide ring position destabilizes switch 1, releasing it to reorganize into the open conformation, leading to destabilization of switch 2 and loss of interactions with magnesium, which contributes to increased nucleotide dissociation.

To confirm that this conformation is not an artifact of crystallization, we performed hydrogen-deuterium exchange mass spectrometry on K-Ras4B^{WT}, K-Ras4B^{G12D}, and K-Ras4B^{A146T} (Fig. 2C, Supplementary Figs. S3-S5). Consistent with an extended switch 1, K-Ras4B^{A146T} demonstrated increased rates of deuterium exchange (relative to K-Ras4B^{WT} and K-Ras4B^{G12D}) in switch 1 and areas normally shielded by switch 1, such as the P loop (Fig. 2C). Of note, we also saw increased exchange in loops beneath the nucleotide ring, likely related to rapid nucleotide exchange.

Similar to the K-Ras4B^{A146T}:GDP structure, the SOS1:H-Ras co-crystal structure exhibits an extension of switch 1 away from the body of H-Ras (Fig. 2D), consistent with the role of SOS in catalyzing nucleotide exchange (20). The K-Ras4B^{A146T} structure suggests that enhanced SOS1-mediated GDP dissociation is related to altered switch 1 dynamics that enable interactions between SOS1 and K-Ras^{A146T}. In short, these studies provide a structural mechanism for the increased rate of both intrinsic and GEF-induced nucleotide exchange seen for K-Ras4B^{A146T}.

Tissue-specific phenotypes of K-Ras alleles

Distinct mechanisms of activation likely lead to differences in the steady state and kinetic levels of activated K-Ras in a cell, but it is difficult to predict whether this translates into distinct oncogenic behaviors. To study the contextual responses to K-Ras alleles activated by different mechanisms, we generated a mouse strain carrying a Cre-inducible allele (*Kras*$^{\text{LSL-A146T}}$) in the endogenous *Kras* locus (Supplementary Fig. S6A-D). This new mouse model – with the exception of the activating missense mutations – is identical to the extensively characterized *Kras*$^{\text{LSL-G12D}}$ allele (4,21), allowing us to directly compare the molecular and cellular effects of the two mutations.

We first crossed *Kras*$^{\text{LSL-A146T}}$ mice to *Fabp1-Cre* mice, which express Cre recombinase in the colonic and distal small intestinal epithelia (22). We found that colons expressing K-Ras$^{\text{A146T}}$ displayed moderate crypt hyperplasia that was intermediate between K-Ras$^{\text{WT}}$- and K-Ras$^{\text{G12D}}$-expressing colonic epithelium (Fig. 3A-C). Using immunohistochemistry for phosphorylated histone H3 (PH3) to quantify crypt proliferation, we found that colons expressing K-Ras$^{\text{A146T}}$ had an intermediate hyperproliferative phenotype compared to those expressing K-Ras$^{\text{WT}}$ or K-Ras$^{\text{G12D}}$ (Fig. 3D). Thus, compared to the strong effect that K-Ras$^{\text{G12D}}$ has on basal colonic epithelial homeostasis, expression of K-Ras$^{\text{A146T}}$ caused intermediate cellular and tissue-level phenotypes.

To determine whether K-Ras$^{\text{A146T}}$ could alter hematopoietic homeostasis, we crossed *Kras*$^{\text{LSL-A146T}}$ to the *Mx1-Cre* driver, which directs polyinosinic:polycytidylic acid (pI/pC)-inducible Cre expression to hematopoietic stem cells (23). Induction of K-Ras$^{\text{A146T}}$ led to myelodysplastic syndrome / myeloproliferative neoplasm (MDS/MPN) that was qualitatively similar to that induced by K-Ras$^{\text{G12D}}$ mice, but with a significantly delayed onset. White blood cell counts slowly rose and hemoglobin levels declined over time in animals expressing K-Ras$^{\text{A146T}}$, with death occurring with severe anemia and splenomegaly at an older age than in animals expressing K-Ras$^{\text{G12D}}$ (Fig. 3E,F and Supplementary Fig. 7A). Flow cytometry of tissues from moribund mice revealed that the end-stage phenotypes in animals expressing K-Ras$^{\text{G12D}}$ and K-Ras$^{\text{A146T}}$ were similar, with expansion of immature myeloid cells in the bone marrow and spleen

(Supplementary Fig. 7B). We also tested whether K-Ras^{A146T}, like K-Ras^{G12D}, caused a cell-intrinsic hyperproliferative phenotype in myeloid progenitor cells. K-Ras^{A146T} had only a modest effect on colony growth, with only a few spontaneous colonies and a nearly normal GM-CSF dose response (Fig. 3G).

Although codon 146 mutations comprise ~10% of the *KRAS* mutations that occur in CRC, they rarely occur in other epithelial cancers, like pancreatic ductal adenocarcinoma (PDAC). To determine whether forced expression of K-Ras^{A146T} in the pancreas alters homeostasis, we crossed *Kras*$^{LSL\text{-}A146T/+}$ mice to *Pdx1-Cre* mice, which express Cre recombinase in the developing pancreas (6). Unlike K-Ras^{G12D}, which promoted the development of pancreatic intraepithelial neoplasia (PanIN) by 8 weeks-of-age, pancreases expressing K-Ras^{A146T} did not exhibit any histologic evidence of PanIN at this age (Fig. 3H). We followed a cohort of *Pdx1-Cre* ; *Kras*$^{LSL\text{-}A146T}$ beyond 2 months-of-age, including 4 animals beyond 300 days of age, and did not see any cases of PanIN. To determine whether acute pancreatitis could cooperate with K-Ras^{A146T} to promote PanIN, we treated animals with cerulein. Although this treatment accelerated the formation of PanIN in pancreases expressing K-Ras^{G12D}, it did not do the same in the context of K-Ras^{A146T} (Supplementary Fig. 8), indicating that this allele is not able to alter pancreatic homeostasis on its own.

K-Ras cooperation in tumor models

KRAS mutations do not drive cancer formation on their own, but rather cooperate with tumor suppressor gene mutations to promote malignant progression. As such, we posited that combining mutant K-Ras with appropriate secondary mutations might unmask the oncogenic potential of weak alleles like A146T. To explore this idea, we first crossed *Fabp1-Cre* carrying a *Kras*^{LSL} allele to those with a conditional allele of the *Apc* tumor suppressor gene, allowing us to determine the relative effects of K-Ras^{A146T} and K-Ras^{G12D} on colonic tumorigenesis. We found that colonic tumors expressing different K-Ras alleles shared common histologic features (Fig. 4A), but that they were significantly different in their abilities to reduce overall survival (Fig. 4B), just like in human CRC patients (11,13). The significant difference in survival among the different genotypes was likely due to a difference in the number of tumors, as the size of the tumors and the general histology – all were characterized as adenomas with low-grade dysplasia – were not significantly different (Supplementary Fig. S9A-B).

Next, we crossed our two K-Ras alleles to *Pdx1-Cre* ; *Tp53*$^{LSL\text{-}R270H/+}$ animals so that we could determine how each contributes to the development of pancreatic cancer. Interestingly, 9/12 of the *Pdx1-Cre* ; *Tp53*$^{LSL\text{-}R270H/+}$ animals expressing K-Ras^{A146T} developed pancreatic tumors, as did all of the controls expressing K-Ras^{G12D} (Fig. 4C). Although all of the cancers were invasive PDAC, the median lifespan of tumor-bearing mice expressing K-Ras^{A146T} was approximately 120 days, which represented a significantly extended survival relative to animals expressing K-Ras^{G12D} (Fig. 4D). Together, our studies demonstrate for the first time that different mutant forms of K-Ras exhibit distinct oncogenic properties that are tissue-specific and reflect the incidence of codon 146 mutations in human cancer.

Allelic imbalance is reported to be common in cancers expressing mutant K-Ras and to play a key role in regulating the response to MAPK inhibition (24). To determine whether allelic imbalance plays a role in cancers expressing K-Ras^{A146T}, we first analyzed mutant and wild-type frequency in primary human cancers from the TCGA (25). Although we were able to identify cancers with *KRAS* allelic imbalance, we found no significant difference in the frequency of mutant or wild-type allele copy number variation when cancers were expressing *KRAS*^{A146T} or *KRAS*^{G12D}

(Supplementary Fig. 10A,B). We also measured *KRAS* copy number in human cancer cell lines and primary mouse tissues by droplet digital PCR (ddPCR). Most of the colon cancer cell lines we analyzed had two copies of *KRAS*, although LS1034, which expresses K-Ras^{A146T}, had multiple copies of the gene, as previously reported (11) (Supplementary Fig. 10C). Similarly, we found that the majority of mouse tumors from colon and pancreas carried just two copies of *Kras* (Supplementary Fig. 10D). The lone exception was a PDAC expressing K-Ras^{A146T}, which had four copies. In summary, it does not appear that allelic imbalance or amplification plays is required to enhance the tumorigenic properties of K-Ras^{A146T}.

Global proteomic analysis of K-Ras alleles

Since mutant K-Ras functions as the initiator of many different signal transduction cascades, we posited that the distinct *in vivo* oncogenic activities exhibited by mutant forms of K-Ras either derive from distinct signaling properties of the individual mutants themselves or from regulatory mechanisms incident to GAP- or GEF-mediated activation. To address this hypothesis, we performed multiplexed mass spectrometry on colons, pancreases, and spleens from mice expressing K-Ras^{WT}, K-Ras^{G12D}, or K-Ras^{A146T} to generate total proteome, phosphoproteome, and scaled phosphoproteome datasets, which normalize phosphopeptide levels to total protein levels (Fig. 5A, Supplementary Tables S2-S10). All tissues were collected from animals that were approximately 2-months-old. At this age, colons expressing mutant K-Ras are hyperplastic, but do not differ from wild-type in the proportions of epithelial cell types. For pancreases and spleens, there is likely to be some difference in cellular representation between wild-type and mutant tissues at this age. Principal component analysis (PCA) of data from each tissue revealed that allele replicates clustered together, but that different genotypes separated in principal component space

(Supplementary Figs. S11-S13). PCA of all of the collected data revealed two general aspects of signaling: (1) samples from the same tissue tended to cluster together regardless of *Kras* genotype and (2) samples expressing K-Ras^{A146T} tended to cluster closer to samples expressing K-Ras^{WT} than they did to samples expressing K-Ras^{G12D} (Supplementary Fig. S14).

To uncover pathway-level biological insights from our MS data, we first used total protein data to perform gene-set enrichment analysis (GSEA) to identify Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched in either K-Ras^{A146T} or K-Ras^{G12D} mutant tissues relative to WT. In the colon dataset, we found that proteins corresponding to many KEGG pathways were significantly enriched or de-enriched in colons expressing K-Ras^{G12D} or K-Ras^{A146T} relative to WT (Fig. 5B, Supplementary Table S11). Only two pathways were significantly ($P < 0.1$ or FDR < 0.25) enriched in the presence of both K-Ras mutants (Fig. 5B). Importantly, we observed that the majority of the KEGG pathways analyzed were not coordinately regulated between the two K-Ras genotypes and several pathways were significantly affected in opposite directions by the two K-Ras mutants (Fig. 5B). Similar to the analysis of our colon dataset, GSEA applied to the pancreas dataset identified only a single KEGG pathway significantly enriched in the presence of both mutants (Fig. 5C, Supplementary Table S12). Similar to the colon, the majority of the KEGG pathways analyzed were discordantly regulated by the two K-Ras genotypes in pancreas and several pathways were significantly affected in opposite directions by the two K-Ras mutations (Fig. 5C). Interestingly, KEGG_NITROGEN_METABOLISM, which was coordinately up-regulated by K-Ras^{A146T} and K-Ras^{G12D} in the colon, was coordinately down-regulated in the pancreas (Fig. 5B,C). In spleen, K-Ras^{G12D} and K-Ras^{A146T} exhibited more similar signaling properties and there were no KEGG pathways that were discordantly regulated by the two mutant alleles relative to WT (Fig. 5D, Supplementary Table S13). In short, the two mutant forms of K-Ras induced distinct downstream signaling effects that were strongly influenced by the tissue in which they were expressed. Importantly, although K-Ras^{A146T} is a weak mutant form of the oncoprotein, it significantly altered the proteome, even in pancreas, where it did not induce a histologic phenotype.

When bound to GTP, K-Ras can directly bind to and activate downstream kinases. Additionally, multiple kinases regulate GEF or GAP activity and cells may utilize these mechanisms differentially in response to different modes of K-Ras activation (26,27). Nevertheless, it can be difficult to directly determine kinase activation state from MS data, especially when activating phosphorylation sites are not in the phosphopeptide dataset. To indirectly infer kinase activation downstream of mutant K-Ras within our scaled phospho-peptide dataset, we used an approach that combines the GSEA algorithm with custom 'gene sets' constructed of kinases and their confirmed substrates. Similar to our GSEA analysis of KEGG pathways, we found that there was differential activation of kinases induced by the two K-Ras genotypes in the three different tissues (Fig. 5E-G). In colon, there were more kinases predicted to be activated by K-Ras^{A146T} than by K-Ras^{G12D}. The only kinase predicted to be activated by both alleles was Erk2, although the substrates that contributed to enrichment were distinct in the two genotypes (Fig. 5E). There were fewer kinases identified in the pancreas dataset (Fig. 5F), presumably because the number of phosphorylation sites was smaller. As in our analysis of the total proteomes, the ability of the two K-Ras mutants to activate specific kinases was more similar in the spleen and there were no kinases that were predicted to be regulated in opposite directions (Fig. 5G). Of note, there was not a single pathway identified in the total proteome GSEA or the phosphoproteome SKAI that correlated, across all tissues, with the ability of mutant K-Ras to alter homeostasis. These results suggest that the signaling properties of K-Ras are allele- and tissue-specific and that mutant forms of K-Ras may be able to promote tumorigenesis through a variety of different signaling mechanisms.

Allele-specific effects in human cancers

We next sought to determine whether K-Ras mutants were associated with distinct signaling properties in human cancers. We focused on CRC since it is the only cancer type in which the overall frequency of codon 146 mutation is high enough to provide statistical power to approach this question. Although there are published proteomic data from primary human CRCs, there are not enough codon 146 mutants in the dataset to perform a robust statistical analysis (28). Instead, we first performed multiplexed MS on a panel of CRC cell lines expressing different mutant alleles of *KRAS* (Supplementary Tables S14-S16). Here, we directly compared protein abundance or phosphorylation between cell lines comparing K-Ras^{G12D} or K-Ras^{A146T} and then compared the results to a similar analysis of our data from the mouse colon. Of the 6,279 proteins detected in both human and mouse datasets, we found that 156 (2.5%) were concordantly dysregulated between the two genotypes (Supplementary Fig. 15A, Supplementary Table S17). We found a similar pattern for 1,854 phosphopeptides that were detected in both human and mouse datasets, of which 306 (16.5%) were concordantly dysregulated (Supplementary Fig. 15A, Supplementary Table S17). This analysis demonstrates that there is a certain degree of allele-specific signaling that is conserved from mouse to human.

Given that K-Ras signaling was conserved in cancer cell lines, we took this opportunity to ask whether the different mechanisms of activation of the K-Ras mutants translated into different responses to upstream perturbation even in the presence of a heterogeneous genetic background. SHP2 (encoded by the *PTPN11* gene) is a phosphatase that functions upstream of Ras proteins by controlling the activity of GEF (29). Since K-Ras^{A146T} is activated by virtue of increased intrinsic nucleotide exchange and is also sensitive to GEF-induced exchange, we explored whether SHP2 inhibition would be more effective in the context of this mutant than when *KRAS* is wild-type or mutated at codon 12. On the contrary, cell lines expressing K-Ras^{A146T} were not more sensitive to SHP2 inhibition (Supplementary Fig. S15B), suggesting that the increase in intrinsic exchange is sufficient to drive the oncogenic properties of K-Ras^{A146T}.

In order to address the question of allele-specific molecular properties in primary human CRCs, we utilized gene expression data from the TCGA (25). When compared to cancers expressing WT K-Ras, cancers expressing K-Ras^{G12D} had numerous differentially expressed genes, however no obvious pattern was detected in the pathways or cellular processes affected (Supplementary Fig. 16A). For cancers expressing K-Ras^{A146T}, there were no genes differentially expressed relative to WT (Supplementary Fig. 16A). Thus, as in our proteomic analysis of mouse tissues, human cancers expressing K-Ras^{A146T} are more similar to WT than they are to cancers expressing K-Ras^{G12D}.

Human CRCs expressing different mutant forms of K-Ras appeared to be much less similar to one another than were tissues from genetically engineered mice and this is probably due to the fact that human cancers acquire many mutations that normalize the signaling networks between samples with distinct *KRAS* mutations. With this in mind, we hypothesized that if *KRAS* mutations arise early and confer distinct signaling properties on a developing cancer, then each allele might accumulate mutations in different sets of genes that will complement its specific oncogenic properties. Using available genome-wide mutation data from the TCGA CRC dataset, we asked whether cancers with different *KRAS* alleles accumulated mutations in similar or distinct secondary genes. Indeed, we identified 37 genes that were mutated specifically in the context of a specific *KRAS* allele (Supplementary Fig. 16B). For example, *PIK3CA*, which encodes the p110α catalytic subunit of Phosphoinositide 3-kinase (PI3K), is statistically co-mutated with *KRAS* and this is largely due to its specific co-mutation with G12D alleles. There were more genes specifically mutated in the context of K-Ras^{A146T} (28) than K-Ras^{G12D} (10), however the overall number of mutations in cancers with different *KRAS* alleles was not different (Supplementary Fig. 16B). There were no genes that co-mutated with both K-Ras^{A146T} and K-Ras^{G12D}. The observation of distinct mutational profiles associated with different *KRAS* alleles supports a model in which mutant K-Ras proteins exhibit oncogenic properties that are allele-specific.

Translation of signaling properties to tissue phenotypes

Our next goals were to validate the results of our MS analysis and to understand how K-Ras allele signaling regulates cellular behaviors in the context of intact tissues. We focused on proteins that are known to function downstream of oncogenic Ras, as many of these appeared to vary between genotypes. For example, in our total protein analysis from colon, we detected a graded increase in Spred1 and Spred2, negative regulators of MAPK signaling, from WT to K-Ras^{G12D} and we were able to confirm allele-specific Spred2 up-regulation by western blotting (Supplementary Fig. 17A,B). Moreover, GSEA analysis predicted up-regulation of p90Rsk activity in colons expressing K-Ras^{A146T}, but down-regulation in colons expressing K-Ras^{G12D} (Fig. 5E, Supplementary Fig. 17C). Consistent with this prediction, we found that phosphorylation of ribosomal protein S6 on Ser240/Ser244 – a p90Rsk substrate – was down-regulated in colons expressing K-Ras^{G12D} (Supplementary Fig. 17D). Finally, GSEA identified negative enrichment for Akt1 substrates in colons expressing K-Ras^{G12D} and we confirmed by western blotting that phosphorylation of Akt on Ser473 – a readout of Akt activity – was reduced (Supplementary Fig. 17E,F).

Erk2 was the only kinase predicted by GSEA to be activated by both K-Ras^{A146T} and K-Ras^{G12D} in the colon (Fig. 5E). We used quantitative western blotting to assess phosphorylation of Erk1/2 in colonic epithelia and found that colons expressing K-Ras^{A146T} and K-Ras^{G12D} exhibited higher levels of phosphorylated Erk1/2 than did WT colons, although K-Ras^{G12D} had a significantly stronger effect (Fig. 6A,B). Given the relatively low level of MAPK pathway activation in the context of K-Ras^{A146T}, we next evaluated whether, like K-Ras^{G12D}, it promotes hyperproliferation through MAPK signaling. Indeed, we found that inhibition of Mek with trametinib reduced the level of crypt proliferation in colons expressing K-Ras^{G12D} or K-Ras^{A146T} to the level of WT (Fig. 6C), indicating that even low levels of MAPK signaling are sufficient to increase the rate of proliferation in this tissue.

K-Ras^{G12D} also promotes proliferation in the small intestine, but its primary phenotypic manifestation in this tissue is inhibition of Paneth cell differentiation, which is due to MAPK activation (30). To determine the relationship between K-Ras genotype, MAPK signaling, and differentiation in the small intestinal epithelium, we measured Erk1/2 phosphorylation by quantitative western blotting and Paneth cell differentiation by staining for lysozyme. Similar to the colon, we found that K-Ras^{A146T} increased Erk1/2 phosphorylation, although not to the extent of K-Ras^{G12D} (Fig. 6D,E). Interestingly, we found that Paneth cells were intact in colons expressing K-Ras^{A146T}, but absent in colons expressing K-Ras^{G12D} (Fig. 6F). To confirm that the Paneth cell phenotype in animals expressing K-Ras^{G12D} is related to MAPK pathway activation, we treated animals with trametinib for four days and then performed immunofluorescence for lysozyme. We found that Paneth cell differentiation was restored in colons expressing K-Ras^{G12D} following Mek inhibition (Fig. 6F). These collective results suggest that different mutant forms of K-Ras produce different levels of MAPK signaling, which translate into both quantitatively and qualitatively different phenotypes. Low-levels of MAPK activation cross the threshold required to alter proliferative kinetics, but a higher level of MAPK activation is required to cross the threshold required to interfere with differentiation.

Since PanIN is a phenotype driven by defective differentiation induced by high MAPK signaling in the pancreatic duct (31), we explored whether K-Ras^{A146T} was unable to induce PanIN because its MAPK activation level is too low to affect differentiation. We noted from our MS data that the substrates downstream of Erk1/2 were differentially regulated by K-Ras^{G12D} and K-Ras^{A146T} and, in particular, that Erk1/2 substrates were much more highly phosphorylated in colon and spleen than they were in pancreas (Supplementary Fig. S18). At the cellular level, while p-Erk1/2 was essentially undetectable by immunohistochemistry in pancreatic ducts expressing K-Ras^{WT}, it was highly expressed in the ducts and associated PanIN when K-Ras^{G12D} was expressed (Fig. 6G). Ducts expressing K-Ras^{A146T} exhibited detectable p-Erk1/2, but at a level much lower than that seen in ducts expressing K-Ras^{G12D} (Fig. 6G), suggesting that it fails to pass the MAPK threshold required to disrupt pancreatic homeostasis.

DISCUSSION

We have combined structural biology with mouse modeling and multiplexed MS to establish a new paradigm for oncoprotein signaling. We find that different mutant forms of K-Ras can have distinct biochemical mechanisms of activation that translate into distinct global signaling properties. Moreover, we have found that the ability of mutant K-Ras to impinge upon different aspects of tissue homeostasis (*i.e.* proliferation and differentiation) is dependent upon the absolute steady state level of MAPK activation, and thus the allele-specific signaling properties translate into allele-specific phenotypic manifestations.

Initially, our crystal structure of K-Ras4B^{A146T} provided a molecular mechanism for the increased nucleotide exchange that defines A146T as an activating missense mutation (Figs. 1-2). This unique mechanism of activation translated into distinct *in vivo* oncogenic properties relative to K-Ras^{G12D}, which is activated by virtue of a defect in GTP hydrolysis, and these distinct oncogenic properties are the result of distinct allele-specific and tissue-specific signaling. In the intestinal epithelium and in blood, tissues where codon 146 mutations comprise a significant portion of *KRAS* alleles, K-Ras^{A146T} was able to disrupt basal homeostasis, although its effects were consistently weaker than those of K-Ras^{G12D} (Fig. 3). These graded effects are likely due to the different biochemical mechanisms of activation conferred by the G12D and A146T mutations (Fig. 1). K-Ras^{G12D} likely transmits a constant high-level signal to downstream pathways, while K-Ras^{A146T}, because of its reduced affinity to nucleotide, likely transmits an unstable weak downstream signal. This hypothesis is supported by our analysis of MAPK signaling in the intestinal epithelia, which demonstrated that the steady state level of pathway activation is higher in colons expressing K-Ras^{G12D} than in colonic and small intestinal epithelia expressing K-Ras^{A146T} (Fig. 6A-B,D-E). This gradient of MAPK pathway activation (WT < A146T < G12D) produced a proportional proliferative response in the colonic epithelium (Fig. 3D, 6C). The effect of mutant K-Ras on differentiation was not proportional, however, as mice expressing K-Ras^{G12D} lacked Paneth cells in the small intestine because of high MAPK signaling, while K-Ras^{A146T}-expressing mice had intact Paneth cells (Fig. 6F). These observations suggest that the MAPK threshold for inhibition of differentiation is higher than the threshold for promoting proliferation. Consistent with this model, K-Ras^{A146T} exhibited relatively low MAPK activation in the pancreas and did not induce PanIN (Fig. 3H, 6G), a phenotype that requires dysregulation of differentiation.

Our global analysis of allele- and tissue-specific signaling of mutant forms of K-Ras highlights the complex nature of signaling pathways, with MAPK serving as a prime example. Erk2 was predicted to be activated by both mutant forms of K-Ras in both colon and spleen, however the kinase-substrate relationships were not consistent across these contexts. In colon, the majority of Erk2 substrates that were up-regulated in mutant tissue relative to WT were specific to one allele (Fig. 5E). In the spleen, the substrate patterns of the two alleles were similar, and they resembled the pattern seen in colons expressing K-Ras^{A146T} (Fig. 5G). Another example demonstrating that the MAPK output from each K-Ras allele is distinct is exemplified by p90Rsk substrates. We observed that phosphorylation of p90Rsk substrates was up-regulated in colons expressing K-Ras^{A146T}, but down-regulated in colons expressing K-Ras^{G12D} (Fig. 5E). This observation suggests that K-Ras^{G12D} induces a previously unrecognized feedback loop at the level of p90Rsk, downstream of Erk. The differential activity of p90Rsk itself could also exert a negative feedback onto MAPK signaling through phosphorylation of SOS1 (a Ras GEF) on Ser1134 and Ser1161, which allows binding of 14-3-3 and disrupts signaling from receptor tyrosine kinases to Ras (32). Of note, we detected differential phosphorylation of Ser1120 (the mouse Ser1134 equivalent) on Sos1 by K-Ras^{A146T} and K-Ras^{G12D} (Fig. 5E). This result suggests that this negative feedback mechanism might be disrupted in K-Ras^{G12D}-mutant tissue, which could contribute to sustained MAPK signaling, or could be up-regulated by K-Ras^{A146T} to confer some advantage in cellular fitness. These observations suggest that pathway activation can assume many forms and that the ultimate functional output of a pathway is not dependent upon activation of a single component, but by the cumulative effects of pathway flux and feedback mechanisms.

While our studies focus specifically on K-Ras, the results illustrate a number of general principles that will relate to other oncogenes. First, variations in the biological properties of distinct alleles of the same oncogene can explain, at least in part, the tissue-specific mutational profiles seen in human cancers. Next, even in tissues where distinct alleles have oncogenic phenotypes, different biochemical mechanisms of oncoprotein activation can lead to profound differences in signaling. Since the phenotypic output of an activated oncoprotein depends on the state of the cellular signaling network as a whole, distinct alleles can produce different phenotypes. The implication of this paradigm is that different mutant forms oncoproteins can make equivalent functional contributions to cancer initiation and progression, yet have allele-specific therapeutic vulnerabilities in downstream effector pathways. The combination of structural biology, mouse genetics, and global signaling analysis deepens our understanding of how the oncogenic output of K-Ras relates to its biochemical mechanism of activation and provides the first demonstration that tissue-specific allelic selection is driven by the distinct signaling properties of mutant K-Ras alleles.

METHODS

Hydrogen Deuterium Exchange Mass Spectrometry (HDX MS)

HDX MS experiments were performed similarly to previous reports (16,33). For comparison between WT and A146T, proteins were diluted into sample buffer (10 mM HEPES pH 8.0, 5 mM $MgCl_2$, 5 μM GDP, 150 mM NaCl) to a concentration of 10 μM prior to labeling. Samples were independently labeled with deuterium in duplicate, using identical experimental conditions to allow direct comparisons between K-Ras WT and A146T. HDX was initiated by diluting 2.0 μl of protein sample (10 mM HEPES, pH 8.0, 5 mM $MgCl_2$, 5 μM GDP, 150 mM NaCl) 18-fold in labeling buffer (10 mM HEPES, pH 8.0, 5 mM $MgCl_2$, 5 μM GDP, 150 mM NaCl) containing 99% deuterium oxide at room temperature. The labeling reaction was quenched at six time points

(5 s, 10 s, 1 m, 10 m, 1 h, 4 h) through the addition of 38.0 µl quench buffer (300 mM sodium phosphate, pH 2.47) at 0°C. For experiments comparing WT and G12D, protein sample buffer (20 mM HEPES pH 7.5, 150 mM NaCl), labeling buffer (20 mM HEPES pH 7.5, 150 mM NaCl), and quench buffer (0.8%FA; 0.8M GdnHCl; pH 2.4) were slightly different than in the A146T experiment. Quenched samples were immediately analyzed. Deuterium measurement with MS was performed as previously described (16,33) using a Synapt G2si HDMSE, online digestion Enzymate column, 1x50 1.8 µm HSS T3 separation, 1.7 µm BEH trap. Deuterium incorporation graphs (Supplementary Fig. S4,S5) were generated using DynamX 3.0 software (Waters) by subtracting the centroid of the isotopic distribution at each labeling time point from the centroid of the isotopic distribution of the undeuterated reference species. Since the data were not corrected for back-exchange, each data point represents the relative deuterium level at each time point for each peptide (34).

GTPase Activity Assay

Intrinsic and GAP-stimulated GTPase activity for WT and mutant K-Ras4B proteins were measured using EnzCheck phosphate assay system (Life Technologies) as previously described (17). In brief, K-Ras4B proteins (2.5 mg/ml) in buffer (20 mmol/L Tris pH 8.0, 50 mM NaCl) were loaded with GTP at RT for 2 hours by exposing to exchange buffer containing EDTA. Proteins were buffer exchanged to assay buffer (30 mmol/L Tris pH 7.5, 1 mmol/L DTT) and the concentration adjusted to 2 mg/ml. GTP loading was verified by back extraction of nucleotide using 6M urea and evaluation of nucleotide peaks by HPLC using an ion-exchange column as described previously (33). The assay was performed in a clear 384-well plate (Costar) by combining GTP-loaded K-Ras4B proteins (50 µM final) with MESG (200 µM final), and purine nucleotide phosphorylase (5 U/ml final). GTP hydrolysis was initiated by the addition of $MgCl_2$ at a working concentration of 40 μM. For GAP stimulation P120GAP was included at 50 μM. Absorbance at 360 nm was read every 8 to 15 seconds for 3600 seconds at 20^{o}C. All measurements were performed in triplicate.

Nucleotide Dissociation Assay

Nucleotide dissociation reactions were performed using a fluorescent guanine nucleotide analogue (Mant-GTP) that shifts absorbance when bound to Ras (15). K-Ras4B proteins (100 μM) were loaded with Mant-GDP by incubating with 300 μM Mant-GDP and 5 mM EDTA in PBS at 25°C for 2 hours. Reactions were terminated with 10 mM $MgCl_2$, and then buffer was exchanged to remove EDTA and unbound nucleotides by using 2 mL Zeba™ (ThermoFisher Scientific) desalting cartridges (7,000 Da MWCO) into buffer PBS supplemented with 2 mM $MgCl_2$ and 1 mM EDTA. Reactions were started by mixing RAS and other reagents using a two channel injection system mounted on a Synergy Neo reader (BioTek). Fluorescence was measured every 0.5 seconds for 5 minutes at 360nm/440nm (excitation/emission) in a 384-well plate at ambient temperature. Final reaction conditions for intrinsic dissociation (K-Ras4B alone) were 2 μM K-Ras4B, 4 μM GDP in PBS with 2 mM $MgCl_2$, 1 mM EDTA. Final reaction conditions for SOS-mediated dissociation were 2 μM K-Ras4B, 4 μM GDP, 2 μM SOS1 in the same buffer. Data were exported and analyzed using Prism 7 (GraphPad Software, Inc.). All measurements were performed in triplicate.

Microscale Thermophoresis (MST)

K-Ras4B proteins (130 μM) in 1X PBS and 200 μM GDP were incubated at room temperature for 1hr. The solution was then desalted into 1X PBS with 2 mM $MgCl_2$ and 1% tween 20 via a Zeba™ (Thermo Fisher) spin desalting column. 16-point serial dilutions of K-Ras4B were prepared and mixed 1:1 with GFP-^{556}SOS^{1049} in final volume of 20 μL. The reaction mixtures were loaded into premium treated capillaries and analyzed by a Monolith NT.115 (Nanotemper Technologies) at 40% MST power and 40% LED power with a laser-on time of 30 s. The K_D was calculated by taking the average of duplicate F_{norm} measurements at each concentration and fitting the data by using Palmist (35) and plotting in Prism 7.

RAF Interaction Assay

K-Ras4B:RAF kinase interaction assays were performed as previously described (16). Purified RAF-1 RBD was labeled with maleimide PEG biotin (Pierce) following the manufacturer's recommended protocol. Purified flag-tagged K-Ras4B (1 mg/mL) were loaded with GMPPNP (Sigma-Aldrich) by incubating for 2 hours at 25°C with a 50-fold excess of nucleotide in the presence of alkaline phosphatase (Thermo-Fisher). RAF–RBD–biotin was diluted to a final concentration of 40 nM and Flag-K-Ras4B to 10 nM in assay buffer (20 mM Tris pH 7.5, 100 mM NaCl, 1 mM $MgCl_2$, 5% glycerol, 0.5% BSA) and added to individual wells of a low-volume white 384-well plate (PerkinElmer). Complexes were disrupted by addition of a dilution series (2,000 nM to 0.5 nM) of each mutant K-Ras4B protein. The assay was developed by addition of streptavidin donor and anti-flag acceptor AlphaScreen beads (10 μg/mL). Alpha signal was measured after overnight incubation at 4°C. All readings were performed in triplicate.

Structure Determination and Refinement

K-Ras4B^{A146T} protein was concentrated to 13 mg/ml prior to crystallization trials using an Amicon® Ultra Centrifugal Filter with a molecular weight cutoff of 10,000 Daltons. Single protein crystals were grown using hanging drop vapor diffusion at 20°C from a reservoir solution containing 1.4 M sodium malonate pH 7.0, 0.1 M BIS-TRIS propane pH 7.0. Prior to diffraction crystals were soaked briefly in cryo-protectant containing reservoir solution supplemented with 25% (v/v) glycerol. X-ray diffraction data was collected at the Advanced Photon Source beamline 19-ID at 100K. The initial diffraction data were indexed, integrated, and scaled using HKL2000. The K-Ras4B^{A146T}:GDP crystal structure was determined by molecular replacement using wild-type (WT) K-Ras4B (PDB entry 4OBE) as an initial search model with Phaser in the CCP4 Suite (36). The model was further built and refined using Phenix (37) and COOT (38) by manual model correction. Waters were added and the final refinement performed by Phenix Phenixrefine (37). Stereochemical parameters of the final model were analyzed with the assistance of Molprobity (39). Data collection and refinement statistics are listed in Supplementary Table S1.

Animal Studies

Animal studies were approved by the Institutional Care and Use Committee at Massachusetts General Hospital, Beth Israel Deaconess Medical Center, and the University of California, San Francisco. Mice were fed *ad libitum* and housed in a barrier facility with a temperature-controlled environment and twelve-hour light/dark cycle. *Pdx1-Cre* and *Mx1-Cre* mice were obtained from The Jackson Laboratory (Strain No. 014647). *Fabp1-Cre* (Strain 01XD8), *Apc*^{2lox14} (Strain 01XP3), *K-Ras*$^{LSL\text{-}G12D}$ (Strain 01XJ6), and *Tp53*$^{LSL\text{-}R270H}$ (Strain 01XM1) mice were obtained from the NCI Mouse Repository. Experimental animals were maintained on genetic background that was 80-95% C57BL/6. Survival curves for tumor-bearing animals were analyzed using log-rank test (Prism 7).

The *Kras*$^{\text{LSL-A146T-Neo}}$ allele was generated in TL1 embryonic stem (ES) cells (129/Sv genetic background) (40). Using embryonic stem (ES) cells with a transcriptional stop element (lox-stop-lox, LSL) inserted into the endogenous *Kras* locus *(Kras*$^{\text{LSL-WT}}$), we engineered an alanine to threonine mutation at codon 146 in exon 4 to generate *Kras*$^{\text{LSL-A146T}}$ mice. ES cells were injected into C57BL/6J embryos in the Harvard Medical School Transgenic Mouse Core and chimeric males were backcrossed to C57BL/6 females. The PGK-Neo selection cassette was deleted from the *Kras* locus by crossing *Kras*$^{\text{LSL-A146T-Neo}}$ mice to a Flp deleter strain (Jackson Laboratory, stock number 007844). The *Kras*$^{\text{LSL-A146T}}$ allele was subsequently maintained in a primarily C57BL/6 background by continuous backcrossing. Age-matched *Fabp1-Cre or Pdx1-Cre* littermates were used as K-Ras WT controls in all experiments. For experiments analyzing coon and pancreas, all animals were 8-12 weeks-of-age. For analysis of hematopoietic phenotypes, mice were maintained in a C57BL/6 × 129SvJae F1 background. Mx1-Cre was induced by i.p. injection with 250 μg polyinosinic-polycytidylic acid (pIpC) at four weeks. Peripheral blood counts were measured using a Hemavet 950FS instrument (Drew Scientific). Serial white blood cell counts were analyzed as described in (41).

Tissue Staining, IHC, and IF

Tissue was harvested from mice, fixed in 10% formalin overnight at room temperature, and processed into paraffin blocks. Tissue sections (5 μm) were deparaffinized in a standard xylene and ethanol series. Standard H&E staining protocols were used. Histology of colon and pancreas was review by a board-certified gastrointestinal pathologist (Dr. Nowak). For crypt height analyses, crypts were measured using Olympus VS-ASW version 2.7 software as described in (42).

For IHC, deparaffinized sections were subjected to antigen retrieval in a pressure cooker in Target Retrieval Solution, pH 6 (DAKO®). IHC was performed using the EnVision+ HRP Kit (DAKO®). Images were acquired using an Olympus BX-UCB slide scanner. Primary antibodies were diluted in Antibody Diluent (DAKO®) and incubated overnight at 4°C. For IF, following antigen retrieval, sections were incubated in Protein Block, Serum-Free (DAKO®) for twenty minutes at room temperature. Primary antibodies were diluted in Antibody Diluent (DAKO®) and incubated overnight at 4°C. Slides were then washed in PBS, and secondary antibodies diluted in Antibody Diluent were filtered through 0.45 μM filters (Millipore), and incubated for one hour at room temperature in the dark. Coverslips were mounted using ProLong® Diamond with DAPI (Thermo Fisher) and images were acquired using a Zeiss Axio Imager Z2. Primary antibodies included: anti-phospho-histone-H3 (Ser10; CST), anti-E-cadherin (BD), anti-β-catenin (CST), and anti-lysozyme (Thermo). Fluorescent secondary antibodies included: anti-mouse IgG2a Alexa Fluor 488 (ThermoFisher) and anti-rabbit Alexa Fluor 594 (ThermoFisher). Proliferation results were analyzed using Mann-Whitney U tests (Prism 7).

Western Blotting

Tissue was harvested from mice and flash frozen in liquid nitrogen immediately. Tissue was lysed in Bio-Plex® Lysis Buffer (Bio-Rad) supplemented with Factor I, Factor II, PSMF (2 nM), and cOmplete™ protease inhibitor cocktail (Sigma-Aldrich), or Mg2+ Lysis/Wash Buffer (Millipore), supplemented with cOmplete™ protease inhibitor cocktail (Sigma-Aldrich) and phosphatase inhibitor cocktails 2 and 3 (Sigma). Following clearing by centrifugation, protein lysates were quantified using a bicinchoninic acid assay (BCA, Pierce) and samples were equally loaded for SDS-PAGE. Western blotting was performed according to standard protocols and analysis was performed on an Odyssey® CLx Infrared Imaging System (LI-COR®). Western blot images were quantified using Image Studio Software (LI-COR®). Primary antibodies included: anti-Ras (Millipore, 05-516), anti-α-tubulin (Sigma, T6074), anti-p-Erk1/2 (Thr202/Tyr204; Cell Signaling Technologies [CST], 4377), anti-Erk1/2 (CST, 4696), anti-GAPDH (CST, 5174), anti-p-Akt (Ser473; CST, 4060), anti-Akt (CST, 9272), and anti-pS6 (Ser240/244; CST, 5364). Secondary antibodies included: anti-mouse IgG Alexa Fluor 680 (ThermoFisher, A21058) and anti-rabbit IgG Alexa Fluor 800 (ThermoFisher, A32735). Western blot results were analyzed using Mann-Whitney U tests (Prism 7).

Inhibitor and Cerulein Treatments

For analysis of proliferation, 8–12-week-old mice were treated once every 12 h for 24 h with 0.25 mg/kg trametinib or 10% DMSO vehicle. For analysis of Paneth cell differentiation, 8-12 week old mice were administered 1 mg/kg trametinib twice per day for four days or 4% DMSO in corn oil. For cerulein treatment, 8–12-week-old mice were treated with 250 µg/kg cerulein (in normal saline, 0.9% NaCl) or with normal saline vehicle once daily for five days. Mice were sacrificed and tissues were harvested 30–31 days after the first treatment.

Mass Spectrometry

Mouse tissues or human cell line pellets were homogenized using a polytron at medium speed followed by 10 passes through a 21-gauge needle in 100 mM EPPS, 8 M Urea, pH 8.5, supplemented with 1X cOmplete™ protease inhibitor cocktail (Sigma-Aldrich) and 1X PhosSTOP phosphatase inhibitors (Sigma-Aldrich). The homogenate was sedimented at 21,000 x g for 5 min and the supernatant was transferred to a new tube. Protein concentrations were determined using a BCA assay (ThermoFisher Scientific). Proteins were subjected to disulfide bond reduction with 5 mM Tris (2-carboxyethyl)phosphine (room temperature, 30 min) and alkylation with 10 mM iodoacetamide (room temperature, 30 min in the dark). Excess iodoacetamide was quenched with 10 mM dithiotreitol (room temperature, 15 min in the dark). Methanol-chloroform precipitation was performed prior to protease digestion. Four parts of neat methanol were added to each sample and vortexed, one part chloroform was added to the sample and vortexed, and three parts water was added to the sample and vortexed. The sample was centrifuged at 14,000 RPM for 2 min at room temperature and washed twice with 100% methanol. Samples were resuspended in 200 mM EPPS, pH 8.5, and digested at room temperature for 13 h with Lys-C protease at a 100:1 protein-to-protease ratio. Trypsin was then added at a 100:1 protein-to-protease ratio and the reaction was incubated for 6 h at 37°C. Phosphopeptides were enriched following digestion and prior to TMT labeling using a method based on that of Kettenbach and Gerber and described elsewhere (43,44). TMT and Off-line Basic pH Reversed-phase Fractionation were performed as previously described in (44).

All samples were analyzed on an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific, San Jose, CA) coupled to a Proxeon EASY-nLC 1200 liquid chromatography pump (Thermo Fisher Scientific). Peptides were separated on a 100 μM inner diameter microcapillary column packed with 35 cm of Accucore C18 resin (2.6 μM, 150 Å, ThermoFisher). For each analysis, we loaded approximately 2 μg onto the column. For whole proteome analysis, peptides were separated using a 180 min gradient of 3-25% acetonitrile in 0.125% formic acid with a flow rate of 450 nL/min. Each analysis used an MS3-based TMT method (45). The scan sequence began with an MS1 spectrum (Orbitrap analysis, resolution 120,000, 400−1,400 Th, automatic gain control [AGC] target 5E5, maximum injection time 100 ms). The top ten precursors were then selected for MS2/MS3 analysis. MS2 analysis consisted of: collision-induced dissociation, quadrupole ion trap analysis, AGC 2E4, normalized collision energy (NCE) 35, q-value 0.25, maximum injection time 120 ms), and isolation window at 0.7. Following acquisition of each MS2 spectrum, we collected an MS3 spectrum in which multiple MS2 fragment ions are captured in the MS3 precursor population using isolation waveforms with multiple frequency notches (45). MS3 precursors were fragmented by HCD and analyzed using the Orbitrap (NCE 55, AGC 2.5E5, maximum injection time 120 ms, resolution was 60,000 at 400 Th). For MS3 analysis, we used charge state-dependent isolation windows: for charge state z=2, the isolation window was set at 1.3 Th, for z=3 at 1 Th, for z=4 at 0.8 Th, and for z=5-6 at 0.7 Th.

Likewise, for phospho-proteome analysis, peptides were separated using a 120 min gradient. The scan sequence began with an MS1 spectrum (Orbitrap analysis, resolution 120,000, 400−1400 Th, AGC target 5E5, maximum injection time 100 ms). The top ten precursors were then selected for MS2/MS3 analysis. MS2 analysis consisted of the following: collision-induced dissociation, quadrupole ion trap analysis, AGC 2E4, NCE 35, q-value 0.25, maximum injection time 120 ms), and isolation window at 0.7. Following acquisition of each MS2 spectrum, we collected an MS3 spectrum in which multiple MS2 fragment ions are captured in the MS3 precursor population using isolation waveforms with multiple frequency notches (45). MS3 precursors were fragmented by HCD and analyzed using the Orbitrap (NCE 55, AGC 3E5, maximum injection time 300 ms, resolution was 60,000 at 400 Th). For MS3 analysis, we used charge state-dependent isolation windows: for all charge states, the isolation window was set at 1.2 Th.

LC-MS^3 Analysis

Mass spectra were processed using a Sequest-based in-house software pipeline (46). Data analysis was performed as previously described in (44,47). Proteins and phosphopeptides were quantified by summing reporter ion counts across all matching PSMs using in-house software, as described previously (46). PSMs with poor quality, MS3 spectra with at least one TMT reporter ion channels missing, MS3 spectra with TMT reporter summed signal-to-noise ratio that were less than 100, or had no MS3 spectra were excluded from quantification (48). Each reporter ion channel was summed across all quantified proteins and normalized assuming equal protein loading of all 10 samples across the TMT10-plex. Protein quantification values were exported for further analysis as described below.

Scaled phosphopeptides were calculated as a composite ratio of a phosphopeptide and the corresponding total protein value where possible. UniProt IDs were used as the identifier to match phosphopeptides with their protein value.

Analysis of MS Data

GSEA was conducted on the protein dataset. KEGG pathway gene sets were downloaded from MSigDB [GSEA, KEGG1, KEGG2, KEGG3, MSIGDB]. To uniquely identify proteins within our dataset, gene symbols were converted to murine UniProt IDs by programmatically accessing the ID mapping tool (http://www.uniprot.org/mapping/), retaining both reviewed and unreviewed matches. GSEA was run using a pre-ranked list of genes. The ranking metric used is listed below, a non-parametric signal-to-noise value,

$$\frac{m_A - m_B}{MAD_A + MAD_B}$$

where m is the median and MAD is the median absolute deviation when comparing groups A versus B (either G12D versus WT or A146T versus WT). MAD has a minimum value of $0.2 * |m|$. Additional parameters used in the analysis can be found in Data S10-S12. Pathways presented were considered enriched if the enrichment of their substrate set surpassed a threshold of either a nominal p-value < 0.1 or an FDR-q value < 0.25. To visualize results from each comparison, we performed hierarchical clustering as above.

For inference of kinase activity, we performed GSEA on phospho-proteomic data with custom gene sets constructed of kinases and their respective species-specific substrate sites, termed substrate site sets, assembled from datasets available from PhosphoSite (49). A pre-ranked enrichment was conducted as described above, using scaled phospho-peptide data. Kinases were considered significantly active if the enrichment of their substrate set surpassed a threshold of either a nominal p-value < 0.1 or an FDR-q value < 0.25. Results from each comparison were clustered via hierarchical clustering as above.

Droplet Digital PCR (ddPCR)

Genomic DNA was extracted from snap frozen or formalin-fixed paraffin-embedded cell pellets or mouse tissues using QIAGEN DNeasy Blood and Tissue Kit or QIAamp DNA FFPE Tissue Kit. ddPCR was carried out according to Bio-Rad's protocol. In brief, 100 ng of sample DNA was digested with HaeIII at 37°C for 1 hour and added to a PCR reaction master mix containing 2x ddPCR Master Mix (Bio-Rad), 20x target assay mix, and 20x reference assay mix, both made with forward and reverse primers, hybridization probe, and water. Droplets were generated on Biorad's QX200 with droplet generation oil, subject to amplification (95°C 10 min, 94°C 30 sec, 59°C 1 min, repeated 40x, 98°C 10 min, 8°C hold), and read on Bio-Rad's QX200 Droplet Reader running

QuantaSoft software. Copy number (CN) was calculated using the following equations, concentration = -ln(# negative droplet / # total droplets), CN = (target concentration / reference concentration) * # copies of reference in genome. Mouse *Kras* was amplified with primers 5'-AGCAAGGAGTTACGGGATTC-3' and 5'- TGCCATCAGGAAATGAAGAT-3' with hybridization probe 5'- FAM-TGCTCCAGATGGTGTTCGTCA-Black Hole Quencher-3', while the reference gene *(Xist)* was amplified with 5'- TGTTTGAACTCCCAGACCTCTT -3' and 5'-AGGACATTCAGGCAAAAGAAAA-3' with hybridization probe 5'-HEX-CAACCTGGCTCCATCTTCTCTGTTCC-Black Hole Quencher-3'. Mouse oligonucleotides were ordered from IDT. Human *KRAS* copy number was detected with Bio-Rad's validated copy number assay mix (dHsaCP2507077, HEX) with reference RPP30 (dHsaCP1000485, FAM).

SHP2 inhibition in cell lines

DiFi cells were a gift from Robert Coffey (Vanderbilt University), CCCL18 and HCC2998 were provided by David Solit (Memorial Sloan Kettering Cancer Center), and the remaining cell lines were purchased from ATCC. Cell lines were validated by STR profiling and screened for mycoplasma by Genetica .

All cells were cultured in recommended media at 37°C, 5% CO_2. SHP099 (ChemieTek) was dissolved in DMSO to 10 mM for storage. The day before drug treatment, 4000-6000 cells per well were plated in 96-well plates in 100 μL of complete media, including a plate with all cell lines for day 0 readout for GR value calculations (day 0 plate). The next day (day 0), SHP099 was serially diluted in culture media in the range of 2 nM - 100 μM and 100 μL of drug dilution added to respective wells, resulting in a final drug concentration range of 1 nM – 50 μM. DMSO-only wells received 2% DMSO in culture media. 100 μL of regular culture media was added to the day

0 plate, which was returned to 37°C to equilibrate before adding MTT (Sigma) to determine baseline viability. Drug-treated plates were cultured at 37°C, 5% CO_2 for 72 hours and cell viability determined with the MTT assay. GR values were calculated from absorbance values with the online tool GR Calculator (http://www.grcalculator.org/grcalculator/) and graphs generated by GraphPad Prism software.

Analysis of TCGA Gene Expression

RNA-sequencing data of human CRC tumors was obtained from TCGA Colon Adenocarcinoma and Rectum Adenocarcinoma projects (25). WT cancers were defined as those with no activating mutations in *KRAS*, *NRAS*, or *BRAF*. Differential gene expression analysis was conducted by fitting a linear model to each gene with empirical Bayes smoothing of the standard error using the *limma* package in R (50,51). *P* values were controlled for multiple hypothesis testing using the Benjamini-Hochberg procedure.

Analysis of TCGA Allele Frequencies

Whole-exome sequencing data of the *KRAS* gene from the TCGA-COAD and TCGA-READ projects were downloaded from Genomic Data Commons and indexed using *samtools* (version 1.9) (52). Read counts data were extracted using *bcftools* (version 1.9) *mpileup* and mutations were assigned using *bcftools call* and annotated with *ANNOVAR* (June 1, 2017) (52,53). Tumor sample purity data and genomic copy number data were downloaded from the PanCanAtlas Publications website (https://gdc.cancer.gov/about-data/publications/pancanatlas). Only samples with purity of at least 0.25 were used for further analysis.

The copy number of the *KRAS* locus was adjusted for the purity of the sample (the fraction of the sample that are cancer cells). The measured copy number of the sample n_s is a linear combination of the copy number of the tumor cells n_t and normal cells coincidentally extracted n_n, and can expressed as

$$n_s = \rho n_t + (1 - \rho) n_n \tag{1}$$

where ρ is the purity of the tumor sample. Assuming the normal cells have a copy number of two, the copy number of the cancer cells can be solved for using Equation (2).

$$n_t = \frac{1}{\rho}(n_s - 2(1 - \rho)) \tag{2}$$

The copy number of the mutant *KRAS* allele was adjusted using the same formulation as described by (54). Briefly, the relationship between the number of reads of the mutant allele r_m, the total number of sequencing reads of the locus r_l, the copy number of the mutant allele n_m, and the copy number of the locus n_l can be expressed as

$$\frac{r_m}{r_l} = \frac{n_m}{n_l} \tag{3}$$

To account for the purity of the sample, n_l is replaced with the copy number of the sample n_s using equation (1), and the total number of sequencing reads is weighted by the purity

$$\frac{r_m}{\rho r_l} = \frac{n_m}{\rho n_t + (1 - \rho) n_n} \tag{4}$$

which, assuming the normal cells have two copies of the wild type allele, solves to

$$n_m = \frac{r_m}{\rho r_l}(\rho n_t - 2(1 - p)) \tag{5}$$

Negative final values were rounded to zero. Data were plotted using the *ggplot2* package in R.

Co-mutation Analysis

DNA sequence data for 224 colorectal tumors were obtained through cBioPortal for Cancer Genomics for TCGA colorectal carcinoma cohort. A one-sided Fisher's Exact Test with Benjamini-Hochberg FDR correction was applied to assess the significance of co-occurring mutations with *KRAS* and specific *KRAS* mutant alleles. Candidate *KRAS* allele co-mutation events were selected at a threshold of FDR $q < 0.10$. Co-mutation results were visualized using Cytoscape 3.4.0 to create a statistical network with edges between genes and the *KRAS* allele node with which they significantly co-mutate.

Authors' contributions

Conception and design: E. Poulin, J. Engen, B. Braun, S. Gygi, D. Lauffenburger, K. Westover, and K. Haigis

Acquisition of data: E. Poulin, A. Bera, J. Lu, Y. Lin, J. Paulo, T. Huang, C. Morales, W. Yan, J. Nowak, R. DeStefanis, P. Ghazi, S. Gondi, T. Wales, R. Iacob, L. Bogdanova, J. Gierut, Y. Li

Computational analysis: S. Strasser, B. Joughin, C. Johnson, J. Lu, W. Yan, A. Bera, J. Cook, and D. Brubaker

Contribution of reagents: P. Perez-Mancera

Writing, review, and/or revision of manuscript: E. Poulin, K. Westover, and K. Haigis

Study supervision: K. Haigis and K. Westover

Acknowledgments: We thank the staff at the structural biology laboratory at UT Southwestern Medical Center and at beamline 19ID of Advanced Photon Source for facilitating X-ray data collection and processing. Results are derived from work performed at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source, operated by the University of Chicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. This work was supported by grants from the National Institutes of Health: R01CA178017 and R01CA195744 to K. Haigis; U01CA215798 to K. Haigis and D. Lauffenburger; R01CA173085 and P30CA082103 to B. Braun; and K01DK098285 to J. Paulo. This work was also supported by grants from the Department of Defense (W81XWH-16-1-0106 to K. Westover) and the Cancer Prevention and Research Institute of Texas (RP170373 to K. Westover). E. Poulin and J. Gierut were supported by postdoctoral fellowships from the American Cancer Society. Y. Lin was supported by a fellowship from the Landry Cancer Biology Consortium. S. Strasser was supported by a National Science Foundation Graduate Research Fellowship (Grant No. 1122374). D. Brubaker was funded by a grant from Boehringer-Ingelheim as part of the SHINE program.

Data and materials availability: The accession numbers for the atomic coordinates and structure factors for the crystal structure of K-Ras^{A146T}:GDP have been deposited in Protein Data Bank under accession code 6BOF.

References

1. Haigis KM. KRAS Alleles: The Devil Is in the Detail. Trends Cancer **2017**;3:686-97.
2. Gibbs JB, Sigal IS, Poe M, Scolnick EM. Intrinsic GTPase activity distinguishes normal and oncogenic ras p21 molecules. Proc Natl Acad Sci USA **1984**;81:5704-8.
3. Johnson L, Mercer K, Greenbaum D, Bronson RT, Crowley D, Tuveson DA*, et al.* Somatic activation of the K-ras oncogene causes early onset lung cancer in mice. Nature **2001**;410:1111-6.
4. Jackson EL, Willis N, Mercer K, Bronson RT, Crowley D, Montoya R*, et al.* Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. Genes Dev **2001**;15:3243-8.
5. Braun BS, Tuveson DA, Kong N, Le DT, Kogan SC, Rozmus J*, et al.* Somatic activation of oncogenic Kras in hematopoietic cells initiates a rapidly fatal myeloproliferative disorder. Proc Natl Acad Sci USA **2004**;101:597-602.
6. Hingorani SR, Petricoin EF, Maitra A, Rajapakse V, King C, Jacobetz MA*, et al.* Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. Cancer Cell **2003**;4:437-50.
7. Haigis KM, Kendall KR, Wang Y, Cheung A, Haigis MC, Glickman JN*, et al.* Differential effects of oncogenic K-Ras and N-Ras on proliferation, differentiation and tumor progression in the colon. Nat Genet **2008**;40:600-8.
8. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA*, et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov **2012**;2:401-4.
9. Feig LA, Cooper GM. Relationship among guanine nucleotide exchange, GTP hydrolysis, and transforming potential of mutated ras proteins. Mol Cell Biol **1988**;8:2472-8.
10. Edkins S, O'Meara S, Parker A, Stevens C, Reis M, Jones S*, et al.* Recurrent KRAS codon 146 mutations in human colorectal cancer. Cancer Biol Ther **2006**;5:928-32.
11. Janakiraman M, Vakiani E, Zeng Z, Pratilas CA, Taylor BS, Chitale D*, et al.* Genomic and biological characterization of exon 4 KRAS mutations in human cancer. Cancer Res **2010**;70:5901-11.
12. Montalvo SK, Li L, Westover KD. Rationale for RAS mutation-tailored therapies. Future Oncol **2017**;13:263-71.
13. Imamura Y, Lochhead P, Yamauchi M, Kuchiba A, Qian ZR, Liao X*, et al.* Analyses of clinicopathological, molecular, and prognostic associations of KRAS codon 61 and codon 146 mutations in colorectal cancer: cohort study and literature review. Mol Cancer **2014**;13:135.
14. Misale S, Yaeger R, Hobor S, Scala E, Janakiraman M, Liska D*, et al.* Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer. Nature **2012**;486:532-6.
15. Eccleston JF, Moore KJ, Brownbridge GG, Webb MR, Lowe PN. Fluorescence approaches to the study of the p21ras GTPase mechanism. Biochem Soc Trans **1991**;19:432-7.
16. Lim SM, Westover KD, Ficarro SB, Harrison RA, Choi HG, Pacold ME*, et al.* Therapeutic targeting of oncogenic K-Ras by a covalent catalytic site inhibitor. Angew Chem **2014**;53(1):199-204.

17. Hunter JC, Manandhar A, Carrasco MA, Gurbani D, Gondi S, Westover KD. Biochemical and Structural Analysis of Common Cancer-Associated KRAS Mutations. Mol Cancer Res **2015**;13:1325-35.
18. Dharmaiah S, Bindu L, Tran TH, Gillette WK, Frank PH, Ghirlando R*, et al.* Structural basis of recognition of farnesylated and methylated KRAS4b by PDEdelta. Proc Natl Acad Sci USA **2016**;113:E6766-E75.
19. Maurer T, Garrenton LS, Oh A, Pitts K, Anderson DJ, Skelton NJ*, et al.* Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated nucleotide exchange activity. Proc Natl Acad Sci USA **2012**;109:5299-304.
20. Boriack-Sjodin PA, Margarit SM, Bar-Sagi D, Kuriyan J. The structural basis of the activation of Ras by Sos. Nature **1998**;394:337-43.
21. Tuveson DA, Shaw AT, Willis NA, Silver DP, Jackson EL, Chang S*, et al.* Endogenous oncogenic K-ras(G12D) stimulates proliferation and widespread neoplastic and developmental defects. Cancer Cell **2004**;5:375-87.
22. Wong MH, Saam JR, Stappenbeck TS, Rexer CH, Gordon JI. Genetic mosaic analysis based on Cre recombinase and navigated laser capture microdissection. Proc Natl Acad Sci USA **2000**;97:12601-6.
23. Kuhn R, Schwenk F, Aguet M, Rajewsky K. Inducible gene targeting in mice. Science **1995**;269:1427-9.
24. Burgess MR, Hwang E, Mroue R, Bielski CM, Wandler AM, Huang BJ*, et al.* KRAS Allelic Imbalance Enhances Fitness and Modulates MAP Kinase Dependence in Cancer. Cell **2017**;168:817-29 e15.
25. Cancer Genome Atlas N. Comprehensive molecular characterization of human colon and rectal cancer. Nature **2012**;487:330-7.
26. Pamonsinlapatham P, Hadj-Slimane R, Lepelletier Y, Allain B, Toccafondi M, Garbay C*, et al.* p120-Ras GTPase activating protein (RasGAP): a multi-interacting protein in downstream signaling. Biochimie **2009**;91:320-8.
27. Yin G, Kistler S, George SD, Kuhlmann N, Garvey L, Huynh M*, et al.* A KRAS GTPase K104Q Mutant Retains Downstream Signaling by Offsetting Defects in Regulation. J Biol Chem **2017**;292:4446-56.
28. Zhang B, Wang J, Wang X, Zhu J, Liu Q, Shi Z*, et al.* Proteogenomic characterization of human colon and rectal cancer. Nature **2014**;513:382-7.
29. Grossmann KS, Rosario M, Birchmeier C, Birchmeier W. The tyrosine phosphatase Shp2 in development and cancer. Adv Cancer Res **2010**;106:53-89.
30. Feng Y, Bommer GT, Zhao J, Green M, Sands E, Zhai Y*, et al.* Mutant KRAS promotes hyperplasia and alters differentiation in the colon epithelium but does not expand the presumptive stem cell pool. Gastroenterology **2011**;141:1003-13.
31. Collins MA, Yan W, Sebolt-Leopold JS, Pasca di Magliano M. MAPK signaling is required for dedifferentiation of acinar cells and development of pancreatic intraepithelial neoplasia in mice. Gastroenterology **2014**;146:822-34.
32. Saha M, Carriere A, Cheerathodi M, Zhang X, Lavoie G, Rush J*, et al.* RSK phosphorylates SOS1 creating 14-3-3-docking sites and negatively regulating MAPK activation. Biochem J **2012**;447:159-66.
33. Harrison RA, Lu J, Carrasco M, Hunter J, Manandhar A, Gondi S*, et al.* Structural Dynamics in Ras and Related Proteins upon Nucleotide Switching. J Mol Biol **2016**;428:4723-35.

34. Wales TE, Engen JR. Hydrogen exchange mass spectrometry for the analysis of protein dynamics. Mass Spectrom Rev **2006**;25:158-70.
35. Scheuermann TH, Padrick SB, Gardner KH, Brautigam CA. On the acquisition and analysis of microscale thermophoresis data. Anal Biochem **2016**;496:79-93.
36. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr **2007**;40:658-74.
37. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N*, et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr **2010**;66:213-21.
38. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr **2010**;66:486-501.
39. Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ*, et al.* MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr **2010**;66:12-21.
40. Labosky PA, Barlow DP, Hogan BL. Mouse embryonic germ (EG) cell lines: transmission through the germline and differences in the methylation imprint of insulin-like growth factor 2 receptor (Igf2r) gene compared with embryonic stem (ES) cell lines. Development **1994**;120:3197-204.
41. Akutagawa J, Huang TQ, Epstein I, Chang T, Quirindongo-Crespo M, Cottonham CL*, et al.* Targeting the PI3K/Akt pathway in murine MDS/MPN driven by hyperactive Ras. Leukemia **2016**;30:1335-43.
42. Lyons J, Ghazi PC, Starchenko A, Tovaglieri A, Baldwin KR, Poulin EJ*, et al.* The colonic epithelium plays an active role in promoting colitis by shaping the tissue cytokine profile. PLoS Biol **2018**;16:e2002417.
43. Kettenbach AN, Gerber SA. Rapid and reproducible single-stage phosphopeptide enrichment of complex peptide mixtures: application to general and phosphotyrosine-specific phosphoproteomics experiments. Anal Chem **2011**;83:7635-44.
44. Paulo JA, McAllister FE, Everley RA, Beausoleil SA, Banks AS, Gygi SP. Effects of MEK inhibitors GSK1120212 and PD0325901 in vivo using 10-plex quantitative proteomics and phosphoproteomics. Proteomics **2015**;15:462-73.
45. Ting L, Rad R, Gygi SP, Haas W. MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. Nat Methods **2011**;8:937-40.
46. Huttlin EL, Jedrychowski MP, Elias JE, Goswami T, Rad R, Beausoleil SA*, et al.* A tissue-specific atlas of mouse protein phosphorylation and expression. Cell **2010**;143:1174-89.
47. Paulo JA, O'Connell JD, Gygi SP. A Triple Knockout (TKO) Proteomics Standard for Diagnosing Ion Interference in Isobaric Labeling Experiments. J Am Soc Mass Spectrom **2016**;27:1620-5.
48. McAlister GC, Huttlin EL, Haas W, Ting L, Jedrychowski MP, Rogers JC*, et al.* Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. Anal Chem **2012**;84:7469-78.
49. Lyons J, Brubaker DK, Ghazi PC, Baldwin KR, Edwards A, Boukhali M*, et al.* Integrated in vivo multiomics analysis identifies p21-activated kinase signaling as a driver of colitis. Sci Signal **2018**;11:eaan3580.
50. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W*, et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res **2015**;43:e47.

51. Phipson B, Lee S, Majewski IJ, Alexander WS, Smyth GK. Robust Hyperparameter Estimation Protects against Hypervariable Genes and Improves Power to Detect Differential Expression. Ann Appl Stat **2016**;10:946-63.
52. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N*, et al.* The Sequence Alignment/Map format and SAMtools. Bioinformatics **2009**;25:2078-9.
53. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res **2010**;38:e164.
54. Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC*, et al.* The landscape of cancer genes and mutational processes in breast cancer. Nature **2012**;486:400-4.

FIGURE LEGENDS

Figure 1. Biochemical characterization of mutant K-Ras. **A,** GDP dissociation curves. Each allele was evaluated +/- SOS1. **B,** Quantification of nucleotide exchange rates. The intrinsic exchange rate of K-Ras4B^{A146T} is significantly higher than K-Ras4B^{WT} and K-Ras4B^{G12D} and it is further enhanced by SOS1. ** $P < 0.01$, unpaired t test; *** $P < 0.01$, unpaired t test. **C,** SOS1:K-Ras4B interaction as measured by microscale thermophoresis. K-Ras4B^{A146T} shows enhanced affinity toward SOS1 relative to K-Ras4B^{WT}. The SOS1 construct includes the REM and CDC25 domains. **D,** Measurement of intrinsic GTP hydrolysis. K-Ras4B^{A146T} and K-Ras4B^{G12D} show slightly decreased hydrolysis relative to K-Ras^{WT}. Inset shows the curve for the hydrolysis reaction. ** $P < 0.01$, unpaired t test; *** $P < 0.01$, unpaired t test. **E,** Measurement of p120GAP-induced GTP hydrolysis. GAP-induced hydrolysis is reduced for K-Ras4B^{G12D}, but not by K-Ras4B^{A146T}. Inset shows the curve for the hydrolysis reaction. * $P < 0.05$, unpaired t test; *** $P < 0.01$, unpaired t test. All assays were done in triplicate.

Figure 2. Structure of K-Ras4B^{A146T}. **A,** Comparison of K-Ras4B^{WT}, K-Ras4B^{A146T}, and K-Ras4B^{G12D} X-ray structures demonstrating extension of switch 1 (yellow) and increased flexibility of switch 2 (highlighted in green) in K-Ras4B^{A146T}. **B,** 2Fo-Fc electron density contoured at 1.4σ for X-ray crystal structure of GDP:K-Ras4B^{A146T}. GDP is colored by element and surrounding active site residues are labeled. The active site Mg from superimposed structure of GDP:K-Ras4BWT (4OBE) is also shown in magenta, demonstrating no density in this location for K-Ras4B^{A146T}. **C,** Regions of K-Ras4B^{A146T} (highlighted in red) demonstrating increased deuterium exchange relative to WT, as measured by HDX-MS. **D,** SOS:H-Ras (PDB 1bkd) structure

demonstrates the SOS-RAS binding interface. Multiple regions of H-Ras would be covered by switch 1 in the closed form.

Figure 3. K-Ras^{A146T} exhibits tissue-specific effects on homeostasis. **A,** H&E images of the colonic epithelium from mice expressing different forms of mutant K-Ras. **B,** Quantification of epithelial crypt height as a function of colonic location in mice with indicated K-Ras genotypes. **C,** Box and whisker plot illustrating the quantification of average epithelial crypt height in the colons of *Fabp1-Cre* mice expressing the indicated K-Ras allele. N=6 for WT, N=5 for A146T, and N=7 for G12D. * $P < 0.05$, Mann-Whitney U test; ** $P < 0.01$, Mann-Whitney U test. **D,** Box and whisker plot illustrating the quantification of average number of PH3-positive cells per crypt in the colons of *Fabp1-Cre* mice with indicated K-Ras genotypes. N=6 for WT, N=7 for A146T, and N=6 for G12D. * $P < 0.05$, Mann-Whitney U test; ** $P < 0.01$, Mann-Whitney U test. **E,** White blood counts (WBC) from *Mx1-Cre* animals expressing different K-Ras alleles in hematopoietic cells. **F,** Hemoglobin (Hb) counts from *Mx1-Cre* animals expressing different K-Ras alleles in hematopoietic cells. Error bars show mean ± s.e.m., and lines show modeled fixed effects from a mixed linear effects model. N=14 for WT, N=6 for A146T, and N=28 for G12D. **G,** Myeloid colony formation by bone marrow plated in methylcellulose with varying concentrations of GM-CSF. Experiments were done in duplicate. **H,** Representative H&E images of pancreases from 8-week-old *Pdx1-Cre* mice expressing WT K-Ras, K-Ras^{A146T}, or K-Ras^{G12D}. Scale bar = 100 μm in all panels.

F**igure 4.** Allele effects in tumor models. **A,** Colonic tumors from *Fabp1-Cre* ; *Apc*^{2lox14}/+ animals expressing different K-Ras alleles. Top row: H&E, Bottom row: IHC for β-catenin. Later stage tumors expressing K-Ras^{A146T} or K-Ras^{G12D} are not distinguishable histologically. **B,** Survival curves for *Fabp1-Cre* ; *Apc^{2lox14}*/+ animals bearing colonic tumors. The penetrance of colonic tumors was 100% in all groups. **C,** Representative histology from pancreatic tumors expressing different mutant forms of K-Ras. **D,** Survival curves for *Pdx1-Cre* ; *Tp53*$^{LSL-R270H/+}$ animals bearing pancreatic tumors. All animals express Trp53^{R270H} in the pancreas. Note that penetrance of PDAC was 100% in G12D, but only 75% in animals expressing A146T. Scale bar = 100 um in all panels. *P* values from survival curves calculated using Log-rank test.

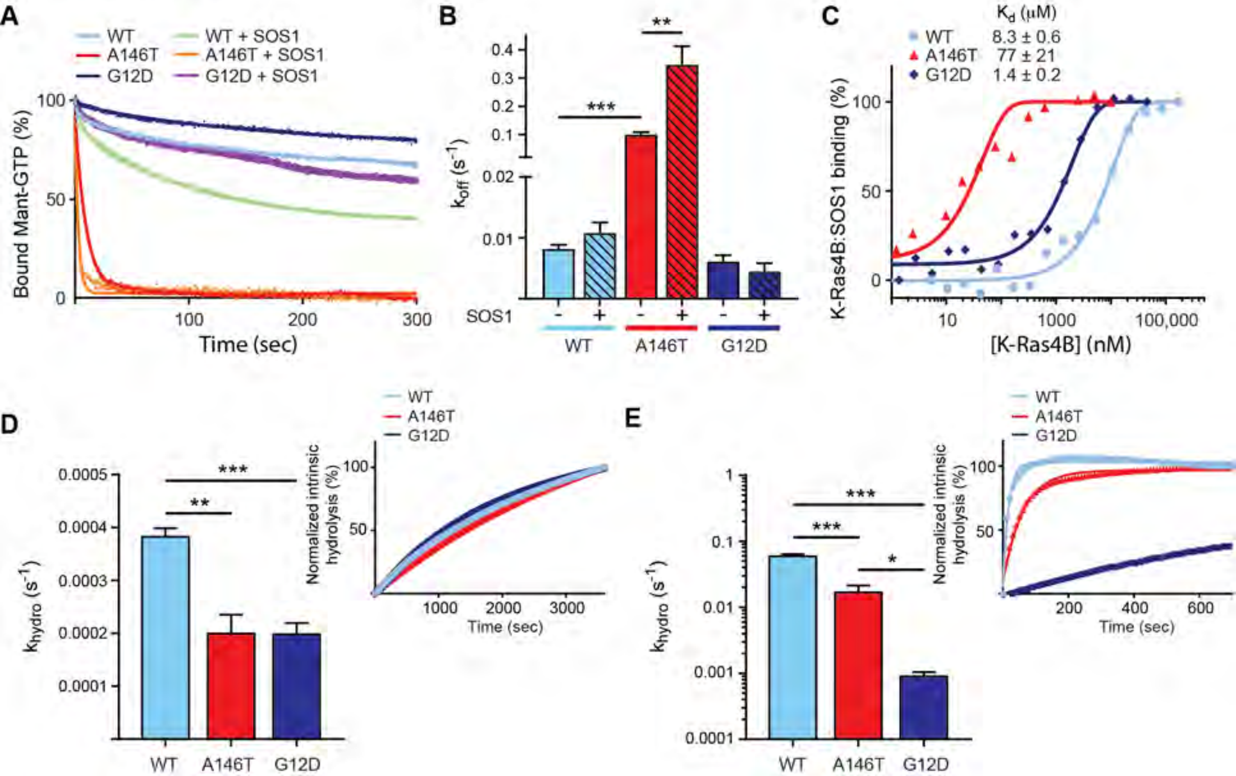
Figure 5. Global proteomics and phospho-proteomics analysis. **A,** Schematic of sample preparation, processing, and data analysis workflow for mass spectrometry. **B,** GSEA results for KEGG pathway enrichment comparing A146T:WT (denoted by red line) and G12D:WT (denoted by blue line) using the colon total protein dataset. Each line in the heat map represents the normalized enrichment score (NES) for a single KEGG pathway. The expanded detail illustrates the rank metrics for proteins contributing to KEGG_NITROGEN_METABOLISM (commonly up-regulated) and KEGG_CALCIUM_SIGNALING_PATHWAY (divergently regulated) enrichment. Each line in the heat map represents the rank metric for an individual protein in the corresponding KEGG pathway. **C,** GSEA results for KEGG pathway enrichment comparing A146T:WT and G12D:WT using the pancreas total protein dataset. The expanded detail illustrates the rank metrics for proteins contributing to KEGG_BASAL_TRANSCRIPTION_FACTORS (commonly up-regulated), KEGG_NITROGEN_METABOLISM (commonly down-regulated), and KEGG_COMPLEMENT_AND_COAGULATION_CASCADES (divergently regulated). **D,** GSEA results for KEGG pathway enrichment comparing A146T:WT and G12D:WT using the spleen total protein dataset. The expanded detail illustrates the rank metrics for proteins

contributing to KEGG_CELL_CYCLE (commonly up-regulated) and KEGG_COMPLEMENT_AND_COAGULATION_CASCADES (commonly down-regulated). **E,** Kinase activity inference results using the colon scaled phosphopeptide dataset for each K-Ras allele compared to WT. Each line in the heat map represents the normalized enrichment score (NES) for a single kinase. Erk2 and p90Rsk substrate enrichments are presented in the expanded detail. The asterisk (*) indicates Ser1120 of Sos1. **F,** Kinase activity inference results using the pancreas scaled phosphopeptide dataset for each K-Ras allele compared to WT. Cdk2, Pkaca, and Ck2a1 substrate enrichments are presented in the expanded detail. **G,** Kinase activity inference results using the spleen scaled phosphopeptide dataset for each K-Ras allele compared to WT. Camk2a and Erk2 substrate enrichments are presented in the expanded detail. For panels B-G, NES values are shown for significantly enriched pathways or kinases (nominal *P*-value < 0.1 or an FDR q-value < 0.25); insignificant pathways are shown with an NES value of 0. Those insignificant in both A146T:WT and G12D:WT are not shown. Expanded detail compares protein rank metrics that contribute to the pathway enrichment; non-contributing substrates shown with a rank metric of 0 unless they do not contribute to either allele in which case they are not shown.

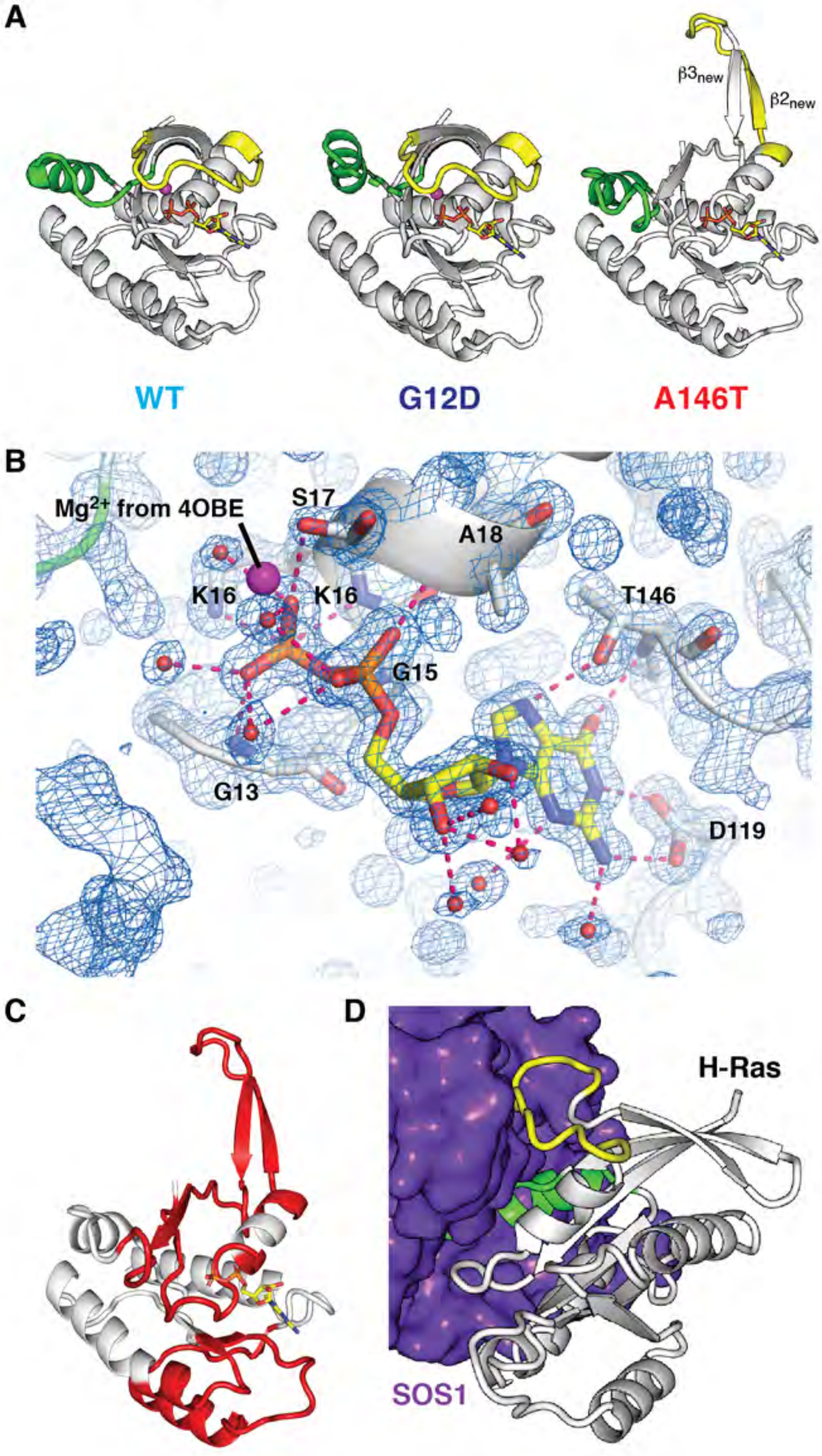
Figure 6. MAPK thresholds determine cellular behaviors in colon and small intestine. **A,** Representative western blots illustrating increased levels of p-Erk1/2 in the colons of *Fabp1-Cre* mice with different K-Ras genotypes. Each lane contains colon tissue lysate from one individual mouse. **B,** Box and whisker plots illustrating quantification of western blot analysis for p-Erk1/2 For each signal, bands were normalized to the corresponding Gapdh intensity, and then normalized values of the phosphorylated form were divided by the normalized values of total Erk1/2. N=6 mice per genotype. * $P < 0.05$, Mann-Whitney U test; ** $P < 0.01$, Mann-Whitney U test. **C,** Box

and whisker plots illustrating the quantification of average number of PH3-positive cells per crypt following trametinib (Meki) treatment of *Fabp1-Cre* mice expressing different K-Ras alleles (0.25 mg/kg, twice in 24 hours). N=5 for WT vehicle- and trametinib-treated, N=4 for A146T vehicle-treated, N=5 for A146T trametinib-treated, and N=5 for G12D vehicle- and trametinib-treated. * $P < 0.05$, Mann-Whitney U test; ** $P < 0.01$, Mann-Whitney U test. **D,** Representative western blot analysis of MAPK signaling components in the ilia of *Fabp1-Cre* mice with different K-Ras alleles. Each lane contains lysate from the distal ileum of an individual mouse. **E,** Box and whisker plots illustrating quantification of western blot analysis for p-Erk1/2. N=9 mice per genotype. * $P < 0.05$, Mann-Whitney U test; ** $P < 0.01$, Mann-Whitney U test. **F,** Representative immunofluorescence of the ilia from *Fabp1-Cre* mice with indicated K-Ras genotypes for E-cadherin (purple) and Lysozyme (green) following treatment with vehicle or trametinib (Meki, 1 mg/kg, twice daily for four days). Lysozyme is a marker for Paneth cells. **G,** Immunohistochemistry for p-Erk1/2 in pancreases from *Pdx1-Cre* mice expressing mutant forms of K-Ras. Insets contain high magnification images of pancreatic ducts. Scale bar = 50 μm in all panels.

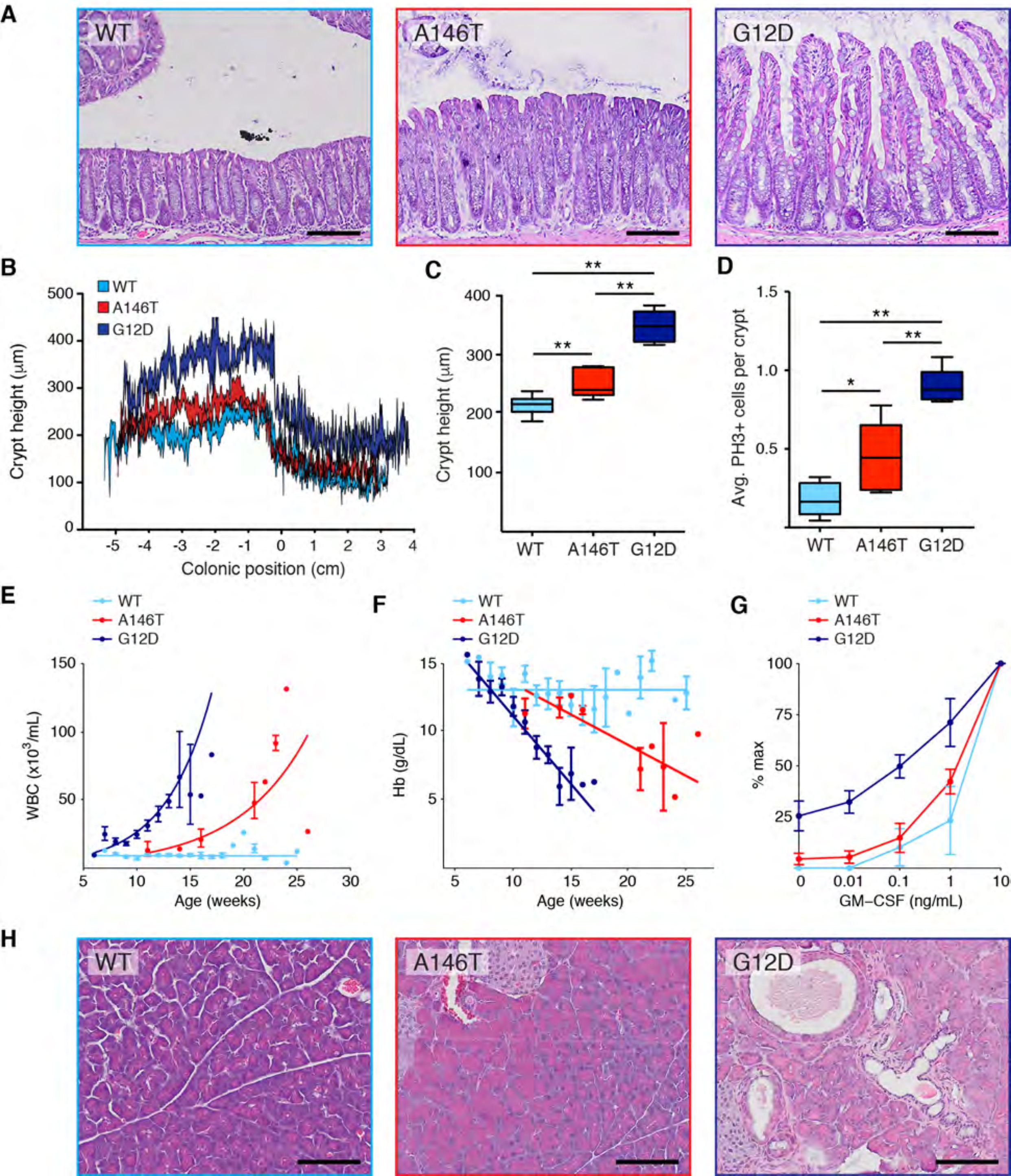
Poulin, Figure 1



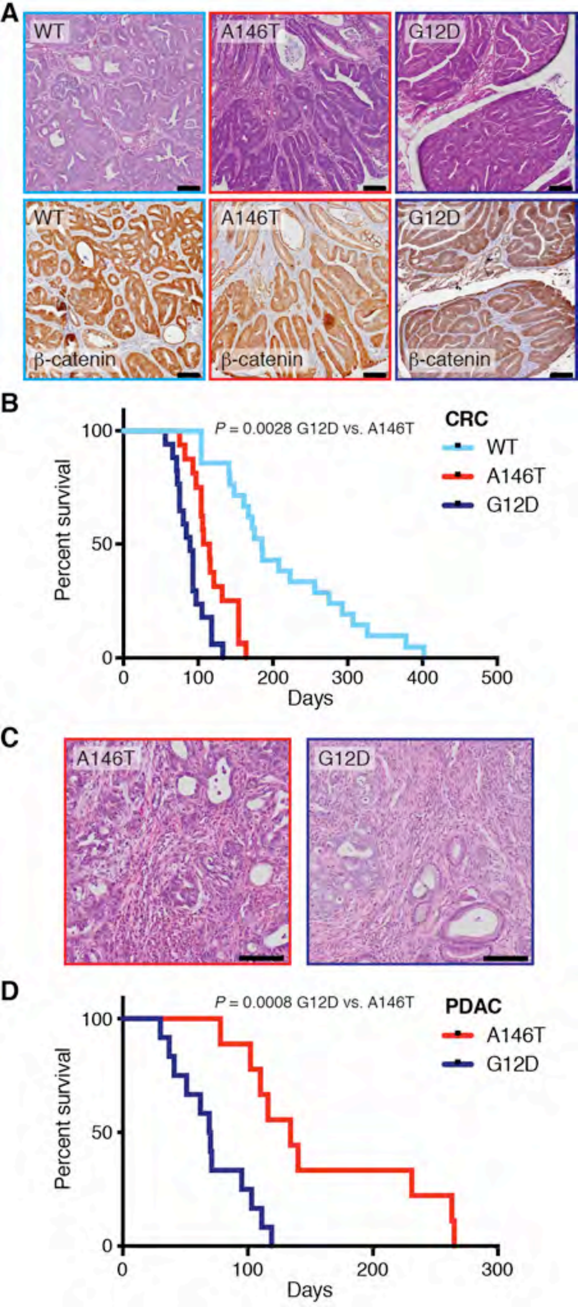
Poulin, Figure 2



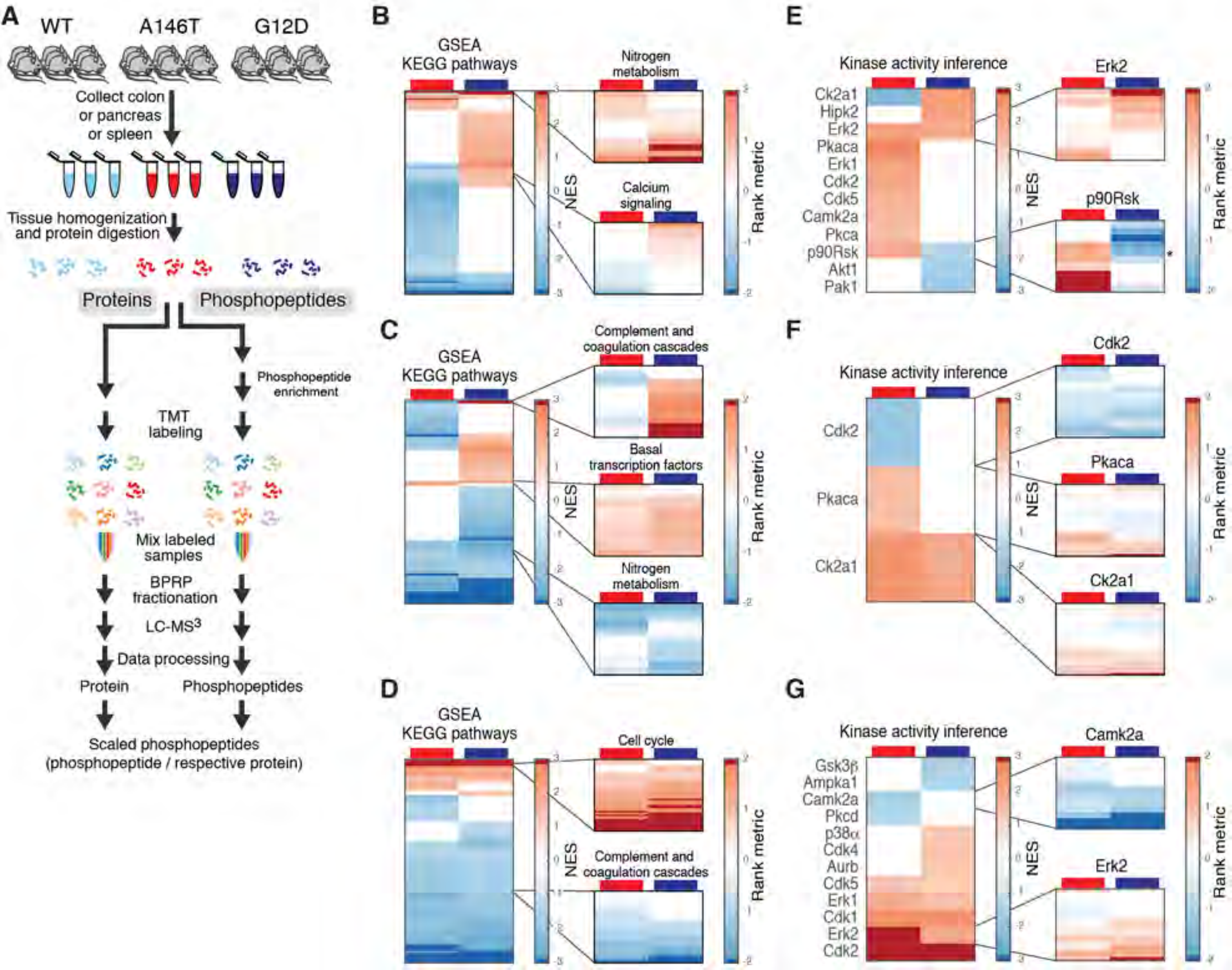
Poulin, Figure 3



Poulin, Figure 4



Poulin, Figure 5



Poulin, Figure 6