RAS internal tandem duplication disrupts GTPase-activating protein (GAP) binding to activate oncogenic signaling

Andrew C. Nelson\textsuperscript{1*}, Thomas J. Turbyville\textsuperscript{2*}, Srisathiyanarayanan Dharmaiah\textsuperscript{2}, Megan Rigby\textsuperscript{2}, Rendong Yang\textsuperscript{1}, Ting-You Wang\textsuperscript{3}, John Columbus\textsuperscript{2}, Robert Stephens\textsuperscript{2}, Troy Taylor\textsuperscript{2}, Drew Sciacca\textsuperscript{1}, Getiria Onsongo\textsuperscript{1}, Anne Sarver\textsuperscript{4}, Subbaya Subramanian\textsuperscript{4}, Dwight V. Nissley\textsuperscript{2}, Dhirendra K. Simanshu\textsuperscript{2,3}, Emil Lou\textsuperscript{5,3}

\textsuperscript{1}Department of Pathology, University of Minnesota, Minneapolis, Minnesota 55455, USA
\textsuperscript{2}NCI RAS Initiative, Cancer Research Technology Program, Frederick National Laboratory for Cancer Research, Leidos Biomedical Research, Inc., Frederick, Maryland 21701, USA
\textsuperscript{3}The Hormel Institute, University of Minnesota, Austin, Minnesota 55912, USA
\textsuperscript{4}Department of Surgery, University of Minnesota, Minneapolis, Minnesota 55455, USA
\textsuperscript{5}Department of Medicine, Division of Hematology, Oncology and Transplantation, University of Minnesota, Minneapolis, Minnesota 55455, USA

*These authors contributed equally to this manuscript

\textsuperscript{3}Corresponding authors:
Andrew Nelson, MD, PhD
E-mail: nels2055@umn.edu

Dhirendra K Simanshu, PhD
Email: dhirendra.simanshu@nih.gov

Emil Lou, MD, PhD
E-mail: emil-lou@umn.edu

Running title: Internal Tandem Duplication in NRAS

Keywords: RAS protein; colon cancer; structural biology; protein structure; oncogene
Abstract

The oncogene RAS is one of the most widely studied proteins in cancer biology, and mutant-active RAS is a driver in many types of solid tumors and hematological malignancies. Yet the biological effects of different RAS mutations and the tissue-specific clinical implications are complex and nuanced. Here, we identified an internal tandem duplication (ITD) in the switch II domain of NRAS from a patient with extremely aggressive colorectal carcinoma. Results of whole-exome DNA sequencing of primary and metastatic tumors indicated that this mutation was present in all analyzed metastases and excluded the presence of any other clear oncogenic driver mutations. Biochemical analysis revealed increased interaction of the RAS ITD with Raf proto-oncogene Ser/Thr kinase (RAF), leading to increased phosphorylation of downstream MAPK/ERK kinase (MEK)/extracellular signal–regulated kinase (ERK). The ITD prevented interaction with neurofibromin 1 (NF1)–GTPase-activating protein (GAP), providing a mechanism for sustained activity of the RAS ITD protein. We present the first crystal structures of NRAS and KRAS ITD at 1.65–1.75 Å resolutions, respectively, providing insight into the physical interactions of this class of RAS variants with its regulatory and effector proteins. Our in-depth bedside-to-bench analysis uncovers the molecular mechanism underlying a case of highly aggressive colorectal cancer and illustrates the importance of robust biochemical and biophysical approaches in the implementation of individualized medicine.
Introduction

Driver mutations in KRAS, primarily at codons 12, 13, 61, and 146, have been well-characterized and are identified in 35-40% of cases of colorectal cancer (CRC). An additional 5-10% of CRC cases harbor a mutation in NRAS at functionally identical codons. RAS mutations are associated with decreased response to epidermal growth factor receptor (EGFR) inhibition and increased incidence of distant metastasis (1). Recent advances in clinical genomic profiling have expanded recommended testing to look for a broader spectrum of mutations in KRAS, NRAS, and BRAF in patients with metastatic CRC to guide therapy planning with anti-EGFR targeted agents. This expansion has led to the identification of rare or previously uncharacterized RAS mutations. Further, the clinical characteristics of NRAS-mutated CRCs have not been well established because of their relatively low prevalence. There is a gap in our knowledge about the biological function and overall oncogenic role of less common forms of RAS mutations, especially within NRAS.

Here, our collaborative research team reports characterization of a newly identified NRAS internal tandem duplication (ITD) of 10 amino acids within the switch II domain, in a patient with exceptionally widespread CRC (>100 individual metastatic sites) that underwent clinical next-generation sequencing (NGS) testing (2). The prevalence of such RAS ITDs in CRC is unclear, and at the time of initial diagnosis (2015), no in-frame insertions of this size or larger were catalogued within NRAS exon 3 in the COSMIC database (8) and the published literature for this type of mutation in CRC was sparse (5); thus, the clinical and biological significance of this alteration was initially uncertain.

This NRAS ITD is located within the switch II domain of the protein, which includes the Q61 amino acid critical for GTP hydrolysis (9); thus, we inferred this mutation might prevent GAP-mediated inactivation of the NRAS protein and lead to constitutive downstream signaling. Given the paucity of clinical information but with a plausible biological mechanism, this mutation was interpreted as a variant with possible clinical significance (10) for predicting resistance to anti-EGFR therapies. No other mutations in BRAF, PIK3CA, KRAS, or HRAS were identified. Unfortunately, the patient’s disease continued to rapidly progress and he passed away two weeks later.

Results

Next-generation sequencing (NGS) identifies an NRAS ITD in cerebellar metastasis

The patient was a 48 year-old man with widely metastatic colorectal cancer. Initial KRAS Sanger sequencing (the institutional standard of care at that time) of his primary tumor was negative for pathogenic mutations. Despite ongoing treatment that included blockade of epidermal growth factor receptor (based on KRAS wild-type status only), the patient experienced further metastatic progression (including to the brain). Therefore a focused clinical NGS panel was used to assess a cerebellar lesion for mutations in KRAS, NRAS, HRAS, BRAF, and PIK3CA. This post-treatment testing panel identified an in-frame duplication of 30 nucleotides in exon 3 of NRAS (NM_002524:c.164_193dup, p.I55_Y64dup; Supplementary Table 1); the variant allele fraction (VAF) was 25% in the context of an estimated tumor cellularity of 50%. At the time of this diagnosis, no in-frame insertions of this size or larger were catalogued within NRAS exon 3 in the COSMIC database (8) and the published literature for this type of mutation in CRC was sparse (5); thus, the clinical and biological significance of this alteration was initially uncertain.

Post-mortem examination and focused NGS testing of multiple tumor sites

An autopsy was performed, which catalogued >100 individual metastatic lesions within nearly
all solid organs. To further characterize the patient’s malignancy, three metastatic sites (cerebellum, lung, and liver) plus the archived primary tumor from the surgical resection were assessed by a clinically-focused solid tumor NGS panel of 13 genes. The NRAS ITD was identified in all four samples, confirming that this mutation was present in the primary tumor and in three independent metastatic sites that developed at different times during the patient’s disease course (Supplementary Table 2). In addition, a pathogenic TP53 mutation (NM_000546: c.824G>T, p.C275F) was identified in the primary tumor and all three metastatic sites. No other pathogenic mutations were identified in the remaining 11 clinically-relevant genes assessed in these four samples at the time of autopsy.

Exome sequencing of primary and metastatic tumors
Given the lack of additional potential driver mutations on the targeted clinical cancer NGS panel, we used whole exome NGS of the archived primary tumor, lung metastasis, and cerebellar metastasis to more extensively explore whether other genomic drivers were responsible for the aggressive clinical course of the disease. Variant annotation and filtering identified 847 potentially significant somatic variants in the primary tumor and an additional 551 potentially significant variants that were unique to the two metastatic lesions (i.e., variants not identified in the primary tumor) (Figure 1A). These variant lists were further prioritized for loss of function variants (frameshift indels or nonsense), in-frame indels, canonical splice site variants, and single nucleotide missense variants with high-to-moderate Mutation Assessor functional impact scores (11); this process left 169, 58, and 72 prioritized variants for review in the primary, lung, and central nervous system (CNS) lesions, respectively (Supplemental Data File 1).

Overall, somatic exome analysis of the patient’s primary colon tumor, lung, and CNS metastasis confirmed that the NRAS ITD was present in the majority of tumor cells in all sites of disease studied. Further, no other clear mechanistic driver mutations were identified by whole exome sequencing. Loss of tumor suppressive functions of TP53 and CSMD1 likely contributed to the aggressive phenotype of the malignancy. Interestingly, the presence of NRAS or CSMD1 mutations, either alone or in combination, was associated with decreased overall survival in the Cancer Genome Atlas (TCGA) CRC dataset (Figure 1C) (12-14).

Interrogation of TCGA datasets to assess prevalence of RAS ITDs in colorectal and other solid cancers.
The original analysis of the TCGA colorectal adenocarcinoma project (12) did not report any similar insertion mutations within RAS genes. We hypothesized that this type of duplication may have been missed by the bioinformatics pipelines that produced the published datasets, so we performed a focused re-analysis of the RAS genes using the primary sequencing files from several TCGA projects using the ScanIndel bioinformatics pipeline (15), which has been optimized to detect longer insertion mutations. We analyzed the colorectal adenocarcinoma (n=634), pancreatic adenocarcinoma (n= 186), lung adenocarcinoma (n=586), and lung squamous cell carcinoma (n=511) TCGA data sets. No clinically significant in-frame insertion mutations within NRAS, KRAS, or HRAS were identified across these studies. This analysis suggests that somatic in-frame insertion mutations (including ITDs) within RAS genes are potentially a rare event in the pathogenesis of RAS-related solid cancers. However, the size of these cohorts are likely not large enough to confidently estimate the prevalence of rare but clinically important cancer gene alterations (16).

Cell-based functional characterization of NRAS ITD confirms constitutive activity
To determine the functional activity of the NRAS ITD protein, we used bioluminescence resonance energy transfer (BRET) to detect protein-protein interactions in live, intact cells (17,18). We co-transfected the human colorectal cancer cell line Caco-2 and the immortalized human embryonic kidney HEK293T cells with plasmids expressing ITD mutant or wildtype RAS proteins linked to a HaloTag acceptor together with plasmids expressing the RAS
effector RAF1 tagged with a NanoLuc (NLuc) bioluminescent donor construct. In saturation experiments, cells were transfected with a constant amount of the RAF1-NLuc donor construct in the presence of increasing amounts of the indicated RAS-HaloTag acceptor constructs. We generated saturation curves for the interactions between RAF1 and wild-type KRAS, wild-type NRAS, ITD mutant KRAS, Q61R mutant KRAS, and ITD mutant NRAS (Figure 2). Curves were produced by a hyperbolic nonlinear regression and the BRET$_{50}$ was calculated (curve fit parameters provided in Supplementary Tables 3-6). The BRET$_{50}$ is the concentration of transfected acceptor DNA that generates a bioluminescence signal at 50% of the maximum (BRET$_{max}$); lower values of BRET50 are consistent with higher affinity of interaction. These experiments showed a greater affinity between RAF1 and the internal duplication mutants compared to the wild-type proteins (Figure 2A, 2B, 2C). Biologically, other known carcinogenic point mutations in RAS (KRAS Q61R) recruit RAF1 with higher affinity, leading to more downstream signaling (Figure 2A). These results suggest that the 10 amino acid ITD (in both NRAS and KRAS) drives interactions with the effector RAF1 at a similar level of potency to the well-characterized Q61R oncogenic missense mutation.

**Downstream MAPK signaling is activated by the RAS ITD**

To further demonstrate the biologic relevance of the increased affinity of NRAS ITD to RAF in cells, we evaluated the effect of expression of the internal duplication mutants on downstream signaling pathways. Transfection of increasing amounts of NRAS ITD plasmid construct into Caco-2 cells showed a significant induction of MAP Kinase signaling, measured by both ERK and MEK phosphorylation in Western blots (Figure 3A). This pathway activation was similar to that of the established oncogenic mutant NRAS Q61R (Figure 3A). Moreover, when overexpressed in HEK293T cells, both the NRAS and KRAS internal tandem duplication mutants showed increases in both ERK and MEK phosphorylation when compared to the wildtype proteins (Figure 3B, 3C). The signaling activation by the KRAS ITD protein was similar to levels induced by the common KRAS G12D oncogenic mutant (Figure 3D). These results were consistently produced across three independent experiments (Supplementary Figure 1), providing clear evidence that a functional consequence of the ITD mutation is constitutive activation of the oncogenic MAP kinase signaling pathway.

**Recombinant RAS ITD binds to RAF1 but not to NF1 (neurofibromin 1) GAP**

To examine the effect of the ITD in RAS on its interaction with GAP and its effector proteins, we purified the recombinant GTPase domain (1-169) of NRAS and KRAS with and without the ITD and carried out binding studies using isothermal titration calorimetry (ITC). We compared the binding of wild-type NRAS and NRAS ITD with NF1 GAP and RAF1 effector proteins. ITC experiments showed that WT NRAS and KRAS bound to NF1 (GRD; GAP-related domain) with high affinity (Kd = 1 µM) as reported previously (19), whereas ITD mutants of NRAS and KRAS showed complete loss of binding to NF1, suggesting that the ITD resulted in fully impaired GAP-mediated GTPase activity (Fig. 4A and Supplementary Fig. 2B and 2C). Unlike NF1 GAP, both WT and ITD mutants of NRAS and KRAS bound to RAF1 (RBD; RAS-binding domain) with similar affinity in vitro, confirming that the ITD does not negatively affect RAF1 binding (Fig. 4B). This biochemical evidence is consistent with the cell biology data (Figures 2 and 3) showing increased physiologic interaction between RAS ITD and RAF with subsequently increased MAPK signaling.

To explore the lack of NF1-GAP:RAS ITD binding further, we returned to our BRET cell-based expression system. By co-expressing NF1 with the RAF1 and RAS constructs, we demonstrated biologically that NF1 expression significantly decreased the peak mBRET signal for wild-type RAS constructs, but had no effect on the RAS ITD constructs (Fig. 4C). This finding confirmed within an intact cell model system that the switch II ITD prevented effective
interaction of RAS with GAP protein, leading to sustained activity.

Next, we examined the effect of ITD on the intrinsic and GAP-mediated GTP hydrolysis in RAS using a fluorescent-based assay in the absence and presence of NF1, respectively, in the reaction mixture. Interestingly, the presence of NF1 GAP, which has been shown to increase the rate of intrinsic GTP hydrolysis by $10^5$ fold for wild-type RAS, failed to stimulate any increase in the rate of GTP hydrolysis for RAS ITD (Supplementary Fig. 3). These results suggest that ITD disrupts RAS interaction with GAP protein and results in loss of GAP-mediated GTP hydrolysis.

Crystal structure of RAS ITD provides rationale for its oncogenicity

To understand the effect of ITD on RAS tertiary structure and RAS interactions with GAP and effector proteins, we crystallized and solved the structure of GDP-bound NRAS ITD and KRAS ITD to a resolution of 1.65 Å and 1.75 Å, respectively (Fig. 5A, 5B and Supplementary Table 7). In the NRAS ITD structure, a major part of ITD in the switch II region showed no interpretable electron density suggesting that this region is likely to be very flexible or disordered inside the crystal (Fig. 5A). In the KRAS ITD structure, most of the amino acids that are part of the ITD region had interpretable electron density as it is stabilized by crystal contact, which suggests that the observed conformation showing a large conformational change in the switch II region is one of the snapshots of many possible conformations for the highly dynamic ITD (Fig. 5B and Supplementary Fig. 4A and 4B). The region with uninterpretable electron density in the KRAS ITD structure includes the last three residues of the ITD (residues 72-74) and the six residues (residues 75-80) present at the end of switch II. In the KRAS ITD structure, residues E63 and Y64 make a helical turn which is followed by seven residues (I65-Q71) long loop in the middle of the switch II region, which protrudes away from the core of the protein. The B-factor of ITD and switch II region is higher than rest of the protein suggesting that this region is likely to be the most flexible part of the protein. Structural superposition of GDP-bound NRAS ITD with GDP-bound KRAS ITD shows similar tertiary structure with a partially disordered switch II region due to ITD mutation (Fig. 5C). Structural comparison of GDP-bound NRAS ITD with GMPPNP-bound wild-type NRAS highlighted similar tertiary structure including the switch I region and the difference in the switch II because of the ITD (Fig. 5D).

Structural superposition of the KRAS ITD with the previously solved structures of KRAS-NF1 GAP complex (PDB ID: 6OB2), HRAS-PI3Kγ complex (PDB ID: 1HE8), and HRAS-RAF1(RBD) complex (PDB ID: 4G0N) showed a steric clash of RAS ITD in the case of NF1 and PI3Kγ, but not in the case of RAF1 (Fig. 5D, 5E and 5F). This observation is aligned with previous results that have shown the RAS-RAF1 interaction is limited to switch I (20), whereas NF1 and PI3Kγ interact with RAS protein via both switch I and II regions (19,21). To examine the effect of ITD on RAS-PI3Kγ interaction, we measured the binding affinity of wild-type NRAS and NRAS ITD with PI3Kγ using the ITC experiment. Results from the ITC experiments showed that WT NRAS binds to PI3Kγ with a dissociation constant (Kd) of 8.3 µM, whereas NRAS ITD showed complete loss of binding to PI3Kγ. This observation provides additional evidence that the presence of ITD results in loss of interaction between RAS and PI3Kγ. In the surface representation, RAS ITD does seem to clash with RAF1-RBD partially. However, the ITC results that show that RAS ITD binds to RAF1-RBD with similar affinity as that of wild-type RAS, and the flexibility seen for the ITD region in the crystal structures described here, make it clear that the ITD has no effect on RAS interaction with RAF1-RBD (Supplementary Fig. 5). Unlike oncogenic mutations that are limited to a different side chain at a single amino acid, this ITD of 10 residues creates a significant steric clash that is likely to result in complete loss of binding of GAP protein to RAS ITD.

Discussion

We report the biochemical, biophysical, and functional characterization of a NRAS ITD
mutation that disrupts the switch II domain and prevents GAP-mediated GTP hydrolysis. This ITD was identified in a patient with extremely aggressive, therapeutically resistant colonic adenocarcinoma in which whole exome analysis did not reveal any other activating, driver mutations. This NRAS ITD demonstrated increased affinity with BRAF, suggesting decreased hydrolysis of GTP and thus prolonged RAS activation and downstream effects.

In a clinical context, identification of less common variants of oncogenic Ras over the past decade has expanded the pool of testable mutations that help predict biological behavior and predict therapeutic efficacy of drugs such as inhibitors of EGFR in patients with metastatic CRC. Despite expansion of clinical Ras sequencing platforms, concern remains that additional variants are not routinely detected, and can thus affect clinical outcomes. Though interrogation of several TCGA datasets suggest ITDs have a low prevalence in solid malignancies, we have subsequently identified a similar 27 base pair, 9 amino acid insertion mutation in the Switch II domain of KRAS (NM_004985: c. 162_188dup; p. D54_E62dup) in a case of pancreatic adenocarcinoma at our institution. The increasing reports of Ras ITD in a spectrum of malignancies from our institution and others indicates that the ITD is a clinically important, albeit rare, form of Ras mutation.

The biological activity of larger Switch II insertions appear to be substantial, as this ITD facilitates Ras-Raf interaction in a cell biology system that leads to increased MAPK signaling. More specifically, the ITD blocks the interaction of Ras-Gap in vitro, and our cell biology system confirms that Gap co-expression does not diminish sustained Ras ITD:Raf1 affinity/interaction. Furthermore, by solving the crystal structure of GDP-bound Ras ITD, we provide convincing structural evidence that the 10 amino acid ITD blocks Ras-Gap binding, and provide insights into the interactions of the Ras ITD with other effector proteins. Our data suggest that ITDs of this length lead to constitutive activation of Ras with significant biologic and clinical consequences.

The work shown here is consistent with findings seen in other limited reports of Ras ITDs. A prior report of a 7 amino acid Ras ITD identified in a pediatric hematologic malignancy provided functional and biochemical data supporting an oncogenic phenotype (6). Eijkelenboom et al. have reported that Ras ITDs (particularly involving Hras, ranging in size from 7-10 amino acids) are slightly more common in vascular malformation/overgrowth syndromes (2%) prevalence (7). The different biologic properties of Ras isoforms (Kar, Hras, and Nras) are mainly dictated by differences in the hypervariable region (i.e. the last 20 residues at the C-terminal end) as well as preferences for different membrane environments and tissue-specific expression patterns (22). However, the G-domain of these Ras isoforms is very similar, and in biochemical and structural experiments, G12, G13 and Q61 mutants of Kar and Nras often behave very similarly. The accumulating literature on Ras ITDs consistently demonstrates this type of alteration activates downstream signaling; however the composite data suggests the specific amino acid length of the duplication may drive variability of biochemical function and the clinical context in which it is identified (i.e. the specific tissue of origin and the specific Ras isoform affected) may drive variability in disease outcome. Continued study of this new class of clinically-relevant Ras mutation is necessary to clarify these details and define the downstream biologic and clinical consequences.

The Nras ITD and pathogenic TP53 p.C275F missense mutations were verified by DNA exome analysis in all three specimens (Figure 1B). In addition, a splice site mutation was identified in all tested tumor sites that abolished the splice donor site of exon 58 in CSMD1 (which contains 70 total exons). CSMD1 (CUB and Sushi Multiple Domains 1) is a transmembrane protein that is not functionally well-characterized, but studies indicate it acts as a tumor suppressor in multiple tumor types (23-27). Specifically, mutations of CSMD1 appear to be enriched in stage 3 and 4 colorectal carcinomas (24), and loss of CSMD1 function by both genomic and epigenomic mechanisms has
been linked to earlier age of onset (27). Thus, it is possible this mutation shared across all tumor sites contributed to the aggressive phenotype of the tumor.

Two mutations that appeared limited to the CNS and lung metastases were identified. Both were classified as variants of uncertain significance. The p.P169H missense mutation in PTEN occurs in the TI catalytic loop, which is a hotspot for missense mutation, but in vitro functional studies indicate this specific variant has nearly wild-type levels of PIP phosphatase activity (28). Thus, it is unclear whether this mutation would have affected PI3K signaling in our patient’s metastatic lesions. The second mutation of interest restricted to the metastases was a p.L459I missense mutation in MPDZ (multiple PDZ containing protein). No specific information on this variant could be identified in somatic mutation databases or the published literature. Nonetheless, loss of MPDZ function has been associated with metastasis and poor prognosis in breast cancer (29); thus, the identification of this mutation in only metastatic sites was intriguing, but not sufficiently conclusive of a pathogenic role.

Elucidation of mutations in both primary and metastatic tumors demonstrated that the NRAS ITD and a TP53 dominant negative missense mutation were conserved across all disease sites, whereas none of the mutations demonstrating heterogeneity between primary and metastases had clearly defined biologic impact. The conservation of a potential RAS driver mutation with genomic heterogeneity primarily limited to variants with uncertain clinical impact is an important lesson to stress from our in-depth analysis of this patient with aggressive CRC, similar to a previous study of multiple paired primary and metastatic pancreatic cancers with KRAS driver mutations (30). Further, it demonstrates the complexity of clinically interpreting a single patient’s tumor exome data to better understand tumor behavior on a personalized level without the aid of further biochemical and structural biology information.

Taken together, this clinic-to-lab endeavor represents a process that effectively identified an explanation for this aggressive and clinically challenging case through rigorous molecular and biochemical analysis. Had this NRAS ITD been characterized prior to the patient’s diagnosis or at any point during his treatment, EGFR inhibitor therapy would not have been administered even in the context of wild-type KRAS testing because of the lack of an anticipated response (31). Although the ITD subclass of RAS alterations is still relatively rare, our finding, in combination with those of the groups cited above, indicates that RAS ITDs have important clinical impact and warrant further biologic study to understand the context-specific impact on cancer progression and therapeutic response.

Potential limitations of this study include the fact that these extensive exome results from multiple sites of disease represent the underlying genomics of one patient, and thus cannot be assumed to be relevant to all NRAS-mutant CRC tumors. The mutation profile of NRAS is different from that of KRAS; therefore, while they are structurally conserved, the small differences in biologic regulation between these RAS subtypes may have diverse clinical impact. While it would have been ideal to perform in vitro analysis of signaling defects using cell lines derived directly from the patient’s tumors, tumor cells were not viable in the post-mortem setting; thus it was not feasible to generate patient-derived tumor cell lines in this case. Nonetheless, in vitro analysis involving overexpression of the ITD in other cell lines, including a RAS-wildtype colon cancer cell line, demonstrated that this NRAS ITD was biologically activating and consistent with other reported forms of ITD.

In summary, we have identified a NRAS ITD that was conserved between primary and metastatic tumors in vivo and had constitutive biologic activity; its crystal structure provides new insight into this distinct form of mutant NRAS. No other clear oncogenic drivers were identified in association with this NRAS alteration in a clinically aggressive human cancer. Our data provide impetus for continuous quality improvement in pan-RAS molecular
analysis to ensure our ability to identify these alterations adequately in the clinical setting.

**Experimental Procedures**

**Patient samples**
A clinical research autopsy was performed following written informed consent from the patient’s family, and tumor samples were snap frozen or stored in formalin fixed, paraffin embedded (FFPE) blocks.

**Targeted Next Generation Sequencing**
An amplicon-based target enrichment of portions of 13 genes followed by next generation sequencing (NGS) was performed on all specimens using a CLIA-validated workflow (2). A subset of this panel is reported clinically for colorectal carcinoma (KRAS, NRAS, BRAF, PIK3CA, HRAS). DNA was extracted from formalin fixed, paraffin embedded (FFPE) samples using the QIAamp DNA mini FFPE tissue kit and deparaffinization solution (Qiagen, Hilden, Germany) and quantified using a Qubit 2.0 fluorometer (ThermoFisher Scientific, Waltham, MA). Amplicon enrichment was performed using a custom designed primer set on the Fluidigm (San Francisco, CA) Biomark Access Array system followed by sequencing with 2x225 base pair reads on a MiSeq instrument using version 3 chemistry (Illumina, San Diego, CA). FASTQ files were processed through a custom bioinformatics pipeline (ScanIndel) for mapping, indel realignment, and variant calling (15).

**Exome Next Generation Sequencing**
Four samples (3 fresh frozen, 1 FFPE) were prepared for whole exome NGS: metastatic lung, metastatic CNS, archived primary colon adenocarcinoma, and normal tissue. Normal tissue was collected for assessment of germline variation; the patient had concurrent chronic myelogenous leukemia, so peripheral blood could not be used. DNA from snap frozen tissues were extracted using the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany). Libraries were enriched using SureSelect QXT reagents and V6+COSMIC exome hybrid capture baits (Agilent, Santa Clara, CA). Sequencing was performed with 2x125 base pair reads on a HiSeq 2500 with V4 chemistry (Illumina, San Diego, CA). FASTQ files were processed as described above. Average coverage was >200X for all four samples. All variants with variant allele fractions >0.3 in normal tissue were considered germline. To identify somatic alterations in tumors, these germline variants were filtered out of the tumor sample variant call files. To identify unique somatic variants that either arose in metastatic tumors or were very rare subclones not detected in the primary tumor, all variants identified in the somatic primary colon variant call file were then further filtered out of the somatic variant call files of the lung and CNS metastases.

**Variant call files were annotated using Cravat (32,33) and Annovar (34). Annotated variant files were filtered to remove synonymous variants not affecting splice sites and any passenger somatic variants that had been recorded as polymorphisms in healthy human population databases with minor allele frequencies > 0.01. Variants were priority filtered for all stopgain, frameshift, in-frame indel, and canonical splice site variants. Missense variants were prioritized by Mutation Assessor scores (11). Alternative variant prioritization lists were also created using SnpEff/SnpSift (35,36). Functional pathway enrichment analysis and annotation were performed with ToppGene (37) and Panther (38).**

**Targeted Re-analysis of TCGA Data for RAS Indels**
TCGA colorectal (n=634), pancreatic (n=186), lung adenocarcinoma (n=586), and the lung squamous cell carcinoma (n= 511) tumor whole exome sequencing bam files were downloaded from the NIH Genomic Data Commons (GDC) database after appropriate regulatory approval. Indel detection was performed with ScanIndel (15). Briefly, we extracted all reads mapped to HRAS (chr11:530,242-537,567), NRAS (chr1:114,702,464-114,718,894), and KRAS (chr12:25,202,789-25,252,931) gene regions (GRCh38). Then BWA-MEM aligned reads were used as input for ScanIndel to find any size...
of indels with minimum 10x coverage and 1% variant allele frequency thresholds.

Cell Culture
HEK293T cells (ATCC, CRL-3216) were cultured in complete phenol red free DMEM (Life Technologies, 31053-028, 4.5g/L Glucose, 3.7g/L Sodium Bicarbonate), 2mM L-Glutamine (Life Technologies, 25030081), and 10% FBS (Hyclone, SH30071.03) at 37°C and 5% CO2.

Transient Transfections
HEK293T cells were seeded at 1.25e5/mL in 12-well plates (Falcon, 353043). After 24 hours, 8µL of Fugene-6 (Promega, E2691) was mixed with various amounts of plasmid DNA and allowed to incubate at room temperature for 15 minutes before addition to cells. After 24 hours, the transfection mixture/media was replaced with complete media and allowed to incubate for another 24 hours.

NanoBRET Assay
Bioluminescence resonance energy transfer (BRET) is a commonly used imaging assay that uses resonance energy transfer to quantify protein-protein interactions between a bioluminescent donor and fluorescent acceptor (39). For the energy transfer to occur, the donor and acceptor molecules must be within 10 nm of each other and in the correct orientation. The advantage of this system is that full length proteins tagged with the relevant donor acceptor pair can be expressed in cells, and interactions between the proteins of interest can be monitored in live, intact cells. For specific interactions between the donor and acceptor fusion proteins, the BRET ratio increases hyperbolically as a function of increasing acceptor/donor ratio. Saturation is reached when all donor molecules are associated with acceptors. Saturation curves are used to compare the relative affinity of proteins for each other; higher magnitude and steeper increases in signal with lower amounts of acceptor are indicative of greater energy exchange, indicating more molecules are interacting, or that the proximity and geometry of the interaction is favorable for the energy transfer to occur between the donor and acceptor, or both.

In preparation for NanoBRET measurement, transfected cells were incubated in cell dissociation buffer with 0.25% Trypsin-EDTA (Life Technologies, 25200-056). Clumps were then broken up by pipetting with the addition of complete media. After pelleting and discarding the supernatant, cell pellets were dispersed in DMEM with 0.1% FBS, counted, and plated in 384-well white Opti-Plates (Perkin Elmer, 6007290) at a density of 10,000 cells/well. Cells were diluted to 5.0E6 cells/mL in Recovery Cell Culture Freezing Medium (Life Technologies, 12648-010). Halo618 ligand (Promega, N1663) was added and incubated at 37°C for 5 hours. The NanoLuc substrate Furimazine (Promega, N1663) was added at a final concentration of 10µM and incubated for 10 minutes at room temperature. Measurements were made in a Perkin-Elmer Envision plate reader equipped with a 460/80nm band-pass filter (donor) and a 600nm long-pass filter (acceptor).

NanoBRET Calculations
milliBRET values (mBRET) were calculated by subtracting the A/D ratio, 600nm/460nm, of the donor-only channels from the A/D ratio of the sample channels and then multiplying by 1000.

\[
mBRET\ Value = 1000 \times (\frac{\text{Acceptor channel emission of sample/donor channel emission of sample}}{\text{Acceptor channel emission of donor only/donor channel emission of donor only}})
\]

mBRET values were plotted on GraphPad Prism using the Hyperbola nonlinear equation with concentration of acceptor on the X axis; these curves were used for calculation of BRET50 and BRETmax. BRET50 is the concentration at which 50% of the maximal mBRET value is achieved, and is considered a relative measure of the affinity of the molecules for each other. Curve fit parameters for Figure 2 panels A-C are summarized in Supplementary Tables 3-6.

Western blotting
Western blots were performed using standard experimental approaches to determine induction of MAPK signaling by the ITD mutants as compared to wild-type RAS isoforms. Three replicates were performed for each blot, as described here in detail. After transfection and a
48 hour incubation, HEK293T cells were lysed in a detergent buffer containing 20mM Tris HCl, 150mM NaCl, 1mM EDTA + EGTA, 1% Triton, and Halte protease and phosphatase inhibitor cocktail (Thermo Scientific, no. 78440). Lysates were clarified with centrifugation (16,000xg for 15 minutes at 4°C), protein concentration was analyzed (BCA kit, Thermo Scientific no. 23227), and 25µg protein was loaded per lane onto a Bolt 4-12% 15-well Bis-Tris gel (Invitrogen, no. NW00105BOX). Protein transfer was performed using Thermo Fisher Scientific’s iBlot 2 Dry Blotting System (no. IB21001) and transfer stacks (no. IB23001). After blocking in Odyssey Blocking buffer (LI-COR, no. 927-50003) for one hour at room temperature, membranes were then incubated at 4°C overnight in Odyssey Blocking buffer containing 0.1% Tween 20 and the following antibodies: ERK 1/2 (mouse monoclonal, Cell Signaling Technology no. 4696; 1:1000 dilution), pERK 1/2 (rabbit monoclonal, Cell Signaling Technology no. 4370; 1:2000 dilution), MEK 1/2 (mouse monoclonal, Cell Signaling Technology no. 4694; 1:1000 dilution), pMEK 1/2 (rabbit monoclonal, Cell Signaling Technology no. 9154; 1:1000 dilution), vinculin (mouse monoclonal, Sigma-Aldrich no. V9131; 1:1000 dilution), and HaloTag (mouse monoclonal, Promega no. G9211; 1:1000 dilution). Membranes were incubated for one hour at room temperature with IRDye secondary antibodies (Goat anti-Mouse 680RD, LI-COR 926-68070; Goat anti-Rabbit 800CW, LI-COR 926-32211; 1:10,000 dilutions). Membranes were washed, and images captured using the LI-COR Odyssey CLx Imaging System.

Cloning, expression, and purification of recombinant proteins
DNA constructs for the expression of G-Hs.KRAS4b(1-169) and Hs.RAF1(52-131) were previously described (40). DNA constructs for the expression of G-Hs.KRAS4b(1-169)-ITD, G-Hs.NRAS(1-169), G-Hs.NRAS(1-169)-ITD, and NF1GAPiso2(1198-1530) in the format of His6-MBP-tev-POI (MBP, maltose-binding protein; tev, tobacco etch virus protease recognition sequence; POI, protein of interest) were generated using previously outlined protocols in Taylor et al. (41) and additional details available at addgene.org/11517.

The BL21 STAR (rnel31) _E. coli_ strain carrying the DE3 lysogen and rare tRNAs (pRare plasmid CmR) was transformed with the expression plasmids (AmpR) except for NRAS and NRAS ITD which were used to transform _V. natriegens_ (42). G-Hs.KRAS4b(1-169), G-Hs.KRAS4b(1-169)-ITD, and NF1GAPiso2(1198-1530) were expressed using the Dynamite protocol and Hs.RAF1(645-1131) was expressed using the auto-induction protocol, both as described in Taylor et al. 2017 (41). G-Hs.NRAS(1-169) and G-Hs.NRAS(1-169)-ITD were expressed in auto-induction media ZYM-20052 as outlined in Taylor et al. (41), with modifications. Specifically, media was amended with 1.5% NaCl or Instant Ocean (Aquarium Systems) w/v, no lactose was added, IPTG induction began at an OD<sub>600</sub> of 4-5, induction temperature was 30°C, and cells were harvested ~6-8 hr after induction.

All expressed proteins were purified as outlined for G-Hs.KRAS4b(1-169) in Kopra et al. (43). Briefly, the expressed proteins of the form His6-MBP-TEV-target were purified from clarified lysates by IMAC, treated with His6-TEV protease to release the target protein, and the target protein separated from other components of the TEV protease reaction by a second round of IMAC. Proteins were further purified by gel filtration chromatography in a buffer containing 20 mM HEPES, pH 7.3, 150 mM NaCl, 2 mM MgCl<sub>2</sub> (GTPases only), and 1 mM TCEP. The peak fractions containing pure protein were pooled, flash frozen in liquid nitrogen, and stored at -80°C. PIK3Cg(144-1102)-His6 was received from Dr. Simon Messing.

Plasmid DNA Generation
Fusion constructs for BRET and mammalian expression were generated using combinatorial Multisite Gateway (44). For N-terminal Halotag-RAS fusions, three components were mixed in a Multisite LR reaction: a strong CMV51 promoter (att4-att5), a HaloTag (Promega) fusion protein with upstream Kozak initiation sequence and lacking a stop codon (att5-att1), and a standard Gateway Entry clone of the
various downstream RAS constructs partners (att1-att2). In all reactions, a generic mammalian expression Gateway Destination vector of the appropriate format (att4-att2 or att4-att3) was used. Correct recombinants were isolated and verified by restriction digest, and transfection-ready DNA was prepared using Qiagen plasmid preparation kits. NanoLuc-RAF1 fusions were made in a similar fashion using an att5-att1 NanoLuc (Promega) construct and a RAF1 Entry clone (Addgene #70497). For C-terminal RAF1-NanoLuc fusions, a different configuration of Multisite Gateway was used incorporating a CMV51 promoter (att4-att1), a RAF1 Entry clone lacking a stop codon (Addgene #70498) and a C-terminal NanoLuc fusion (att2-att3). Entry clones for general RAS constructs were from the RAS mutant entry clone collection (Addgene #1000000089)

Isothermal titration calorimetry measurements
Binding affinities of GMPNP-bound wild-type and ITD mutant of NRAS and KRAS (1-169) with NF1-GRD (1098-1530) and RAF1-RBD (52-131) were measured using isothermal titration calorimetry (ITC). Protein samples were prepared by extensive dialysis in a buffer (filtered and degassed) containing 20 mM HEPES (pH 7.3), 150 mM NaCl, 5 mM MgCl2, and 1 mM TCEP. For the ITC experiment, 60 μM of KRAS and 600 μM of NF1-GRD or RAF1-RBD were placed in the cell and syringe, respectively. ITC experiments were carried out in a MicroCal PEAQ-ITC instrument (Malvern) at 25 °C using an initial 0.4 μL injection and 18 subsequent injections of 2.2 μL each at 150-second intervals. Data analysis was performed based on a binding model containing “one set of sites” using a nonlinear least squares algorithm incorporated in the MicroCal PEAQ-ITC analysis software (Malvern).

Intrinsic and GAP-mediated GTP hydrolysis assay
The rate of intrinsic and NF1-mediated GTP hydrolysis was measured using a fluorescence-based assay. KRAS ITD nucleotide exchange from the GDP to GTP bound state was carried out using a protocol described previously (45). The efficiency of the GTP exchange was measured by HPLC (46). For this assay, ITD mutant of KRAS (3 μM final concentration) was diluted into 50 mM Tris-HCl pH 7.5, 1 mM DTT, and 1 mM MgCl2 containing 4.5 μM phosphate-binding protein labeled with a phosphate sensor (MDCC, Fisher Scientific) for intrinsic GTPase activity or containing 1 nM NF1-GAP for GAP mediated GTPase activity. Reactions were performed in a black 384 well plate and read in a Tecan M1000 Infinite plate reader at ambient temperature using excitation and emission wavelengths of 430 nm and 550 nm with data collection every 30 sec. All reactions were carried out in quadruplicate. Fluorescence traces were fitted into an exponential associative function, and the rate constants were extracted using Mathematica’s nonlinear regression analysis.

Crystallization, data collection, and structure determination of GDP-bound RAS ITD
To crystallize NRAS ITD and KRAS ITD bound to GDP, we carried out crystallization screenings using the sitting-drop vapor diffusion method using sparse matrix screens. The GDP-bound NRAS ITD was crystallized in 0.1 M Na Citrate pH 5.0, 15% PEG 4000 and 0.1 M MgCl2 whereas crystals of KRAS ITD were observed in crystallization condition containing 200 mM lithium acetate and 2.2 M ammonium sulfate. Crystals were harvested for data collection and cryoprotected with a 30% (v/v) solution of glycerol, before being flash-cooled in liquid nitrogen. The diffraction dataset was collected on 24-ID-C and 24-ID-E beamlines at the Advanced Photon Source (APS), Argonne National Laboratory. Crystallographic datasets were integrated and scaled using XDS (47). Structures of GDP-bound NRAS ITD and KRAS ITD were solved by molecular replacement using the program Phaser with GDP-bound wild-type NRAS (PDB ID: 3CON) and wild-type KRAS (PDB ID: 6MBT) as search models (48). The initial structure was refined using iterative cycles of manual model building using COOT (49) and refinement using Phenix.refine (50). Placement of ligands was followed by identification of potential sites of solvent molecules by the automatic water-picking algorithm in COOT and Phenix.refine. The positions of these automatically picked waters were checked manually during model
building. The data collection and refinement statistics are summarized in Supplementary Table 7. Figures were generated with PyMOL (Schrödinger, LLC).

**Protein Data Bank accession numbers**
The atomic coordinates and structure factors of the GDP-bound NRAS ITD and KRAS ITD have been deposited in the Protein Data Bank and assigned accession numbers 6WGH, and 6PQ3, respectively.

**Study Approval**
The patient and family requested that the patient’s body be donated for cancer research upon death. Full written consent was obtained from both parents prior to patient autopsy and harvesting of the patient’s organs and tumors for research. The University of Minnesota Institutional Review Board reviewed and approved the case and provided a research exemption. These studies and experimental procedures fully abide by the Declaration of Helsinki principles.

**Statistics**
Kaplan-Meier estimate of survival in the 2012 TCGA Colorectal Carcinoma cohort (12) was analyzed on the publicly available cBioPortal web interface (13,14) with the P-value estimated by a log-rank test using default parameters as published. Error bars in all figures represent the data mean +/- SEM. Two-way ANOVAs with Tukey multiple comparison tests were run in GraphPad to report significant differences among nanoBRET measurements.

**Data Availability:**
All data are contained within the manuscript.
Acknowledgements
The authors thank the patient and his parents for donating his remains for the advancement of cancer research, with the hope that this work will be of benefit to future patients. We also thank Dr. Christine Henzler for her guidance regarding bioinformatic analysis, and Michael Franklin for assistance with editorial review of the manuscript. We thank members of the FNLCR Protein Expression Laboratory for help with cloning (Jennifer Mehalko, Vanessa Wall, Dom Esposito), protein expression (John-Paul Denson, Jennifer Mehalko, Nitya Ramakrishnan, Bill Gillette), protein purification (John-Paul Denson, Peter Frank, Jose Sanchez Hernandez, Shelley Perkins), and QC (Matt Drew, Peter Frank). The authors also wish to thank Dr. Simon Messing for the generous gift of PIK3CG(144-1102)-His6. We are thankful to Albert Chan, Timothy Tran, and Timothy Waybright for their help with the structural and biochemical studies. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, and the mention of trade names, commercial products, or organizations does not imply endorsement by the US Government.

Funding and additional information
This research was supported in part by contributions from the patient’s parents, and research funding from the Department of Laboratory Medicine and Pathology, University of Minnesota (A.C.N.); Eastern Star Scholars, University of Minnesota Masonic Cancer Center (A.C.N.); American Cancer Society Clinical Scientist Development Grant (CSDG-18-139-01-CSM; A.C.N.); the Mezin-Koats Colon Cancer Research Award (E.L.); The Randy Shaver Cancer Research and Community Fund (E.L.); the Litman Family Fund for Cancer Research; a Central Society for Clinical and Translational Research Early Career Development Award (E.L.); Institutional Research Grant #118198-IRG-58-001-52-IRG94 from the American Cancer Society (E.L.); Minnesota Masonic Charities; Minnesota Medical Foundation/University of Minnesota Foundation (E.L.); the Masonic Cancer Center and Department of Medicine, Division of Hematology, Oncology and Transplantation, University of Minnesota (E.L.); and the NIH Clinical and Translational Science KL2 Scholar Award 8UL1TR000114 (to E.L.). Structural work is based upon research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P30 GM124165). The Eiger 16M detector on the 24-ID-E beam line is funded by a NIH-ORIP HEI grant (S10OD021527). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. This project was funded in whole or in part with federal funds from the National Cancer Institute, NIH Contract HHSN261200800001E.

Conflict of interest:
The authors declare that they have no conflicts of interest within the contents of this article.
References

1. De Roock, W., De Vriendt, V., Normanno, N., Ciardiello, F., and Tejpar, S. (2011) KRAS, BRAF, PIK3CA, and PTEN mutations: implications for targeted therapies in metastatic colorectal cancer. *The lancet oncology* 12, 594-603

2. Henzler, C., Schomaker, M., Yang, R., Lambert, A. P., LaRue, R., Kincaid, R., Beckman, K., Kemmer, T., Wilson, J., Yohe, S., Thyagarajan, B., and Nelson, A. C. (2018) Optimization of a microfluidics-based next generation sequencing assay for clinical oncology diagnostics. *Ann Transl Med* 6, 162

3. Guedes, J. G., Veiga, I., Rocha, P., Pinto, P., Pinto, C., Pinheiro, M., Peixoto, A., Fragoso, M., Raimundo, A., Ferreira, P., Machado, M., Sousa, N., Lopes, P., Araujo, A., Macedo, J., Alves, F., Coutinho, C., Henrique, R., Santos, L. L., and Teixeira, M. R. (2013) High resolution melting analysis of KRAS, BRAF and PIK3CA in KRAS exon 2 wild-type metastatic colorectal cancer. *BMC Cancer* 13, 169

4. Riihimaki, M., Hemminki, A., Sundquist, J., and Hemminki, K. (2016) Patterns of metastasis in colon and rectal cancer. *Sci Rep* 6, 29765

5. Wojcik, P., Kulig, J., Okon, K., Zazula, M., Mozdzioch, I., Niepsuj, A., and Stachura, J. (2008) KRAS mutation profile in colorectal carcinoma and novel mutation–internal tandem duplication in KRAS. *Pol J Pathol* 59, 93-96

6. Eijkelenboom, A., van Schaik, F. M. A., van Es, R. M., Ten Broek, R. W., Rinne, T., van der Vleuten, C., Flucke, U., Ligttenberg, M. J. L., and Rehmann, H. (2019) Functional characterisation of a novel class of in-frame insertion variants of KRAS and HRAS. *Sci Rep* 9, 8239

7. Rajalingam, K., Schreck, R., Rapp, U. R., and Albert, S. (2007) Ras oncogenes and their downstream targets. *Biochim Biophys Acta* 1773, 1177-1195

8. Li, M. M., Datto, M., Duncavage, E. J., Kulkarni, S., Lindeman, N. I., Roy, S., Tsimberidou, A. M., Vnencka-Jones, C. L., Wolff, D. J., Younes, A., and Nikiforova, M. N. (2017) Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J Mol Diagn* 19, 4-23

9. Reva, B., Antipin, Y., and Sander, C. (2011) Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res* 39, e118

10. Cancer Genome Atlas, N. (2012) Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 487, 330-337

11. Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., Jacobsen, A., Byrne, C. J., Heuer, M. L., Larsson, E., Antipin, Y., Reva, B., Goldberg, A. P., Sander, C., and Schultz, N. (2012) The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2, 401-404
14. Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., Cerami, E., Sander, C., and Schultz, N. (2013) Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 6, pl1

15. Yang, R., Nelson, A. C., Henzler, C., Thyagarajan, B., and Silverstein, K. A. (2015) ScanIndel: a hybrid framework for indel detection via gapped alignment, split reads and de novo assembly. *Genome Med* 7, 127

16. Lawrence, M. S., Stojanov, P., Mermel, C. H., Robinson, J. T., Garraway, L. A., Golub, T. R., Meyerson, M., Gabriel, S. B., Lander, E. S., and Getz, G. (2014) Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* 505, 495-501

17. Bacart, J., Corbel, C., Jockers, R., Bach, S., and Couturier, C. (2008) The BRET technology and its application to screening assays. *Biotechnol J* 3, 311-324

18. Machleidt, T., Woodrofe, C. C., Schwinn, M. K., Mendez, J., Robers, M. B., Zimmerman, K., Otto, P., Daniels, D. L., Kirkland, T. A., and Wood, K. V. (2015) NanoBRET--A Novel BRET Platform for the Analysis of Protein-Protein Interactions. *ACS Chem Biol* 10, 1797-1804

19. Rabara, D., Tran, T. H., Dharmaiah, S., Stephens, R. M., McCormick, F., Simanshu, D. K., and Holderfield, M. (2019) KRAS G13D sensitivity to neurofibromin-mediated GTP hydrolysis. *Proc Natl Acad Sci U S A* 116, 22122-22131

20. Thapar, R., Williams, J. G., and Campbell, S. L. (2004) NMR characterization of full-length farnesylated and non-farnesylated H-Ras and its implications for Raf activation. *J Mol Biol* 343, 1391-1408

21. Pacold, M. E., Suire, S., Perisic, O., Lara-Gonzalez, S., Davis, C. T., Walker, E. H., Hawkins, P. T., Stephens, L., Eccleston, J. F., and Williams, R. L. (2000) Crystal structure and functional analysis of Ras binding to its effector phosphoinositide 3-kinase gamma. *Cell* 103, 931-943

22. Castellano, E., and Santos, E. (2011) Functional specificity of ras isoforms: so similar but so different. *Genes Cancer* 2, 216-231

23. Escudero-Esparza, A., Bartoschek, M., Gialeli, C., Okroj, M., Owen, S., Jirstrom, K., Orimo, A., Jiang, W. G., Pietras, K., and Blom, A. M. (2016) Complement inhibitor CSMD1 acts as tumor suppressor in human breast cancer. *Oncotarget* 7, 76920-76933

24. Farrell, C., Crimm, H., Meeh, P., Croshaw, R., Barbar, T., Vandersteenhouven, J. J., Butler, W., and Buckhaults, P. (2008) Somatic mutations to CSMD1 in colorectal adenocarcinomas. *Cancer Biol Ther* 7, 609-613

25. Kamal, M., Holliday, D. L., Morrison, E. E., Speirs, V., Toomes, C., and Bell, S. M. (2017) Loss of CSMD1 expression disrupts mammary duct formation while enhancing proliferation, migration and invasion. *Oncol Rep* 38, 283-292

26. Ma, C., Quesnelle, K. M., Sparano, A., Rao, S., Park, M. S., Cohen, M. A., Wang, Y., Samanta, M., Kumar, M. S., Aziz, M. U., Naylor, T. L., Weber, B. L., Fakharzadeh, S. S., Weinstein, G. S., Vachani, A., Feldman, M. D., and Brose, M. S. (2009) Characterization CSMD1 in a large set of primary lung, head and neck, breast and skin cancer tissues. *Cancer Biol Ther* 8, 907-916

27. Shull, A. Y., Clendenning, M. L., Ghoshal-Gupta, S., Farrell, C. L., Vangapandu, H. V., Dudas, L., Wilkerson, B. J., and Buckhaults, P. J. (2013) Somatic mutations, allele loss, and DNA methylation of the Cub and Sushi Multiple Domains 1 (CSMD1) gene reveals association with early age of diagnosis in colorectal cancer patients. *PLoS One* 8, e58731
28. Rodriguez-Escudero, I., Oliver, M. D., Andres-Pons, A., Molina, M., Cid, V. J., and Pulido, R. (2011) A comprehensive functional analysis of PTEN mutations: implications in tumor- and autism-related syndromes. Human molecular genetics 20, 4132-4142
29. Martin, T. A., Watkins, G., Mansel, R. E., and Jiang, W. G. (2004) Loss of tight junction plaque molecules in breast cancer tissues is associated with a poor prognosis in patients with breast cancer. Eur J Cancer 40, 2717-2725
30. Makohon-Moore, A. P., Zhang, M., Reiter, J. G., Bozie, I., Allen, B., Kundu, D., Chatterjee, K., Wong, F., Jiao, Y., Kohutek, Z. A., Hong, J., Attiyeh, M., Javier, B., Wood, L. D., Hruban, R. H., Nowak, M. A., Papadopoulos, N., Kinzler, K. W., Vogelstein, B., and Iacobuzio-Donahue, C. A. (2017) Limited heterogeneity of known driver gene mutations among the metastases of individual patients with pancreatic cancer. Nat Genet 49, 358-366
31. Hsu, H. C., Thiam, T. K., Lu, Y. J., Yeh, C. Y., Tsai, W. S., You, J. F., Hung, H. Y., Tsai, C. N., Hsu, A., Chen, H. C., Chen, S. J., and Yang, T. S. (2016) Mutations of KRAS/NRAS/BRAF predict cetuximab resistance in metastatic colorectal cancer patients. Oncotarget 7, 22257-22270
32. Douville, C., Carter, H., Kim, R., Niknafs, N., Diekhans, M., Stenson, P. D., Cooper, D. N., Ryan, M., and Karchin, R. (2013) CRAVAT: cancer-related analysis of variants toolkit. Bioinformatics 29, 647-648
33. Masica, D. L., Douville, C., Tokheim, C., Bhattacharya, R., Kim, R., Moad, K., Ryan, M. C., and Karchin, R. (2017) CRAVAT 4: Cancer-Related Analysis of Variants Toolkit. Cancer Res 77, e35-e38
34. Yang, H., and Wang, K. (2015) Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. Nat Protoc 10, 1556-1566
35. Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., and Ruden, D. M. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 6, 80-92
36. Cingolani, P., Patel, V. M., Coon, M., Nguyen, T., Land, S. J., Ruden, D. M., and Lu, X. (2012) Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. Front Genet 3, 35
37. Chen, J., Bardes, E. E., Aronow, B. J., and Jegga, A. G. (2009) ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic Acids Res 37, W305-311
38. Mi, H., Muruganujan, A., and Thomas, P. D. (2013) PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic Acids Res 41, D377-386
39. Xie, Q., Soutto, M., Xu, X., Zhang, Y., and Johnson, C. H. (2011) Bioluminescence resonance energy transfer (BRET) imaging in plant seedlings and mammalian cells. Methods Mol Biol 680, 3-28
40. Dharmaiah, S., Tran, T. H., Messing, S., Agamasu, C., Gillette, W. K., Yan, W., Waybright, T., Alexander, P., Esposito, D., Nisley, D. V., McCormick, F., Stephen, A. G., and Simanshu, D. K. (2019) Structures of N-terminally processed KRAS provide insight into the role of N-acetylation. Sci Rep 9, 10512
41. Taylor, T., Denson, J. P., and Esposito, D. (2017) Optimizing Expression and Solubility of Proteins in E. coli Using Modified Media and Induction Parameters. *Methods Mol Biol* **1586**, 65-82

42. Weinstock, M. T., Hesek, E. D., Wilson, C. M., and Gibson, D. G. (2016) Vibrio natriegens as a fast-growing host for molecular biology. *Nat Methods* **13**, 849-851

43. Kopra, K., Vuorinen, E., Abreu-Blanco, M., Wang, Q., Eskonen, V., Gillette, W., Pulliainen, A. T., Holderfield, M., and Harma, H. (2020) Homogeneous Dual-Parametric-Coupled Assay for Simultaneous Nucleotide Exchange and KRAS/RAF-RBD Interaction Monitoring. *Anal Chem* **92**, 4971-4979

44. Wall, V. E., Garvey, L. A., Mehalko, J. L., Procter, L. V., and Esposito, D. (2014) Combinatorial assembly of clone libraries using site-specific recombination. *Methods Mol Biol* **1116**, 193-208

45. Shutes, A., and Der, C. J. (2006) Real-time in vitro measurement of intrinsic and Ras GAP-mediated GTP hydrolysis. *Methods Enzymol* **407**, 9-22

46. Agamasu, C., Ghirlando, R., Taylor, T., Messing, S., Tran, T. H., Bindu, L., Tonelli, M., Nissley, D. V., McCormick, F., and Stephen, A. G. (2019) KRAS Prenylation Is Required for Bivalent Binding with Calmodulin in a Nucleotide-Independent Manner. *Biophys J* **116**, 1049-1063

47. Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* **66**, 133-144

48. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674

49. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501

50. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221
Figures and Figure Legends

A

| Samples            | Potentially Relevant Variants | Prioritized Variants |
|--------------------|------------------------------|----------------------|
| Primary tumor      | 847                          | 169                  |
| Lung metastasis    | 156*                         | 58*                  |
| CNS metastasis     | 393*                         | 72*                  |

B

| Gene  | Mutation    | Mutation type           | Interpretation | Site       |
|-------|-------------|-------------------------|----------------|------------|
| NRAS  | p.I55_Y64dup| Internal tandem duplication | Pathogenic   | All        |
| TP53  | p.C275F    | Missense                | Pathogenic    | All        |
| CSMD1 | c.9100+1G>T | Splicing                | Pathogenic    | All        |
| PTEN  | p.P169H    | Missense                | VUS           | Lung, CNS  |
| MPDZ  | p.L459I    | Missense                | VUS           | Lung, CNS  |

C

Figure 1: Whole exome sequencing of a clinically aggressive CRC suggests the NRAS ITD is the primary oncogenic driver mutation.
A) Overview of filtered variants prioritized for molecular pathology review in whole exome NGS of the patient’s primary tumor, lung, and CNS metastases; * indicates variants unique to the metastases vs. the primary tumor.

B) Molecular pathologist interpretation of clinically relevant variants, including tumor site(s) at which each variant was identified.

C) Overall survival of patients with NRAS or CSMD1 mutations (red, N=48, median OS 38.9 months) was significantly (p=0.002) shorter than patients without mutations in these genes (blue, N=162) in the CRC TCGA cohort.
Figure 2: Switch II Internal Tandem Duplication of both NRAS and KRAS proteins increases the affinity for RAF effector. NanoBRET saturation curves of transiently transfected RAF1 NanoLuc donor constructs (constant) and titrated, transiently transfected Halotag-RAS acceptor constructs. BRETmax and BRET50 values represent the maximum number of protein-protein interactions and protein affinity, respectively. A (top) Halotag-NRAS ITD fusion proteins show increased interaction with effector RAF1 compared to the wildtype protein when overexpressed in Caco-2 cells. *Interactions with NRAS wildtype to be significantly different to both NRAS Q61R and NRAS ITD (p<0.0005 from 0.0625-4µg DNA; p<0.02 at 8µg DNA). A (bottom) Both KRAS ITD and NRAS ITD mutants display increased RAF1 interaction compared to KRAS wildtype in HEK293T cells. *NRAS ITD mBRET values are significantly different to KRAS WT (p<0.0001 from 0.0094-0.3µg DNA, p<0.008 at 0.6 µg DNA) **KRAS ITD mBRET values are significantly different to KRAS WT (p<0.0001 from 0.0094-0.075µg DNA, p<0.03 at 0.15µg DNA) B) nanoBRET saturation curves of NRAS ITD and NRAS wildtype in HEK293T cells. *RAF1 interactions are significantly different (p<0.0005) between 0.00195-0.00781µg DNA C) nanoBRET saturation curves of KRAS ITD and KRAS wildtype. *RAF1 interactions are significantly different (p<0.0001) between 0.00195-0.25µg DNA. Nonlinear regressions were performed in GraphPad Prism (see Supplementary Tables 2-5 for curve fit parameters). Error bars represent n=3 technical replicates. Statistical significance of the differences was determined by 2-way ANOVA in GraphPad Prism.
Figure 3. Expression of ITD mutant RAS leads to MEK and ERK activation. A, B) Halotag-NRAS ITD fusion proteins over-activate the MAP Kinase cascade when transiently overexpressed in titrated amounts in (A) Caco-2 cells and (B) HEK293T cells. Pathway induction, as measured by phospho-ERK and phospho-MEK using Western blots, is comparable to oncogenic NRAS Q61R constructs, and less than NRAS wildtype. KRAS ITD constructs are similarly able to induce pathway overactivation when overexpressed in HEK293T cells C) at a greater level than KRAS4b wildtype and D) at similar levels to KRAS G12D.
Figure 4: The ITD mutation in RAS does not affect RAS-RAF1 interaction but blocks RAS-RasGAP binding. 

A, B) Isothermal titration calorimetry experiments to measure the dissociation constant for GMPPNP-bound WT NRAS and NRAS ITD with (A) RasGAP NF1 (GRD) and (B) effector RAF1 (RBD). Differential power (DP) is a measure of energy required to maintain isothermal conditions between the reference cell and the sample cell. C) mBRET saturation values assessing the ability of NF1 GAP co-expression to squelch RAF-RAS interaction for wild-type and ITD mutant KRAS and NRAS.
Figure 5: Crystal Structure of GDP-bound NRAS ITD and KRAS ITD provide insights into the effect of ITD on RAS interaction with GAP and effector proteins. A) The tertiary structure of GDP-bound NRAS ITD. B) The tertiary structure of GDP-bound KRAS ITD. C, D) Structural superposition of GDP-bound NRAS ITD with (C) GDP-bound WT KRAS and (D) GMPPNP-bound WT NRAS. D, E, F) Models of ITD mutants of NRAS and KRAS in complex with (E) NF1-GRD (PDB ID: 6OB2), (F) PI3Kγ (PDB ID: 1HE8), and (G) RAF1-RBD (PDB ID: 4G0N) generated using the structural superposition of NRAS ITD on K/HRAS present in KRAS-NF1, HRAS-PI3Kγ, and HRAS-RAF1 (RBD) complexes. These models suggest that the ITD (shown in red) in RAS would sterically clash with NF1 GAP and PI3Kγ and not with RAF1.
RAS internal tandem duplication disrupts GTPase-activating protein (GAP) binding to activate oncogenic signaling
Andrew C. Nelson, Thomas J. Turbyville, Srisathiyarayanan Dharmaiah, Megan Rigby, Rendong Yang, Ting-You Wang, John Columbus, Robert Stephens, Troy Taylor, Drew Sciacca, Getiria Onsongo, Anne Sarver, Subbaya Subramanian, Dwight V. Nissley, Dhirendra K. Simanshu and Emil Lou

J. Biol. Chem. published online May 11, 2020

Access the most updated version of this article at doi: 10.1074/jbc.RA119.011080

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts