SI Appendix

Materials. DLD-1, HCT116, and SW48 cell lines were obtained from Horizon Discovery and maintained at 37 °C and 5% CO_2 in RPMI 1640 supplemented with 10% FBS, 2 mM L-glutamine, 25 mM sodium bicarbonate, and 5% penicillin-streptomycin per manufacturer's instructions. COR-L23 and H1792 cell lines were provided by the Frederick National Laboratory for Cancer Research (FNLCR) RAS Initiative program as frozen cell pellets. Cell lines were authenticated by STR profiling, both performed by the manufacturer and confirmed in-house at time of purchase using ATCC guidelines. Primary colorectal tumor samples were obtained from the NCI Biospecimen Core Resource with prior Institutional Review Board approval. Fully processed, recombinant KRAS4b protein standards were provided by the FNLCR RAS Initiative program. A rabbit monoclonal antibody (EP1125Y) reactive with HRAS, NRAS and KRAS was purchased from Abcam and used for immunoblotting, along with an HRP-conjugated goat anti-rabbit IgG purchased from Millipore Sigma. Immunoprecipitation (IP) was initially performed using an anti-RAS rat monoclonal antibody (Y13-259) agarose conjugate purchased from Millipore Sigma, and later using agarose beads (Thermo Fisher Scientific) conjugated in-house to the Abcam monoclonal antibody using the dimethyl pimelimidate (DMP) crosslinking protocol described in (1). MS solvents were purchased from Fisher Scientific. All other reagents were purchased from Sigma-Aldrich, unless otherwise noted.

Preparation of solutions. Lysis buffer comprised 50 mM Tris, pH 7.5, containing 150 mM NaCl, 1% Nonidet P-40, two Complete mini EDTA-free Protease Inhibitor Cocktail tablets (Roche) and one PhosSTOP Phosphatase Inhibitor Cocktail tablet, or with Pierce Halt Protease and Phosphatase Inhibitor Cocktail (Thermo Fisher Scientific). Wash Buffer 1 comprised 50 mM Tris, pH 7.5, containing 150 mM NaCl, 1% Nonidet P-40, one Complete mini EDTA-free Protease Inhibitor Cocktail tablet and one PhosSTOP Phosphatase Inhibitor Cocktail tablet, or with Pierce Halt Protease and Phosphatase Inhibitor Cocktail. Wash Buffer 2 comprised 50 mM Tris, pH 7.5, containing 150 mM NaCl and no protease or phosphatase inhibitors. Elution Buffer comprised 0.2% (vol/vol) trifluoroacetic acid (TFA, HPLC-grade, Sigma Aldrich) in water. All solutions were prepared fresh and chilled on ice before each use.

Immunoblot analysis. Protein samples were resolved by SDS-PAGE on Any kD Mini-PROTEAN TGX Precast Protein Gels (BioRad) and transferred to PVDF membranes. Membranes were blocked with 5% dehydrated milk in TBS + 0.1% (vol/vol) Tween-20 (TBS-T) prior to incubation with antibodies diluted in TBS-T. Total RAS was visualized with anti-RAS rabbit monoclonal primary antibody (EP1125Y) at 1:1,000 dilution. Blots were then incubated with HRP-conjugated goat anti-rabbit secondary antibody at 1:2,000 dilution, followed by incubation with Luminata Crescendo Western HRP Substrate (Millipore) before imaging and quantification of bands of interest using the GelDoc XR imaging system (BioRad).

Cell lysis. Cells were detached from culture flasks by incubating with Versene (Thermo Fisher Scientific) for 15 minutes at 37 °C, counted via Cellometer (Nexcelom) according to the manufacturer's protocol, centrifuged at 300 ×*g* at 4 °C for 5 minutes, and washed with PBS to remove residual dissociation reagent. Cell pellets could be stored at -80 °C or subjected to lysis

immediately. Cell pellets were incubated with lysis buffer (600 µL per 2e6 cells) for 30 minutes on ice, with vortexing or resuspension every 10 minutes, and then sonicated for 30 seconds (amplitude 50%, 2 seconds on, 1 second off) with a Sonic Dismembrator (Fisher Scientific). Tissue samples were first cryopulverized using a Retsch Mixer Mill MM400 (frequency 30 Hz, 2 minutes) twice, keeping samples immersed in liquid nitrogen in between milling sessions. Lysis buffer (in sufficient volumes to resuspend all material, which varied from 1-12 mL) was added to the frozen tissue powder, and the same lysis procedure performed as described above. The lysates from cells and tissues were clarified by centrifugation at 12,000 ×*g* for 30 minutes at 4 °C. Lysates were stored at -80 °C or immediately subjected to RAS immunoprecipitation. A bicinchoninic acid (BCA) assay kit (Thermo Fisher) was used to quantify protein concentrations.

Optimization of RAS protein immunoprecipitation. The IP method optimization used cultured colorectal cancer cell lines and was based on two criteria: RAS protein amount isolated and compatibility with TD LC-MS analysis. Different lysis buffers were evaluated, resulting in the observation that a detergent-containing buffer was necessary for raising the yield of RAS and resulting in high quality TD spectra using LC-MS. RAS proteins were eluted using 0.2% TFA, which facilitated subsequent desalting prior to downstream MS analysis. RAS enrichment levels were visualized by immunoblot and quantified by densitometric analysis of the RAS-specific bands, demonstrating that the eluted RAS proteins were concentrated by more than 10-fold in the elution step when compared to the lysate used as input (Fig. 1*B*). With the optimized IP procedure, KRAS, NRAS and HRAS could be detected, with LC-MS measurements further optimized for precise detection of KRAS proteoforms with good signals in intact (MS1) and fragmentation (MS2) spectra. Consistent with transcriptomic data and previous studies on cancer cell lines (2), KRAS4a is only a small fraction of total KRAS (<10%) relative to the 4b isoform and was not detected in the TD assay.

Cell lysates containing 1-8 mg of total input protein (in the cell extract obtained from the lysis step described above) were thawed on ice and diluted with lysis buffer to a final volume of 900 µL. Antibody-conjugated agarose (100 µL slurry) was added and the suspension incubated at 4 °C overnight (minimum of 12 hours) with constant rotation. The agarose beads were collected in a Mobicol (MoBiTec) connected to a vacuum manifold or by centrifugation for 5 min. at 800 rpm and 4˚C. The collected beads were then washed three times with Wash Buffer 1 (5 mL) and three times with Wash Buffer 2 (5 mL). The RAS proteins were eluted off the beads by incubation with 200 µL Elution Buffer for 30 minutes at 25 °C and 650 rpm on a Thermomixer (Eppendorf). The elution step was repeated twice and eluent from both steps was combined prior to desalting using a C4 ZipTip (Millipore) according to manufacturer instructions.

***In vitro* nitrosylation of KRAS4b.** *In vitro* nitrosylation was performed as previously described (3) with minor modifications. Five micrograms of fully processed, recombinant KRAS4b (4) were mixed with a 100-molar excess of *S*-nitrosoglutathione (GS-NO) in HEN buffer (250 mM HEPES-NaOH, pH 7.7, 1 mM EDTA, 0.1 mM neocuproine). The reaction was carried out at room temperature for 30 minutes in the dark. The nitrosylated protein was desalted using a C4 ZipTip (Millipore) prior to MS analysis. To reverse nitrosylation, nitrosylated protein was incubated with 50 mM DTT at room temperature for 15 minutes, followed by C4 ZipTip desalting prior to MS analysis.

Optimization of LC-MS/MS parameters. To provide sensitivity for the low quantities of purified KRAS4b obtained after IP from cells or non-renewable clinical samples, we used nanocapillary LC separation connected to a Q Exactive HF mass spectrometer (5, 6). To chromatographically resolve isolated RAS proteins, samples (5 µL) were injected onto a trap column (150 µm ID × 3 cm) using an autosampler (Thermo Fisher Scientific). A nanobore analytical column (75 µm ID × 15 cm) was coupled to the trap in a vented tee setup, with a 15 µm spray tip from New Objective connected downstream. The trap, analytical column, and spray tip were packed with polymeric reverse phase (PLRP-S, Agilent) media (5 µm d_p, 1,000 Å pore size). The Dionex Ultimate 3000 LC system was operated at a flow rate of 2.5 µL/minute for loading samples onto the trap. Proteins were separated on the analytical column and eluted into the mass spectrometer using a flow rate of 300 nL/minute and the following gradient: 5% B at 0 minutes; 15% B at 5 minutes; 50% B at 25 minutes; 95% B from 28-31 minutes; 5% B from 34 to 50 minutes. Solvent A consisted of 94.8% water, 5% acetonitrile, and 0.2% formic acid, while Solvent B consisted of 5% water, 94.8% acetonitrile, and 0.2% formic acid.

The mass spectrum of a single purified proteoform can appear complex due to the presence of multiple charge states (e.g., Fig. 1*C*). To simplify analysis and establish a robust quantitative assay, we targeted individual charge states and studied their fragmentation behavior as explained in the main text and Fig. S2. We developed a selected ion monitoring (SIM) method for three of the most abundant charge states (23+ 24+, and 25+) and one less abundant with somewhat lower charge (20+). By selecting a narrow window around charge states of KRAS4b, we detected lower amounts of the protein with SIM than when scanning the full range of mass-to-charge ratios (*m/z*). Precursors isolated using the resolving quadrupole can be fragmented in a targeted fashion (tMS2) to provide diagnostic fragment ions, which can also be used to parse and quantify the relative abundance of proteoforms and their precursors (7). MS parameters were adapted from previous work (5); briefly, the instrument was operated in "protein mode", i.e., with reduced extended trapping of ions in the HCD cell wherein the N_2 gas pressure was reduced to ~2 mbar. Other parameters were set as follows: heated transfer capillary temperature of 320 °C, S-lens RF amplitude of 50%, and 15 V in-source dissociation to favor adduct removal and further desolvation.

Acquisition parameters varied depending on scan type. Full MS scans were acquired within a 500−2000 *m/z* window, averaging 4 microscans per scan, applying a resolving power (r.p.) of 120,000 (at 200 *m/z*), with a target value for the automatic gain control (AGC) of 1e6 charges, and a maximum injection time of 50 milliseconds. Selected ion monitoring (SIM) scans, with width set to 10 *m/z* units, were recorded averaging 4 microscans, with r.p. of 120,000 (at 200 *m/z*), an AGC target of 5e4 and a maximum injection time of 400 milliseconds. Tandem MS (MS2) was performed via higher-energy collision dissociation (HCD) by using different energies in order to optimize KRAS4b fragmentation (i.e., to maximize sequence coverage at higher collisional energies and retain labile PTMs at lower energies). Applied normalized collision energy (NCE) varied within the 19-25% range, in steps of 2%, with the "stepped-collision energy" option enabled. MS2 was triggered under either data-dependent acquisition mode, or alternatively, under tMS2 mode by adding species of interest to an inclusion list. Precursor selection was obtained through the resolving quadrupole with either a wide isolation window of 5-8 *m/z* (to relatively quantify the proteoforms of different alleles) or a narrow one of 3 *m/z* (for determining

the stoichiometry of methylation within proteoforms derived from the same allele). All MS2 scans were collected with an r.p. of 60,000 (at 200 *m/z*) and minimum *m/z* value set to 400, an AGC target of 1e6, and a maximum injection time of 800 milliseconds. Data were processed and analyzed using ProSightPC 4.0, ProSight Lite (8, 9) and QualBrowser/FreeStyle, part of the Xcalibur software packaged provided with the Q-Exactive HF. Quantitation of KRAS4b used the same methods reported for MS1 and MS2 signals in Pesavento, *et al.* (10). In brief, the protein ion relative ratios (PIRRs) from a particular proteoform are used to determine the ratios of proteoforms present. When MS2 scans are used, the fragmentation ion relative ratios (FIRRs) faithfully report the ratios of different modified forms present in the sample when several are isolated together prior to ion fragmentation.

Recombinant GG-KRAS4b WT control study. To determine if the IP-TDMS workflow contributes artefactual changes in observed PTM levels (i.e., differential COOMe levels observed in certain samples), 5 µg of the recombinant, fully processed GG-KRAS4b WT FMe standard was spiked into a DLD-1 MUT whole cell lysate after lysis, sonication, and clarification by centrifugation. Following addition of the recombinant standard, the IP was performed using the exact protocol described above, and the resulting purified material was injected in triplicate and analyzed with the targeted MS method. Diagnostic y_{12} ions retaining (469.306 *m/z* for 3^+) or lacking the methyl group (464.630 *m/z* for 3^+) were used for quantitative comparisons.

Supplemental Results

IP-MS results from H1792, COR-L23, and SW48 (G12D/+) cell lines. H1792 are a heterozygous *KRAS* (G12C/+) lung carcinoma cell line, COR-L23 are a homozygous *KRAS* (G12V/G12V) lung carcinoma cell line, and the isogenic SW48 (G12D/+) colorectal cell line was derived by introducing a single G12D KRAS gene into the parental SW48 KRAS (+/+) cell line (similar to SW48 G13D Mut cells in main text). In all three cell lines, the predominant KRAS4b proteoform was the N-terminally acetylated, Cys185 farnesylated, and C-terminally carboxymethylated species for both wild type and mutant KRAS4b. Within the context of COR-L23 cells, only the mutant KRAS4b G12V protein was detected as calculated from the relative abundances of the diagnostic b_{53} fragment ions (5862.01 Da for the WT b_{53} ion, 5904.02 Da for the mutant G12V b_{53} ion) (Table S1). For the H1792 cells, the relative amounts of wild type and mutant G12C KRAS4b expressed were not equal; rather, 79% of the detected KRAS4b was the mutant form. This is in stark contrast to the colorectal cell line data, where WT and mutant KRAS4b proteoforms were detected in equal amounts within the context of WT/G13D cells (e.g. DLD-1 Par) and G12D/WT cells (SW48). For this data acquisition, a narrower isolation window of 3 *m/z* was utilized to specifically isolate each wild type and mutant proteoform individually, allowing for more unambiguous assignment of the relative amounts of C-terminal COOMe to the appropriate precursor species. With this method, we were able to determine that the relative amount of C-terminal COOMe were roughly equivalent for both wild type and mutant KRAS4b in the three cell lines analyzed (Table S1).

Recombinant GG-KRAS4b WT control experiment. Regarding the C-terminal COOMe, a control experiment performed by spiking recombinant, fully processed GG-KRAS4b WT standard into DLD-1 MUT whole cell lysates resulted in no significant loss of C-terminal COOMe (i.e., 3% or lower) in comparison to fresh recombinant standard (Table S4). This demonstrated that the significantly altered COOMe levels observed in clinical tumors and in G12 mutant cell lines were not artefacts of the IP-TDMS assay itself.

IP-MS results from additional primary tumor samples. As the first four tumor samples (Subjects 1-4) showed variable KRAS4b C-terminal COOMe patterns, the optimized IP-TDMS method with a narrower isolation window of 3 *m/*z was utilized to determine with high precision the relative amount of this dynamic PTM on wild type and mutant KRAS4b within two additional tumor samples. Of these, one was heterozygous *KRAS* (G12V/WT) (Subject 5), while the other was homozygous *KRAS* (WT/WT) (Subject 6). As with the initial tumor samples, the majority of the KRAS4b protein in both tumors was detected as the fully processed, N-terminally acetylated, C-terminally carboxymethylated, and Cys185 farnesylated proteoform for both wild type and mutant G12V KRAS4b (Fig. S6 and S7). Consistent with the data from the first two mutant *KRAS* tumors, KRAS4b G12V comprised 18% of the total amount of KRAS4b protein in these samples. With the narrower isolation window, complete assignment of the relative amount of C-terminal COOMe was possible for each wild type and mutant G12V proteoform from Subject 5, where 49% of WT KRAS4b and 9% of G12V KRAS4b retained the modification, respectively (Table S3).

Supplemental Discussion

KRAS4b proteoforms from H1792, COR-L23, and SW48 (G12D/+) cell lines. Despite the different genetic and tissue backgrounds, the predominant KRAS4b proteoforms identified in the COR-L23 and H1792 lung carcinoma cell lines were the N-terminally acetylated, C-terminally carboxymethylated, and Cys185 farnesylated forms of both WT and mutated (either G12V or G12C) KRAS4b (Fig. S6). For the COR-L23 cell line, the only KRAS4b proteoforms detected were those consistent with the G12V genotype and with the PTMs mentioned above. However, unlike the heterozygous colorectal cancer cell lines analyzed in the main text (DLD-1 Par, HCT116 Par, and SW48 (G13D/WT)), the H1792 lung cancer cell line did not have equal amounts of wild type and mutant KRAS4b G12C expression levels. The relatively high expression level of the KRAS4b G12C proteoform may be due to G12C being a weakly activating mutation of KRAS when compared to other KRAS mutations (11). Notably, both cell lines displayed drastically different C-terminal COOMe levels compared to the colorectal cancer cell lines analyzed in the main text (compare Table S1 to Table 1), while the overwhelming majority of KRAS4b proteoforms detected in the colorectal SW48 (G12D/WT) cell line were C-terminally carboxymethylated. While the lung cell lines reflect more closely the C-terminal COOMe status in the primary human colorectal tumors, further studies within the context of the *in vitro* model systems will be required to establish the link between the expression levels of C-terminal COOMe for both WT and mutant forms of both KRAS4b and KRAS4a and the impact on KRAS4b activity, function, and localization.

Differential C-terminal COOMe of KRAS4b proteoforms in colorectal tumor from Subject 5. A striking observation with the additional colorectal tumors analyzed with the narrower isolation window was that a small proportion (9%) of the mutant KRAS4b G12V proteoform in Subject 5 had the C-terminal COOMe, while 49% of the wild-type KRAS4b had this modification (Table S3). Therefore, the G12V proteoform in this tumor sample appears to have been selectively modified to remove the C-terminal COOMe *in vivo*. This observation underscores the importance of the proteoform centric approach described here and raises new questions regarding differential modulation of oncogenic KRAS alleles via post-translational modification.

Screening for KRAS4a proteoforms within cell lines and tumors. The KRAS4b MS method was modified to measure the theoretical *m/z* precursor charge states of fully processed KRAS4a (with C-terminal methylation, Cys180 palmitoylation, and Cys186 farnesylation, with a 10 *m/z* window centered on 942.36 *m/z* for the 23+ charge state) and for partially processed KRAS4a (with C-terminal methylation and Cys186 farnesylation, with a 10 *m/z* window centered on 932.01 *m/z* for the 23+ charge state), as the palmitoylation modification is linked to KRAS4a with a thioester that is possibly lost during sample preparation. Interrogation of these *m/z* windows during analysis of the IP elutions generated from both cell lines and tumor samples did not result in the unambiguous and MS2-verified detection of KRAS4a proteoforms. We therefore concluded that the relative abundance of KRAS4a within our starting whole-cell lysates was not sufficient to reach the limit of detection for our optimized IP-TDMS workflow at this time. Further studies, particularly when KRAS4a-specific antibodies become available, will be required in order to explore the KRAS4a proteoform space and determine what complementary interactions might exist between KRAS4a and KRAS4b proteoforms within the context of colorectal and other tumors.

Complementarity of TD and BU strategies for RAS protein analysis (Fig. S8). The 4 RAS proteins (KRAS4a, KRAS4b, HRAS, and NRAS) share >80% total amino acid sequence identity, and 100% sequence identity within the first 86 amino acids. Consequently, BU proteomic analysis for specific characterization of KRAS proteins is complicated by the many identical tryptic peptides that map to one or more *RAS* isoforms (Fig. S8*B*). In a previous analysis of the DLD-1 isogenic cell lines by targeted, quantitative BU proteomics to specifically measure the relative amounts of mutant and wild-type KRAS proteins (Fig. S8*B*), a specific tryptic peptide comprising amino acids Gln150 to Arg161 found in HRAS and NRAS (but not KRAS4a or KRAS4b) was used to determine the amount of the tryptic peptide spanning amino acids Leu6 to Lys16 (with Gly13, or “WT”) produced from HRAS and NRAS (12). Quantitation of the same peptide sequence containing the Gly13Asp substitution (“G13D”), which could only be present within KRAS, allowed the stoichiometry of the mutated form of KRAS to be determined within the total RAS population (reproduced here in Fig. S8*B*). Interestingly, the derived isogenic cell lines upregulated the expression of the single remaining *KRAS* gene to generate total KRAS protein levels comparable to those present in the DLD-1 PAR cells. However, while quantitative BU proteomics can be used to precisely measure the amount of mutant and wild type protein expressed in these cell lines, this approach cannot determine the relative amounts of KRAS4a and KRAS4b expression. In contrast, TD proteomics can make this determination to precisely identify and quantify which isoforms and proteoforms are present within a KRAS population (Fig. S8*C*).

The TD approach utilized here focused on KRAS4b proteoforms, as the KRAS4a isoform is expressed at levels that are currently below the detection threshold of this assay. The strength of the TD approach is in the detection of unique KRAS4b proteoforms, namely the nitrosylated KRAS4b proteoforms only found in the wild type isogenic cell lines (Fig. 2) and the KRAS4b proteoforms with loss of C-terminal methylation in the cell lines and tumor samples (Fig. 3, Fig. S6, and Fig. S7). Note that these specific proteoform assignments could not have been discovered with BU proteomics, suggesting the complementary use of the two technologies (Fig. S8*D*).

As TD and BU protein analyses provide different yet complementary information about peptides and proteoforms, both approaches could be combined to provide deeper insight into cancer biology (Fig. S8*D*). For example, the BU data reproduced here in Fig. S8*B* from Hutton *et al.* 2016 (12) indicate significant changes in KRAS protein expression levels as a result of single *KRAS* allele knockout, but do not address isoform specificity. Improvements in the TD assay described here could address this limitation and examine whether an increase in wild-type and mutant KRAS in DLD-1 WT and DLD-1 MUT cells, respectively, is due to the upregulation of KRAS4a and/or KRAS4b isoforms. Furthermore, while the TD assay only measured KRAS4b nitrosylation, a targeted, quantitative BU proteomics study could be deployed to monitor Cys118 nitrosylation on KRAS4a and KRAS4b simultaneously (as the tryptic peptide containing Cys118 is shared between the two KRAS isoforms, but is unique for both HRAS and NRAS). Although this BU approach would obviate the proteoform specificity afforded by TD proteomics, it would provide highly specific measurements of KRAS nitrosylation, and thus generate new hypotheses for additional proteoform-resolved assays using primary patient tumor samples. The combined new qualitative and quantitative information about RAS regulation through post-translational modification could not only identify key roles of individual RAS proteoforms in colorectal cancer but also potentially reveal new prognostic indicators or therapeutic targets.

References

1. DeHart CJ, Chahal JS, Flint SJ, & Perlman DH (2014) Extensive Post-translational Modification of Active and Inactivated Forms of Endogenous p53. *Molecular & Cellular Proteomics* 13(1):1-17.
2. Tsai FD*, et al.* (2015) K-Ras4A splice variant is widely expressed in cancer and uses a hybrid membrane-targeting motif. *Proc Natl Acad Sci U S A* 112(3):779-784.
3. Wang Y, Liu T, Wu C, & Li H (2008) A strategy for direct identification of protein S-nitrosylation sites by quadrupole time-of-flight mass spectrometry. *J Am Soc Mass Spectrom* 19(9):1353-1360.
4. Gillette WK*, et al.* (2015) Farnesylated and methylated KRAS4b: high yield production of protein suitable for biophysical studies of prenylated protein-lipid interactions. *Sci Rep* 5:15916.
5. Fornelli L*, et al.* (2017) Advancing Top-down Analysis of the Human Proteome Using a Benchtop Quadrupole-Orbitrap Mass Spectrometer. *J Proteome Res* 16(2):609-618.
6. Scheltema RA*, et al.* (2014) The Q Exactive HF, a Benchtop mass spectrometer with a pre-filter, high-performance quadrupole and an ultra-high-field Orbitrap analyzer. *Mol Cell Proteomics* 13(12):3698-3708.
7. Toby TK, Fornelli L, & Kelleher NL (2016) Progress in Top-Down Proteomics and the Analysis of Proteoforms. *Annu Rev Anal Chem (Palo Alto Calif)* 9(1):499-519.

8. DeHart CJ, Fellers RT, Fornelli L, Kelleher NL, & Thomas PM (2017) Bioinformatics Analysis of Top-Down Mass Spectrometry Data with ProSight Lite. *Methods Mol Biol* 1558:381-394.
9. Fellers RT*, et al.* (2015) ProSight Lite: graphical software to analyze top-down mass spectrometry data. *Proteomics* 15(7):1235-1238.
10. Pesavento JJ, Mizzen CA, & Kelleher NL (2006) Quantitative analysis of modified proteins and their positional isomers by tandem mass spectrometry: human histone H4. *Anal Chem* 78(13):4271-4280.
11. Miller MS & Miller LD (2011) RAS Mutations and Oncogenesis: Not all RAS Mutations are Created Equally. *Front Genet* 2:100.
12. Hutton JE*, et al.* (2016) Oncogenic KRAS and BRAF drive metabolic reprogramming in colorectal cancer. *Mol Cell Proteomics*.

Supplemental Tables

Table S1. Detection of KRAS4b proteoforms in colorectal and lung cancer cell lines.[a]

Cell lines	Cell type	*KRAS* genotype	% Mut KRAS4b detected[b]	% C-terminal WT COOMe[c]	% C-terminal Mut COOMe[b]
NCI-H1792	Lung	G12C/WT	79% (±3)	67% (±1)	68% (±2)
COR-L23	Lung	G12V/G12V	99% (±1)	N.D.	76% (±1)
SW48 G12D	Colorectal	G12D/WT	42% (±4)	99% (±1)	99% (±1)

[a]each average is listed as the mean ±standard error from that mean (±SEM)
[b]percentage of mutant KRAS4b listed is the average of 3 experimental replicates
[c]percentage of COOMe listed is the average of 3 experimental replicates using the specific MS2 events for each specific proteoform with narrow (3 *m/z*) isolation window method

Table S2. KRAS4b proteoforms identified in this study; PFR # is the proteoform record number, a unique accession in the human proteoform atlas (http://repository.topdownproteomics.org/).

PFR #	Mutational status	N-terminal modification	C-terminal modification	PTMs	Cell line/tumors identified
249914	WT	Acetylated	COOMe	Cys185 Farnesyl	DLD-1 PAR, DLD-1 WT, HCT116 PAR, HCT116 WT, SW48 PAR, SW48 G13D, SW48 G12D, NCI-H1792, Subjects 1-6
249915	WT	Acetylated	None	Cys185 Farnesyl	NCI-H1792, Subjects 5 and 6
249916	WT	Acetylated	COOMe	Cys185 Farnesyl Cys118 Nitrosylated	DLD-1 WT, HCT116 WT
249917	G13D	Acetylated	COOMe	Cys185 Farnesyl	DLD-1 PAR, DLD-1 Mut, HCT116 PAR, SW48 G13D, Subjects 3 and 4
249918	G13D	Acetylated	None	Cys185 Farnesyl	DLD-1 Par, DLD-1 Mut, Subjects 3 and 4
249919	G12C	Acetylated	COOMe	Cys185 Farnesyl	NCI-H1792
249920	G12C	Acetylated	None	Cys185 Farnesyl	NCI-H1792
249921	G12V	Acetylated	COOMe	Cys185 Farnesyl	COR-L23, Subject 5
249922	G12V	Acetylated	None	Cys185 Farnesyl	COR-L23, Subject 5
249923	G12D	Acetylated	COOMe	Cys185 Farnesyl	SW48 G12D
249924	G12D	Acetylated	None	Cys185 Farnesyl	SW48 G12D

Table S3. Allele-specific detection of KRAS4b proteoforms in two additional colorectal tumor samples using narrow (3 *m/z*) isolation of precursor ions.

Tumor sample	Disease stage	Total % C-terminal COOMe[a]	Allele	Allelic KRAS4b detected	% C-terminal COOMe[b]	% C-terminal COO^{-}[b]
Subject 5 (WT/G12V)	IIIC	47%	WT	82%	49%	51%
			G12V	18%	9%	91%
Subject 6 (WT/WT)	IIIC	90%	WT	100%	90%	10%

[a]percentage of COOMe listed calculated from MS2 events from wide isolation window method
[b]percentage of COOMe listed calculated from the specific MS2 events for each specific proteoform via the narrow isolation window method

Table S4. Relative amounts of C-terminal carboxymethylation from spiking recombinant GG-KRAS4b WT standard into whole cell extracts of DLD-1 Mut.

Cell lines	Cell type	*KRAS* genotype	% C-terminal COOMe[a]
N/A	N/A	GG-KRAS4b WT	100%
DLD-1 Mut	Colorectal	G13D/- (GG-KRAS4b WT Spike)	97% (±1)

[a]percentage of COOMe listed is the average of 3 experimental replicates

Supplemental Figures

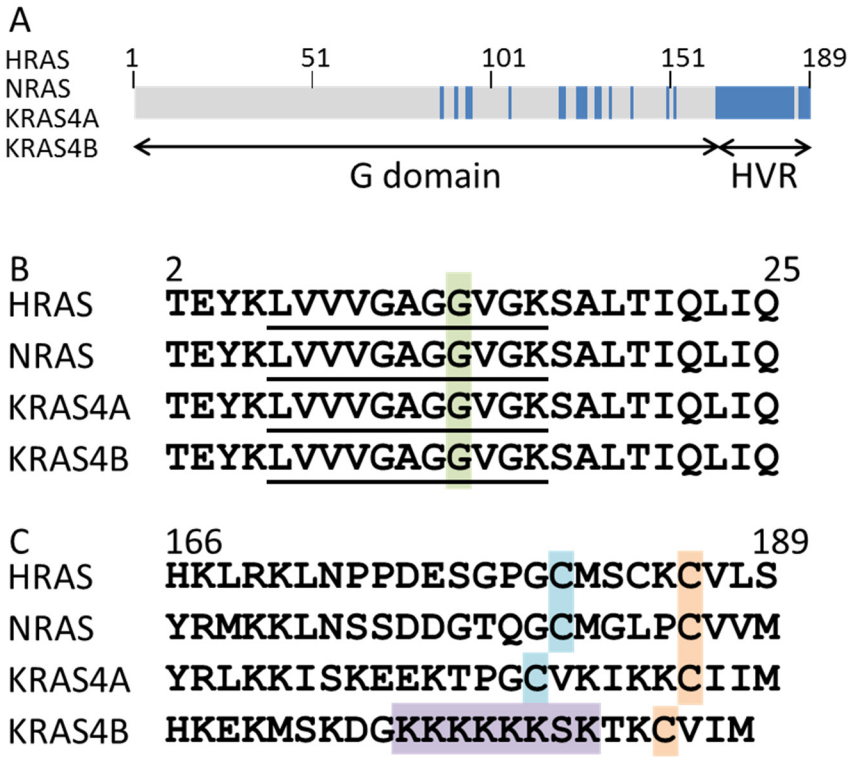

Fig. S1. Human RAS protein isoforms exhibit high sequence identity in their G domains and diverge significantly in their hypervariable regions. (*A*) Schematic representation of the sequence identity among RAS proteins. Vertical blue lines correspond to sites of sequence divergence. (*B*) The first 86 amino acids are identical among the 4 RAS proteins. Amino acids 2-25 are shown here, with the tryptic peptide containing the G13D site underlined, and the G13 residue highlighted across the 4 sequences. (*C*) The last 24 amino acids constitute the hypervariable region of RAS proteins and contain modification sites important for membrane targeting and association. Highlighted in orange is the farnesylated cysteine of the CAAX motif, highlighted in blue are sites of palmitoylation, and highlighted in purple is the polybasic domain of KRAS4b.

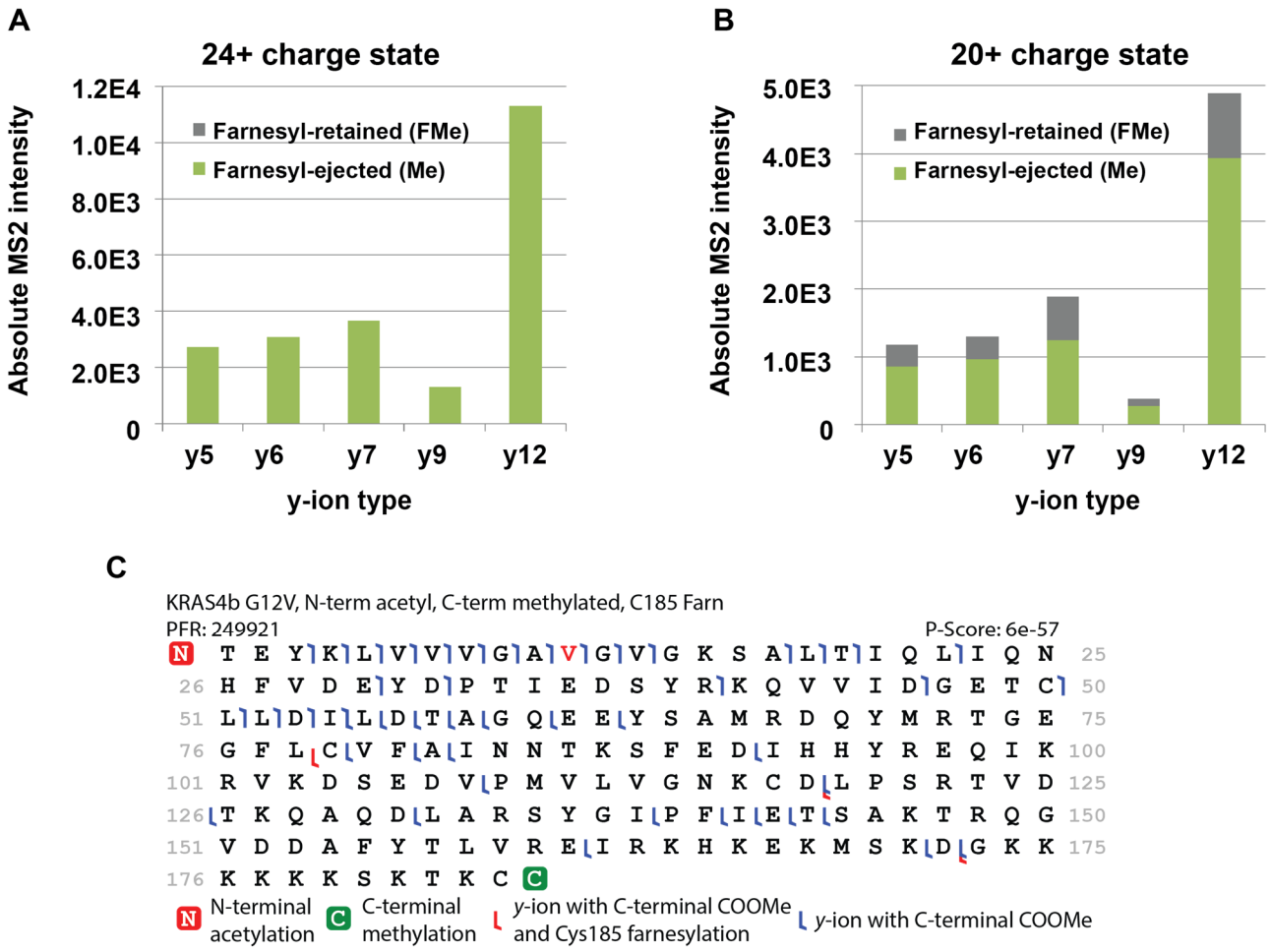

Fig. S2. Effect of precursor charge state on the relative intensities of several *y*-ions during KRAS4b fragmentation. (*A*) Bar chart illustrating ejection of the farnesylation during fragmentation of the 24+ charge state. (*B*) Bar chart illustrating detection of *y*-ions maintaining the farnesylation (FMe) and *y*-ions without the farnesylation (Me). (*C*) Example graphical fragment map of endogenous KRAS4b G12V from IP-TDMS of COR-L23 cells highlighting additional *y*-ions with retention of farnesylation at Cys185 (red flags). The 23+ precursor was selected for fragmentation for this data, and fragmentation was performed with a stepped collision energy of 18 and 23 (NCE). The *p*-score is a metric of matching fragment ions by chance and was calculated from fragment ions without the retained farnesylation.

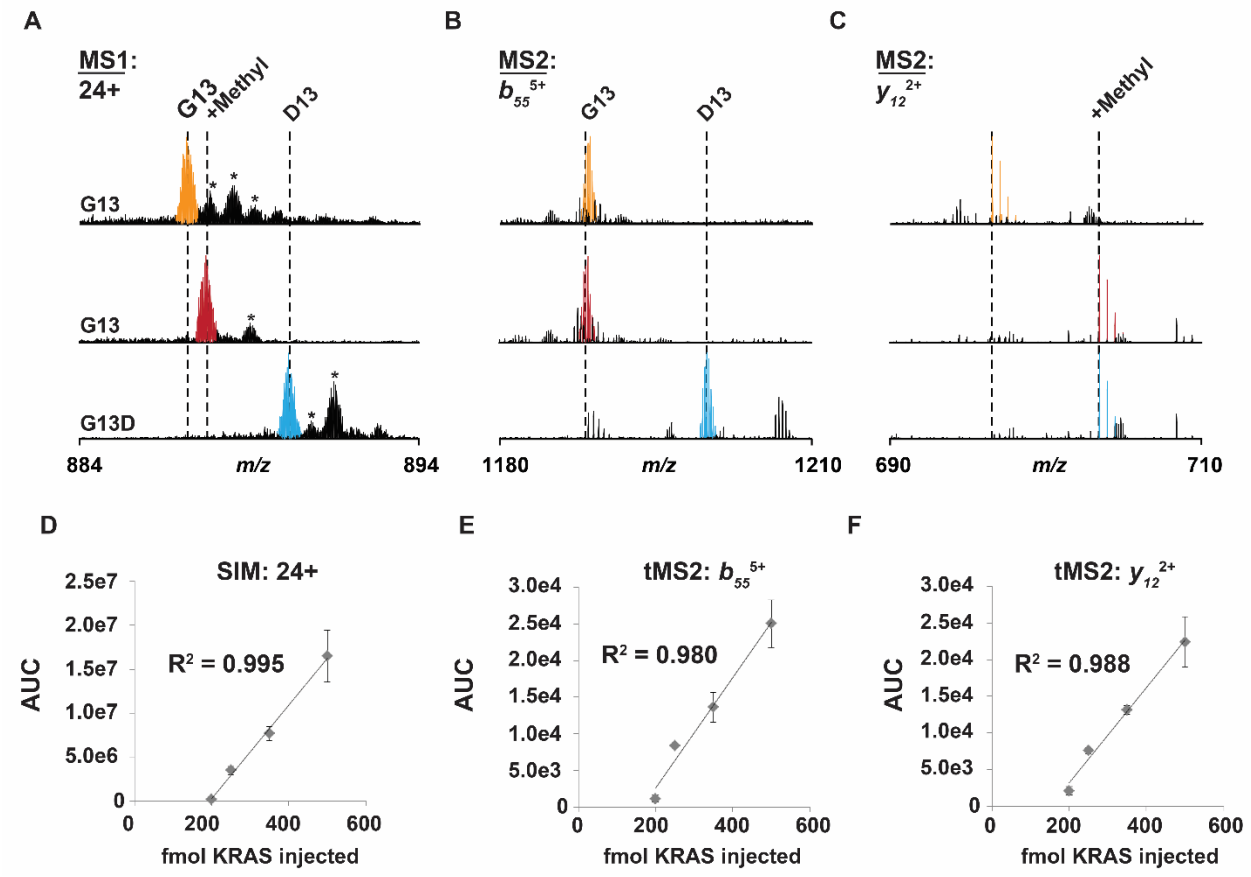

Fig. S3. Representative mass spectra of KRAS4b recombinant standards. (*A*) The MS1 of the 24+ charge state of KRAS4b standards clearly shows a 58 Da mass shift from WT to G13D mutant KRAS and a 14 Da mass shift due to differences in C-terminal COOMe. Asterisks denote known ESI oxidation products. (*B*) The 58 Da mass shift can also be seen in N-terminus containing *b*-ions, such as the b_{55} ion pictured here. (*C*) Fragment ions containing the C-terminus, such as the y_{12}, can be used to confirm the absence of COOMe. (*D*) Relationship between the amount of KRAS standard injected into the mass spectrometer and the peak area of the major proteoform (from SIM scans). (*E*) Relationship between amount of KRAS standard injected and the peak area of fragment ion b_{55} (tMS2 scan mode). (*F*) Relationship between amount of KRAS standard injected and the peak area of fragment ion y_{12} (tMS2 scan mode).

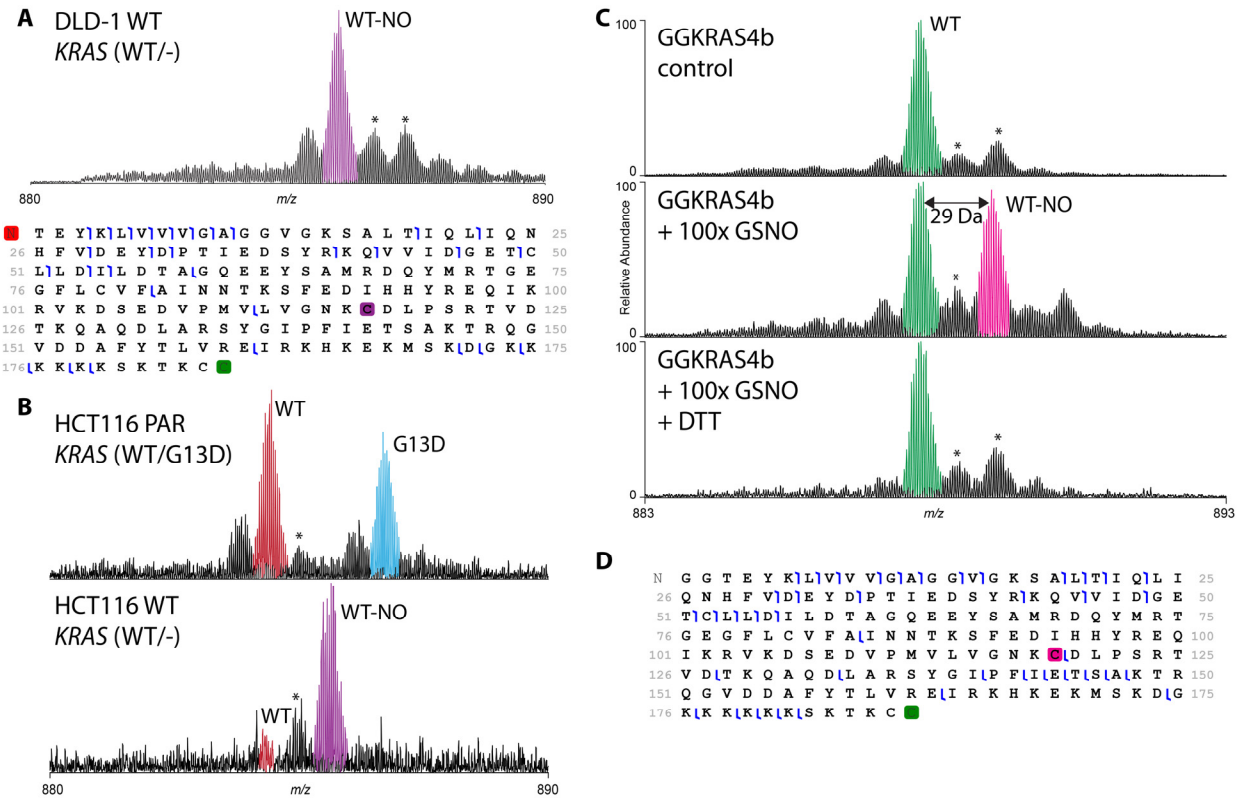

Fig. S4. Detection of nitrosylated KRAS proteoforms. (*A*) Top: Mass spectrum of the 24+ charge state of endogenous WT KRAS4b bearing a nitrosylation from DLD-1 WT cells. Bottom: Fragment map identifying the position of +29 Da nitrosylation to cysteine 118. Blue flags represent *b*- and *y*-ions detected and mapped to the primary structure. This proteoform contains the following modifications: N-terminal acetylation (red square), C-terminal farnesylation and methylation (green square) and nitrosylation (purple square). Asterisks denote known oxidation products due to electrospray ionization. (*B*) Mass spectrum of 24+ charge state of endogenous KRAS4b from HCT116 PAR (top) and HCT116 WT (bottom) cells, confirming nitrosylation of only the WT KRAS4b proteoform in HCT116 WT cells. Asterisks denote known oxidation products due to electrospray ionization. (*C*) Mass spectra and fragment map of a KRAS4b standard (top) treated to add (middle), and then remove (bottom), Cys118 nitrosylation. The standard is labeled as GG-KRAS4b, as it only differs from endogenous KRAS by an N-terminal GG tag, and contains both C-farnesylation and methylation. Upon incubation with nitrosoglutathione (GS-NO), a new proteoform (pink, middle panel) appears that differs from the standard by 29 Da, corresponding to the mass of the nitrosyl group. The nitrosylation can then be removed (bottom) by incubation with DTT. Asterisks denote known oxidation products due to electrospray ionization. (*D*) Graphical fragment map generated from the tandem mass spectrum of the GG-KRAS4b proteoform nitrosylated at Cys118 (pink species from the middle of Panel *C*).

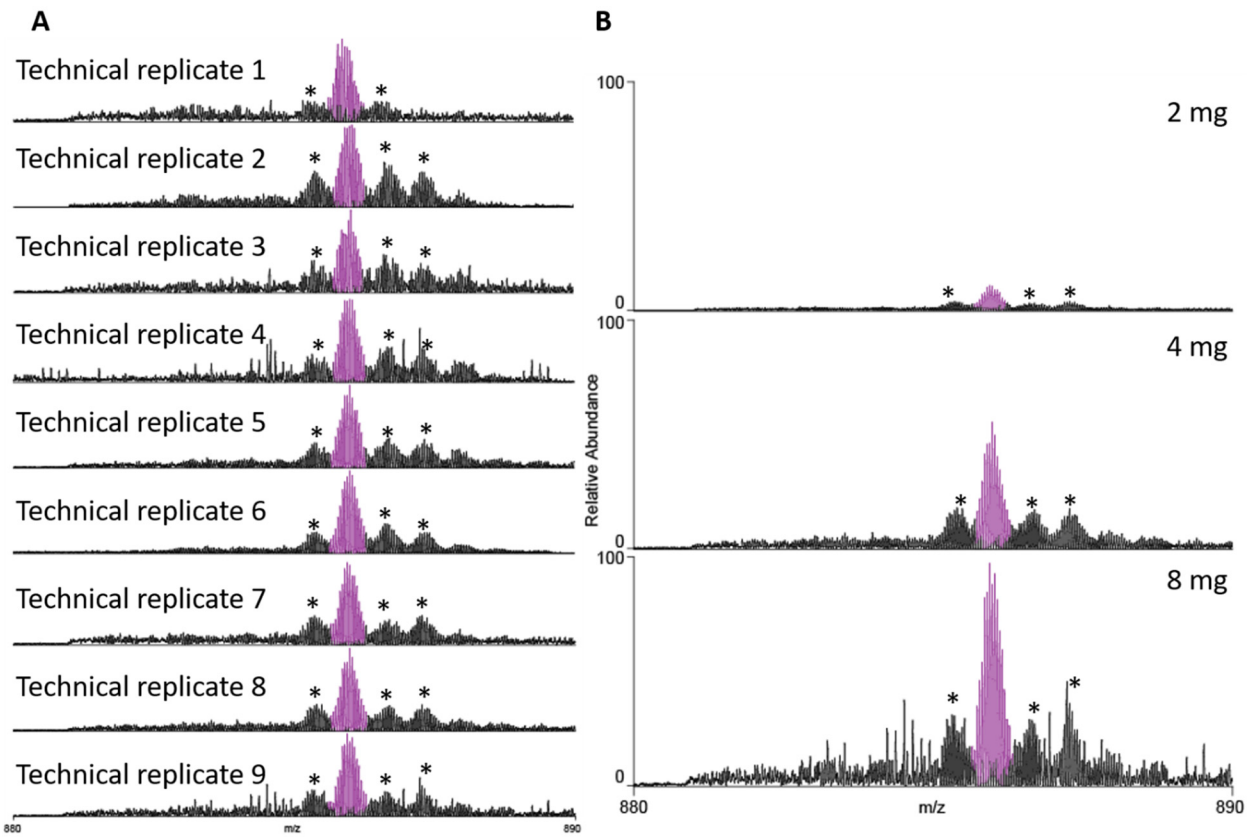

Fig. S5. Detection of Cys118 nitrosylation of KRAS4b in DLD-1 WT cells. Asterisks denote known oxidation products due to electrospray ionization. (*A*) Mass spectra from nine experimental replicates (i.e., replicated over a several month period) confirm >90% nitrosylation of wild type KRAS4b expressed in DLD-1 WT cells. Asterisks denote ESI-MS artifacts of oxidation (right of major peak) and an ammonia neutral loss (left of major peak). (*B*) MS spectra of KRAS4b isolated using 2 mg, 4 mg and 8 mg of lysate from DLD-1 WT cells during the IP step, illustrating that nitrosylation levels detected are independent of input lysate.

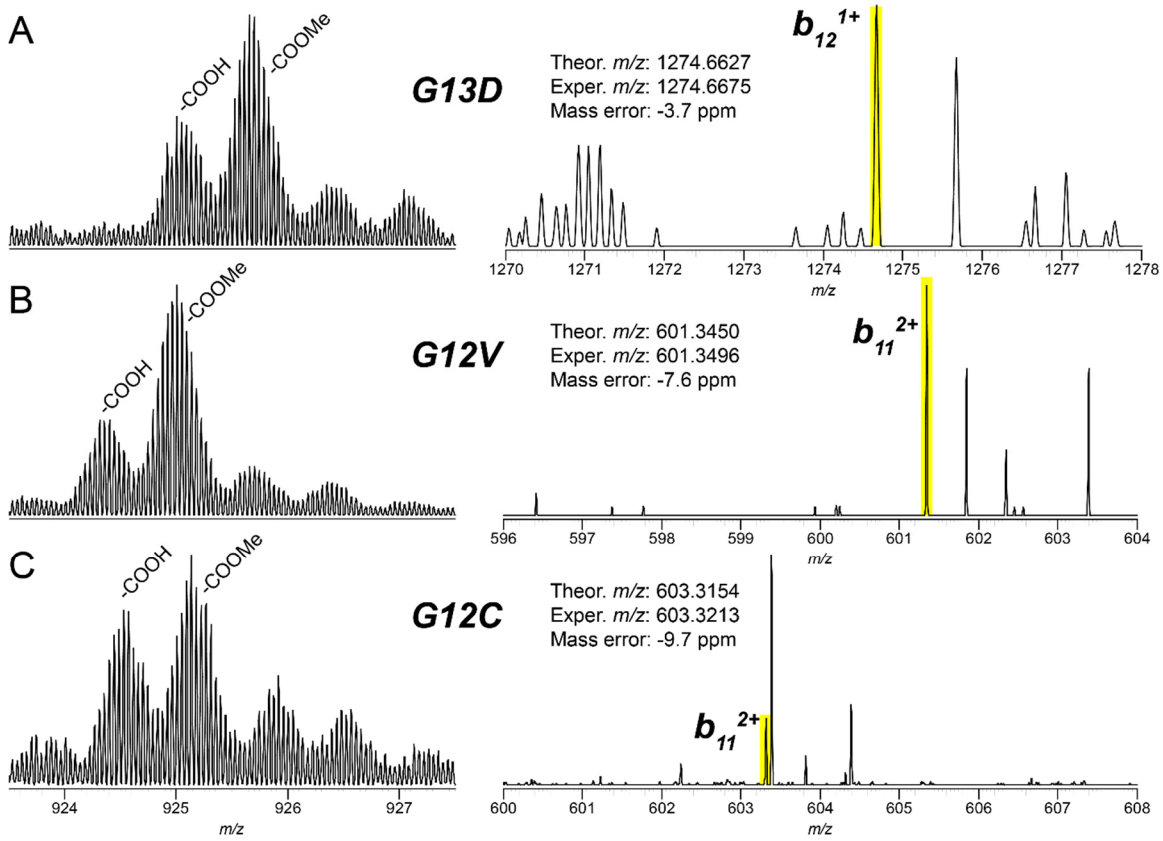

Fig. S6. Allele-specific characterization of KRAS4b proteoforms using 3 *m/z* isolation of precursor ions. (*A*), (*B*) and (*C*) show examples of the allele-specific characterization of oncogenic KRAS4b proteoforms derived from the G13D, G12V and G12C alleles, respectively. The panels at left show the intact protein ions for the three that each correspond to fully farnesylated proteoforms, with or without the C-terminal COOMe (-FMe and –Farn, respectively). The panels at right show the position (monoisotopic peak highlighted in yellow) of diagnostic b_{11}^{2+} or b_{12}^{1+}ions (the initial Met is cleaved off) whose masses differ for the three alleles: the b_{12}^{1+} ion for the sequence with G13D is detected at *m/z* 1274.65 (*A*), the b_{11}^{2+} ion for G12V is detected at m/z 601.34 (*B*), and the b_{11}^{2+} ion corresponding to the G12C sequence falls at m/z 603.31 (*C*).

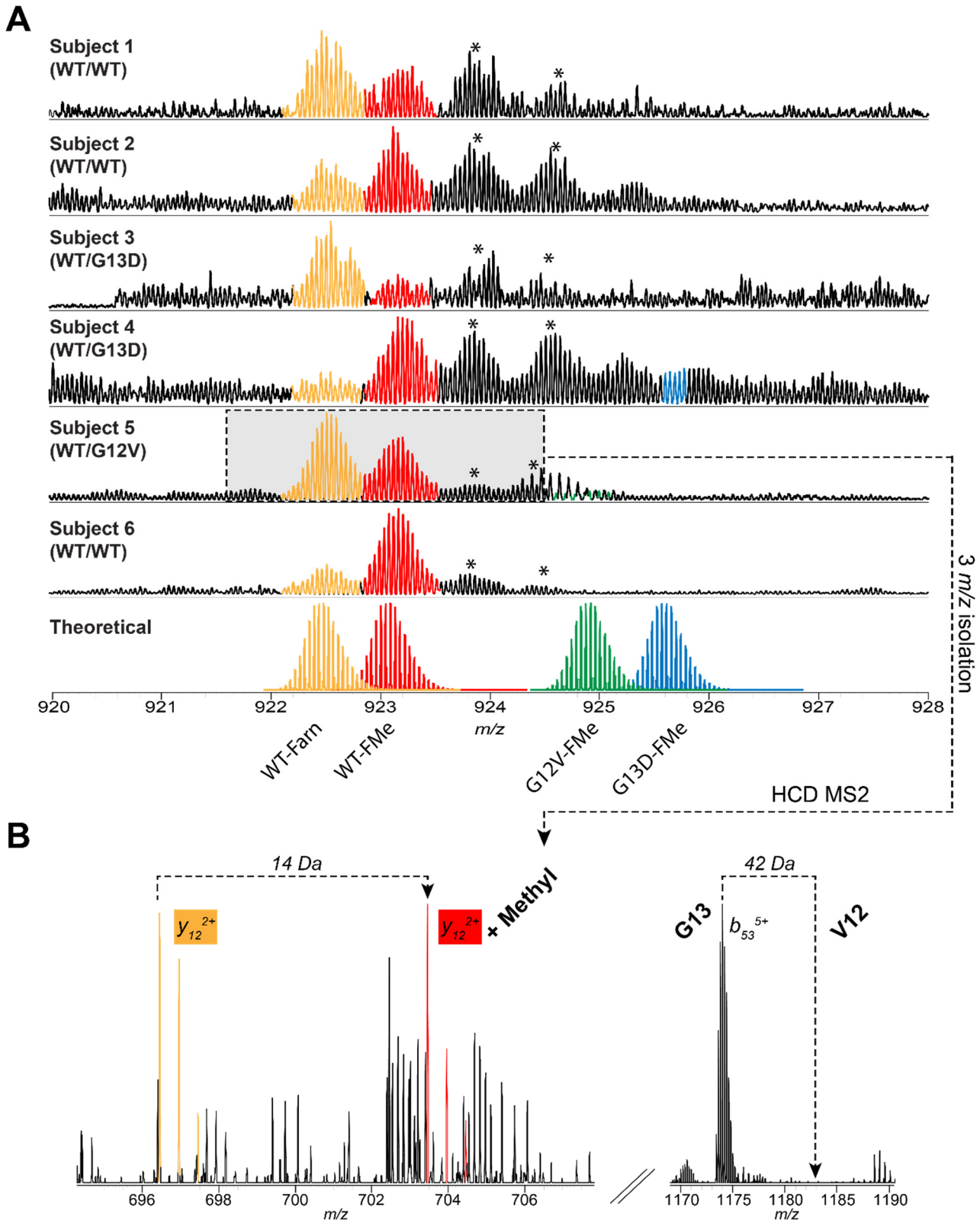
A
Subject 1
(WT/WT)
Subject 2
(WT/WT)
Subject 3
(WT/G13D)
Subject 4
(WT/G13D)
Subject 5
(WT/G12V)
Subject 6
(WT/WT)
Theoretical
920
921
922
923
924
925
926
927
928
m/z
WT-Farn
WT-FMe
G12V-FMe
G13D-FMe
3 m/z isolation
HCD MS2
B
14 Da
y_{12}^{2+}
y_{12}^{2+} + Methyl
696
698
700
702
704
706
m/z
42 Da
G13
b_{53}^{5+}
V12
1170
1175
1180
1185
1190
m/z

Fig. S7. Precise KRAS4b detection in tumor samples. (*A*), the 23+ charge state of KRAS4b as detected in the 6 analyzed tumor samples; the bottom panel shows theoretical isotopic distributions of the major proteoforms: highlighted in orange is WT KRAS4b-Farn, highlighted in red is WT KRAS4b-FMe, in green G12V KRAS4b-FMe, and in blue is the position of the G13D KRAS4b-FMe proteoform (not asserted to be present in this study by MS1 based detection, but detected by MS2 diagnostic fragment ions in Fig 4*B*). Asterisks denote known oxidation products due to electrospray ionization. (*B*) illustrates the capabilities offered by state-of-the-art quadruople-Orbitrap FTMS instrumentation in selectively determining the relative amounts of methylation in KRAS4b proteoforms derived from specific alleles. In this example, only WT KRAS4b was isolated from Subject 5 with an isolation window of 3 *m/z*, as demonstrated by the absence of any b_{53} ion corresponding to the G12V mutant allele (mass shift: +42 Da, Panel *B*, right side). The relative intensity between (i) farnesylated and (ii) farnesylated and carboxymethylated KRAS4b proteoforms derived exclusively from the WT allele can be calculated using the ratio of y_{12} ions (Panel *B*, left side). Colored lines indicate theoretical location of y_{12} ions and do not reflect observed ion intensities.

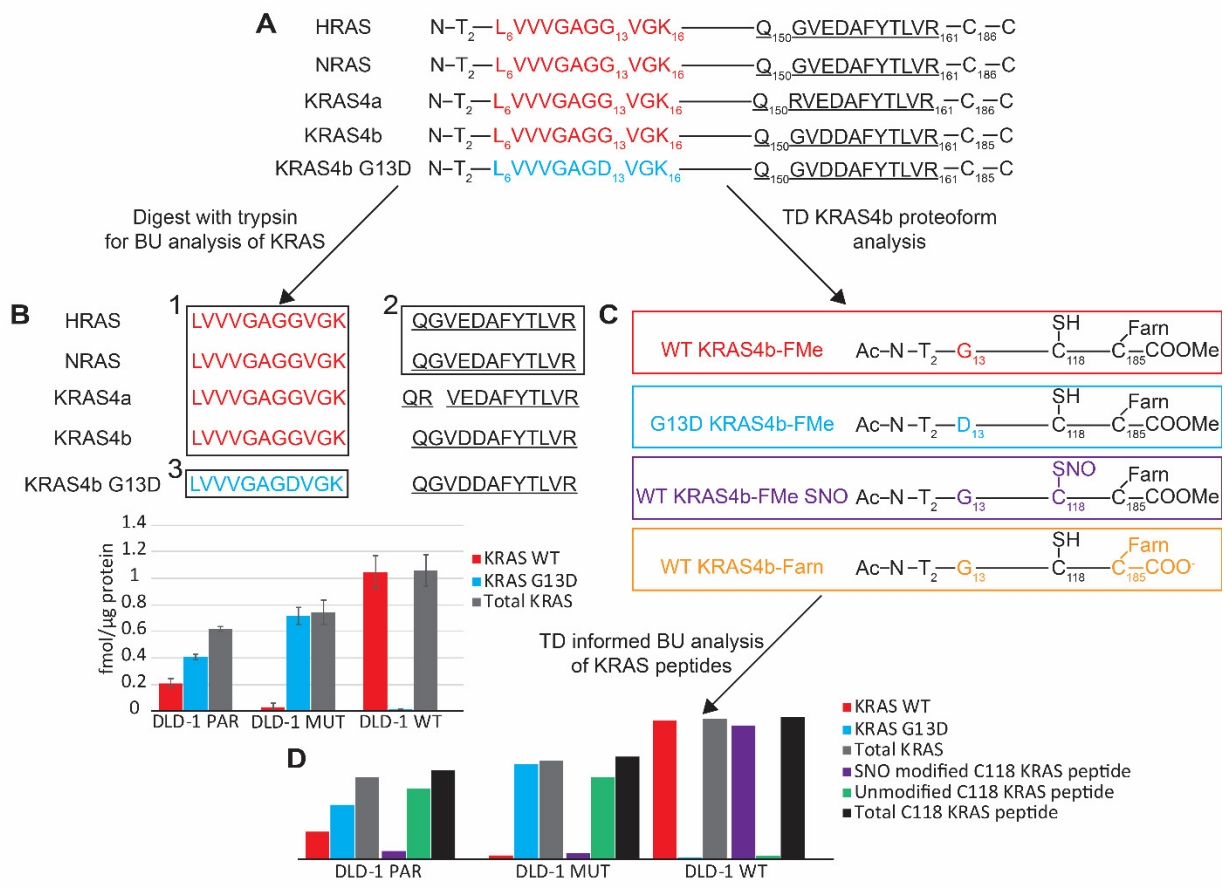

Fig. S8. Complementarity of TD and BU proteomics for the analysis of KRAS. (*A*) Relevant amino acid sequences for the 4 RAS proteins and KRAS4b with the G13D amino acid substitution. Red amino acid sequences denote tryptic peptides containing the Gly13 amino acid ("WT"), while the blue sequence denotes the peptide containing the substitution of Asp13 for Gly13 ("G13D"). Underlined sequences denote C-terminal tryptic peptides that differentiate the three RAS proteins. (*B*) BU quantitative analysis of the WT peptide (Box 1), the distinguishing HRAS and KRAS peptide (Box 2), and mutant peptide (Box 3) can be used to determine the amount of KRAS gene dosage. The amount of WT KRAS can be determined in a sample (red bar in bar chart) by measuring the Box 1 peptides and subtracting the measurement of the Box 2 peptides, the amount of mutant protein (blue bar in bar chart) can be determined by directly measuring the mutant peptide, and the amount of total KRAS is the sum of these two measurements (grey bar in bar chart). (*C*) Direct analysis of KRAS4b proteoforms by TD proteomics can be used to specifically survey which KRAS4b proteoforms are present in a sample, such as wild type KRAS4b-FMe (red), mutant G13D KRAS4b-FMe (blue), wild type KRAS4b-FMe nitrosylated (purple), and wild type KRAS4b-Farn (orange). (*D*) Theoretical combination of BU and TD proteomics. TD proteoform information can be used to generate novel BU quantitative assays, such as quantitatively measuring the tryptic peptide containing the unmodified (green) and the nitrosylated Cys118 residue (purple) in addition to measuring the WT peptide (blue) and the mutant peptide (red).