Differential Reprogramming of Isogenic Colorectal Cancer Cells by Distinct Activating KRAS Mutations

Dean E. Hammond, Craig J. Mageean, Emma V. Rusilowicz, Julie. A. Wickenden, Michael J. Clague, and Ian A. Prior

INTRODUCTION

KRAS is a member of the highly homologous p21 Ras family of monomeric GTPases. Three isoforms (HRAS, KRAS, and NRAS) are expressed in all mammalian cells and function as molecular switches downstream of cell surface receptors, such as epidermal growth factor receptor (EGFR), to stimulate cell proliferation and cell survival. Mutations of Ras at the conserved codons 12, 13, or 61 result in an impaired intrinsic hydrolysis rate or binding to GTPase activating proteins (GAPs). Despite a high degree of similarity, Ras isoforms hydrolysis rate or binding to GTPase activating proteins conserved codons 12, 13, or 61 result in an impaired intrinsic codon 13 with almost all of the remainder at codon 12. Tumors harbor KRAS mutations, 19% of these mutations are at codon 12 and 13. While mutations at both sites are activating, due to impaired GAP binding, the position of the mutation has functional and clinical relevance.

Metastatic colorectal cancer (mCRC) is one of the leading causes of cancer-related death worldwide. A third of CRC tumors harbor KRAS mutations, 19% of these mutations are at codon 13 with almost all of the remainder at codon 12. Mutations of KRAS at codon 12 are more potent than codon 13 mutations at transforming cells and are associated with a more aggressive metastatic colorectal cancer phenotype. Despite this, patients with codon 13 mutations display a significantly worse prognosis. Furthermore, patients with codon 12 and codon 13 mutations exhibit differential responsiveness to treatment.

These data suggest that each activating KRAS mutation generates a distinctive signaling output. Early Ras research supports these observations by demonstrating that variant amino acid codon mutations are not equally transforming. The mechanistic basis for this is unclear but may relate to differences in nucleotide hydrolysis rates that could translate into differential coupling with and activation of Ras effectors. For example, G12D and G12V, exhibit different GTP hydrolysis rates (G12D ~40% and G12V ~10% of wild-type respectively). Alternatively, the mutations may differently affect the distribution of GTP-Ras between conformational states that differ in effector recognition.

Various omic approaches have been previously used to identify KRAS signatures, typically using cell lines harboring oncogenic Ras variants. A drawback with some of these studies is the variability of the genetic background between cell lines that confounds attribution of results directly to the presence of oncogenic Ras. One strategy to overcome this has
been the use of isogenic cells. However, almost all models employed so far have involved, either stable overexpression of oncogenic Ras randomly inserted into the genome on an isogenic background or genetic ablation of a wild type or oncogenic KRAS allele.

We exploit recently developed model cell culture systems that accurately recapitulate the genetic changes present in human CRCs. Specifically, we are using isogenic human SW48 CRC cell-lines in which targeted homologous recombination with the endogenous KRAS gene has been used to knock-in a panel of KRAS codon 12 and 13 mutations commonly found in CRC. The G12D, G12V, and G13D KRAS mutations present in our isogenic cell panel are the three most abundant mutations representing 75% of all cases of CRC harboring a KRAS mutation. We have used quantitative proteomic approaches to determine (phospho)proteomic signatures associated with each KRAS mutation. This combination of cell model and experimental approach represents the contemporary gold standard for precise analysis of endogenous oncogenic KRAS signaling. Importantly, we find that each of the activating mutations that we have investigated display distinct output signatures. We identified a subset of proteins and phosphosites associated with codon 12 versus codon 13 responses. Among these are the kinase proteins DCLK1 and MET which show the same patterns of KRAS-dependent overexpression across a broad panel of codon 12 mutant isogenic SW48 cells.

\section*{MATERIALS AND METHODS}

\subsection*{Cell-Lines and SILAC}

Isogenic SW48 cells were obtained from Horizon Discovery. The clones used were heterozygous knock-in (G12V+/−) of K-Ras activating mutation KRAS\textsuperscript{G12V} (cat. no. HD 103-007 0395), heterozygous knock-in (G12D+/−) of K-Ras activating mutation KRAS\textsuperscript{G12D} (HD 103-011 00436) and heterozygous knock-in (G13D+/−) of K-Ras activating mutation KRAS\textsuperscript{G13D} (HD 103-002 0025). These were referenced to homozygous KRAS\textsuperscript{WT} expressing cells (HD PAR-006 0027), hereafter referred to as Parental cells. For KRAS knock down studies, SW48 PAR and G12D cells containing doxycycline inducible shRNA targeting KRAS were generated. The following sequences were used: shRNA#A top strand: CCAGGGATACAAGCTTAATCCAGATCCCGATATGGTGACAGCTTTGTATCGTTTTTTT, bottom strand: AATTAAGAACGCAGTACAATGATCGTTTTTTT; shRNA#B top strand: CCAGGGATACACTTGTTTTAGAAGACGTCTGCTGGTGTTTTTAT, bottom strand: AATTAAAACCGGTGAGTCAAAGGTGGTCTGAAGGTTTCTTTTTT. shRNA#B from each experimental configuration was used. Each cell line was expanded in McCoy’s 5A medium supplemented with 10% dialyzed FBS (Dundee CRC). To knock down KRAS, the cells were grown in media containing 100 ng/μL doxycycline for 1 week. All cells were maintained in McCoy’s SA medium supplemented with 10% dialyzed FBS (Dundee Cell Products). To generate light, medium and heavy stable isotope-labeled cells, arginine- and lysine-free McCoy’s medium was supplemented with 200 μg/mL l-proline and either l-lysine (Lys0) together with l-arginine (Arg0), l-lysine-H\textsubscript{4} (Lys4) with l-arginine-U\textsuperscript{13}C\textsubscript{6} (Arg6) or l-lysine-U\textsuperscript{13}C\textsubscript{6}-N\textsubscript{2} (Lys8) with l-arginine-U\textsuperscript{13}C\textsubscript{6}-N\textsubscript{2} (Arg10) at final concentrations of 28 mg/L for the arginine and 146 mg/L for the lysine until fully metabolically labeled. The extent of isotope incorporation was assessed using an R-script as described. Cell lysates were prepared, quantified, subjected to SDS PAGE and in-gel tryptic digest as described previously. At least three biological replicate data sets representative of each KRAS\textsuperscript{MUTANT} versus Parental SW48 were obtained (n = 4 for KRAS\textsuperscript{G12D} versus Parental).

\subsection*{Sample Preparation}

For phosphopeptide (pSer/Thr/Tyr) isolation, we used filter-aided sample preparation (FASP)\textsuperscript{27} followed by fractionation using strong cation exchange (SCX) chromatography and TiO\textsubscript{2}-based phosphopeptide isolation (based on refs 28 and 29 and described previously in refs 25 and 30). In parallel, quantitative SW48 isogenic cell-line proteome analyses were carried out by resolving a 50 μg aliquot of each SILAC mixture by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) on a 4−12% NuPAGE gel (Invitrogen), prior to protein visualization by Colloidal Blue staining (Invitrogen). Gel lanes were then cut into 48 bands each, according to protein content, in-gel digested overnight at 37 °C with trypsin (4 ng/μL working concentration; Trypsin GOLD, sequencing grade, Promega) to cleave C-terminal to arginine and lysine residues, dried, and redissolved in 0.05% TFA prior to LC-MSMS analysis of each gel slice.

\subsection*{LC−MS/MS and Data Processing}

A total of 5 μL of each sample was fractionated by nanoscale C18 high performance liquid chromatography (HPLC) on a Waters nanoACQUITY UPLC system coupled to an LTQ-OrbitrapXL (Thermo Fisher) fitted with a Proxeon nanoelectrospray source. Peptides were loaded onto a 5 cm × 180 μm trap column (BEH-C18 Symmetry; Waters Corporation) in 0.1% formic acid at a flow rate of 15 μL/min and then resolved using a 25 cm × 75 μm column using a 20 min linear gradient of 3 to 62.5% acetonitrile in 0.1% formic acid at a flow rate of 400 nL/min (column temperature of 65 °C). The mass spectrometer acquired full MS survey scans in the Orbitrap (R = 30 000; m/z range 300−2000) and performed MSMS on the top five multiple charged ions in the linear quadrupole ion trap (LTQ) after fragmentation using collision-induced dissociation (30 ms at 35% energy). Full scan MS ions previously selected for MSMS were dynamically excluded for 180 s from within a rolling exclusion list (with n = 1). Phosphopeptides were also analyzed using multistage activation (R = 60 000, neutral loss mass list: 49.0, 65.3, 98.0) for the top six multiply charged ions, using a 60 min linear gradient of 3 to 62.5% acetonitrile in 0.1% formic acid, all other conditions as above. All spectra were acquired using Xcalibur software (version 2.0.7; Thermo Fisher Scientific).

Raw MS peak list files from each experimental configuration were searched against the human IPI database (version 3.77) using the Andromeda search engine\textsuperscript{31} and processed with the MaxQuant software suite\textsuperscript{32} (version 1.2.2.5) as described previously. The minimum required peptide length was set to 6 amino acids and two missed cleavages were allowed. Cysteine carbamidomethylation (C) was set as a fixed modification, whereas oxidation (M) and S/T/Y phosphorylation were considered as variable modifications. The initial precursor and fragment ion maximum mass deviations were set to 7 ppm and 0.8 Da, respectively, for the search of the ipi_HUMAN_y3.77.fasta database containing 89 709 entries. The results of the database search were further processed and statistically evaluated by MaxQuant. Peptide and protein false discovery rates were set to 0.01. Proteins with at least one peptide unique to the protein sequence were considered as
valid identifications. For protein quantitation, only proteins with at least three peptides (one unique) were selected. In addition, all experiments were also analyzed together using Andromeda and MaxQuant in a single iteration of the pipeline.

Data obtained from MaxQuant analyses were evaluated using Excel and MeV (version 4.8.1; www.tm4.org/mev). To compare the interexperimental correlation between biological replicate experiments peptides or phosphopeptides present in two or more of each experimental configuration (2/4 or 2/3 [KRASG12D versus Parental]) were log_{10} transformed, plotted on scatter plots and the R² correlation(s) visualized as heatmaps. Hierarchical clustering was performed using MeV on the R² data. Principle component analysis on covariances was performed with JMP10. Peptide data were included for analysis if ratios were available for every Par/mutant condition. No imputation was performed. Peptides with missing ratios were excluded from the analysis.

Cluster Analysis, GO Analysis and Linear Kinase Motif Analysis

GProX analysis of log_{2} transformed MaxQuant data sets using unsupervised fuzzy c-means clustering was performed as described previously to identify corresponding genes in the proteome and phosphoproteome data sets.26,33 Gene Ontology analysis using DAVID Bioinformatics Database34 was performed using the Entrez Gene ID identifiers of shortlisted proteins and phosphopeptides. Over-represented terms within the short-lists were calculated using a background list comprising all genes identified across our experiments (threshold count = 2; EASE score = 1). Terms with a p-value < 0.1 in at least one cluster were selected, log_{10} transformed, hierarchically clustered and plotted as a heatmap. Phosphopeptides within the data set with a phosphorylation localization probability ≥ 0.75 (class 1) were analyzed for common motifs and their putative regulatory kinases using MotifX35 and NetworKIN v2.036 as described previously.26

Western Blotting

A 25 μg cell lysate was run on SDS gels, and subsequently transferred to nitrocellulose membrane. Membranes were blocked and then blotted using the indicated primary antibody. Membranes were then incubated with IR dye coupled secondary antibodies and detected using the Odyssey system (Licor). Within this study the following primary antibodies were used: polyclonal anti-AKAP12 (C3, Gene Tex), monoclonal anti-AKAP12, polyclonal anti-α-tubulin (Sigma-Aldrich), rabbit monoclonal pan-RAS (Epitomics), and rabbit monoclonal anti-ALDH3A1 (Abcam).

RNA Extraction and QPCR

Total RNA was extracted from SW48 cells using a Qiagen RNaseasy kit. cDNA was made by reverse transcription of 1 μg of RNA using RevertAid H-minus M-MuLV reverse transcriptase (Fermentas) and oligo(dT) primer (Promega). Quantitative real-time PCR (QPCR) was performed using a real-time PCR detection system (Bio-Rad) using IQ SYBR Green Supermix. QPCR was conducted in triplicate with 1 μL of cDNA and 150 nmol primers. Samples underwent 40 cycles of amplification at 94 °C (30 s) and 60 °C/62 °C (60 s), fluorescence was read at 60 °C/62 °C, and melt curves analyzed. For each sample, the Ct values for DCLK1 and KRAS were normalized to the reference gene ACTB and the control sample and represented as 2^{-ΔΔCt}.

RESULTS

Quantitative Proteomic Analysis Reveals KRAS Mutation-Specific Network Responses

Isogenic SW48 colorectal cancer cell lines harboring either wild type or a G12D, G12V, or G13D mutated KRAS allele were used to investigate the effects of amino-acid substitution specificity (G12D vs G12V) or codon-specificity (G12D vs G13D) on KRAS signaling. Stable isotope labeling of amino acids in cell culture (SILAC) allows different cell populations to be selectively labeled with isotopes of arginine and lysine and analyzed by mass spectrometry in a triplexed configuration (Figure 1;28). Following SILAC labeling, cell lysates were either run directly on SDS-PAGE gels or subjected to TiO₂-based phosphoprotein enrichment procedures. High-resolution mass-spectrometry of gel slices or peptide fractions allowed us to compare their proteome and signaling network responses downstream of each KRAS mutant (Figure 1). All KRAS mutants were compared to a parental wild-type KRAS control in each triplex configuration with an n of 3 or 4 biological repeats for each comparison.
In total, across all experiments, responses were measured from 2359 unique proteins in the proteome data set and 3971 unique phosphopeptides from the TiO2 purifications (3311 phosphosites unique by sequence; Supporting Information Table 1). A total of 65% of proteins and 35% of phosphopeptides were sampled at least twice across biological replicates (Supporting Information Figure 1). A total of 3727 phosphosites could be assigned to a specific position within the protein with a probability of at least 0.75 (class 1 sites). The 3727 class 1 phosphosites were composed of 3030 pSer, 632 pThr, and 65 pTyr sites mapped to 1288 proteins.

To examine experimental reproducibility and intermutation response variability, we performed cross-correlation analysis between all experimental pairs across biological replicates (Figure 2A). For both proteome and phosphoproteome levels there is consistent biological reproducibility but a high degree of variability between cell lines containing different KRAS mutations. Isogenic SW48 cells harboring codon 12 KRAS mutations share greater correlation compared to cells harboring a G13D mutation. Values for peptides/phosphopeptides shared between each experiment were used for $R^2$ cross-correlation analysis. Pearson correlation coefficient ($r$) indicates linkage strength and relationship between experimental conditions. (B) Principle component analysis of data following combination of biological replicates. Codon 12 mutant KRAS cell lines share similar projections at both proteome and phosphoproteome levels.

**Figure 2.** Oncogenic KRAS variants display mutation-specific changes to proteome and phosphopeptide networks. (A) At both proteome and phosphoproteome levels there is consistent biological reproducibility but a high degree of variability between cell lines containing different KRAS mutations. Isogenic SW48 cells harboring codon 12 KRAS mutations share greater correlation compared to cells harboring a G13D mutation. Values for peptides/phosphopeptides shared between each experiment were used for $R^2$ cross-correlation analysis. Pearson correlation coefficient ($r$) indicates linkage strength and relationship between experimental conditions. (B) Principle component analysis of data following combination of biological replicates. Codon 12 mutant KRAS cell lines share similar projections at both proteome and phosphoproteome levels.

In total, across all experiments, responses were measured from 2359 unique proteins in the proteome data set and 3971 unique phosphopeptides from the TiO2 purifications (3311 phosphosites unique by sequence; Supporting Information Table 1). A total of 65% of proteins and 35% of phosphopeptides were sampled at least twice across biological replicates (Supporting Information Figure 1). A total of 3727 phosphosites could be assigned to a specific position within the protein with a probability of at least 0.75 (class 1 sites). The 3727 class 1 phosphosites were composed of 3030 pSer, 632 pThr, and 65 pTyr sites mapped to 1288 proteins.

To examine experimental reproducibility and intermutation response variability, we performed cross-correlation analysis between all experimental pairs across biological replicates (Figure 2A). For both proteome and phosphoproteome data sets, hierarchical clustering indicates good experimental reproducibility with a least a 5-fold higher cross-correlation coefficient between biological repeats compared to any of the interisogenic cell type correlations. Importantly, the consistent responses of each cell line allow us to clearly observe KRAS mutation-specific signaling signatures. More specifically, both hierarchical clustering and principal component analysis (Figure 2B) indicate that although changes in the proteome and phosphoproteome outputs are similar between the G12D/ G12V mutants there is a divergence between codon 12 and codon 13 mutants. Closer inspection of the responses within our data sets reveals that G13D mutant cells exhibit more prevalent protein and phosphopeptide up-regulation than the G12 mutant cells (Supporting Information Figure 1). For example, almost 50% of G13D phosphopeptides are up-regulated versus <10% of phosphopeptides in G12 mutant cell lines (Supporting Information Figure 1A and 1C). Therefore, we observe that both the type of amino acid substitution and the codon positioning of the mutation, influence the outputs of oncogenic KRAS, with codon position having the greatest effect.

Notably, there were very few proteins or phosphosites that showed significant pan-mutation responses (Supporting Information Table 4). One protein and 29 phosphosites exhibited ≥1.5 fold up- or down-regulation versus wild type Ras across all three G12V, G12D, and G13D cell lines. These included some of the codon-specific responders such as AKAP12 and DCLK1 where though increases were seen in each of the mutant cell lines, there was a significant bias in favor of one or more KRAS variant. Almost all of the genes showing pan-mutation responses are involved in mRNA processing and transcriptional regulation.

**Differential Responses between Codon 12 and 13 Mutant KRAS Cells**

We clustered proteins and phosphopeptides according to their mutation-specific responses versus parental controls using
proteome meta-analysis software GProX. These clusters represent groups of proteins or phosphosites that show matching profiles across each of the cell lines. Six distinct phosphopeptide and proteome clusters were identified (Figure 3A, Supporting Information Figure 2 and Table 2). Of particular interest were those representing codon 12- versus codon 13-specific responses that are most marked in clusters 5 and 6.

Analysis of the proteome data revealed 115 proteins in clusters 5 and 6 for which changes in abundance are associated with KRAS codon 12 versus 13-mutant signaling (Supporting Information Figure 2). In this data set, Gene Ontology (GO) analysis revealed that certain mitochondrial proteins involved in oxidative phosphorylation are enriched in cluster 5, that is, these proteins are decreased in G13D relative to codon 12 mutant cells (Supporting Information Figure 2). For example, within the proteome data set, we observe decreases in abundance of peptides from 5 of the 11 members of the cytochrome bc1 complex (complex III) and succinate dehydrogenase of complex II in G13D cells but not other components of the respiratory chain. In contrast, metabolic enzymes including those involved in gluconeogenesis are enriched in cluster 6 (Supporting Information Figure 2); however, the majority of enzymes within the proteome data set that are associated with glycolysis, including pyruvate kinase M2 do not significantly change between the cell lines. The most

Figure 3. Phosphopeptides displaying codon 12 vs 13 mutant KRAS responses. (A) GProX clustering of changes seen in the phosphoproteome. Ratios for proteins that exhibit a change in expression level within a KRAS-mutated environment were subjected to unsupervised clustering with the Fuzzy c means algorithm. Clusters corresponding to six different response patterns were identified. The number (n) of phosphopeptides in each pattern is indicated. Clusters 5 and 6 contain phosphoproteins that are likely to be signatures of codon 12 and codon 13 KRAS mutations in colorectal cancer. (B) GO analysis indicates that proteins associated with the cytoskeleton and cell adhesion are significantly enriched in clusters 5 and 6. (C) Phosphopeptides and selected proteins are highlighted within a scatter graph of G12D versus G13D responses.
significant response was seen for the aldehyde dehydrogenase ALDH3A1 that is increased in G13D cells and decreased in G12D cells. We also note the presence in clusters 5 and 6 of the membrane trafficking and organizing proteins SEC23B, ANXA1, ANX3, and ANX11 and the increased expression of the stage-specific colorectal cancer biomarker SERPINB5 in G12D cells.38

A total of 274 phosphopeptides from 198 proteins representing 17% of the total number of unique sites were observed in clusters 5 and 6 of the phosphoproteome data set (Figure 3A). In total, 56 out of 274 phosphopeptides are associated with 27 proteins linked by GO analysis to cell adhesion or cytoskeletal function in clusters 5 and 6 of the phosphoproteome data set (Figure 3B). Among these, MAP1B and TJP2/ZO-2 are represented by multiple phosphopeptides (Figure 3C). Other notable phosphopeptide representatives of clusters 5 and 6 are the HGF-receptor MET Thr995 and Caveolin-1 Ser37 sites that both exhibit >10-fold increases in abundance in G12D versus G13D cells and the Ras effector BRAF Ser729 site that is decreased in G12D versus G13D cells. In the case of caveolin and MET, a component of this change is due to the higher levels of protein expression observed in G12D and G12V cells as judged by Western blotting (Figure 6).

Phosphopeptide members of clusters 5 and 6 that originated from proteins with multiple phosphorylation sites were curated to examine the patterns of response across all detected sites within these proteins (Figure 4 and Supporting Information Figure 3). Where available, proteome data are also presented (squares) to see the extent to which phosphopeptide responses were influenced by changes in protein abundance rather than a proportional increase in phosphorylation. In the majority of cases where comparisons could be made, proteome changes were a minor influence on phosphopeptide ratios. Interestingly, most phosphosites within a protein trended in a similar direction for both cluster 5 and cluster 6 members, indicating coordinated increase or decrease of phosphorylation at multiple sites within a protein.

Given the central role of kinases in mediating Ras responses and modulating phosphonetworks, we examined their contribution to our data sets. In total, we detected peptides from 38 kinase (Supporting Information Table 3). A total of 35 kinase phosphopeptides out of 96 were responsive (≥1.5-fold change compared to parental control) to the presence of at least one of the oncogenic Ras mutations (Figure 5A and Supporting Information Table 3). These included core growth factor receptor-Ras pathway members EGFR, MET, MAP2K2 (MEK2) and ERK2 as well as CDC42BPB, NEK9 and PAK4 (Figure 5A and B). Phosphopeptides from eight kinases were present in clusters 5 and 6 (Supporting Information Table 3). Among these is Thr185 that becomes phosphorylated during activation of ERK2 (Figure 5A). This phosphosite shows specific down-regulation in G13D vs codon 12 cell lines suggesting that KRAS G13D is impaired in its ability to activate ERK2. To investigate the wider context of the kinases regulating the phosphosites in clusters 5 and 6, we used NetworKIN analysis that integrates consensus substrate motifs and contextual modeling to predict potential kinases for each phosphosite.36 A significant number of cluster 5 and 6 members are potential targets of kinases that regulate the cell cycle and promote proliferation (cyclin-dependent, casein, MAP and MOK kinases; Figure 5C).

DCLK1 Up-Regulation Is Only Observed in Codon 12 Mutant KRAS Cells

Upon inspection of our data, we have chosen to follow up several proteins based on enrichment factor, biological relevance, and availability of reagents. Within our proteome data set we saw a number of proteins that were highly expressed in KRAS codon 12 mutant cell lines. The pre-eminent examples of this were doublecortin-like kinase-1 (DCLK1) and A-kinase anchor protein 12 (AKAP12) that were up-regulated at least 8-
fold in KRAS G12D versus parental cells (Supporting Information Table 1). The distinctive pattern of expression of these proteins was confirmed by Western blotting (Figure 6A) and was recapitulated in a second independent clone of each cell line (Supporting Information Figure 2A). All kinases identified in the proteome data set except DCLK1 are collated with their respective phosphopeptide values (B). MotifX analysis identified significantly over-represented linear phosphorylation motifs from the set of sites present in GProX phosphopeptide clusters 5 and 6. Candidate kinases that regulate these sites were predicted using NetworKIN. Analysis of all of the phosphosites in our data set is provided for comparison (long list). Average responses of the sites associated with each motif in response to the presence of KRAS mutations are indicated within the Output column. The total number of submissions used in the MotifX analysis and average response heatmap are indicated (n). Kinases that had constituent peptides or phosphopeptides detected in our data sets are labeled (●).

Our observation that DCLK1 is highly overexpressed in KRAS G12D cells is striking, given recent data that this protein is a colon cancer tumor stem cell marker.99 Up-regulation of DCLK1 is also seen in G12V cells (Figure 6A). To examine the extent to which this is a codon 12-specific phenomenon, we extended our analysis to a broader panel of isogenic SW48 cells expressing codon 12 variants. Strikingly, we observe that DCLK1 expression is significantly increased with all codon 12 mutants while codon 13 is equivalent to parental cells (Figure 6B). MET exhibits a corresponding pattern of up-regulation across the codon 12 mutant panel potentially indicating a common mechanism promoting up-regulation. QPCR analysis

Figure 5. Kinase responses and predicted kinase regulators of phosphosites associated with codon 12 versus codon 13 KRAS outputs. Peptide (blue) and phosphopeptide (red) responses of all kinases present in our data sets are depicted (A). All kinases identified in the proteome data set except DCLK1 are collated with their respective phosphopeptide values (B). MotifX analysis identified significantly over-represented linear phosphorylation motifs from the set of sites present in GProX phosphopeptide clusters 5 and 6. Candidate kinases that regulate these sites were predicted using NetworKIN. Analysis of all of the phosphosites in our data set is provided for comparison (long list). Average responses of the sites associated with each motif in response to the presence of KRAS mutations are indicated within the Output column. The total number of submissions used in the MotifX analysis and average response heatmap are indicated (n). Kinases that had constituent peptides or phosphopeptides detected in our data sets are labeled (●).
indicates that the increased levels of DCLK1 observed in codon 12 KRAS mutant cell lines are due to transcriptional up-regulation rather than via regulation of translation or protein stability (Figure 6C). The molecular weight of DCLK1 observed in our Western blotting experiments and the distribution of peptides identified by mass spectrometry indicate that this represents isoform 3 or isoform 4 of DCLK1. These consist of an active kinase domain but lack the N-terminal doublecortin-like domains required for binding to microtubules (Figure 6D).

To investigate whether transcriptional up-regulation of DCLK1 seen in G12D cells was directly KRAS-dependent rather than via a change in translation or protein stability (Figure 6C). The molecular weight of DCLK1 observed in our Western blotting experiments and the distribution of peptides identified by mass spectrometry indicate that this represents isoform 3 or isoform 4 of DCLK1. These consist of an active kinase domain but lack the N-terminal doublecortin-like domains required for binding to microtubules.

To investigate whether transcriptional up-regulation of DCLK1 seen in G12D cells was directly KRAS-dependent rather than via an increase in the expression of DCLK1 protein expression in KRAS G12D cells (Figure 7A). QPCR-based analysis of KRAS and DCLK1 transcripts revealed proportional reductions in KRAS and DCLK1 expression in G12D cells in response to KRAS knockdown (Figure 7B).

**DISCUSSION**

The combination of isogenic cell lines and large-scale quantitative proteomics has resulted in unprecedented depth of coverage of pathways specifically engaged by oncogenic KRAS variants. Our first and perhaps most striking observation was that each type of activating codon mutation specifies a distinct KRAS signaling output. This is an important insight because to date almost all Ras studies and Ras-related clinical trials have treated Ras mutations as being equivalent.

We were interested in the mechanisms by which KRAS codon 12 and codon 13 mutations may differentially impact upon cell status in CRC tumors. A total of 274 phosphopeptides and 115 proteins differentially responded to the presence of codon 12 versus codon 13 KRAS mutants. Numerous proteins that we have identified have prominent links to colon cancer or properties associated with malignant cells. Among these were the cell adhesion associated protein AKAP12 and the cell surface organizer Caveolin 1 that is a...
tumor suppressor implicated in regulating Ras signaling and KRAS mediated colorectal cancer cell migration.\(^{41-47}\) Both proteins exhibited higher expression in KRAS codon 12 mutant cells versus G13D. A similar pattern was also observed for the HGF receptor c-MET. MET has a well-established role in Ras-mediated tumorigenesis where it is up-regulated;\(^{48,49}\) further, MET has a well-established role in Ras cells versus G13D. A similar pattern was also observed for the

---

**Figure 7.** KRAS G12D drives transcriptional upregulation of DCLK1. The induced expression of two independent shRNAs specific for KRAS results in significant decreases in KRAS and DCLK1 protein (A) and proportional decreases in DCLK1 mRNA (B). A pan-Ras antibody is used in (A); the upper band of the doublet corresponds to KRAS (arrow).

---

Our analysis revealed DCLK1 to be the most amplified of all proteins in any of our KRAS mutant isogenic cells compared to wild type parentals, and this up-regulation is also reflected in the mRNA transcript levels. Importantly, this amplification is reversed upon suppression of KRAS expression. This indicates a continued direct role for KRAS, rather than an irreversible adaptive response, or selection pressure, in regulating DCLK1 expression.

Although DCLK1 is a relatively poorly understood kinase, we note that analysis of gene coexpression across almost 1000 cell lines reveals that the microtubule stabilizing protein MAP1B that is a prominent phosphosite responder in the codon 12 cells is among the top 20 nearest neighbor genes with DCLK1 suggesting functional cooperation between these proteins.\(^{52}\) DCLK1 is frequently overexpressed in colorectal cancer and associated with poor prognosis.\(^{53}\) A genome wide mutant KRAS synthetic lethality screen previously identified the related kinase DCLK2 as a stringent hit in colorectal DLD1 cells.\(^{54}\) These data suggest that DCLK1 is biologically relevant to colorectal cancer cell survival. Furthermore, DCLK1 is also a CRC tumor stem cell specific marker;\(^{39}\) ablation of DCLK1+ tumor stem cells results in regression of CRC polyps; however, there was no formal linkage with KRAS status established in these in vivo studies.

Overexpression of DCLK1 is seen with all variants of codon 12 KRAS mutant cells but not in G13D cells. However, this relationship is likely to be highly context-dependent because, in this case, interrogation of the Cancer Cell Line Encyclopaedia reveals no significant correlation between the presence of a codon 12 mutated KRAS allele and DCLK1 levels in a panel of 275 cancer cell lines or within the subset of 61 colorectal cell lines (Supporting Information Figure 5).\(^{52}\) Our data show expression of the short forms of DCLK1 containing the kinase domain but not the microtubule binding double cortin domains. Although none of the previous colorectal studies have discriminated between which DCLK isoforms are contributing to their results, data from studies of brain function reveal specific up-regulation of short C-terminal DCLK1 transcripts in adult brain that are associated with modulating memory and cognitive abilities.\(^{55,56}\)

Our study represents the first unbiased global screen of signaling pathways downstream of endogenous oncogenic KRAS. Our experimental approach enabled differences in outputs emanating from each KRAS mutant to be identified without the confounding effects of significant differences in genetic background. The majority of nodes within the immediate Ras signaling network displayed differential responses at the proteome and phosphoproteome level (Figure 8). The mechanistic basis for this is currently unclear; however, it vividly illustrates the importance of factoring precise mutation status into the designs and interpretation of experiments comparing Ras function. For example, several recent studies identified genes that are synthetically lethal when depleted or inhibited in cells harboring oncogenic KRAS\(^ {20,54,57-60}\) Each study used a different panel of cell lines with a variety of codon 12 or codon 13 mutations and responsiveness between cell types was inconsistent. Our data predict that synthetic lethality would likely vary, depending upon which specific mutation is present, and suggest that an isogenic cell line approach will be important for identifying contingencies of drug responsiveness on mutation status.
In summary, we have found that each of the three main KRAS mutations generates a distinct signaling network signature and proteome expression profile. Furthermore, we have demonstrated that a key collection of genes with known functions in promoting oncogenic colorectal cancer signaling and tumorigenesis exhibit codon-specific KRAS dependence for their expression and/or phosphorylation. Among these is the colon cancer stem cell marker and kinase DCLK1. Our analysis revealed that a key collection of genes with known functions in promoting oncogenic colorectal cancer signaling and tumorigenesis exhibit codon-specific KRAS dependence for their expression and/or phosphorylation. Among these is the colon cancer stem cell marker and kinase DCLK1. Our analysis provides fundamental insights into basic Ras biology with significant implications for the design and interpretation of large-scale studies of oncogenic Ras signaling across cell panels.

ACKNOWLEDGMENTS

This work was supported by European Union Seventh Framework Programme funding (grant no. 259015 COLThERES) and by the Wellcome Trust. We would like to thank Chris Torrance at Horizon Discovery for his initial conceptual input and continued support. We also thank Andy Jones and Da Qi at the University of Liverpool for help with publicly accessible data deposition.

REFERENCES


Author Contributions

These authors contributed equally to this study.

Notes

The authors declare the following competing financial interest(s): Julie Wickenden is an employee of Horizon Discovery.


