Supplementary Methods

Expression and purification of full-length STK33

STK33 (NM_030906) was cloned incorporating a TEV-cleavable, N-terminal His7–tag into a pFastBac vector and expressed using the Bac-to-Bac baculovirus/insect cell expression system (Invitrogen). This pFastBac construct, MAH7-TEV-STK33 (amino acids 2-514), was generated by PCR using an Incyte Collection cDNA clone (LIFESEQ90205785, Open Biosystems) as template. The resulting recombinant baculovirus stock was used to infect *Trichoplusia ni* cells (Orbigen) in ExCell 405 Medium (Sigma) with incubation at 27°C for 50 hours post-infection before the cells were harvested by centrifugation. The protein was purified by passing the cell lysate through a Ni-NTA column, subjecting the eluate to TEV protease cleavage, Ni-NTA subtraction, and Superdex 75 size-exclusion chromatography (SEC). Two STK33 peaks were fractionated on the SEC column consistent with the 60 kDa monomer and 120 kDa dimer. The peak containing the 60 kDa STK33 was used for the high throughput screen to identify small molecule STK33 kinase inhibitors.

High throughput siRNA screens

SiRNAs from Qiagen Inc. (Valencia, CA) or from Thermo Scientific (Dharmacon Products, Lafayette CO) were used to create libraries with 4-8 siRNAs for each gene (there were 8 siRNAs for KRAS and for STK33). Each siRNA was individually transfected into cells using Lipofectamine RNAiMAX transfection reagent (Life Technologies, Carlsbad CA). SiRNAs from a library plate were diluted in serum-free media to a volume of 6 µl. Transfection reagents diluted in serum-free media to a volume of 5 µl were added to each well using a BiomekFx Robot (Beckman Coulter). After a 20-minute room temperature incubation, cells were added to the plates using a Multidrop (ThermoScientific). After 96 or 120 hours, cell viability was determined with CellTiterGlo™ (Promega, Madison, WI) and luminescence was measured on a luminometer according to the manufacturer's instructions. The final siRNA concentrations (10-30 nM) and RNAiMAX volume used per well (0.02-0.1 µl) and plating cell density (500-1500 cells/well) varied by cell line. Most cell lines were screened using multiple transfection conditions. Results from the viability assays were

processed through Screener® (Genedata, Basel Switzerland). Viability measurements for the siRNAs for each gene were compared to the remainder of the siRNA using Stouffer's method (details below) in order to generate a p-value for each gene.

Calculation of gene-based p-values

Genes were evaluated for their role in cell viability/proliferation by comparing the siRNA targeting the gene of interest (e.g. STK33 or KRAS) to a reference set of siRNA targeting predominantly non-essential genes (generated as described below). P-values were calculated using the Inverse normal method of Stouffer (Hedges 1985). Briefly, each siRNA's effect on viability was ranked against the reference set, and this ranking was converted to a fraction by dividing by (N+1), where N is the number of siRNA in the reference set. Then, the fractions for all the siRNA targeting a given gene were combined using Stouffer's method, resulting in a p-value for each gene. The Stouffer method converts the fractions to Z scores and pools the Z scores using the formula

$$Z = \frac{\sum_{i=1}^{k} Z_i}{\sqrt{k}}$$

where k is the number of siRNA targeting the gene of interest, $Z_i = \Phi^{-1}(f_i)$, where Φ^{-1} is the inverse of the standard normal cumulative distribution function, f_i is the rank of $siRNA_i$ targeting the gene of interest expressed as a fraction, and Z is the Z score for the gene. The Z score for each gene is then converted to a p-value using the standard normal cumulative distribution function $p = \Phi(Z)$.

In large scale siRNA screens, the entire siRNA library is typically used as the reference set, since the number of true hits in the library is normally expected to be low. However, if a library is enriched for essential genes, the statistical power will be diminished. To generate a reference set with fewer true positives, we preferentially removed siRNA targeting essential genes from the reference set as follows. For each siRNA in the library, we assigned what we termed an orthogonal gene average (OGA) p-value. The OGA p-value is identical to the gene p-value described above, except that it is assigned to a siRNA rather than a gene, and instead of being calculated using every siRNA against

the gene of interest, it uses only the siRNA other than the siRNA under consideration. For example, to calculate the OGA p-value for a $siRNA_x$ that targets $gene_y$, every siRNA targeting $gene_y$ is used in the calculation except for $siRNA_x$ itself. This results in the OGA p-value for $siRNA_x$ being correlated with the on-target effects of $siRNA_x$ on $gene_y$, while being completely uncorrelated with the off-target effects of $siRNA_x$. Removal of siRNA from the reference set based on OGA p-value can therefore be used to preferentially remove essential genes without significantly changing the distribution of siRNA targeting non-essential genes in the reference set. In the experiments shown, the OGA p-values were used to calculate a false discovery rate (FDR) by the Benjamini-Hochberg method, and siRNA were removed from the reference set if FDR < 0.5. The top 0.5n siRNA with the highest OGA p-values were also removed, where n is the number of siRNA with FDR < 0.5, so that siRNA were removed from both the high and low OGA p-value tails. The OGA method was also used to generate a reference set on each individual plate in the library for plate normalization purposes.

Quantitation of RNA knockdown and cell viability for HCT-116 and MDA-MB-231 cells

Cells were transfected in reverse, in duplicate 96-well plates, using a Bravo robotic work station (Agilent Technologies). Briefly, 3µl 1 µM siRNA (Qiagen, Valencia, CA) diluted with 7 µl serum-free media was incubated with either 0.1 µl/well or 0.2 µl/well RNAiMAX (Life Technologies, Carlsbad CA) in an equal volume of serum-free media for 20 min in a 96-well plate. Four thousand cells per well in 80 µl volume were added using a Multidrop™ (ThermoScientific) giving a final siRNA concentration of 30 nM. RNA expression levels were quantified from one of the two 96-well plates 24 hours after transfection using the QuantiGene branched DNA assay as noted in the Materials and Methods section. Viability was quantified from the other 96-well plate 96 hours after transfection using CellTiterGlo™ (Promega, Madison, WI) according to the manufacturer's protocol.

The STK33 probe set detects RNA spanning nucleotides 738-1408 (relative GenBank REFSEQ NM_030906). The KRAS probe set detects RNA spanning nucleotides 668-

1521 (relative GenBank REFSEQ NM_004985). The HPRT probe set detects RNA spanning nucleotides 530-1048 (relative GenBank REFSEQ NM_000194). The cyclophilin B (PPIB) probe set detects RNA spanning nucleotide 69-629 (relative GenBank REFSEQ NM_000942).

Supplementary Table 1. siRNA Information

Gene symbol (Entrez Gene ID#)	Name	Vendor (product ID)	Target sequence
KRAS (3845)	KRAS_1Q	Qiagen (custom)	CCCGGTCCTTAGGTAGTGCTA
	KRAS_2Q	Qiagen (custom)	ACCTATGGTCCTAGTAGGAAA
	KRAS_3Q	Qiagen (custom)	TTACATAGACTTAGGCATTAA
	KRAS_4Q	Qiagen (custom)	AAGAATAGTCATAACTAGATT
	KRAS_5Q	Qiagen (SI03101903)	GACGATACAGCTAATTCAGAA
	KRAS_6Q	Qiagen (SI03106824)	GTGGACGAATATGATCCAACA
	KRAS_7Q	Qiagen (SI02634191)	CTCCTAATTATTGTAATGTAA
	KRAS_8Q	Qiagen (SI02662051)	AAGGAGAATTTAATAAAGATA
	KRAS_238.ST	Invitrogen (HSS105871, Stealth)	GACGATACAGCTAATTCAGAATCAT
STK33 (65975)	STK33_1Q	Qiagen (SI00139741)	CAGTGGTTAACAGGCAATAAA
	STK33_2Q	Qiagen (SI00139748)	ATGTGTCTAGTGCATCCTTAA
	STK33_6Q	Qiagen (SI02660203)	TCCATAAGTGACTGTGCTAAA
	STK33_7Q	Qiagen (SI02660210)	GAGCATAGGCGTCGTAATGTA
	STK33_3D	Dharmacon (D-005383-03)	CTAAGGAACTACTAGATAA
	STK33_4D	Dharmacon (D-005383-04)	TCACAGACATCAAGCATTG
	STK33_5D	Dharmacon (D-005383-05)	ACATAAAGGTGACTGATTT
	STK33_6D	Dharmacon (D-005383-06)	GAGCATAGGCGTCGTAATG
PLK1 (5347)	PLK1_6	Qiagen (SI02223837)	CCGGATCAAGAAGAATGAATA
	PLK1-4	Qiagen (SI00071638)	CCCGAGGTGCTGAGCAAGAAA
EIF4A3 (9775)	DDX48	Qiagen (custom)	AAAGAGCAGATTTACGATGTA
PPIB (5479)	Cyclophilin B	Qiagen (custom)	TGGTGTTTGGCAAAGTTCT
n/a	Random1_4	Qiagen (custom)	AAGAGGGACATAAAGCTTTCA
n/a	Random2_4	Qiagen (custom)	AACGCAGAGTTCGACCGTTTA
n/a	Random6_1	Qiagen (custom)	AAGGGCAACATCAAGGTTTAT
n/a	M13_4	Qiagen (custom)	AATGCGCTTCCCTGTTTTTAT
n/a	Luciferase	Qiagen (custom)	ACGTACGCGGAATACTTCG
n/a	Scrambled	Invitrogen (12935-300, Stealth)	Proprietary

n/a, not applicable

Supplementary Figure Legends

Figure S1. SiRNA oligonucleotides STK33_1Q and STK33_7Q siRNAs Reduce STK33 RNA in NOMO-1 and SKM-1 cells. STK33 RNA levels were determined in each group by bDNA assay 24 hours after transfection (same transfection as Figure 1B and 1D – see methods) of control siRNAs (scrambled - scram, cyclophilin B – cycl B, luciferase - lucif) or with siRNAs against STK33, KRAS or PLK1. Cyclophilin B RNA levels were also determined as a control. The data are expressed as a ratio of STK33 levels over cyclophilin B levels (arbitrary units ± standard deviation, n=2).

Figure S2. Efficient Knockdown of STK33 and KRAS with siRNAs in HCT-116 Cells. A, STK33 and KRAS RNA levels were determined 24h after siRNA transfection in HCT-116 cells using a bDNA assay. Data are expressed as a ratio of STK33 levels over cyclophilin B levels. B, In a replicate 96-well plate, the effects of the siRNA on cell viability were quantified using a homogeneous luminescent assay for ATP (CellTiterGlo™). Arrows indicate off-target activities of some STK33 siRNAs. All experiments were done in triplicate and the data are expressed as mean ± standard deviation.

Figure S3. Efficient Knockdown of STK33 and KRAS with siRNAs in MDA-MB-231 Cells. A, STK33 and KRAS RNA levels were determined 24h after siRNA transfection in MDA-MB-231 cells using a bDNA assay. Data are expressed as a ratio of STK33 levels over cyclophilin B levels. B, In a replicate 96-well plate, the effects of the siRNA on cell viability were quantified using a homogeneous luminescent assay for ATP (CellTiterGlo™). Arrows indicate off-target activities of some STK33 siRNAs. All experiments were done in triplicate and the data are expressed as mean ± standard deviation.

Figure S4. STK33 RNA and Protein Levels are Variable across Cell Lines. A, STK33 levels were determined by western blotting. The same cell lysates were used to determine the levels of S6K (p85 and p70) and RPS6 phosphorylation on a separate gel. B, STK33 RNA expression levels in breast, pancreas and colon cancer cell lines. STK33 RNA levels were determined by a bDNA assay from cell lysates. For each line, the expression of STK33 was normalized to the level of HPRT and the ratio is reported as arbitrary units ± standard deviation (n=2). Cell lines that carry a KRAS mutation are marked with an asterisk.