STAR Methods section

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-FLAG	Sigma	F3165
Anti-V5	Life technologies	R96025
Anti-YAP1	Cell Signaling	14074
Anti-H3K27 Acetylated	Cell Signaling	8173
Anti-SRP19	Abcam	Ab50932
Anti-PIK3CA	Cell signaling	4249
Anti-CTNNB1	Cell signaling	8480
Anti-H2A-K5-Acetylated	Cell signaling	2527
Anti-H4-K8-Acetylated	Cell signaling	2594
Anti-H3-K9-Acetylated	Cell signaling	9649
Anti-SMARCB1	Cell signaling	11966
Anti-MAP2K1	Cell signaling	12671
Anti-MED12	Cell signaling	14360
Anti-ZNF217	Abcam	ab124927
Deposited Data		
Classification of β-catenin activity in 1032 cancer cell lines	This study	Table S2
mRNA expression following CRISPR mediated gene suppression	This study	Table S5
Proliferation following CRISPR mediated gene suppression.	This study	Table S7
Affinity-based mass spectrometry performed with 57 genes	This study	Table S8
Genetic interactions following simultaneous CRISPR-mediated deletion of 52 genes	This study	Table S9

Proliferation following shRNA mediated gene suppression in 216 cancer cell lines	Project Achilles	https://portals.broadinstitute.org/achilles
Protein interactions curated from publicly available databases	Inweb	http://www.cbs.dtu.dk/suppl/dgf/heart_developmental_networks/
Experimental Models: Cell Lines		
DLD1	ATCC	CCL-221
HT29	ATCC	HTB-38
HCT116	ATCC	CCL-247
RKO	ATCC	CRL-2577
SW480	ATCC	CCL-228
MIAPACA2	ATCC	CRL-1420
Recombinant DNA		
TCF luciferase reporter	Addgene	24308
Sequence-Based Reagents		
Primers for Illumina sequencing of double CRISPR library	This study	Table S10
Primers for qPCR of WNT target genes	This study	Table S10
Seq of pooled sgRNA library	This study	Table S6
Software and Algorithms		
ICS algorithm	This study	Data S1

Contact for reagents and resource sharing

The Lead Author: William C. Hahn
Dana-Farber Cancer Institute
450 Brookline Avenue
Dana 1538
Boston, MA 02215
william_hahn@dfci.harvard.edu
617-632-2641

Experimental models

All cell lines used in this study were maintained in a 37°C incubator with 5% CO_2. The following cell lines were maintained in DMEM containing 10% of Fetal Bovine Serum (FBS); DLD1, HT29, SW480, HCT116, GP2D, MIAPACA2, NCIH1975 and RKO. The following cell lines were maintained in RPMI containing 10% of FBS: LS411N, LS513.

Methods details

β-catenin activity reporter assay. Cell lines were generated by infection with a lentiviral β-catenin/TCF4 reporter (Fuerer and Nusse, 2010). Following puromycin selection (2 µg/ml) 50,000 cells were plated on a 96 well plate and 24 h later luciferase activity was measured using the Luc-Screen detection kit (Applied Biosystems).

CRISPR-Cas9 loss-of-function screen. Lentiviral particles containing the pooled sgRNA library were transduced at low MOI (4 replicates/cell line) into 10 Cas9-expressing cell lines (DLD1, LS411N, LS513, HCT116, SW480, HT29, GP2D, RKO, MIAPACA2, NCI-H1975). Each sgRNA was stably integrated into at least 600 cells. In addition, parental DLD1 cells that did not express Cas9 were also infected with the same library. DNA extracted at 3 or 28 d post infection was used for massively parallel sequencing as described (Shalem et al., 2014).

Combinatorial CRISPR-Cas9 loss of function screen. To minimize recombination events between repetitive sequences, we inserted a variant U6 promoter downstream of the U6 promoter in pXRP003 (Supplemental Fig. 4A). To generate a pooled double sgRNA library, 115 sgRNAs were individually PCR amplified together with the *S. Pyogenes* tracer sequence and inserted into position 1 (AgeI/EcoRI restriction sites). Vectors containing an sgRNA in position 1 were pooled and digested with BsmBI (Thermo Scientific) and a pool of the same 115 sgRNAs was ligated into the BsmBI cloning sites (exactly as described above for single sgRNA pooled library). Following ligation the library was electroporated into Stbl4 cells (Life Technologies) grown at 30^{o}C. Lentiviral particles containing the combinatorial CRISPR-Cas9 library were transduced (3 replicates/cell line) into 4 Cas9-expressing cell lines (DLD1, HCT116, HT29 and RKO). Following puromycin selection (2 µg/ml), genomic DNA (5µg) extracted at 3 or 28 DPI was PCR amplified using NEBNext® High-Fidelity 2X PCR Master Mix (New England Biolabs) and F/RPCR1_2sg primers (Supplementary Table 10). This mixture (5µl) was used for a second PCR amplification using barcoded-staggered primers (Supplementary Table 10). Paired end Illumina sequencing was used for sequencing of double sgRNAs. The forward and reverse sequences were aligned to the original sgRNA sequences using Bowtie I suite (Langmead et al., 2009).

Analysis of double CRISPR-Cas9 screen. sgRNA combinations with less then 50 reads at t=0 were discarded. Read counts from sgRNA combinations were combined into one combination score and normalized using equation 1: Y = (Read count + 1)/(total read count) normalized read counts were then normalized to control containing sgRNA combinations using equation 2: $Z = \mathrm{Log}_2((Y_{combo}/Y_{Control})*1*10^6)$. The fold change of every combination was calculated using equation 3: $FC_{combo} = Z_{combo\_t=0} - Z_{combo\_t=28}$. For calculating the genetic interaction score we used a previously reported S score that takes into account both consistency and magnitude (Collins et al., 2006) $S = (FC_{Experimental} - FC_{Calculated})/\mathrm{sqrt}(S_{Var}/N_{Experimental} + S_{Var}/N_{Calculated})$ where $S_{Var} = (var_{Experimental} \times (N_{Experimental} - 1) + var_{Calculated} \times (N_{Calculated} - 1))/(N_{Experimental} + N_{Calculated} - 2)$.

Generation of stable Cas9-expressing cell lines. WT Cas9 was cloned into a lentiviral vector driven by EIFα promoter containing blasticidine resistance (pLX311, (Doench et al., 2014)). Following infection and blasticidine selection Cas9 expression was confirmed using an anti-FLAG antibody (Sigma). Cas9 activity was evaluated by measuring GFP fluorescence following introduction of a GFP targeting sgRNA (Doench et al., 2014). Cas9-expressing cell lines were maintained in blasticidine containing media (10 µg/ml).

mRNA profiling following CRISPR-mediated gene suppression. DLD1 cells stably expressing Cas9 were infected with sgRNAs targeting the indicated genes. 8 DPI RNA was extracted using RNeasy kit (Qiagene) and cDNA libraries for massively parallel sequencing were prepared as previously described (Shishkin et al., 2015). Barcoded samples were sequenced on a HiSeq2500 machine (Illumina). Sequence alignment and transcript quantification was obtained using the RSEM software package (Li and Dewey, 2011).

Construction of lentiviral pooled sgRNA library. 6-10 sgRNA targeting exons 1 or 2 of the indicated genes (Supplementary Table 5) were selected using sgRNA design tool (http://crispr.mit.edu/). 96 well plates containing forward and reverse oligonucleotide sequences (Supplementary Table 5) were obtained from Integrated DNA Technologies (IL, USA) and resuspended in 50 µl of TE. Oligonucleotides were annealed by incubation at 95^{o}C for 5 min and then rapidly cooled on ice. Oligonucleotide pairs were pooled and a 1:10 dilution was used for ligation into the BsmBI restriction sites of the sgRNA expressing vector (pXRP003 (Sanjana et al., 2014)). Electrocompetent DH5α cells (Life technologies) were used for propagation and expansion of the library.

Draft PPI. V5 tagged ORFs were cloned into a lentiviral expression vector containing a puromycin selection marker and an EIF1α promoter (pLX307 (addgene catalog number 41392)). DLD1 cells were grown for 2 weeks in SILAC media containing either 15N4,13C6 arginine/15N2,13C6 lysine (R10/K8 “heavy”) or their natural occurring analogs (R0/K0 “light”), supplemented with 10% dialyzed fetal bovine serum (Ong et al., 2002). 3.75x106 R10/K8 SILAC-labeled cells (1 x 10 cm dish) per bait were infected with a V5-tagged ORF. DLD1 cells cultured in R0/K0 SILAC “light” media were infected with non-tagged GFP virus and used as control. Following 24 h incubation in 37oC cells were treated with puromycin (2 µg/ml) for 48 h and were cultured for 5 additional days on ice. Cells were harvested by addition of ice-chilled ModRIPA buffer (low stringency buffer LS) containing 1% Nonidet P-40, 0.1% Na deoxycholate,150 mM NaCl, 1 mM EDTA, 50 mM Tris, pH 7.5, and protease inhibitors (Complete tablets, Roche Applied Science, Indianapolis, IN). Lysate protein concentration was assayed using the A660 assay (Pierce, Rockford, IL) according to the manufacturer’s instructions. 400 µg of protein from light and heavy lysates were immunoprecipitated using anti-V5 magnetic beads (MBL, Woburn, MA) overnight at 4oC with 800 rpm shaking on a ThermoMixerRT. All subsequent steps were performed using an Agilent BravoLH automated liquid handler equipped with a magnetic capture attachment in parallel on all samples simultaneously. The supernatant was aspirated, and the beads were washed first in lysis buffer, then in detergent-free lysis buffer. Heavy and light lysates from each bait were combined before the final wash. Immunoprecipitation eluates were denatured on beads with 2 M urea/50 mM Tris pH 8.0 and then predigested with sequencing-grade modified trypsin (V5113; Promega) in a presumed (expecting 5-10% yield) enzyme-to-substrate ratio of 1:50 for 30 min at room temperature with on a shaker. The supernatant was collected, and protein disulfide bonds of the combined lysates were reduced for 45 min with 5 mM DTT (no. 20291; Thermo Scientific). Cysteines were subsequently alkylated for 45 min with 10 mM iodoacetamide. Samples were digested overnight at room temperature with trypsin in a 1:50 enzyme-to-substrate ratio on a shaker. Peptide mixtures were acidified to a final volumetric concentration of 1% trifluoroacetic acid prior to reversed phase desalting. Subsequent biochemical steps were conducted using an Agilent BravoAM liquid hander equipped with the AssayMAP system. The digested protein complexes were desalted using Agilent RP-S cartridges on the AssayMAP Bravo head according to the recommended procedures. Desalted samples were then separated by strong cation

exchange chromatography at pH cuts of 3.5, 7, and 11, also using the AssayMAP Bravo head with SCX cartridges (different pH solvents were formulated empirically using formic acid/ammonium formate mixtures in 80% acetonitrile). The resulting fractions were again desalted using AssayMAP RPS cartridges. Samples were dried down by vacuum centrifugation and reconstituted in 5 µL 3% acetonitrile/0.1% formic acid. Digests (2 µL) were separated on a Proxeon Easy nLC 1000 UHPLC (2-hour gradient, 15-20 cm x 75 µm column at 50^{o}C packed with 1.9 µm ReproSil (Dr. Maisch, GMBH)) coupled to a Thermo Q-Exactive mass spectrometer. A "Top 12" data acquisition method was employed. Quantitative enrichment ratios for all identified proteins were determined utilizing the SILAC-based design of the experiment with the protein bait in the heavy channel and an empty vector in the light channel. MS raw files were processed for protein identification and quantitation using the MaxQuant software package, version 1.2.2.5. MS/MS spectra were searched against the human Uniprot database downloaded on 27 June 2012 (complete isoforms included). The precursor mass tolerance used in the search was 7 ppm and fragment mass tolerance was 20 ppm. Carbamidomethylation of cysteines was searched as a fixed modification and oxidation of methionines and acetylation of protein N termini were searched as variable modifications. Enzyme specificity was set to trypsin and a maximum of two missed cleavages was allowed for searching. Only proteins with a minimum of 2 quantifiable peptides were included in our dataset. The false positive rate for protein and peptide identification was 1%, as determined using a decoy database. SILAC ratios were corrected by empirically measured "heavy" incorporation ratios, false positive and contaminant hits were filtered out, and log2-transformed ratios were re-centered around the median. Proteins were rank-ordered as potential interactors by their corrected ratios.

Interaction Credibility Scoring (ICS). Interaction Credibility Scoring (ICS) uses the Random Forest (RF) methodology (Breiman, 2001) to utilize both network architectural metrics and the quantitative proteomics metrics to predict real physical interactions amongst the 58 baits (Supplementary Table 8). Known physical interactions were obtained from In-Web (Lage et al., 2007) and Consensus Path-DB (Kamburov et al., 2011) databases. We used two well established edge metrics from graph theory: 1) The edge-betweenness centrality (Girvan and Newman, 2002), which we use to measure of how many shortest paths between proteins a physical interaction exists on; and 2) the Jaccard metric (Jaccard, 1912), which we use a measure of shared interactions partners between respective proteins. The third predictor in the model is the median adjusted (by-bait) heavy-to-light ratio, which is an adjusted version of the primary quantitative metric from SILAC based proteomic experiments. The binary response model for assigning interaction probabilities is defined as Y=[1:known physical interaction from databases; 0: not found in gold standard set from databases]. For each of the 15,206 potential physical interactions, we determine whether they were found in the gold-standard set of known interactions (i.e., the response of the model), and also compute each of the edge architectural metrics after integrating the potential interactions into the larger network of known interactions. Thus, the model is designed to use the adjusted heavy-to-light ratio along with the edgebetweenness and Jaccard measure to try and differentiate between known gold standard interactions found in the bait-prey observations and those that are not known gold-standard interactions. After building the model, we can thereby use the three metrics for each potential new interaction and assign it a probability that it is a true physical interaction. The random forest model is trained on 70% of this data and the Area Under the ROC Curve (AUC) is used to assess the classification power using the 30% holdout. The AUC for the ICS method, 96.9 C.I. (95.7,98.1), was shown to be significantly better than just using the corrected heavy-to-light ratio to predict gold-

standard interactions (Supplementary 3B). The number of trees is a parameter of the RF method and here we used 5,000 trees. The R code used for ICS including instructions of how to run ICS could be found in supplemental data S1.

BAF complex immunoprecipitation. Cells were lysed and homogenized on ice in Buffer A (10 mM HEPES (pH 7.6), 25 mM KCl, 1 mM EDTA, 10% glycerol, 0.1% NP-40, 1 mM DTT and protease inhibitors (Roche) supplemented with 1 mM PMSF)). Nuclei were sedimented by centrifugation (1,000g), re-suspended in 600ul Buffer C (10 mM HEPES (pH 7.6), 3 mM MgCl2, 100 mM KCl, 0.1 mM EDTA, 10% glycerol, 1 mM DTT and protease inhibitors) and lysed by addition of 0.3M ammonium-sulfate. Soluble nuclear proteins were separated from the insoluble chromatin fraction by ultracentrifugation (100,000g, 10mins) and precipitated with 0.3 mg/ml ammonium-sulfate for 20 min on ice. Protein precipitate was isolated by ultracentrifugation (100,000g, 10 mins) and re-suspended in IP buffer (150 mM NaCl, 50 mM Tris-HCl (pH 8.0), 1% NP-40, 0.5% deoxycholate, 1 mM DTT and 1 mM PMSF with protease inhibitors). 1 mg of lysate was incubated with 50µl of Dynabeads® Protein A (Life Technologies) that was pre-incubated with the indicated antibody at 4^{o}C for 16 hours. Beads were washed three times in buffer C and eluted by addition of 2x LDS sample buffer (Life Technologies).

Competition assays. GFP expressing cell lines were infected with control sgRNAs and cell lines stably expressing Cas9 were infected with the indicated sgRNAs. Following Puromycin selection control and experimental sgRNA expressing cell lines were mixed at a 1:1 ratio. Proliferation was monitored using FACS analysis.

Data and Software availability

All datasets generated in this study can be found in Data S1 (http://dx.doi.org/10.17632/jmxfmzkt6f.1). The source code and instructions of how to run ICS can also be found in Data S1 (http://dx.doi.org/10.17632/jmxfmzkt6f.1).