Supplemental Figure Legends

Figure S1: Characterization of β-catenin co-dependencies that scored in RNAi screens (related to Fig. 1). (A) β-catenin activity measured in 5 cell lines using a β-catenin/TCF4 reporter. **(B)** β-catenin expression signature was used to classify 561 tumor samples. Continuous variable permutation was used to identify mutations that correlate with β-catenin activity.

Figure S2: Classification of β-catenin co-dependency candidates (related to Fig. 2). (A) Comparison of the proliferation changes following suppression of frequently mutated genes that scored as β-catenin co-dependencies (red dots represent cell lines harboring mutations in the indicated oncogene). **(B)** Diagram describing CYCLOPS genes. **(C)** The dependency and CN profile of β-catenin-associated CYCLOPS genes. Quantification of immunoblot images shown in Fig. 2E **(D)** or 2F **(E).** Error bars represent the standard deviation of two independent replicate experiments.

Figure S3: CRISPR-Cas9-mediated suppression of β-catenin co-dependencies (related to Fig. 3). Protein and transcript levels of **(A)** β-catenin or **(B)** MAP2K1 following expression of β-catenin- or MAP2K1-targeting sgRNAs in DLD1 cells. **(C)** Pearson correlation was used to assess replicate consistency in CRISPR loss-of-function screen. **(D)** Proliferation changes of 148 negative control sgRNAs (targeting non-human genes) using CRISPR-Cas9 (black circles) or a similar sized pool of shRNAs (gray circles). We note that due to normalization, the negative controls score with a positive score. **(E-H)** Proliferation changes following CRISPR-Cas9-mediated deletion of known oncogenes. Distribution of Pearson correlations in **(I)** CRISPR-Cas9 or **(J)** shRNA screens, between sgRNAs or shRNAs targeting different sequences of the same gene (Inter-gene, light line) or correlation among all sgRNA or shRNAs (dark line) are shown. sgRNAs targeting the same gene are more correlated than shRNAs targeting the same gene. **(K-N)** Distribution of proliferation changes following CRISPR-Cas9-mediated deletion of candidate CYCLOPS genes from (Wang et al., 2015) or **(O-P)** following extensive mutagenesis in haploid cell lines (Blomen et al., 2015).

Figure S4: Draft-PPI credibility scoring (ICS) (related to Fig. 4). (A-D) Distribution of PPI parameters used for scoring of draft-PPI interactions. **(E)** ROC curve plot of PPI identified by

draft-PPI and scored using ICS or heavy/light ratio **(F)** 5-fold cross validation was used to evaluate the prediction power of the ICS modal. **(G)** GAPDH Ct values measured in DLD1 following CRISPR-Cas9-mediated deletion of *TRIP4* or β-catenin.

Figure S5: Combinatorial CRISPR-Cas9-mediated deletion of β-catenin co-dependencies (related to Fig. 5). (A) Depiction of vector for cloning of double sgRNAs. **(B)** Replicate correlation across cell lines used in combinatorial CRISPR-Cas9 screen. **(C)** Proliferation changes induced by double sgRNAs targeting control (non-human) genes. **(D)** Comparison between the proliferation changes induced by expression of an individual sgRNA or a combination of control and gene targeting sgRNA. **(E)** Observed and expected proliferation phenotypes of combinations containing two different sgRNAs targeting the same gene (diagonal in Fig. 6A). Protein levels of CTNNB1 **(F)** or YAP1 expression of 1 or 2 targeting sgRNAs in HT29 cells. **(G)** Protein levels of MED12 and YAP1 in HT29 cells following CRISPR-Cas9 mediated deletion.

Figure S6: GI and proteomic profiling identify a role for YAP1 in regulation of histone modifications (related to Fig. 6). (A) Protein interactions identified by draft-PPI between ZNF217 and components of the HDAC CoRest complex. **(B)** Protein levels of acetylated histone H3 (K27) following CRISPR-Cas9-mediated deletion of *ZNF217*. (C) Quantification of immunoblots in Fig. 6E. Error bars represent the standard deviation of two replicate experiments.

Supplemental Tables

Table S1: Genes used in β-catenin activity expression signature. Symbols of genes used in the β-catenin activity expression signature.

Table S2: β-catenin activity in 1032 CCLE cancer cell lines (http://www.broadinstitute.org/ccle/home). β-catenin activity score generated using the β-catenin gene expression signature.

Table S3: β-catenin co-dependencies in RNAi loss-of-function screen. 177 genes that score as β-catenin co-dependencies using shRNA loss-of-function screen in 216 cancer cell lines.

Table S4: β-catenin co-dependencies associated with loss of expression. Pearson correlation obtained by comparing the shRNA dependency score and expression or CN across all 216 cancer cell lines or within β-catenin active cell lines.

Table S5: Changes in RNA following CRISPR-Cas9-mediated suppression. RSEM values in DLD1 cells 8 DPI with sgRNAs targeting the indicated genes.

Table S6: Sequences of sgRNAs used in this study.

Table S7: Proliferation changes induced by CRISPR-Cas9-mediated gene inactivation. The median of four replicates measuring the proliferation change (Log_2[fold change]) following expression of the indicated sgRNAs.

Table S8: Protein interactions identified by draft-PPI.

Table S9: Proliferation changes induced by Double CRISPRs. The sum of proliferation changes in combinations that contain a control + sgRNA was used for the calculated proliferation changes. The GI score was calculated using the S score (Collins et al., 2006).

Table S9: Primer sequences. Sequence of primers used in this study.

Data S1: Folder containing code and instructions of how to run ICS.

Supplemental Experimental Methods

Datasets used and analyzed.

The following datasets were generated in the current study:

1. Proliferation following CRISPR-Cas9 mediated gene suppression (Table S7).
2. Global expression changes following CRISPR-Cas9 mediated deletion of 15 genes (Table S5).
3. Affinity-based mass spectrometry performed with 57 genes (Table S8).
4. Genetic interactions following simultaneous CRISPR-mediated deletion of 52 genes (Table S9).

The following previously published datasets were generated in previous studies and reanalyzed in the current study.

1. Genome scale shRNA proliferation screens in 216 cancer cell lines (Project Achilles; http://www.broadinstitute.org/achilles/).
2. Inweb protein interaction dataset (http://www.cbs.dtu.dk/suppl/dgf/heart_developmental_networks/).

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 re-suspended in 50 µl of TE. Oligonucleotides were annealed by incubation at 95°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 level 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 $^{15}N_4$,$^{13}C_6$ arginine/$^{15}N_2$,$^{13}C_6$ lysine (R10/K8 "heavy") or their natural occurring analogs (R0/K0 "light"), supplemented with 10% dialyzed fetal bovine serum (Ong et al., 2002). $3.75x10^6$ 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 37°C 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 4°C 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 RP-S 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 edge-betweenness 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 (1,000*g*), resuspended 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 (100,000*g,* 10mins) and precipitated with 0.3 mg/ml ammonium-sulfate for 20 min on ice. Protein precipitate was isolated by ultracentrifugation (100,000*g, 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) at 4°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.

Literature Cited

Blomen, V.A., Majek, P., Jae, L.T., Bigenzahn, J.W., Nieuwenhuis, J., Staring, J., Sacco, R., van Diemen, F.R., Olk, N., Stukalov, A.*, et al.* (2015). Gene essentiality and synthetic lethality in haploid human cells. Science *350*, 1092-1096.

Breiman, L. (2001). Random Forests. Machine Learning *45*.

Collins, S.R., Schuldiner, M., Krogan, N.J., and Weissman, J.S. (2006). A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. Genome biology *7*, R63.

Doench, J.G., Hartenian, E., Graham, D.B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B.L., Xavier, R.J., and Root, D.E. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nature biotechnology *32*, 1262-1267.

Girvan, M., and Newman, M.E. (2002). Community structure in social and biological networks. Proceedings of the National Academy of Sciences of the United States of America *99*, 7821-7826.

Jaccard, P. (1912). The distribution of the flora in the alpine zone. The New Phytologist *11*.

Kamburov, A., Pentchev, K., Galicka, H., Wierling, C., Lehrach, H., and Herwig, R. (2011). ConsensusPathDB: toward a more complete picture of cell biology. Nucleic acids research *39*, D712-717.

Lage, K., Karlberg, E.O., Storling, Z.M., Olason, P.I., Pedersen, A.G., Rigina, O., Hinsby, A.M., Tumer, Z., Pociot, F., Tommerup, N.*, et al.* (2007). A human phenome-interactome network of protein complexes implicated in genetic disorders. Nature biotechnology *25*, 309-316.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC bioinformatics *12*, 323.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Molecular & cellular proteomics : MCP *1*, 376-386.

Sanjana, N.E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. Nature methods *11*, 783-784.

Shishkin, A.A., Giannoukos, G., Kucukural, A., Ciulla, D., Busby, M., Surka, C., Chen, J., Bhattacharyya, R.P., Rudy, R.F., Patel, M.M.*, et al.* (2015). Simultaneous generation of many RNA-seq libraries in a single reaction. Nature methods *12*, 323-325.

Wang, T., Birsoy, K., Hughes, N.W., Krupczak, K.M., Post, Y., Wei, J.J., Lander, E.S., and Sabatini, D.M. (2015). Identification and characterization of essential genes in the human genome. Science *350*, 1096-1101.