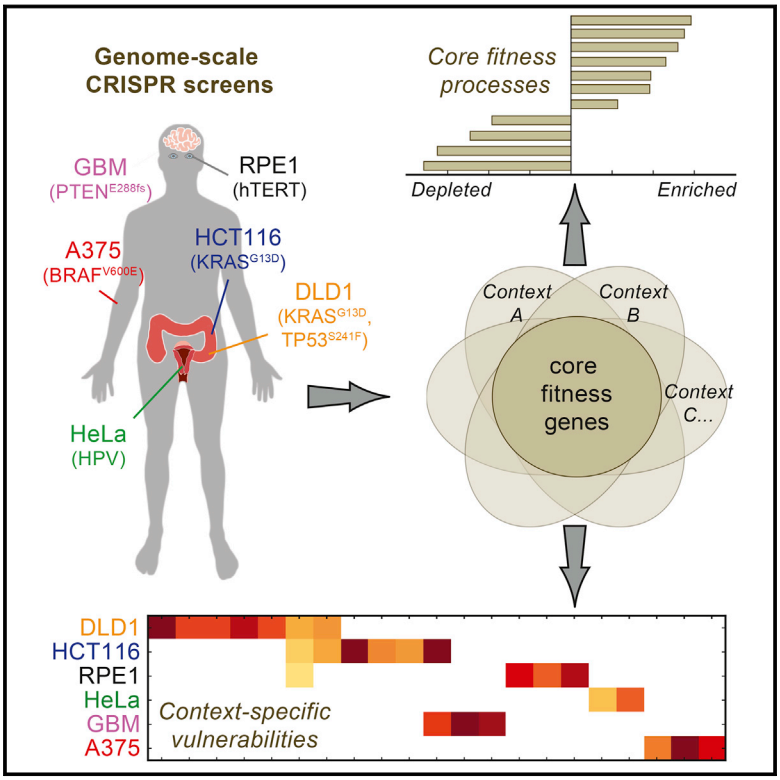
Resource

Cell

High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities

Graphical Abstract



Highlights

- Identification of 5-fold more fitness genes in human cells than previously observed
- Core fitness genes are highly enriched for ancient protein complexes
- Context-specific fitness genes illuminate biological differences between cell types
- Distinct genetic signatures can be used to predict differential drug response

Authors

Traver Hart, Megha Chandrashekhar, Michael Aregger, ..., Daniel Durocher, Stephane Angers, Jason Moffat

Correspondence

j.moffat@utoronto.ca

In Brief

CRISPR knockout screens for essential genes reveal oncogenic drivers specific to different cancer cell lines as well as unexpected metabolic and signaling dependencies that may guide future therapeutic targeting.

Accession Numbers

GSE75189

Hart et al., 2015, Cell *163*, 1515–1526
December 3, 2015 ©2015 Elsevier Inc.
http://dx.doi.org/10.1016/j.cell.2015.11.015

Resource

Cell

High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities

Traver Hart,[1,12] Megha Chandrashekhar,[1,2,12] Michael Aregger,[1,12] Zachary Steinhart,[3] Kevin R. Brown,[1] Graham MacLeod,[3] Monika Mis,[3] Michal Zimmermann,[5] Amelie Fradet-Turcotte,[5,13] Song Sun,[1,2,7,9] Patricia Mero,[1] Peter Dirks,[2,8] Sachdev Sidhu,[1,2] Frederick P. Roth,[1,2,5,9,10,11] Olivia S. Rissland,[2,6] Daniel Durocher,[2,5] Stephane Angers,[3,4] and Jason Moffat[1,2,10,*]

[1]Donnelly Centre, 160 College Street, Toronto, ON M5S3E1, Canada
[2]Department of Molecular Genetics
[3]Department of Pharmaceutical Sciences and Leslie Dan Faculty of Pharmacy
[4]Department of Biochemistry
University of Toronto, Toronto, ON M5S1A1, Canada
[5]The Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, ON M5G1X5, Canada
[6]Molecular Structure and Function Program, The Hospital for Sick Children Research Institute, 686 Bay Street, Toronto, ON M5G0A4, Canada
[7]Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala SE-75123, Sweden
[8]Program in Developmental and Stem Cell Biology, Division of Neurosurgery, Hospital for Sick Children, 555 University Avenue, Toronto, ON M5G1X8, Canada
[9]Department of Computer Science, University of Toronto, Toronto, ON M5G1X8, Canada
[10]Canadian Institute for Advanced Research, Toronto, ON M5G1Z8, Canada
[11]Center for Cancer Systems Biology, Dana-Farber Cancer Institute, Boston, MA 02215, USA
[12]Co-first author
[13]Present address: CHU and St-Patrick Research Group in Basic Oncology, Laval University, Quebec City, QC G1R3S3, Canada
*Correspondence: j.moffat@utoronto.ca
http://dx.doi.org/10.1016/j.cell.2015.11.015

SUMMARY

The ability to perturb genes in human cells is crucial for elucidating gene function and holds great potential for finding therapeutic targets for diseases such as cancer. To extend the catalog of human core and context-dependent fitness genes, we have developed a high-complexity second-generation genome-scale CRISPR-Cas9 gRNA library and applied it to fitness screens in five human cell lines. Using an improved Bayesian analytical approach, we consistently discover 5-fold more fitness genes than were previously observed. We present a list of 1,580 human core fitness genes and describe their general properties. Moreover, we demonstrate that context-dependent fitness genes accurately recapitulate pathway-specific genetic vulnerabilities induced by known oncogenes and reveal cell-type-specific dependencies for specific receptor tyrosine kinases, even in oncogenic KRAS backgrounds. Thus, rigorous identification of human cell line fitness genes using a high-complexity CRISPR-Cas9 library affords a high-resolution view of the genetic vulnerabilities of a cell.

INTRODUCTION

The Human Genome Project has yielded a fairly complete catalog of cellular components, and a major goal moving forward will be to classify all genetic elements involved in normal biological processes and disease (Lander, 2011). With advances in gene editing enabled by the CRISPR-Cas system (Cho et al., 2013; Cong et al., 2013; Jinek et al., 2012; Mali et al., 2013), it is no longer quixotic to seek a comprehensive picture of cellular circuitry for human cells.

Yeast research has helped to establish a conceptual framework for systematic genetics, providing a glimpse of the basic modular organization of a cell and the corresponding genetic landscape (Costanzo et al., 2010). Importantly, the basis for many advances in yeast genetics was the discovery and classification of essential and non-essential genes through systematic genetic knockout studies (Giaever et al., 2002; Winzeler et al., 1999). Yet, this binary classification was later demonstrated to be an artifact of laboratory conditions—while some genes do indeed seem to be constitutively essential, many others revealed fitness defects only under environmental stress or in different genetic backgrounds (i.e., context-dependent essentiality) (Hillenmeyer et al., 2008; Ramani et al., 2012).

The distinction between core and context-dependent essential genes is particularly relevant in humans, where a typical cell expresses perhaps two-thirds of the genome's complement of genes (Hart et al., 2013). Distinguishing a tissue's essential genes, and delineating them from the constitutive essentials shared across all tissues, may hold the key to unlocking tissue-specific disease. In tumors, this is the foundation for the concept of synthetic lethality (Hartwell et al., 1997): genes essential in tumor cells but not in adjacent normal tissues should make ideal therapeutic targets with high effectiveness and minimal side effects. However, identifying these context-specific

essentials has been no easy feat. To date, this field has been constrained by technology: RNA interference (RNAi) has been the best available tool but, despite improvements in reagent design and analytical approaches, its utility is limited by imperfect mRNA knockdown and confounding off-target effects (Echeverri et al., 2006; Moffat et al., 2007).

Moreover, incomplete knockdown adds a further complication to negative selection growth screens using pooled library shRNA reagents, where the final readout is typically a fold-change relative to an initial population. While a strong negative fold change implies a growth defect, the scale of the defect is difficult to quantify: a 50% reduction in proliferation rate could be driven by anything from partial knockdown of a gene whose total loss of function results in arrested or inviable cells to full knockdown of a gene with only a 50% fitness defect. In this study, we define a fitness gene as any gene whose perturbation causes a proliferation defect; essential genes are, by definition, a subset of fitness genes.

Pooled CRISPR-Cas approaches present a major opportunity for systematic classification of human genetic elements into functional categories and biological processes. Proof-of-concept genome-wide knockout screens in human cell lines (Shalem et al., 2014; Wang et al., 2014) suggested that mammalian cells harbored far more fitness genes than RNAi experiments had previously uncovered (Hart et al., 2014). Screens for genes whose loss of function enables drug and toxin resistance (Koike-Yusa et al., 2014; Shalem et al., 2014; Wang et al., 2014), accelerates metastasis (Chen et al., 2015), or influences immune response (Parnas et al., 2015) confirm the vast improvement that the CRISPR technology offers in mammalian genetic screens.

We recently described the ''Daisy Model'' of gene essentiality (Hart et al., 2014), wherein the set of fitness genes for each cell line or tissue context are conceptually represented by the petals of a flower. Petals overlap to varying degrees, but all contexts share a common core of fitness genes that should be present in every cell. We sought to harness the power of CRISPR screens to expand our knowledge of both core and context-dependent fitness genes by generating a high-complexity, second-generation lentiviral library targeting virtually all human protein-coding genes. We used the library to screen five human cell lines, representing a cross-section of immortalized and cancer tissues, to identify genes whose knockouts induced significant fitness defects. This work lays the foundation for understanding the genetic interaction network that underlies human cells at very high resolution and how genetic and epigenetic variation across cell types gives rise to differential genetic vulnerabilities that may represent therapeutic opportunities in cancer and other diseases.

RESULTS

High-Resolution Detection of Human Cell Line Fitness Genes with an Ultra-complex CRISPR Library

We sought to confirm the utility of CRISPR as an improved screening technology to extend the catalog of both the core and context-dependent human fitness genes. In this study, we define a fitness gene as any gene whose perturbation decreases cell growth and proliferation. We relied on lessons learned from first-generation CRISPR genetic screens (Shalem et al., 2014; Wang et al., 2014) and other pooled screening approaches to design a high-complexity, genome-scale guide RNA (gRNA) library targeting human protein coding genes and random controls. Briefly, we took the set of reference essentials and nonessentials from Hart et al. (2014) and identified all gRNAs in the Shalem et al. (2014) data targeting these genes. As expected, the fold-change distribution of all guides targeting essential genes (''essential gRNAs'') is strongly shifted relative to the distribution of guides targeting nonessential genes (Figures S1A and S1B). To identify sequence biases that differentiated high-performing essential gRNAs from low-performing ones, we measured the frequency of each base at each position in the guide of the top quintile of essential gRNA relative to the bottom quintile. We noted a strong bias against uridine in the last four positions of the guide sequence (Figure S1C), consistent with a similar analysis performed by Wang et al. using their own library of sequences targeting ribosomal genes (Wang et al., 2014). Thus, we excluded all candidate guide sequences with U at those positions, as well as guides with GC content < 45% or > 70% (Figure S1D) based on findings in Wang et al. (2014). Finally, to minimize potential off-target effects, we excluded all candidate gRNAs with more than one genomic off-target site within two mismatches of the guide-plus-PAM sequence (Figure S1E).

From the remaining pool of candidate gRNAs, we selected the library based on a two-pass approach. For each gene, coding exons were ordered 5′ to 3′ and the candidate guides for each exon identified. For the first pass, only guides with no genomic off-target sites were used. We iterated over the exons in order and randomly selected one guide per exon. The process was repeated until six guides were assigned, or until there were no more candidate guides (Figure S1F). The resulting library targeted 17,232 genes with a modal number of six guides per gene (Figure S1G). In addition, this library also includes control sequences targeting *LacZ*, *EGFP*, and luciferase, as well as two sets of gRNAs targeting random loci on Chr10: 584 with high specificity (no predicted off-target sites), and 796 ''promiscuous gRNAs'' with 20 or more predicted perfect binding sites throughout the genome. These sequences were designed to sample the phenotypic effects of off-target Cas9 cleavage. In total, this library contains 91,320 sequences, hereafter referred to as the ''90k library.''

For the second pass, the process was repeated, this time allowing candidate gRNA with one potential off-target hit (Figure S1F). We iterated over exons in order until up to 12 guides were assigned to each gene. This supplemental library, in conjunction with the 90k library, brings the total to 176,500 sequences targeting 17,661 protein-coding genes (Figure S1H), which we call the TKO (Toronto KnockOut) library (Table S1).

We validated effectiveness of designed guides in Cas9-expressing cell lines by targeting several genes individually and testing protein knockdown and cellular proliferation defects (Figures S2A–S2D). As a proof of concept, we also performed a positive selection screen with HCT116 cells cultured in an excess of thymidine to look for suppressors of the imposed G1/S arrest, wherein we recovered 11 of 12 gRNAs targeting thymidine kinase 1 (*TK1*) (Figures S2E–S2H), as expected (Meuth, 1989).

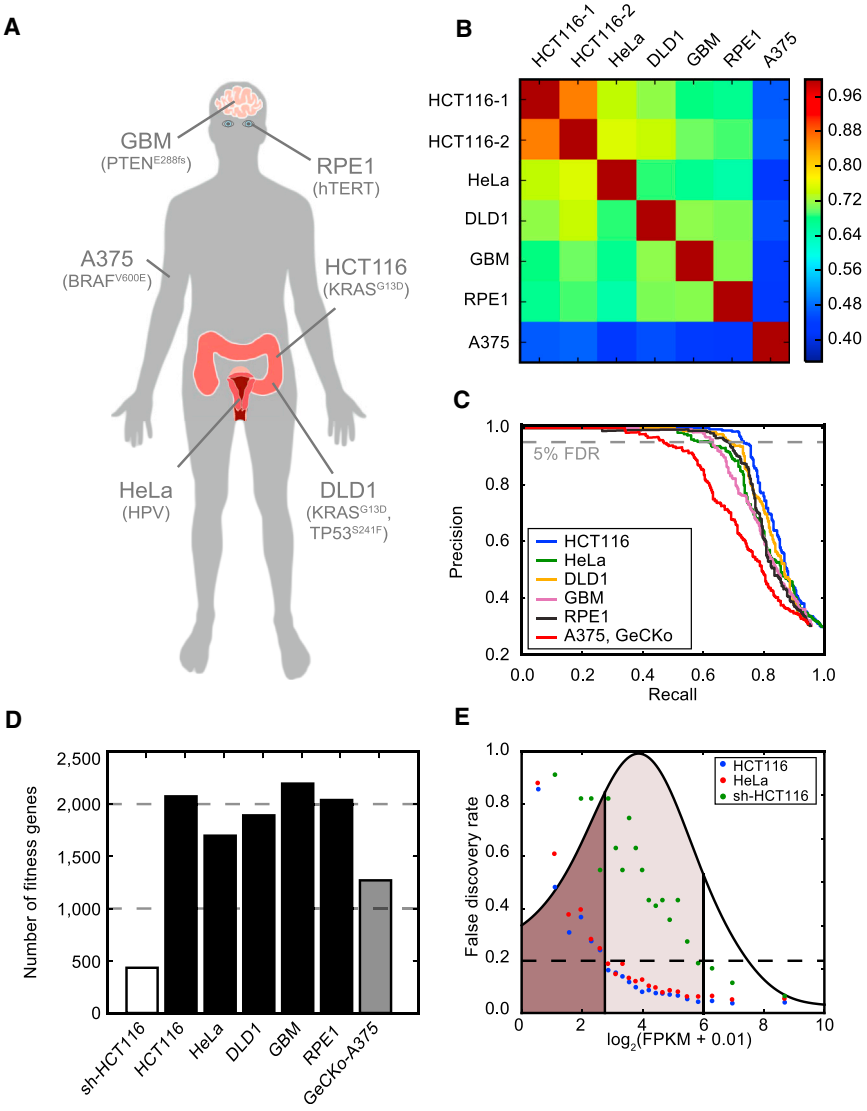


Figure 1. CRISPR-Cas9 Fitness Screens

(A) Cell lines screened and/or analyzed for this study. GBM, patient-derived glioblastoma cell line. RPE1, retinal epithelial cells. HCT116 and DLD1, colorectal carcinoma. HeLa, cervical carcinoma. A375, melanoma, screened in Shalem et al. (2014) and reanalyzed here.

(B) Heatmap of Pearson correlation coefficients among all pairs of screens.

(C) Precision-recall plots of fitness genes for each cell line. Dashed line denotes 5% false discovery rate (FDR).

(D) Number of fitness genes detected in each cell line at 5% FDR. Black, CRISPR screen with TKO library. White, pooled library shRNA screen. Gray, CRISPR screen with GeCKo library.

(E) False discovery rate of screens versus expression level of target genes. Genes are ranked by expression level and binned (n = 500). The false discovery rate (13) of fitness gene calls in each bin (y axis) is plotted against the mean expression level ($\log_2$(fpkm+0.01)) for genes in the bin. CRISPR targets genes at moderate expression levels (light shaded region) at low FDR, while pooled library shRNA achieves this FDR only for high expression genes (white region).

To identify human fitness genes in cell lines, we applied the 177k TKO library to HCT116 colorectal carcinoma cells stably expressing the Cas9 nuclease. After selection for infected cells and population sampling at time zero (T0), we conducted parallel negative selection screens for essential genes and sampled the evolving populations every 3 days from day 6 to day 18, or ~20 cell doublings. Genomic DNA from each time point was isolated and gRNA abundance was measured by deep sequencing of the integrated gRNA cassettes in order to monitor the change in abundance of each gRNA between the initial cell population and each of the subsequent five time points (Figures S2E and S2F). gRNAs targeting fitness genes were expected to drop out of the initial population and result in a lower proportion of total sequences, while the proportion of gRNAs targeting non-fitness genes is maintained. We repeated the screens in triplicate in HCT116 cells and also screened HeLa cervical carcinoma cells with the 177k library in triplicate. Using the same approach, we further screened DLD1 colorectal carcinoma cells, a patient-derived glioblastoma cell line (GBM) and hTERT immortalized RPE1 retinal epithelial cells with the first-pass 90k library (Figure 1A).

We adapted the analytical pipeline described previously in Hart et al. (2014) to classify fitness genes and to evaluate the quality of the screens. Using the gold-standard sets of 360 essential and 927 nonessential genes defined in that study, we observed that the fold-change distribution of gRNAs targeting essential genes is significantly shifted relative to those targeting nonessential genes, and that the shift increases with time (Figures S3A–S3E). We then made improvements to the algorithm described in (Hart et al., 2014), which we call the Bayesian Analysis of Gene Essentiality ("BAGEL") approach, to calculate a log Bayes factor (BF) for each gene. More positive scores indicate higher confidence that a given gene's knockout causes a decrease in fitness but are not necessarily a measure of the severity of the phenotype. Unlike most other analytical methods for large-scale gene perturbation studies, BAGEL uses the data from all reagents and all samples in a screen and therefore provides a ready framework for the integration of experiments with many reagents per gene and multiple time points (see Supplemental Information).

The TKO Library Screens Identify More Fitness Genes than Any Other RNAi or CRISPR Screen

For each of the screens that we carried out, BFs were calculated for all targeted genes (Table S2). We also used BAGEL to calculate

BFs for a previous screen in the A375 melanoma cell line performed with the first generation 65k gRNA GeCKo library (Shalem et al., 2014). To evaluate the consistency of our results, we measured the Pearson correlation coefficient of the BF distributions of all pairs of screens (Figure 1B). All TKO library screens showed very high correlation, with most having correlation coefficients between 0.65 and 0.75, while the replicate HCT116 screens had a correlation coefficient of 0.86. The A375 screen showed lower but consistent correlation with the TKO library screens, suggesting that the BAGEL algorithm could be used to estimate screen quality across different libraries and experiments.

The use of gold-standard reference sets enables the unbiased evaluation of screen performance. After merging the replicate HCT116 screens, we calculated precision-recall curves for the six cell lines (Figure 1C). All TKO library screens showed very high performance, with ~2,000 fitness genes identified in each cell line at a 5% false discovery rate (FDR; Figure 1D). Both 90k and 177k library screens showed similar sensitivity, with HeLa cells yielding the lowest number of fitness genes despite using the larger library, possibly due to high ploidy in HeLa cells. Overall, TKO library screens identified ~40% more fitness genes than the GeCKo screen in A375 (n = 1,270 genes), likely due to the combination of improved design and larger library size, but all CRISPR screens captured 4- to 5-fold more fitness genes than a high-quality genome-wide pooled library shRNA screen in the HCT116 cell line previously performed in our lab (Vizeacoumar et al., 2013), in which only 433 fitness genes were identified at the same FDR. The key driver for the difference in sensitivity appears to be expression level of the target gene: lentiviral-mediated pooled library shRNA screens show high specificity only for high expression targets, while CRISPR screens accurately detect fitness genes across a broad range of expression levels (Figures 1E and S3I).

Functional Characterization of Core Fitness Genes

A common aim of fitness screens in cell lines is to identify those genes whose loss of function is lethal only in a specific genetic background or environmental context. These ''synthetic lethal'' or context-specific essentials must be differentiated from core essentials, which should in principle be detectable in every screen for cell line fitness genes. We previously introduced the concept of the Daisy Model of gene essentiality (Hart et al., 2014), where the fitness genes in each tissue or context are represented by a petal (Figure 2A). Petals overlap to varying degrees but all share a common set of core genes that should be detectable in every cell line. We estimated the core fitness genes by considering the frequency with which each fitness gene was observed across the five TKO cell lines. Of the 4,054 total hits, 1,580 were observed in three or more of the five cell lines, with more than half (829, 52%) observed in all five (Figure 2B).

We defined these 1,580 hits as core fitness genes and analyzed their functional and genomic properties. Perhaps the most striking property was the strong enrichment of protein complexes among this gene set. The number of fitness genes among subunits of selected large (Figure 2C) and small (Figure 2D) protein complexes is compared to the number of subunits among 823 essentials identified across 72 genome-scale shRNA screens of cancer cell lines (Hart et al., 2014). In virtually every case, the five CRISPR screens identified more subunits of a given complex than all 72 shRNA screens, and several complexes are identified exclusively by CRISPR. On average, CRISPR classified 68% of complex members as hits versus only 35% for RNAi. The only exception where RNAi outperformed CRISPR was the ribosome, whose subunits tend to be targeted by a smaller number of guides in the TKO library (Figure S3J).

Among the complexes detected exclusively by TKO, 111 subunits are classified as fitness genes; of these, over 60% (n = 68; Figure 2E) are expressed at moderate levels, where RNAi is less sensitive. Notably, fitness genes that are highly expressed, but still found only by CRISPR, are enriched for genes whose proteins are localized to mitochondria. It was previously observed that some nuclear-encoded mitochondrial proteins have relatively short mRNA half-lives (Tani et al., 2012). Comparing high-expression fitness genes detected by RNAi and CRISPR in HCT116 cells to mRNA half-life data from HeLa cells (Tani et al., 2012) shows that those detected only by CRISPR tend to have much shorter half-lives than those detected by RNAi (Figure 2F), which may partially explain the lack of sensitivity of RNAi screens in detecting these important genes linked with mitochondria.

The Daisy Model predicts that core fitness genes are enriched for fundamental cellular processes such as transcription, translation, and replication. Analyzing the core fitness genes for biological process annotation enrichment confirms this prediction: genes involved in RNA splicing, translation, and DNA replication and repair are more than 4-fold overrepresented among these genes (Figure 2G). Conversely, genes involved in cell-cell communication and organismal development are more than 4-fold depleted, consistent with expectations when screening homogeneous cell cultures.

We then examined a number of physiological and evolutionary properties of the core fitness genes (Figures 2H and 2I). First, we observed that they show strong overlap with human orthologs of both yeast cell and fly cell-line essentials, as well as whole-organism worm and mouse essentials (Figure 2H and Supplemental Information). In contrast, non-core fitness genes (observed in 1–2 cell lines) include less than half as many essential orthologs. Second, core fitness genes show more than 10-fold higher average expression across the ~1,000 cell lines in the Cancer Cell Line Encyclopedia (Barretina et al., 2012) than nonfitness genes and display lower expression variance (Figure 2I), indicating that these genes exhibit constitutive, invariant expression across tissues. Third, and as with previous studies in model organisms, core fitness genes show roughly 4-fold more interaction partners in protein-protein and functional interaction networks. Fourth, core fitness genes also show a lower ratio of nonsynonymous to synonymous mutations when compared to all mouse orthologs, indicating purifying selection on fitness genes. Consistent with this observation, core fitness genes also show lower frequency of deleterious variation across human exomes. However, core fitness genes are neither enriched nor depleted for disease genes, contrary to the depletion for disease genes observed in the smaller set of RNAi-derived, high-expression core fitness genes described in Hart et al. (2014). Lastly, while core fitness genes have similar transcript length and coding sequence length to non-fitness genes, they tend to have

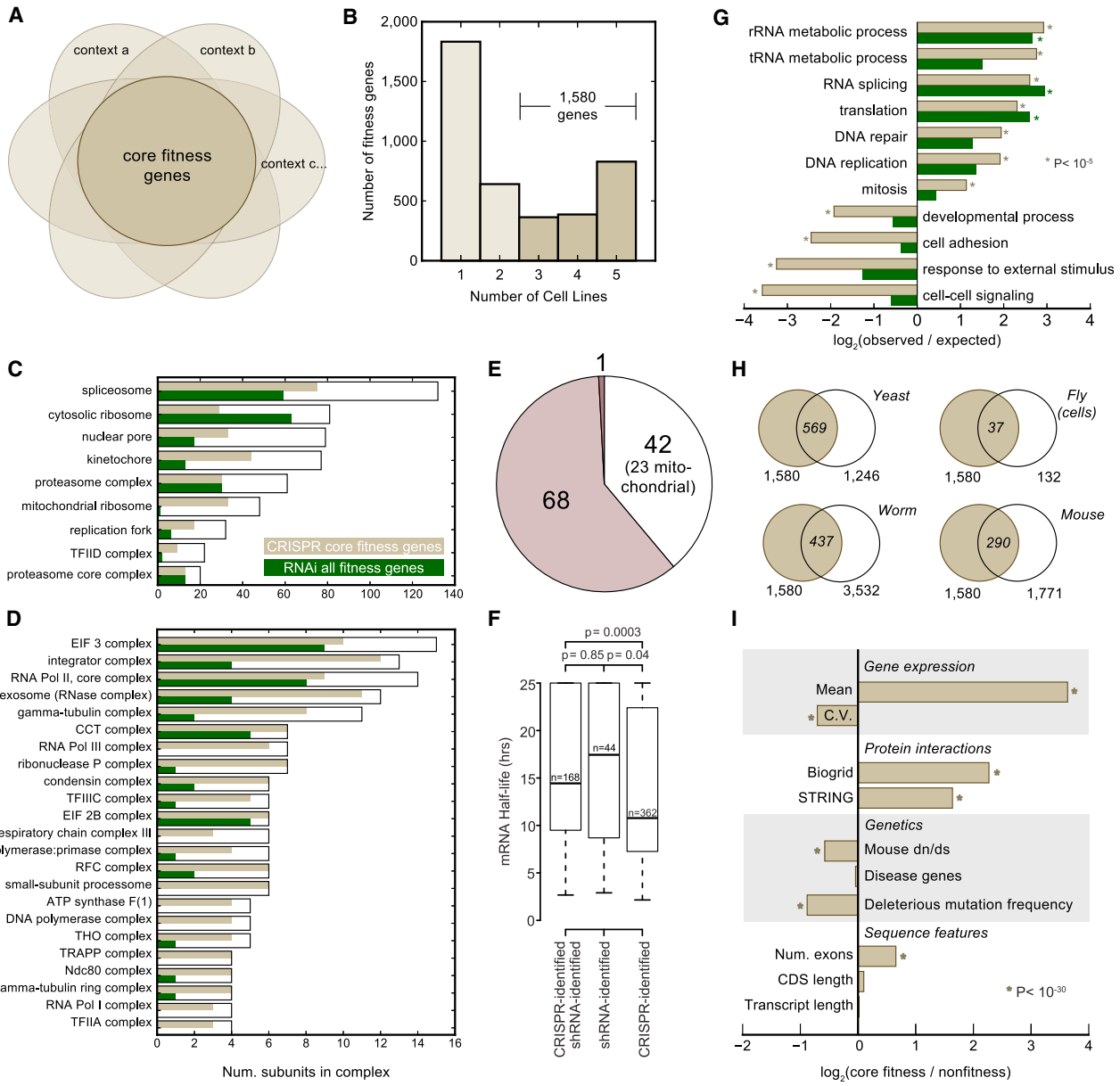


Figure 2. Core Fitness Genes

(A) The Daisy Model of gene essentiality, where each tissue/context is represented by a petal on the flower. Petals overlap to varying degrees but all share a common core of housekeeping essential genes.

(B) Distribution of hits at 5% FDR across the five cell lines assayed with the TKO library. We define hits observed in 3 or more cell lines as core essential genes (n = 1,580).

(C) Essential protein complexes. Number of subunits in the core TKO fitness genes (beige) is compared to the number of subunits in the set of 823 global RNAi essentials from (15) (green) and the total size of the complex (white box).

(D) Small essential complexes, shaded as in (C).

(E) For complexes detected exclusively by TKO (not RNAi), nearly 2/3 show moderate gene expression (shading as in Figure 1E). Highly expressed genes are enriched for mitochondrial proteins.

(F) Distribution of mRNA half-life of fitness genes detected by CRISPR and/or RNAi.

(G) Selected biological processes enriched or depleted in TKO core essentials (beige), depicted as fold-change relative to random expectation. Green bars are functional enrichment/depletion in RNAi global essentials (13).

(H) Intersection of core fitness genes with human orthologs of essential genes in yeast, worm, fly, and mouse.

(I) Other physiological and evolutionary properties of core fitness genes, plotted as fold-change of core fitness genes relative to non-fitness genes.

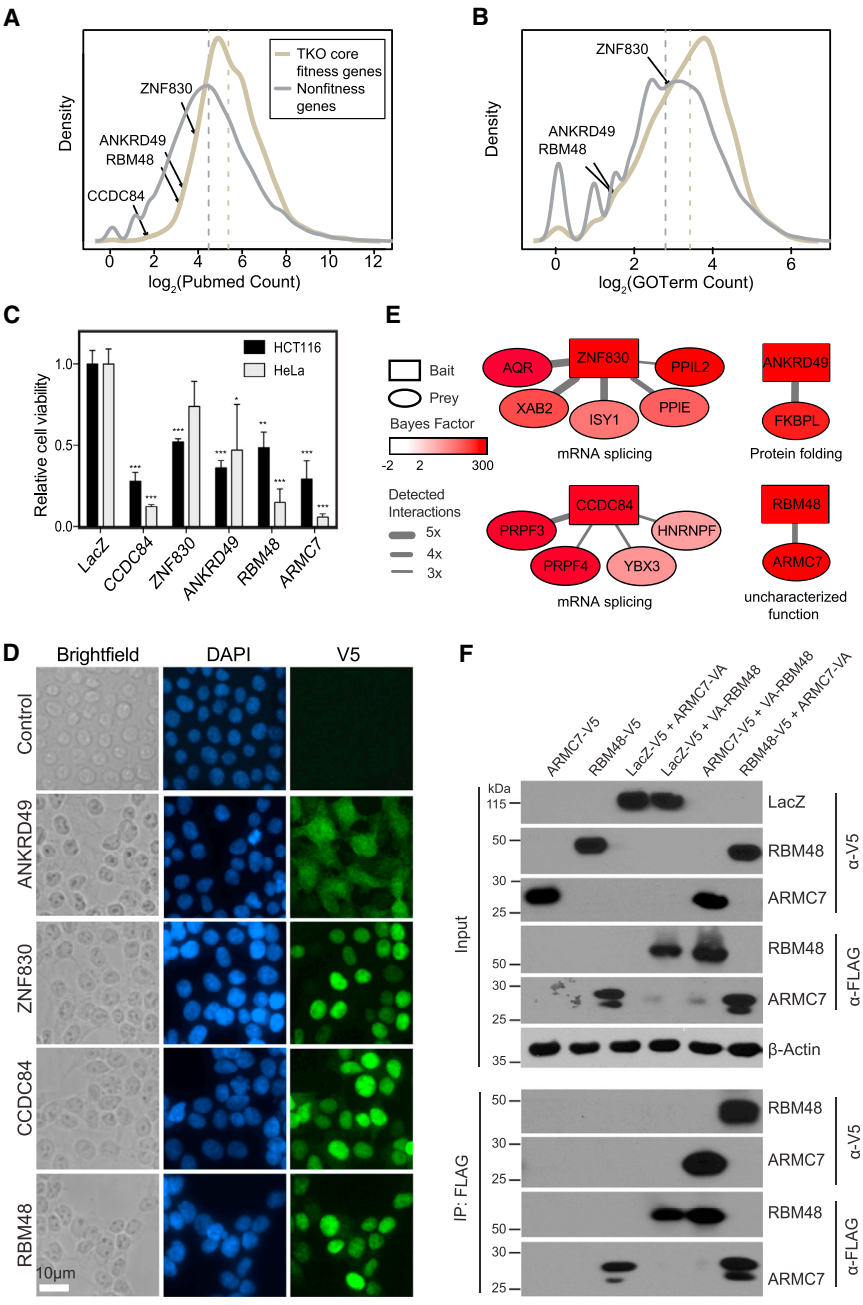


Figure 3. Novel Uncharacterized Fitness Genes

(A and B) Distributions of the number of PubMed citations (A) or GO terms (B) linked to each gene symbol in Entrez Gene for core fitness genes (gold line) versus non-fitness genes (gray line). Dotted lines indicate median value for each distribution. Position of selected uncharacterized fitness genes in the distribution is shown. *CCDC84* had no associated GO terms.

(C) Cell viability assays in HCT116 and HeLa cells infected with lentiviral-based gRNA expression cassettes, as measured by AlamarBlue staining 10 days post-selection. Three different gRNA expression cassettes were used per gene. Data are represented as means ± SD of 3 replicates (n = 3 cassettes x triplicates). ***p < 0.001, **p < 0.01, and *p < 0.05; two-tailed unpaired t test.

(D) Immunofluorescence of 293T cells stably expressing V5-tagged ORFs. Subcellular localization of ORFs was detected with an anti-V5 antibody (green), and nuclei were stained using DAPI (blue).

(E) Summary of the protein-interaction network of ANKRD49, ZNF830, CCDC84, and RBM48 detected by AP-MS analysis. V5- or 3xFLAG-tagged ORFs of the indicated proteins were immunoprecipitated from HEK293T and HCT116 cells, and interactors were identified by mass spectrometry analysis. Bait proteins are indicated by rectangles and prey proteins by circles. The color of the node corresponds to the Bayes Factor essentiality score in HCT116 cells, and the width of the edges corresponds to the number of times the interaction was detected. Results from five independent AP-MS experiments are shown for ANKRD49 and ZNF830 and four independent AP-MS experiments for CCDC84 and RBM48.

(F) Co-immunoprecipitation of RBM48 and ARMC7 from HCT116 cell lysates. HCT116 cells stably expressing V5-tagged RBM48, ARMC7, or LacZ (control) were transfected with the indicated VA (3xFLAG)-tagged ORFs of ARMC7 or RBM48 and immunoprecipitated with anti-FLAG M2 antibodies and then immunoblotted with anti-V5 antibody and anti-FLAG antibody. β-actin was used as a loading control for input samples.

more exons (median 11 for fitness genes versus 7 exons for non-fitness genes).

Inferring the Function of Uncharacterized Core Fitness Genes

There still remains a significant portion of human genes that are uncharacterized or unstudied; part of the ''dark matter'' of the genome. We selected four predicted core essential genes (*ANKRD49*, *ZNF830*, *CCDC84*, and *RBM48*) that have few PubMed citations and gene ontology (GO) annotations (Figures 3A and 3B) and confirmed that all four genes are critical for cell proliferation in HCT116 and HeLa cells by independent treatment with single gRNAs (Figure 3C; $p < 0.05$, Student's t test; Figure S4A). CRISPR knockout of *ZNF830* and *RBM48* with multiple independent gRNAs showed signs of apoptosis as measured by cleaved Caspase 7 and cleaved PARP (Figure S4B). Notably, all four of these genes are expressed at moderate levels refractory to RNAi perturbation. In order to infer the function of each of these genes, we expressed V5-tagged open reading frames (ORFs) of the four genes in HEK293T and HCT116 cells to examine their subcellular localization and protein interaction patterns. By immunofluorescence,

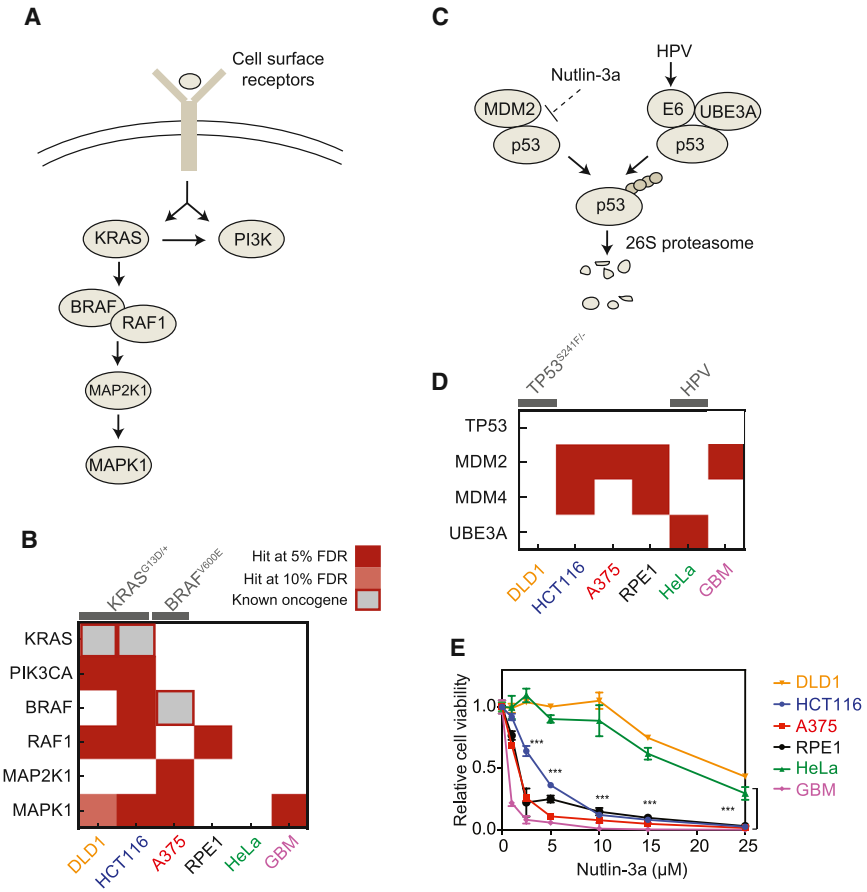


Figure 4. Genotype-Specific Synthetic Lethality

(A) The canonical MAPK pathway.

(B) Context-dependent sensitivity to perturbation of MAPK fitness genes. Cell lines with *KRAS* and *BRAF* mutations show strong pathway bias. Solid, fitness gene at 5% FDR. Shaded, fitness gene at 10% FDR. Gray boxes represent known essential oncogenes missed by the screens.

(C) Canonical and HPV-mediated pP53 degradation pathways.

(D) P53 regulators as fitness genes.

(E) Cell viability assays demonstrating the effect of Nutlin-3a (an inhibitor of MDM2-p53 interaction) treatment at the indicated concentrations.

All data are represented as means ± SD (n = 3 replicates). ***p < 0.001, **p < 0.01, and *p < 0.05; two-tailed unpaired t test.

ANKRD49 showed staining distributed throughout the cell, whereas ZNF830, CCDC84, and RBM48 exhibited enriched nuclear staining (Figure 3D and data not shown). Immunoprecipitations from HEK293T and HCT116 cell lysates expressing V5- and 3×FLAG-tagged ORFs followed by mass spectrometry (IP-MS) identified protein interactions with all four bait proteins (Figure 3E and Table S3). With each of the four bait proteins, robust protein-protein interactions with other core fitness genes were detected, suggesting that these uncharacterized fitness genes may also encode members of essential protein complexes. For example, ANKRD49 was found to associate strongly with FKBPL and together likely function as an Hsp90 co-chaperone to promote aspects of protein folding (Taipale et al., 2014). ZNF830 and CCDC84 interacted with different sets of proteins that are predicted to participate in mRNA splicing (Figure 3E). Notably, ZNF830 was previously found to associate with the XAB2 splicing complex (De et al., 2015; Kuraoka et al., 2008), whereas CCDC84 may be a component of the PRPF splicing complex. Finally, an interaction between RNA-binding motif containing protein RBM48 and the product of the uncharacterized fitness gene *ARMC7* was detected by IP-MS and confirmed by co-immunoprecipitation (Figure 3F). *RBM48* and *ARMC7* have been found to be amplified across several cancer tissues and cell lines (Figures S4C and S4E), and both genes show an enrichment for copy number amplification and a corresponding depletion for copy loss relative to a pool of nonessential genes (Figures S4D and S4F). RBM48 contains an RNA-binding motif, and ARMC7 is an Armadillo-repeat containing protein, which suggests that this RBM48-ARMC7 essential protein complex may have a role in RNA metabolism and/or transcription.

Selective Dependence on Pathways Harboring Known Driver Mutations

Defining core fitness genes provides an important foundation for identifying synthetic lethal genes that illuminate context-specific dependencies on signaling pathways and biological processes, which in turn provide candidate tumor-specific therapeutic targets. The cells assayed here for fitness genes represent a cross-section of common cancer genotypes whose differential genetic vulnerabilities highlight actionable targets and pathways. For example, we expected to observe selective dependence on the mitogen-activated protein kinase (MAPK) pathway, as three of our cell lines harbor driver mutations in this pathway ($KRAS^{G13D}$ in DLD1 and HCT116, and $BRAF^{V600E}$ in A375). The canonical MAPK pathway carries signals transduced by various cell surface receptors (Figure 4A), which regulate many cellular processes, including cell proliferation, cell survival, and cell-cycle control (Lemmon and Schlessinger, 2010). We observe multiple hits downstream of the driver mutations in all three MAPK pathway mutants (Figure 4B), including selective *PI3K* dependency in the KRAS mutant background, consistent with KRAS activation of the PI3K/AKT pathway (Boj et al., 2015; Castellano and Downward, 2011). As the stringent sequence identity filters used to generate the TKO library resulted in only one gRNA targeting the *KRAS* gene, *KRAS* itself does not emerge as a fitness gene. Similarly, the GeCKo library screen does not identify *BRAF* as a hit in the A375 cell line. Nevertheless, the overall trend of oncogene-driven MAPK pathway dependency is strongly affirmed.

In addition to the oncogene-driven MAPK activity, we observed differential p53 regulation consistent with the cells' genotypes. The canonical p53 protein degradation pathway involves ubiquitination by the E3 protein-ubiquitin ligase MDM2 (Figure 4C) (Momand et al., 2000). Three of our wild-type *TP53* cell lines are dependent on *MDM2*, and the related p53 inhibitor *MDM4* is also a hit in two of those three lines (Figure 4D; FDR < 5%). HeLa cells also carry wild-type *TP53* but are driven by human papilloma virus (HPV) infection, where the HPV protein E6 mediates p53 degradation through its interaction with the endogenous UBE3A E3 ubiquitin-protein ligase (Scheffner et al., 1993) (Figure 4C). Importantly, we observed that HeLa cells require *UBE3A* for proliferation, but not *MDM2* or *MDM4*. In contrast, DLD1 cells have an oncogenic $TP53^{S241F/-}$ mutation (Ahmed et al., 2013; Sur et al., 2009), circumventing dependency on these protein degradation pathways. Differential response to treatment with Nutlin-3a, which inhibits MDM2 protein interaction with and subsequent degradation of p53, confirms the selective dependence on this interaction (Figure 4E) and is consistent with prior observations of genotype-dependent response to MDM2 inhibition (Sur et al., 2009). Neither *MDM2* nor *MDM4* are discovered in the HCT116 RNAi screen, demonstrating both the improved sensitivity and specificity that TKO fitness screens provide.

Identification of Novel Context-Specific Vulnerabilities

While the genetic dependencies that emerge as a consequence of the known oncogenic drivers demonstrate the sensitivity of the CRISPR screens, the large number of fitness genes identified in each cell line enables a data-driven approach to discovering context-specific vulnerabilities. We took the fitness genes in each cell line and subtracted the core fitness genes to specify a set of context-specific genes for each line (Figure 5A). We then analyzed each set of genes for enrichment in Gene Ontology biological process or cellular component annotations, using the GOrilla web service (Eden et al., 2009). Comparing enriched terms revealed that each cell line had a highly unique signature of essential biological processes and associated components (Figure 5B). We noted that the "telomere maintenance" pathway was enriched only in RPE1 cells, which is consistent with that cell line having been immortalized by telomerase reverse transcriptase (hTERT; Bodnar et al., 1998). We discovered unexpected differential hits across several receptor tyrosine kinases (RTKs) upstream of the MAPK signaling pathway described earlier (Figures 4B and 5C). To our surprise, the epidermal growth factor receptor *EGFR* was a strong hit in $KRAS^{G13D}$ DLD1 cells, along with signal transduction adapters *SHC1*, *GRB2*, and *SOS1* (Figure 5D). In contrast, $KRAS^{G13D}$ HCT116 cells showed no hits in this pathway. DLD1 cells show a selective response to EGFR inhibitor erlotinib (Figure 5E), confirming this observation. Unlike HCT116, DLD1 cells carry a $TP53^{S241F}$ mutation, which we hypothesized might be related to the *EGFR* dependency; interestingly, HPAF-II cells also carry both *KRAS* (G12D) and *TP53* (P151S) mutations and are also sensitive to erlotinib (Figure S5A). However, further examination of a series of isogenic HCT116 cell lines with varying *TP53* genotypes, and DLD1 cells with varying KRAS genotypes, indicated that *TP53* genotype does not re-sensitize HCT116 cells to erlotinib (Figure S5B). Thus, dependency on signaling components upstream of a well-characterized driver oncogene violates predictions that the driver mutation disconnects cells from upstream signal relays and highlights potential differences in therapeutic options even within well-defined subtypes.

Other RTKs also showed differential sensitivity across our screens. GO biological process term "insulin receptor signaling," enriched in HeLa-specific fitness genes, includes fibroblast growth receptor *FGFR1* and insulin-like growth factor 1 receptor *IGF1R*. HeLa cells showed specific sensitivity to the FGFR1 inhibitor PD173074 (Figure 5F), as predicted by comparing *FGFR1* hits across cell lines (Figure 5D), while treatment with custom synthetic human antibodies generated against IGF1R that block IGF stimulation of IGFR1 (Figure S5C) reduced cell viability in HeLa and A375 but not in HCT116 cells (Figure 5G).

The GO term enrichment analysis also suggested a surprising and specific dependency on mitochondrial activity (Figure 5B). In particular, fitness genes in HCT116 are highly enriched for subunits of electron transport chain complex I; as predicted, growth of HCT116 cells was strongly inhibited by complex I inhibitor rotenone (Figures 5H and 5I) and by biguanide derivatives phenformin (Figure S5D) and metformin (Figure 5J), an inhibitor of oxidative phosphorylation (Bridges et al., 2014) that is commonly prescribed for type 2 diabetes. Similarly, DLD1 fitness genes are enriched for genes encoding subunits of the mitochondrial ribosome (Figure 5B). Several antibacterial drugs, including linezolid (Nagiec et al., 2005), tigecycline (Olson et al., 2006), and chloramphenicol (McKee et al., 2006), have been shown to specifically inhibit mitochondrial ribosome function in mammalian cells (Figure 5H) (Skrtić et al., 2011). As predicted, DLD1 cells show preferential sensitivity to these small molecules (Figures 5K, 5L, and S5E).

In addition to fitness genes that were negatively selected in our screens and conditionally required for cell proliferation, we also identified context-specific genes whose knockout enhanced proliferation. Since BAGEL is designed to identify negatively selected genes, we used the MAGeCK algorithm (Li et al., 2014) and identified a number of positively selected genes from our screens. For example, the HeLa screen identified components of the Hippo tumor suppressor pathway, including *NF2*/merlin (adjusted p = 0.005) and *LATS2* (p = 0.018). The GBM screen identified the tumor suppressor gene TP53 (p = 0.001), as well as other potential candidate tumor suppressor genes, including the nuclear exportin *XPO7* (p = 0.001) and *DHX29*/DEAH-box protein 29 (p = 0.001). The RPE1 screen identified *TP53* (p = 0.005), as well as a number of additional candidate tumor suppressor genes, including *PDCD10*/*CCM3*, mutations in which are one cause of cerebral cavernous malformations. Taken together, our screens identified both negatively and positively selected genes simultaneously and report clear results across different genetic backgrounds.

Guides Targeting Random Loci Suggest Negligible Off-Target Effects

A key design criterion for the TKO library was strict control over potential off-target effects by excluding candidate gRNA with multiple genomic binding sites, even including up to two

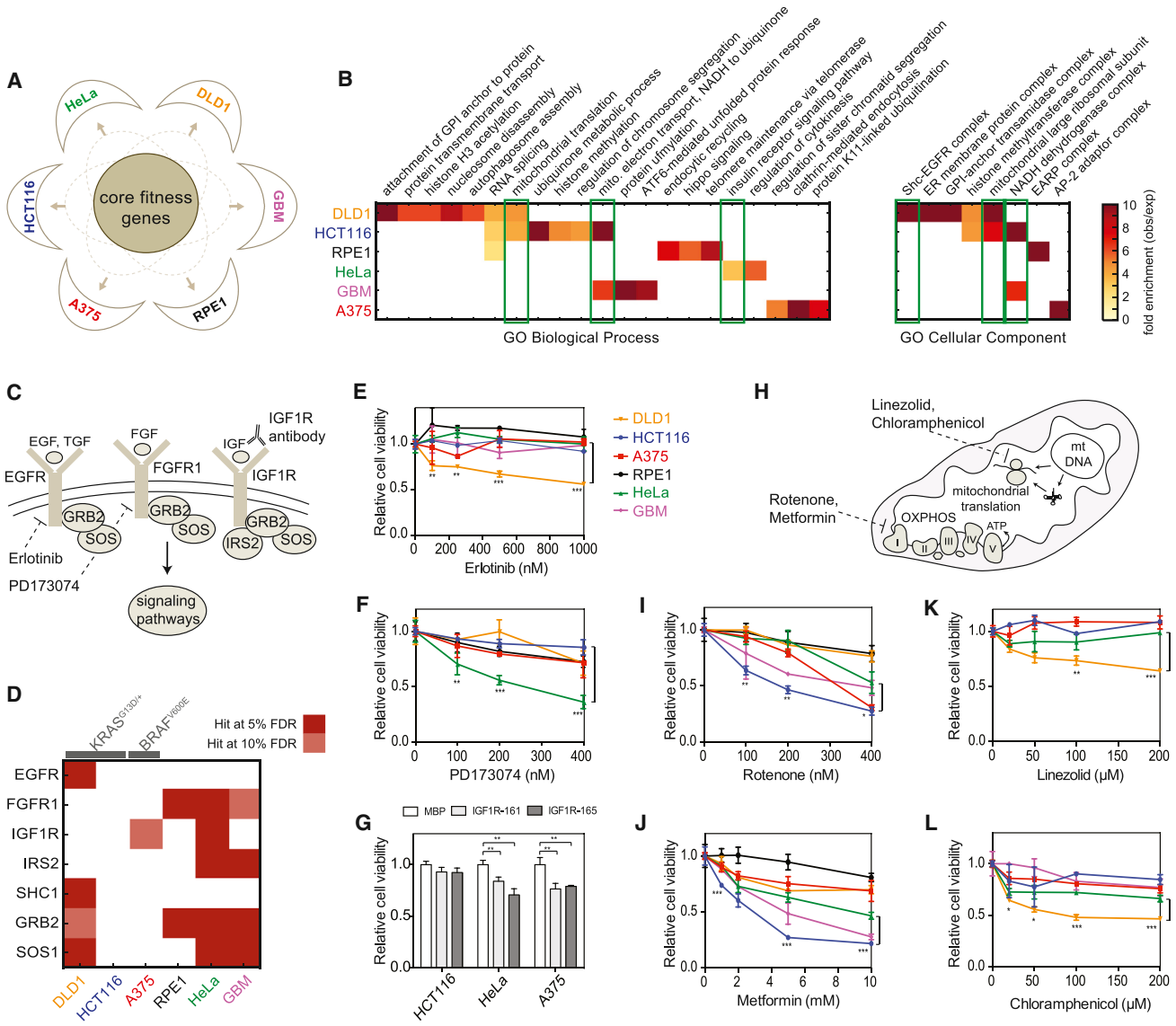


Figure 5. Data-Driven Discovery of Genetic Vulnerabilities

(A) The Daisy Model from Figure 2, here focusing on the petals: each cell line's fitness genes, minus the core.

(B) Selected enriched Gene Ontology Biological Process and Cellular Compartment terms. Color represents fold enrichment for terms meeting a p value threshold ($p < 10^{-3}$). Green box highlights experimentally validated GO terms.

(C) Representative RTK complexes and their inhibitors.

(D) Context-dependent sensitivity to perturbation of RTK fitness genes. Solid, fitness gene at 5% FDR. Shaded, fitness gene at 10% FDR.

(E) Cell viability assays demonstrating the effect of erlotinib (EGFR inhibitor) and (F) PD173074 (FGFR1 inhibitor) treatment at the indicated concentrations in DLD1, HCT116, A375, RPE1, HeLa, and GBM cells.

(G) Cell viability assay demonstrating the effect of custom synthetic IGF1R antibodies (IGF1R-161 and IGF1R-165 at 50 μg/ml) in the indicated cell lines. MBP (Maltose Binding Protein)-IgG was used as a control antibody.

(H) Differentially essential mitochondrial functions and their inhibitors.

(I and J) (I) Cell viability assays demonstrating the effect of rotenone (complex I inhibitor) and (J) metformin (OXPHOS inhibitor) treatment at the indicated concentrations.

(K and L) (K) Cell viability in response to linezolid and (L) chloramphenicol, inhibitors of mitochondrial translation. All data are represented as means ± SD (n = 3 replicates). ***p < 0.001, **p < 0.01, and *p < 0.05; two-tailed unpaired t test.

mismatches. To more systematically explore the phenotypic effect of off-target gRNA binding, we examined the fold changes of the TKO library guides targeting random loci on chromosome 10. Guides with only one predicted target site (n = 584) show negligible fold-change distribution, similar to guides targeting nonessential genes (Figure S1I). However, promiscuous

gRNAs—those targeting at least 20 genomic loci (n = 796)—are more similar to guides targeting essential genes (Figure S1I). This behavior is consistent across all five TKO cell lines and suggests that off-target effects only become an issue in negative selection screens if there are many target sites; the odds of a single random off-target locus causing a significant phenotype appear to be quite low.

DISCUSSION

We developed and used the TKO library to conduct genetic screens to identify fitness genes in human cell lines. We identified ~2,000 high-confidence fitness genes in each cell line, four to five times more genes than an RNAi screen yielded at the same false discovery rate. Using the Daisy Model as a framework, we classified 1,580 genes as core fitness genes (hits in 3 or more of the 5 cell lines). This vastly expands the catalog of human cell line fitness genes; by way of comparison, the number of CRISPR-identified core human fitness genes that have essential yeast orthologs (n = 569) is more than 50% larger than the entire set of core fitness genes identified by systematic analysis of 174 human cell-line shRNA screens (n = 360; Hart et al., 2014).

Of particular interest is the strong enrichment among core fitness genes for well-annotated protein complexes. Enlarging the census of essential protein complexes has potentially major chemotherapeutic implications. In a recent study, Nijhawan et al. (2012) discovered that partial copy loss of subunits of some essential protein complexes—presumably as passenger losses related to genetic deletion of nearby tumor suppressors—renders cells more susceptible to inhibition of those complexes. The study relied on RNAi to identify genetic susceptibility and yielded results enriched for highly expressed ribosome, proteasome, and spliceosome subunits. Consistent with this hypothesis, Liu et al. (2015) report that the RNA polymerase II subunit *POLR2A*, located at cytoband 17p13, is frequently deleted along with neighboring tumor suppressor *TP53*. This copy loss results in selective sensitivity to perturbation of either the *POL2RA* gene, by RNAi, or of RNA polymerase complex function by α-amantin. RNA polymerase II is one of many essential complexes for which specific inhibitors already exist, which could lead to repurposing of existing drugs to target tumor-specific haplo-insufficiencies. With the CRISPR core fitness genes, we expand the hypothesis of passenger-driven genetic vulnerability by extending the list of candidates to dozens of protein complexes comprised of hundreds of subunits.

This set of core fitness genes should prove useful across many disciplines of human genetics and genomics. In cancer genomics, chromosomally unstable tumors should, in principle, show lower frequency of deletion at fitness gene loci than random. In population genetics, these loci should show lower rates of deleterious mutation and copy number variation, based on the Daisy Model hypothesis that cell line fitness genes are a subset of organism-level fitness genes. Furthermore, the census of core fitness genes will be refined as the number of high-quality cell line genetic screens expands and researchers integrate these results with other functional genomics and orthology data. Finally, this refined core will be used as an updated yardstick to evaluate improvements in both screening technologies and analytical methods.

Though preliminary, our census of core fitness genes provides a foundation for identifying context-specific fitness genes that illuminate biological differences between cell types and warrants further research into genotype-driven therapeutic targets for cancer. Each cell line we assayed offered a unique signature of essential biological processes, many of which we confirmed by orthogonal small-molecule or antibody inhibitors. In general, essential pathways were consistent with expectations based on known oncogenic drivers—cell lines with *KRAS* or *BRAF* driver mutations, for example, rely on downstream MAPK pathway elements, and wild-type *TP53* cells showed vulnerability to inhibition of MDM2. However, we also observed counterintuitive differences between ostensibly similar cell lines. Both HCT116 and DLD1 are *KRAS*-mutant colorectal cancer cell lines, but DLD1 also relies on EGFR signaling, contradicting the model that activating mutations decouple biological pathways from upstream signaling. HCT116, on the other hand, demonstrates a specific dependence on ETC complex I activity and selective inhibition in response to metformin, a widely prescribed antidiabetic drug. Lower cancer rates among metformin users have been observed for years, leading to clinical interest in metformin as an anticancer agent (Jalving et al., 2010). This preliminary work strongly supports the idea that oxidative phosphorylation dependency—a clear exception to the Warburg effect—is a targetable weakness of some tumors.

It is tempting to ascribe the differences between these two similar cell lines to other differences among well-characterized cancer genes—for example, *TP53*, which carries an oncogenic mutation in DLD1 but not HCT116. However, such an observation is at best correlative, given only two cell lines. We emphasize that it is the discovery of this unexpected variation, confirmed through orthogonal assays, that represents the true novelty of this aspect of our results, though we cannot currently explain the source of this variation. In fact, we are far from understanding the genetic and epigenetic rewiring of cells that give rise to these selective vulnerabilities. It took a decade of technology development, systematic experimentation, and advanced analysis to map the genetic network of a yeast cell in a single growth condition (Costanzo et al., 2010), and we are only just beginning to understand the dynamics of that network (Bandyopadhyay et al., 2010). Translating that work into humans must take into account the "dimension problem" of more diverse genotypes, more epigenetic states/tissue types, and more environmental conditions (i.e., growth media and required supplements for different cell lines). Nevertheless, with the advent of CRISPR-Cas9 technology, this goal may finally be achievable. Future systematic high-resolution fitness screens across multiple query knockout strains in an isogenic background, analogous to the yeast approach, should yield a first draft of the human genetic network. Concurrent screening of a large, diverse panel of well-characterized cancer cell lines (Boehm and Golub, 2015) may tell us how genetic variation changes that network and how to take advantage of the surprising emergent vulnerabilities of specific tumors. This work advances the current state of the art for those future studies and provides a crucial baseline for their analysis and integration.

EXPERIMENTAL PROCEDURES

Online Resources

TKO library sequences, raw and processed data from the experiments described here, and details on the BAGEL software can be downloaded at the TKO website at http://tko.ccbr.utoronto.ca.

Genome-Scale Lentiviral gRNA Library Design and Construction of the TKO Library

gRNA target sites containing the N20NGG motif were chosen from protein-coding exons and filtered for nucleotide composition and transcript position. gRNAs were synthesized as 58-mer oligonucleotides on two microarrays (CustomArray), each with a density of ~90,000 sequences, and amplified by PCR as a pool. The PCR products were purified and cloned into pLCKO vector. Colonies were scraped off plates and pooled, and the plasmid DNA was extracted. Library virus was produced in HEK293T cells, and MOI was determined for each cell line screened (see Supplemental Experimental Procedures).

Pooled gRNA Depletion Screens

Cas9-expressing cells were infected with the lentiviral TKO library at an MOI ~0.3 such that every gRNA is represented in ~270 cells. 24 hr after infection, infected cells were selected with puromycin for 48–72 hr. After selection, cells were split into three replicates, passaged every 3 days, and maintained at 200-fold coverage. Cells were collected at a 200-fold coverage for genomic DNA extraction at day 0 and every 3 days from day 6 to day 18 post-selection.

Genomic DNA was extracted from cell pellets. gRNA inserts were amplified via PCR using primers harboring Illumina TruSeq adapters with i5 and i7 barcodes, and the resulting libraries were sequenced on an Illumina HiSeq2500 (see Supplemental Experimental Procedures).

Analysis of Negative Selection Screens: The BAGEL Algorithm

The BAGEL algorithm was used to calculate a Bayes factor for each gene representing a confidence measure that the gene knockout results in a fitness defect. The algorithm uses reference sets of essential (n = 360) and nonessential (n = 927) genes from Hart et al. (2014). Gene-level Bayes factors for all screens described in the manuscript are available in the Supplemental Experimental Procedures.

Analysis of Screen False Discovery Rates Using Gene Expression

In addition to using the gold standard reference sets, we evaluated screen quality by an independent method comparing essential genes to gene expression. A background error rate for each screen was calculated by taking all genes with trace/no expression (logFPKM < −2) and measuring the fraction classified as essential, assuming that these trace/no expression genes are false positives (see Supplemental Experimental Procedures).

Identifying Core Fitness Genes

We chose a per-screen false discovery rate (FDR, 1-Precision) of 5% as a cut-off. Genes with BFs above the threshold in the given cell line were considered fitness genes in that cell line. Genes observed in 3 or more of the 5 TKO screens (n = 1,580) were considered "core fitness genes" and retained for further analyses (see Supplemental Experimental Procedures).

ACCESSION NUMBERS

The accession number for the RNA-seq data reported in this paper is GEO: GSE75189.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2015.11.015.

AUTHOR CONTRIBUTIONS

T.H., M.C., M.A., Z.S., S.A., and J.M. designed and constructed the TKO library and designed experiments. M.C., M.A., Z.S., G.M., and M.M. performed the experiments. T.H., M.C., M.A., K.R.B., S. Sun, O.S.R., F.P.R., D.D., S.A., and J.M. analyzed the data. M.Z., A.F.-T., and P.D. provided cell lines; and S. Sidhu provided synthetic antibodies. T.H., M.C., M.A., and J.M. wrote the manuscript with input from all the other authors.

ACKNOWLEDGMENTS

We thank Michael Costanzo, Amy Tong, Charlie Boone, and Brenda Andrews for helpful discussions and comments on the manuscript. We thank Olga Sizova and Amy Tong for technical assistance, Sachin Kumar for providing plasmids, all the members of the Moffat and Angers labs for helpful discussions, Peggy Wang for graphic design, and Aaron Schimmer for chemicals. We thank the Exome Aggregation Consortium and the groups that provided exome variant data for comparison (http://exac.broadinstitute.org/about). We also thank Dax Torti at the Donnelly Sequencing Centre for assistance with sequencing. M.A. holds a postdoctoral fellowship from the Swiss National Science Foundation, A.F.-T. holds a CIHR postdoctoral fellowship, M.Z. is an EMBO Long-Term postdoctoral fellow, and S. Sun holds an international postdoc grant from the Swedish Research Council. This work was supported by grants from the Ontario Research Fund (J.M.), Ontario Institute for Cancer Research ITV Program (J.M.), Canadian Foundation for Innovation (J.M.), Canada Research Chairs program (J.M., S.A., and D.D.), Krembil Foundation (D.D.), a Canadian Institutes for Health Research operating grant to J.M. (CIHR-342551), and foundation grant to D.D. (FDN 143343). D.D. is a Tier 1 Canada Research Chair in Molecular Mechanisms of Genome Integrity, S.A. is a Tier 2 Canada Research Chair in Functional Architecture of Signal Transduction, and J.M. is a Tier 2 Canada Research Chair in Functional Genomics of Cancer.

Received: July 21, 2015
Revised: September 1, 2015
Accepted: October 30, 2015
Published: November 25, 2015

REFERENCES

Ahmed, D., Eide, P.W., Eilertsen, I.A., Danielsen, S.A., Eknæs, M., Hektoen, M., Lind, G.E., and Lothe, R.A. (2013). Epigenetic and genetic features of 24 colon cancer cell lines. Oncogenesis *2*, e71.

Bandyopadhyay, S., Chiang, C.Y., Srivastava, J., Gersten, M., White, S., Bell, R., Kurschner, C., Martin, C., Smoot, M., Sahasrabudhe, S., et al. (2010). A human MAP kinase interactome. Nat. Methods *7*, 801–805.

Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehár, J., Kryukov, G.V., Sonkin, D., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature *483*, 603–607.

Bodnar, A.G., Ouellette, M., Frolkis, M., Holt, S.E., Chiu, C.P., Morin, G.B., Harley, C.B., Shay, J.W., Lichtsteiner, S., and Wright, W.E. (1998). Extension of life-span by introduction of telomerase into normal human cells. Science *279*, 349–352.

Boehm, J.S., and Golub, T.R. (2015). An ecosystem of cancer cell line factories to support a cancer dependency map. Nat. Rev. Genet. *16*, 373–374.

Boj, S.F., Hwang, C.I., Baker, L.A., Chio, I.I., Engle, D.D., Corbo, V., Jager, M., Ponz-Sarvise, M., Tiriac, H., Spector, M.S., et al. (2015). Organoid models of human and mouse ductal pancreatic cancer. Cell *160*, 324–338.

Bridges, H.R., Jones, A.J., Pollak, M.N., and Hirst, J. (2014). Effects of metformin and other biguanides on oxidative phosphorylation in mitochondria. Biochem. J. *462*, 475–487.

Castellano, E., and Downward, J. (2011). RAS Interaction with PI3K: More Than Just Another Effector Pathway. Genes Cancer *2*, 261–274.

Chen, S., Sanjana, N.E., Zheng, K., Shalem, O., Lee, K., Shi, X., Scott, D.A., Song, J., Pan, J.Q., Weissleder, R., et al. (2015). Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. Cell *160*, 1246–1260.

Cho, S.W., Kim, S., Kim, J.M., and Kim, J.S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat. Biotechnol. *31*, 230–232.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823.

Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh, J.L., Toufighi, K., Mostafavi, S., et al. (2010). The genetic landscape of a cell. Science *327*, 425–431.

De, I., Bessonov, S., Hofele, R., dos Santos, K., Will, C.L., Urlaub, H., Lührmann, R., and Pena, V. (2015). The RNA helicase Aquarius exhibits structural adaptations mediating its recruitment to spliceosomes. Nat. Struct. Mol. Biol. *22*, 138–144.

Echeverri, C.J., Beachy, P.A., Baum, B., Boutros, M., Buchholz, F., Chanda, S.K., Downward, J., Ellenberg, J., Fraser, A.G., Hacohen, N., et al. (2006). Minimizing the risk of reporting false positives in large-scale RNAi screens. Nat. Methods *3*, 777–779.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics *10*, 48.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., et al. (2002). Functional profiling of the Saccharomyces cerevisiae genome. Nature *418*, 387–391.

Hart, T., Komori, H.K., LaMere, S., Podshivalova, K., and Salomon, D.R. (2013). Finding the active genes in deep RNA-seq gene expression studies. BMC Genomics *14*, 778.

Hart, T., Brown, K.R., Sircoulomb, F., Rottapel, R., and Moffat, J. (2014). Measuring error rates in genomic perturbation screens: gold standards for human functional genomics. Mol. Syst. Biol. *10*, 733.

Hartwell, L.H., Szankasi, P., Roberts, C.J., Murray, A.W., and Friend, S.H. (1997). Integrating genetic approaches into the discovery of anticancer drugs. Science *278*, 1064–1068.

Hillenmeyer, M.E., Fung, E., Wildenhain, J., Pierce, S.E., Hoon, S., Lee, W., Proctor, M., St Onge, R.P., Tyers, M., Koller, D., et al. (2008). The chemical genomic portrait of yeast: uncovering a phenotype for all genes. Science *320*, 362–365.

Jalving, M., Gietema, J.A., Lefrandt, J.D., de Jong, S., Reyners, A.K., Gans, R.O., and de Vries, E.G. (2010). Metformin: taking away the candy for cancer? Eur. J. Cancer *46*, 2369–2380.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821.

Koike-Yusa, H., Li, Y., Tan, E.P., Velasco-Herrera, Mdel.C., and Yusa, K. (2014). Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nat. Biotechnol. *32*, 267–273.

Kuraoka, I., Ito, S., Wada, T., Hayashida, M., Lee, L., Saijo, M., Nakatsu, Y., Matsumoto, M., Matsunaga, T., Handa, H., et al. (2008). Isolation of XAB2 complex involved in pre-mRNA splicing, transcription, and transcription-coupled repair. J. Biol. Chem. *283*, 940–950.

Lander, E.S. (2011). Initial impact of the sequencing of the human genome. Nature *470*, 187–197.

Lemmon, M.A., and Schlessinger, J. (2010). Cell signaling by receptor tyrosine kinases. Cell *141*, 1117–1134.

Li, W., Xu, H., Xiao, T., Cong, L., Love, M.I., Zhang, F., Irizarry, R.A., Liu, J.S., Brown, M., and Liu, X.S. (2014). MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome Biol. *15*, 554.

Liu, Y., Zhang, X., Han, C., Wan, G., Huang, X., Ivan, C., Jiang, D., Rodriguez-Aguayo, C., Lopez-Berestein, G., Rao, P.H., et al. (2015). TP53 loss creates therapeutic vulnerability in colorectal cancer. Nature *520*, 697–701.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823–826.

McKee, E.E., Ferguson, M., Bentley, A.T., and Marks, T.A. (2006). Inhibition of mammalian mitochondrial protein synthesis by oxazolidinones. Antimicrob. Agents Chemother. *50*, 2042–2049.

Meuth, M. (1989). The molecular basis of mutations induced by deoxyribonucleoside triphosphate pool imbalances in mammalian cells. Exp. Cell Res. *181*, 305–316.

Moffat, J., Reiling, J.H., and Sabatini, D.M. (2007). Off-target effects associated with long dsRNAs in Drosophila RNAi screens. Trends Pharmacol. Sci. *28*, 149–151.

Momand, J., Wu, H.H., and Dasgupta, G. (2000). MDM2–master regulator of the p53 tumor suppressor protein. Gene *242*, 15–29.

Nagiec, E.E., Wu, L., Swaney, S.M., Chosay, J.G., Ross, D.E., Brieland, J.K., and Leach, K.L. (2005). Oxazolidinones inhibit cellular proliferation via inhibition of mitochondrial protein synthesis. Antimicrob. Agents Chemother. *49*, 3896–3902.

Nijhawan, D., Zack, T.I., Ren, Y., Strickland, M.R., Lamothe, R., Schumacher, S.E., Tsherniak, A., Besche, H.C., Rosenbluh, J., Shehata, S., et al. (2012). Cancer vulnerabilities unveiled by genomic loss. Cell *150*, 842–854.

Olson, M.W., Ruzin, A., Feyfant, E., Rush, T.S., 3rd, O'Connell, J., and Bradford, P.A. (2006). Functional, biophysical, and structural bases for antibacterial activity of tigecycline. Antimicrob. Agents Chemother. *50*, 2156–2166.

Parnas, O., Jovanovic, M., Eisenhaure, T.M., Herbst, R.H., Dixit, A., Ye, C.J., Przybylski, D., Platt, R.J., Tirosh, I., Sanjana, N.E., et al. (2015). A Genome-wide CRISPR Screen in Primary Immune Cells to Dissect Regulatory Networks. Cell *162*, 675–686.

Ramani, A.K., Chuluunbaatar, T., Verster, A.J., Na, H., Vu, V., Pelte, N., Wannissorn, N., Jiao, A., and Fraser, A.G. (2012). The majority of animal genes are required for wild-type fitness. Cell *148*, 792–802.

Scheffner, M., Huibregtse, J.M., Vierstra, R.D., and Howley, P.M. (1993). The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. Cell *75*, 495–505.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., and Zhang, F. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science *343*, 84–87.

Skrtić, M., Sriskanthadevan, S., Jhas, B., Gebbia, M., Wang, X., Wang, Z., Hurren, R., Jitkova, Y., Gronda, M., Maclean, N., et al. (2011). Inhibition of mitochondrial translation as a therapeutic strategy for human acute myeloid leukemia. Cancer Cell *20*, 674–688.

Sur, S., Pagliarini, R., Bunz, F., Rago, C., Diaz, L.A., Jr., Kinzler, K.W., Vogelstein, B., and Papadopoulos, N. (2009). A panel of isogenic human cancer cells suggests a therapeutic approach for cancers with inactivated p53. Proc. Natl. Acad. Sci. USA *106*, 3964–3969.

Taipale, M., Tucker, G., Peng, J., Krykbaeva, I., Lin, Z.Y., Larsen, B., Choi, H., Berger, B., Gingras, A.C., and Lindquist, S. (2014). A quantitative chaperone interaction network reveals the architecture of cellular protein homeostasis pathways. Cell *158*, 434–448.

Tani, H., Mizutani, R., Salam, K.A., Tano, K., Ijiri, K., Wakamatsu, A., Isogai, T., Suzuki, Y., and Akimitsu, N. (2012). Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals. Genome Res. *22*, 947–956.

Vizeacoumar, F.J., Arnold, R., Vizeacoumar, F.S., Chandrashekhar, M., Buzina, A., Young, J.T., Kwan, J.H., Sayad, A., Mero, P., Lawo, S., et al. (2013). A negative genetic interaction map in isogenic cancer cell lines reveals cancer cell vulnerabilities. Mol. Syst. Biol. *9*, 696.

Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. Science *343*, 80–84.

Winzeler, E.A., Shoemaker, D.D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J.D., Bussey, H., et al. (1999). Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science *285*, 901–906.