

RESEARCH

Open Access



Dynamic rewiring of biological activity across genotype and lineage revealed by context-dependent functional interactions

Eiru Kim^{1,2}, Lance C. Novak³, Chenchu Lin¹, Medina Colic^{1,4}, Lori L. Bertolet¹, Veronica Gheorghe^{1,4}, Christopher A. Bristow³ and Traver Hart^{1,5*} 

* Correspondence: traver@hart-lab.org

⁵Department of Cancer Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

Full list of author information is available at the end of the article

Abstract

Background: Coessentiality networks derived from CRISPR screens in cell lines provide a powerful framework for identifying functional modules in the cell and for inferring the roles of uncharacterized genes. However, these networks integrate signal across all underlying data and can mask strong interactions that occur in only a subset of the cell lines analyzed.

Results: Here, we decipher dynamic functional interactions by identifying significant cellular contexts, primarily by oncogenic mutation, lineage, and tumor type, and discovering coessentiality relationships that depend on these contexts. We recapitulate well-known gene-context interactions such as oncogene-mutation, paralog buffering, and tissue-specific essential genes, show how mutation rewires known signal transduction pathways, including RAS/RAF and *IGF1R-PIK3CA*, and illustrate the implications for drug targeting. We further demonstrate how context-dependent functional interactions can elucidate lineage-specific gene function, as illustrated by the maturation of proreceptors *IGF1R* and *MET* by proteases *FURIN* and *CPD*.

Conclusions: This approach advances our understanding of context-dependent interactions and how they can be gleaned from these data. We provide an online resource to explore these context-dependent interactions at diffnet.hart-lab.org.

Background

Development of genome-wide CRISPR screening half a decade ago facilitated robust determination of genes required for in vitro proliferation across cancer types [1–3]. Ongoing efforts to identify tumor-specific vulnerabilities encompass hundreds of cancer cell lines [4–7] across dozens of tissue types. Previously, we and others demonstrated that functionally coherent modules of genes can be extracted from CRISPR screens by measuring the similarity of gene knockout fitness profiles across all gene pairs, generating a functional interaction network based on coessentiality [8–13]. Many such modules found in the network encode well-annotated biological processes,



© The Author(s). 2022 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

enabling robust functional prediction for uncharacterized genes [14–16]. Some modules, however, are associated with specific tissues or genomic perturbations, such as the *ESR1-FOXA1* pathway in ER+ breast cancer cell lines and the *BRAF*^{V600E} module in melanoma cells [9, 17, 18]. In this latter cluster, gene mutation is so conflated with tissue specificity that b-RAF appears tightly linked with melanocyte-specific transcription factors *MITF* and *SOX10* but does not show significant correlation with its known protein interaction partner c-RAF (encoded by the *RAF1* gene).

We hypothesized that mutations in specific genes, and more broadly the genetic and epigenetic state of cells, could rewire these coessentiality networks and that disentangling this rewiring could elucidate the biological and therapeutic implications of genetic lesions and tissue-specific diseases. Context-specific interaction rewiring or differential network analysis is a useful tool for understanding how fixed genomes can emerge into heterogeneous cellular and tissue morphologies and phenotypes [19–21]. In prior studies, comparison of coexpression between two contexts selected from a pool were introduced to discover context-dependent interactions [22–24]. In the pre-CRISPR era, integrative functional interaction networks were exploited to prioritize disease genes [25–27] and included efforts to identify tissue-specific networks [28, 29], but these approaches lack the power of inferred functional interaction provided by coessentiality networks derived from screens in human cells. Recent examinations of coessentiality networks such as CEN-tools [18] and FIREWORKS [17] offer the ability to browse how correlates of a query gene vary by background, but offer little ability to interpret the results. Moreover, the ability to go from lesion of interest to lesion-associated network rewiring is limited.

In this study, we developed a framework of identifying dynamic relationships in cancer dependency data associated with differing functional contexts such as variations in tissue of origin and/or genomic lesions. We first categorized genomic perturbations using molecular profiles from the Cancer Cell Line Encyclopedia. Then, we investigated which genomic features were associated with synthetic gene essentiality by applying a logistic regression model, and used these features to search for associated network rewiring. Rewired networks can carry information on the efficacy of targeted drugs, the biological impact of specific lesions, and the tissue-specific activity of genes.

Results

Associating cellular context with emergent gene essentiality

Understanding the causal, or at least associative, basis of variation in gene essentiality is important in matching the right anticancer drugs with responsive tumors [30]. For example, *KRAS* is essential when it undergoes oncogenic mutation [31–33], and SWI/SNF complex member *ARID1B* is essential where its paralog *ARID1A* is deleted [34]. To globally decipher these essential gene/context relationships, we first systematically categorized genomic lesions into gain of function/hotspot (GOF) or loss of function (LOF) using mutation data from 808 cell lines in the Cancer Cell Line Encyclopedia [35] that have matching CRISPR screen data from the Cancer Dependency Map [5–7]. We added cell line metadata including lineage/tumor type, as well as a feature describing EMT state based on the ratio of *CDH1* to *VIM* expression (see “Methods”). After de-duplication and filtering, we retained 2918 binary features as predictor variables.

We used a machine learning approach to estimate the effect of genomic perturbations on variably essential genes (Fig. 1A). We chose logistic regression with an elastic net penalty because it handles binary predictors, because gene essentiality is readily binarized for use as a response variable, and because the L1 component of the elastic net forces most coefficients to zero, resulting in a model that is easily interpreted. We used a binary measure of gene essentiality because we judged that predicting whether a gene is or is not essential in a given background/cell line is more relevant than predicting variation in a continuous measure of fitness. Moreover, the log Bayes Factor output from the BAGEL2 algorithm [36], our chosen analysis pipeline, combines effect size (fitness score) and statistical significance into a single metric and, though a powerful classifier of essential and nonessential genes, is not amenable to cross-sample comparisons without normalization [5, 9]. For each gene, we generated a binary essentiality vector (defining essential as $BF \geq 10$; see “Methods”) across 808 good quality cell lines in DepMap 20Q4 data for use as a response variable, then filtered for those genes that are essential in at least 1% but not more than 80% of the 808 cell lines, resulting in $n = 2987$ response genes.

We then generated a matrix of genetic and metadata features for each cell line, for use as a feature table of predictor variables. For each of the 808 cell lines, each gene ($n = 18,111$) was classified as loss of function (LOF) or gain of function (GOF) based on

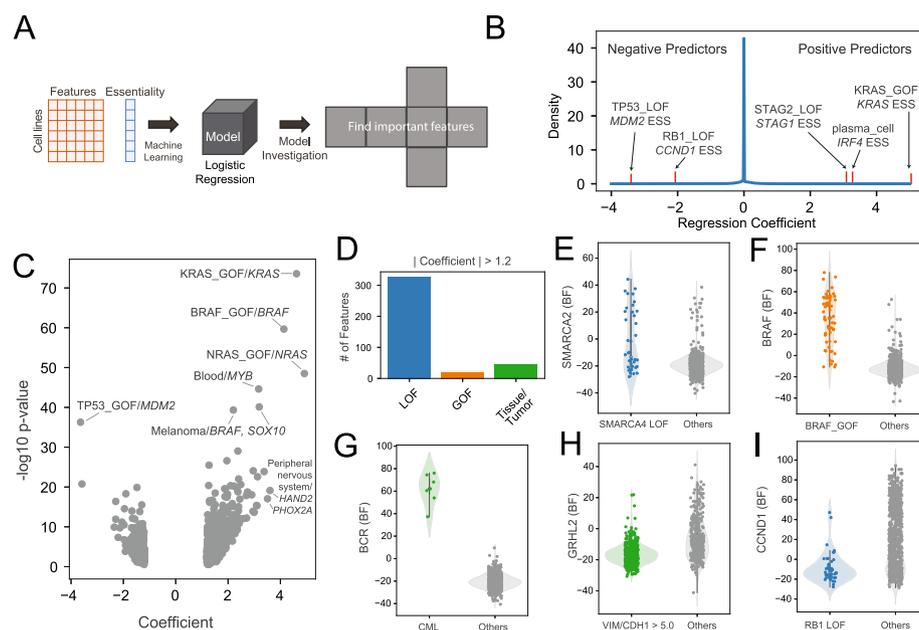


Fig. 1 Underlying drivers of gene essentiality. **A** Schematic of the process for finding important contexts for a gene. We input binary feature vectors across all cell lines and a vector of essentiality of a gene and trained a logistic regression model. Then, we investigated coefficients of features to determine the features important for predicting gene essentiality. **B** Distribution of regression coefficients collected from prediction models of all genes. Positive coefficients reflect gene essentiality in the presence of the feature, and negative coefficients reflect gene essentiality in the absence of the feature. **C** Volcano plot demonstrates consistency between P -values of Fisher exact test (Y axis) and regression coefficients (X axis). **D** Category of 393 features with $|\text{regression coefficient}| > 1.2$. **E–I** Examples of predictions from **E** paralog interaction, **F** oncogene addiction, **G** tissue-specific essential genes, **H** EMT expression signature, and **I** negative association. BF, Bayes Factor from BAGEL2 analysis of DepMap data (positive = essential, negative = nonessential)

mutation calls from the Cancer Cell Line Encyclopedia. Epithelial vs. mesenchymal state was determined by *CDH1/VIM* mRNA expression ratio. Other cell line metadata, including unique labels associated with lineage/disease type, and cell culture conditions, and patient sex, were each transformed into binary feature vectors. All features ($n = 36,406$) were then filtered, removing those describing six or fewer cell lines ($n = 33,475$) and de-duplicating highly correlated features to prevent collinearity. The final de-duplicated binary feature set comprised 2918 predictor variables for the regression model (see “[Methods](#)” for details).

We then used logistic regression with the same table of 2918 features to predict the essentiality profile of each of the 2987 response genes independently. This approach is similar to that used by Lord et al. [37], but our logistic regression is a binary classifier of gene essentiality rather than their linear regression on continuous fitness effects. After calculating all regression models, we investigated which genomic features significantly contributed to predicting essentiality (Fig. 1B, C). We collected coefficients of genomic features across all prediction models of genes, which loosely represent the magnitude of predictive power. A positive coefficient of a genomic feature indicates that the genetic lesion is associated with the gene dependency (e.g., *KRAS_GOF/KRAS* essentiality), and a negative coefficient indicates that gene essentiality is associated with the absence of the lesion (e.g., *TP53_LOF/MDM2* essentiality). Coefficients showed a similar trend with *P*-values derived by the Fisher exact test between a vector of genomic lesions and a vector of gene dependency across cell lines (Fig. 1C). By considering strong coefficients (see “[Methods](#)” and Additional file 1: Figure S1), we identify 393 genomic features strongly predictive of variation in gene essentiality, including 327 loss of function mutations, 20 gain of function mutations, and 46 features derived from cell metadata (Fig. 1D). We recapitulated positive associations in paralog buffering (e.g., *SMARCA4_LOF/SMARCA2^{ess}*, Fig. 1E), oncogenic mutation-induced essentiality (e.g., *BRAF_GOF/BRAF^{ess}*, Fig. 1F), and tumor-specific dependencies (e.g., *CML/BCR^{ess}*, Fig. 1G), as well as negative associations such as epithelial transcription factor *GRHL2* essentiality in cells with low *CDH1/VIM* expression ratio (Fig. 1H) and increased essentiality of cyclin D1 (*CCND1*) in cells with wildtype *RBI* (Fig. 1I). These results are highly concordant with previous predictive models (Additional file 1: Figure S1) [5, 37] and known biology and provide a meaningful set of features for exploring context-dependent functional interaction rewiring. The table of all coefficients is included as Additional file 2.

Measuring context-dependent network rewiring

With the identification of contexts associated with variation in gene essentiality, we expanded our view to functional interaction rewiring associated with genomic and lineage features. In general, genes with correlated fitness profiles across a diverse panel of cell lines tend to operate in the same biological processes, such as enzymatic or signal transduction pathways, often as subunits in the same protein complex. While this correlation provides a powerful framework for inferring gene function and the modular structure of the cell [8–13], it relies on covariation in the underlying gene essentiality vectors but offers no insight into the source of this variation. We hypothesized that, by systematically removing rationally selected subsets of cell lines and measuring the effect

on pairwise correlation, we could identify the drivers of this covariation and, in turn, infer the causal basis for the emergent essentiality of biological processes.

Although differential coessentiality offers a seemingly straightforward way to approach context rewiring of functional interaction networks, the method is beset by complications. One is that, for many genes, physiologically relevant knockout fitness defects are typically only seen in a few cell lines. A related issue is that there are frequently too few cell lines associated with a given context for correlation to be an accurate predictor of functional interaction within the context; e.g., a CML-only coessentiality network has limited meaning if there are only seven CML cell lines. To overcome these difficulties, we designed a strategy comprising a leave-one-out test and bootstrapped network comparison analysis to investigate snapshots of interaction rewiring in essentiality data associated with our features of interest. In the leave-one-out test, interaction rewiring was measured as a differential Pearson correlation coefficient (dPCC) of a context by taking the PCC of knockout fitness profiles (in our case, quantile normalized logBF scores from BAGEL2) from all cell lines (PCC_{all}) and subtracting the PCC using all cell lines except those carrying the feature of interest ($PCC_{\sim context}$) (Fig. 2A). We classified a positive dPCC as a gain of interaction (GOI) because the PCC depends on the presence of cells with the feature of interest (Fig. 2A, top). Conversely,

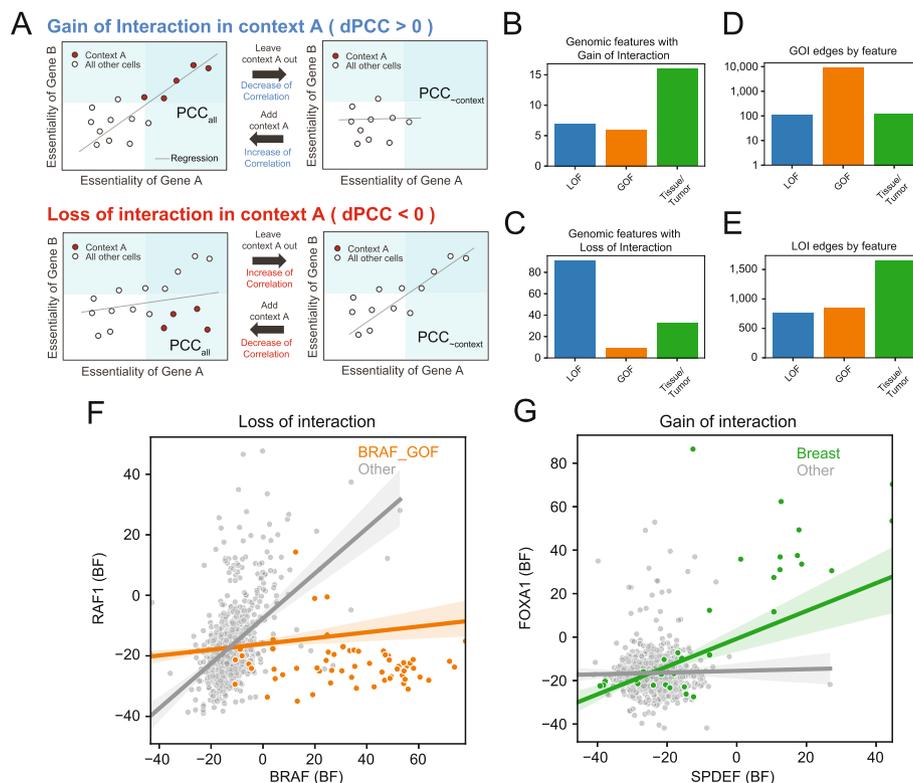


Fig. 2 A framework to identify dynamics in coessentiality network. **A** Classes of context-dependent interaction. In all cases, a differential Pearson correlation coefficient (dPCC) for each gene pair is calculated as the difference between the global network (PCC_{all}) and the PCC from all cells except those harboring the feature of interest ($PCC_{\sim context}$). **B, C** The number of features associated with gain or loss of interaction. **D, E** the number of gained or lost interactions, by feature type. Nearly all GOI edges are associated with TP53_GOF mutations. **F, G** Examples of context-dependent loss and gain of interaction. Axes are Bayes Factors of the genes

we describe a negative dPCC as a loss of interaction (LOI) because the background PCC is improved by removing cells with the feature of interest (Fig. 2A, bottom).

We measured dPCC for the 393 features identified in logistic regression models as significant predictors of gene essentiality, and measured significance by bootstrap re-sampling to estimate a null distribution of dPCC (see “Methods”). We identified 30 features significantly associated ($P \leq 0.001$) with gain of interaction (Fig. 2B) and 133 features, including 91 LOF mutants, associated with loss of functional interaction (Fig. 2C) in the cell lines studied. These features are associated with 9227 gain of interaction and 3261 loss of interaction events (Fig. 2D,E). Notably, of the gain of interaction events, 97% ($n = 8958$) are associated with TP53 gain of function/hotspot mutations, leaving only 269 interactions associated with the remaining 29 features. A list of top dPCC hits is included as Additional file 3.

Two case studies illustrate the effect of context-dependent rewiring of functional interaction networks. RAS and RAF kinases are core members of RTK signaling pathways for proliferation and differentiation [38–42]. Within the canonical MAP kinase signal transduction pathway, b-RAF (*BRAF*) and c-RAF (*RAF1*) form a heterodimer to transmit phosphorylation signals from upstream RAS genes to downstream MEK/ERK targets. In *BRAF*-driven cancers, oncogenic mutation constitutively activates *BRAF* and obviates the need for c-RAF binding. In the global coessentiality network, *BRAF*^{V600E} and *BRAF*^{wt} cell lines are mixed, diluting the ability to discover the wildtype *BRAF*-*RAF1* relationship (background PCC = 0.124). Removing *BRAF*^{V600E} cell lines boosts this correlation to 0.478 (dPCC = -0.363; $P < 0.001$, see “Methods” for details of statistical tests); thus, we infer that BRAF_GOF mutation is causal for a loss of interaction between *BRAF* and *RAF1* (Fig. 2F). Similarly, transcription factor *FOXAI* is essential in a subset of lung and breast cancers but shows a breast-specific functional interaction with ETS family transcription factor *SPDEF* (Fig. 2G).

Reconstruction of directionality in signaling pathways

The dynamic rewiring of functional interactions resulting from cancer-associated mutation has predictable consequences. As previously observed, mutation of the *TP53* tumor suppressor removes the cell’s reliance on *MDM2* and related genes to suppress the proapoptotic activity of wildtype p53 protein. Similarly, constitutive activation of signaling proteins by oncogenic mutation activates downstream signaling partners while removing the need for upstream activation signals. Here we describe two such pathways.

RAS/RAF signaling pathway

Mutations in the RAS pathway are major drivers of numerous cancers. RAS family members *KRAS* and *NRAS* are frequently mutated in colorectal and pancreas adenocarcinoma, and multiple myeloma and melanoma, respectively. *BRAF* is a major driver gene of melanoma and thyroid adenocarcinoma [43, 44]. Oncogenic GOF lesions in these genes caused significant intra- and inter-pathway interaction rewiring (Fig. 3A). *KRAS* and *NRAS* are usually nonessential in CRISPR data; however, mutation or amplification of these genes induces hyperactivation and strong mutually exclusive essentiality. We found that interaction rewiring of GOF mutations of *KRAS*, *NRAS*, and *BRAF*

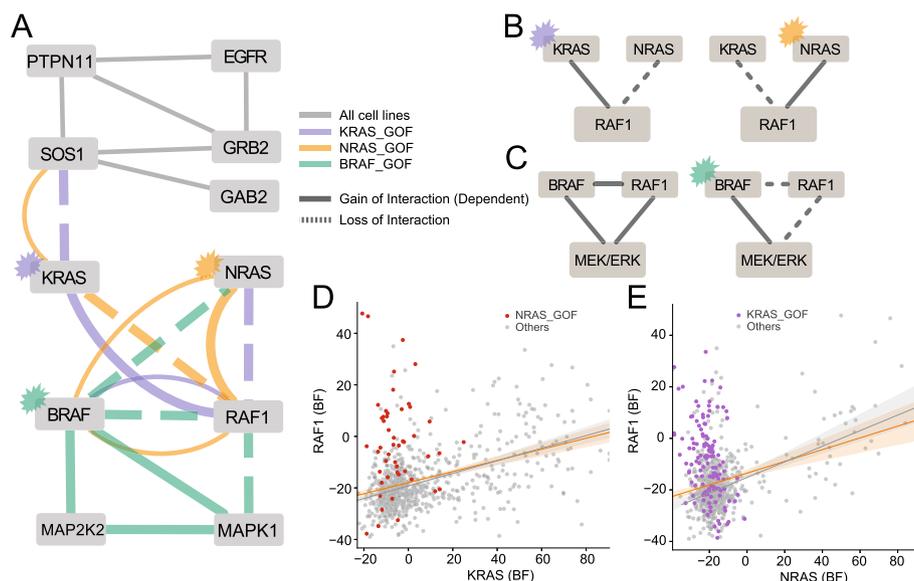


Fig. 3 Interaction rewiring in RAS/RAF pathway. **A** A network diagram of differential interactions in the RTK/ MAP kinase signaling cascade caused by mutations on KRAS, NRAS, and BRAF. **B** KRAS gain of function mutant caused loss of NRAS-RAF1 interaction and NRAS mutant caused loss of KRAS-RAF1 interaction. **C** RAF1 is required for transduction signal to MEK/ERK signal when BRAF is wildtype, while BRAF mutant does not require RAF1 and led to loss of RAF1 interaction with BRAF and the downstream pathway. **D**, **E** Scatter plots of interactions between RAF1 and two RAS proteins

induced gain of interaction with downstream genes and disconnected the network from upstream genes through loss of interaction (Fig. 3A). For example, *KRAS* mutation amplifies the link with downstream signaling partner *RAF1* ($PCC_{all} = 0.390$, $PCC_{\sim KRAS_GOF} = 0.239$, $dPCC = 0.151$, $P < 0.001$), severs the link between *RAF1* and *KRAS* paralog *NRAS* ($PCC_{all} = 0.305$, $PCC_{\sim KRAS_GOF} = 0.472$, $dPCC = -0.167$, $P < 0.001$), and disconnects MAP kinase signaling from the upstream *EGFR* receptor tyrosine kinase signal transduction module ($KRAS-SOS1$ $PCC_{all} = -0.012$, $PCC_{\sim KRAS_GOF} = 0.184$, $dPCC = -0.197$, $P < 0.001$). Similar behavior is seen in *NRAS* and *BRAF* GOF cell lines (Fig. 3B–E). Additional rewiring by *BRAF* mutation abrogates interaction with dimerization partner *RAF1*, consistent with the known biology of *BRAF* mutation. We found that *BRAF* mutation induced loss of interaction between *BRAF* and *RAF1* while *BRAF* and *RAF1* are well correlated in wild-type *BRAF* cell lines (Fig. 3A, C).

IGF1R/PIK3CA signaling pathway

The *PIK3CA* gene encodes a protein kinase that mediates signaling from insulin-like growth factor receptor *IGF1R* to the *AKT* pathway [45] (Fig. 4A), is frequently mutated in a number of cancers, and is itself the target of several chemotherapeutic agents. When mutated, *PIK3CA* is hyperactivated and induces downstream pathways without reliance on *IGF1R* signals [46]. In our differential network analysis, *PIK3CA* gain of function mutation causes not only loss of interaction with *IGF1R* receptor (Fig. 4B) but also the insulin receptor substrate gene *IRS2*, which encodes a protein involved in *IGF1R* signal transduction (Fig. 4C). In addition,

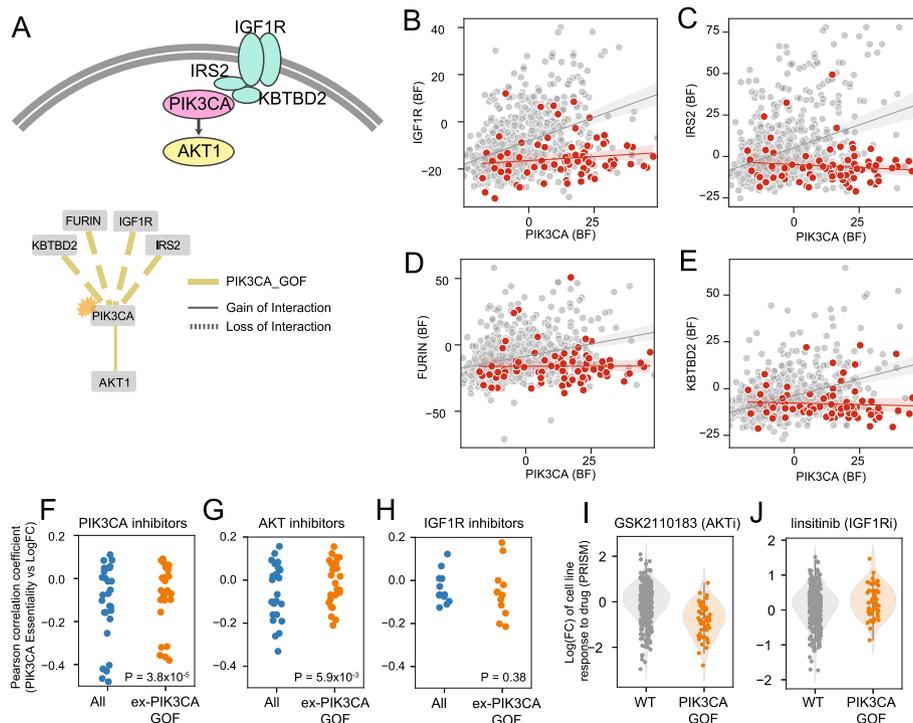


Fig. 4 Dynamics of network rewiring on IGF1R-PIK3CA pathway. **A** A visualization of IGF1R-PIK3CA pathways (upper panel) and a differential network of pathway genes caused by PIK3CA GOF mutation. Upstream pathway genes were disrupted by mutation while correlation with downstream pathway gene AKT1 was boosted. **B–E** Scatter plots of interactions of PIK3CA upstream genes **B** IGF1R and **C** IRS2, plus **D** IGF1R maturation factor FURIN, and **E** KBTBD2. **F–H** Comparisons of effectiveness of **F** PIK3CA inhibitors, **G** AKT inhibitors, and **H** IGF1R inhibitors in all cells (blue) and cells excluding PIK3CA GOF mutations (orange), from PRISM. **I** AKT inhibitor is more effective in PIK3CA GOF, while **J** IGF1R inhibitor is more effective in PIK3CA wt

PIK3CA_GOF causes loss of interaction with *FURIN* (Fig. 4D), a protease required for maturation of the *IGF1R* receptor [47] (*FURIN-IGF1R* $PCC_{all} = 0.540$, $n = 808$ cell lines) and uncharacterized gene *KBTBD2* (Fig. 4E). The high *KBTBD2-IRS2* background correlation ($PCC_{all} = 0.483$), and the loss of interaction with *PIK3CA* upon oncogenic mutation ($PCC_{all} = 0.293$, $PCC_{\sim PIK3CA_GOF} = 0.400$, $dPCC = -0.107$, $P < 0.001$), support a functional interaction between *KBTBD2* and *IRS2*, possibly involving protein maturation or signaling.

We further investigated the impact of interaction rewiring on drug efficacy. We compared the effect of drugs targeting *PIK3CA*, *AKT1*, and *IGF1R*, between wild-type and mutant *PIK3CA* cell lines in PRISM [48] 19Q3 data (Fig. 4F–H), where strong drug efficacy is shown as negative log fold change and negative correlation between BF profiles from CRISPR screens and drug response profiles from PRISM indicates concordance. Cell lines harboring a *PIK3CA* oncogenic mutation were significantly more sensitive to AKT inhibitor GSK2110183 (Fig. 4I). Conversely, *IGF1R* inhibitor linsitinib is effective only in a subset of cell lines with wildtype *PIK3CA* alleles (Fig. 4J). Together, these findings are consistent with the mutational rewiring of the functional interaction network and demonstrate that context-dependent networks can inform drug sensitivity.

Tissue-specific functional interaction

While *IGF1R* is associated with its canonical downstream signaling partner *PIK3CA*, its strongest association in the global coessentiality network is with *FURIN* ($PCC_{all} = 0.540$), reflecting the role of *FURIN* in *IGF1R* receptor maturation [47]. *FURIN* is also strongly associated with carboxypeptidase D (*CPD*), recently reported to be required for pro-*IGF1R* maturation in lung adenocarcinoma [49] (Fig. 5A). Interestingly, the link between *CPD* and *IGF1R* is largely abrogated in glioma cells (Fig. 5B), while the *CPD-FURIN* relationship is maintained (Fig. 5C).

Together these observations suggest that the *CPD-FURIN* activity is required for maturation of a different protein in glioma cells. Notably, hepatocyte growth factor receptor tyrosine kinase *MET* shows a strong gain of interaction with *CPD* in glioma (Fig. 5D). *MET* polypeptide, like *IGF1R*, is proteolytically processed into a mature receptor through cleavage into alpha and beta subunits and linkage via disulfide bonds [50]. Our context-dependent coessentiality network predicts functional linkage between *CPD* and *MET*, specifically in glioma cells, and prior knowledge of *CPD* function elsewhere suggests *CPD* may be involved in posttranslational processing of *MET* proreceptor, a hypothesis also suggested by Han et al. [49].

To test this prediction, we used a dual-guide CRISPR/enCas12a vector (see “Methods”) to knock out either *CPD* or negative control *ADH7* in glioblastoma cell lines and examined *MET* protein. We selected three brain cancer cell lines from PICKLES [51] where the *MET* receptor was expressed but not essential (LN18, U118MG, and M095K). As measured by Western blotting, under normal conditions, *MET* is more abundant than its precursor pro*MET*, indicating efficient

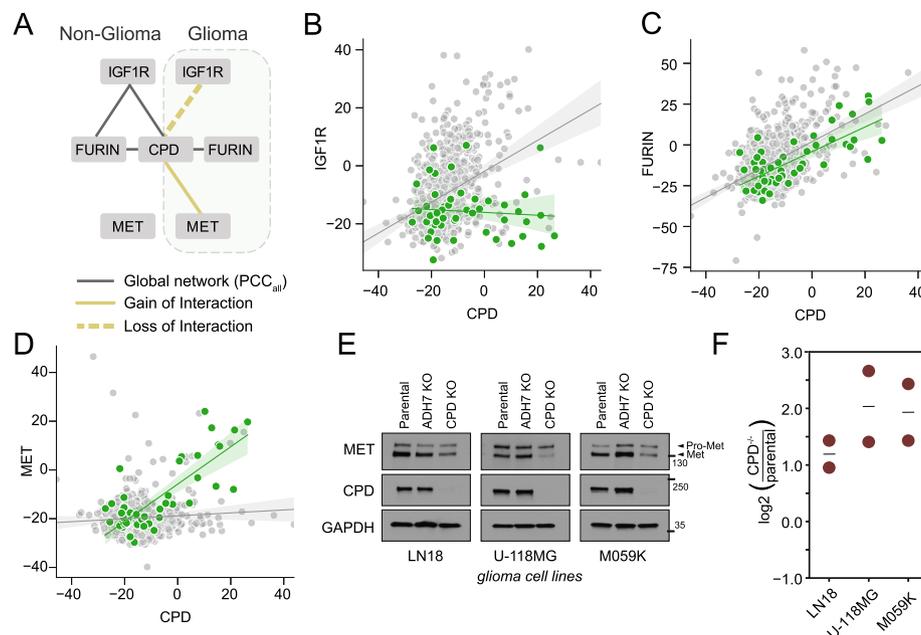


Fig. 5 Tissue-specific network rewiring. **A** A network diagram of interaction rewiring of *CPD* in glioma cell lines. **B** *IGF1R-CPD* loss of interaction was detected in glioma cell lines. **C** *CPD-FURIN* interaction was correlated regardless of tissue type. **D** *CPD* is correlated with *MET* instead of *IGF1R* in glioma cell lines. **E** Abundance of pro*MET* and mature *MET* receptor in three glioma cell lines, in response to *CPD* knockout or negative control (*ADH7ko*). **F** Densitometry of Western blots shows ratio of pro*MET* to *MET* more than doubles in *CPD*^{-/-} cells

posttranslational maturation of the polypeptide (Fig. 5E). In contrast, in *CPD*^{-/-} cells, the ratio of proMET to MET is, on average, 3.6 times that of parental cells (Fig. 5F; Additional file 4). These observations are consistent with *CPD* playing an important role in maturation of the *MET* proreceptor in glioblastoma cells.

Diffnet: explore dynamic functional interactions

A website is available at <https://diffnet.hart-lab.org> for interactive visualization of context-dependent functional interactions. Users can browse or search predefined sets of genetic lesions or tissue/tumor types and see the networks that are associated with differential coessentiality in that context as well as a table of network edges along with key parameters, including PCC, dPCC, and empirical *P*-values. Clicking any edge in the network view will display a scatterplot of logBF scores for the two connected genes across all cell lines in the DepMap data, annotated by whether they are in the selected context. A screenshot of the BRAF/RAF1 change in interaction in the BRAF_GOF context is shown in Additional file 1: Figure S2.

Discussion

In this study, we present a framework for understanding the differential wiring of cells across genotypes and lineages. Using genetic mutation information and cell line metadata, we investigated contexts that cause emergent essentiality by analyzing coefficients obtained from a logistic regression model trained using CRISPR screen data. Using this approach, we identified numerous context-gene associations, including known interactions such as paralog buffering, oncogenes, and tissue-specific essential genes.

We then examined the rewiring of functional interaction networks in the presence of these contexts. We developed a strategy which compares the strength of interaction between two genes, as measured by the Pearson correlation coefficient of their normalized fitness vectors across all samples, to the correlation derived from samples that exclude a specified context, or a “leave-one-out” test. We measured the significance of these changes by bootstrapping a null distribution for every gene pair in every context. Rewiring detected by our approach showed concordance with biological knowledge and discovered new putative context-dependent functional interactions, demonstrating its potential for functional genomics and cancer targeting. We provide a web-based interface for exploring the network edges rewired by mutation and/or lineage at diffnet.hart-lab.org.

Conclusions

The approach derived herein attempts to address one of the key questions arising from the use of coessentiality, or indeed any gene similarity approach, to predict gene co-functionality. These methods offer a powerful approach for predicting gene function, identifying disease genes, and reducing the search space of potential combinatorial gene effects to tractable levels. Coessentiality depends on some underlying variation in gene essentiality across the dataset; genes which are always essential or never essential rarely appear in these networks [9]. In most approaches, the source of this variation is neither known nor questioned; however, the question is of high interest because identifying the causal basis for emergent essentiality would provide direct association of biomarkers

with potential cancer targets as well as insight into the critical biological processes in different cells.

Another feature of raw coessentiality networks is that, as with prior efforts to build integrative functional interaction networks [26], they represent an integration of all contexts present in the cell lines or data sets from which they are derived. Functional interactions between the same genes may differ across contexts [28, 52]; for example, the roles of *FURIN* and *CPD* in promoting the maturation of different cell surface receptors in different tissues. Combining these contexts dilutes the overall correlation between genes with strong but context-dependent functional interaction. The approach we describe here offers a path towards understanding how context-dependent interactions can be gleaned from these data.

Methods

Preprocessing publicly available CRISPR screens

Raw read counts of CRISPR screens used in this study were downloaded from DepMap 20Q4 ($n = 808$ cell lines). For DepMap screens, we removed guide RNAs targeting multiple protein coding genes based on the guide map obtained from the DepMap database and the CCDS protein coding gene annotation to avoid genetic interaction effects. After that, we applied the BAGEL2 pipeline [36] to measure gene essentiality information for each cell line. Gene knockout fitness phenotype is reported as a log Bayes Factor (BF), with positive scores indicating likelihood of essentiality.

Predictive features and response variables for the logistic regression model

Genetic lesions were classified as gain or loss of function based on mutation calls described in the Cancer Cell Line Encyclopedia. For each of the 808 cell lines, each gene ($n = 18,111$) was classified as loss of function (LOF) if it carried a lesion whose “Variant_annotation” annotation was “damaging.” Genes were characterized as gain of function (GOF) if they carried a lesion where either isTCGAhotspot or isCOSMIChotspot = True and if “Variant_annotation” was “other non-conserving.” Other lesions were not classified. See notebook “pre-step01-clean_data” for details.

CCLC Gene expression data in logTPM was downloaded from the CCLC portal at the Broad ($n = 718$ cell lines with matching CRISPR data from Avana 20q4). EMT state was determined by CDH1/VIM expression log ratio ($\log\text{TPM}_{\text{CDH1}} - \log\text{TPM}_{\text{VIM}}$), which resulted in a bimodal distribution across all cell lines. Each cell line was assigned CDH_VIM_hi if the log ratio > 1 , or CDH_VIM_lo for log ratio < -4 (Additional file 1: Figure S3). This CDH_VIM_hi and CDH_VIM_lo were the only gene expression feature used in the classifier. See notebook “step01-merge_features” for details.

Other cell line metadata, downloaded from the CCLC portal at the Broad ($n = 808$ cell lines with matching CRISPR data from Avana 20q4), were transformed into binary features. Each unique entry for “lineage,” “lineage_subtype,” “lineage_sub_subtype,” “culture_type,” and “sex” was considered a unique feature vector. See notebook “step01-merge_features” for details.

All features were merged into one matrix (808 cell lines \times 36,406 features). Features describing six or fewer cell lines ($n = 33,475$) were excluded. Remaining features were then de-duplicated by calculating an all-by-all correlation matrix, and for every pair of

features with $PCC \geq 0.9$, the smaller feature (the one describing fewer cell lines) was discarded. For features with equal size and perfect overlap (e.g., “plasma_cell” and “multiple_myeloma”), one was selected at random and discarded. The final de-duplicated binary feature set comprised 2918 predictor variables for the regression model. See notebook “step02-remove_duplicates” for details.

Gene essentiality was binarized for use as a response variable. Genes with missing values in the Bayes Factor table ($n = 17$) were removed from the dataset. Then, for each cell line, genes with $BF \geq 10$ were classified as essential and all others nonessential, for a set of 18,094 genes in 808 cell lines. We used a binary rather than continuous response variable because, in this study, we seek to identify predictors of whether a gene does or does not have a fitness defect, not to predict variation in that defect. Moreover, while the magnitude of positive log BF may be a proxy for severity of fitness defect, the magnitude of negative log BF is usually not (though see [53] for exceptions). Bayes Factor combines statistical confidence and effect size into a single output, and the dynamic range of BF scores from each screen is highly dependent on that screen’s quality [36, 54]. Bayes Factors are not directly comparable across screens without normalization (see Global normalization, below). See notebook “step02-remove_duplicates” for details on this processing step.

Logistic regression model

Each gene’s binary essentiality vector was used as the response variable in a logistic regression model, using the binary features described above. Genes whose essentiality is invariant across the cell lines are uninteresting; therefore, we excluded genes essential in less than 1% or more than 80% of the samples (Additional file 1: Figure S3). The remaining set included 2987 genes.

A logistic regression model was implemented in Python using the LogisticRegression function in scikit-learn, with an elastic net penalty (L1 ratio = 0.25). Processing time took 78 min on an AMD 5950X CPU with 32 threads. The coefficients for each feature/gene prediction were combined into a single matrix. The intercept of each model closely matched the background probability of gene essentiality (Additional file 1: Figure S1b). For each of the 2987 response genes, the maximum and minimum feature coefficients were determined and, based on histograms of the maximum and minimum values (Additional file 1: Figure S1cd), $|\text{coeff}| > 1.2$ was chosen as a threshold for strong feature-gene association. The feature-gene coefficients matrix was flattened to a list and sorted by coefficient, and this list was used for subsequent analyses. A total of 393 unique features were predictive of gene essentiality with $|\text{coeff}| > 1.2$. See notebook “step03-do_regression” and “step03a-audit_regression” for details. The matrix of all feature-gene coefficients is included as Additional file 2.

Measurement of emergent network rewiring in CRISPR dataset

Global normalization and correlation

To calculate the baseline coessentiality network, we first quantile normalized the log BF table of 18,094 genes with no missing data \times 808 cell lines. Quantile normalization was performed without a reference distribution, setting each gene to its rank mean; e.g., the top-ranked gene in each sample was assigned a value equal to the mean BF of

all top-ranked genes in 808 samples. The global Pearson correlation network of all gene pairs was then calculated using the Python numpy function `corrcoef`, resulting in an $18\text{ k} \times 18\text{ k}$ matrix of Pearson correlation coefficients hereafter referred to as PCC_{all} .

Leave-one-out test

To test for dynamic functional network rewiring, we measured the differential Pearson correlation coefficient (dPCC) between the PCC_{all} global network and the network derived from a subset of cell lines that exclude a given context ($PCC_{\sim\text{context}}$). The formula of dPCC for each gene pair is therefore:

$$dPCC = PCC_{\text{all}} - PCC_{\sim\text{context}}$$

where $PCC_{\sim\text{context}}$ is an $18\text{ k} \times 18\text{ k}$ matrix of Pearson correlation coefficients derived from the $[18\text{ k} \times (808 - n)]$ matrix of quantile normalized BFs, and n is the number of cell lines associated with the “context.” Positive dPCC represents a *gain of interaction (GOI)* because the higher PCC_{all} depends on the presence of cells with the feature of interest, and negative dPCC represents *loss of interaction (LOI)* because excluding cells associated with the feature of interest increases the PCC.

We evaluated the statistical significance of the observed dPCC ($dPCC_{\text{obs}}$) for each gene pair/feature combination empirically. Briefly, we determined n , the number of cell lines associated with the feature. Then we randomly removed n cell lines from the set of all 808 cell lines and calculated the all-by-all PCC matrix for all pairs of genes, PCC_{rand} . For each gene pair, we measured $dPCC_{\text{rand}} = PCC_{\text{all}} - PCC_{\text{rand}}$. We repeated this process 1000 \times and compared the observed $dPCC_{\text{obs}}$ to the bootstrapped distribution of $dPCC_{\text{rand}}$ to calculate a two-sided empirical P -value that $dPCC_{\text{obs}}$ is higher or lower than random expectation. In general, our sampled null distributions of $dPCC_{\text{rand}}$ are not well estimated by classical distributions, and our estimate of P -value is therefore limited to a discrete distribution in steps of 0.001. We repeated this bootstrap resampling to generate null distributions of $dPCC_{\text{rand}}$ for every gene pair for each of the 393 contexts identified by the logistic regression model (a computationally expensive process). See “`calc_diff_coess_pval.py`” for code and details.

It is important to note that this empirical approach does not rely the robust measurement of Pearson correlation—in fact, the approach described here exploits the Pearson’s lack of robustness in the presence of outliers. When the outliers are functionally coherent in some aspect (e.g., breast cancer cells in Fig. 2G, glioma cells in Fig. 5D), their removal in the leave-one-out test and the resulting major change in PCC is precisely the result of the Pearson’s sensitivity to outliers, and would likely not be detectable with nonparametric alternatives such as Spearman’s rank correlation that are designed to be insensitive to such changes. Of further importance is that we do not evaluate the statistical significance of any Pearson correlation, in large part because of these concerns. Our only test statistic is dPCC, measured against null distributions that we sample empirically.

CPD knockout in glioblastoma cells

gRNA plasmid construction

Two gRNAs targeting ADH7 and CPD were designed by CRISPick [55]. Oligonucleotides containing two enCas12a crRNAs (crRNA-DR1-crRNA) (crRNA sequences were listed in Table 1; DR1 is drawn from [56]) were synthesized by Integrated Device Technology and cloned to the pRDA_550 one-component enCas12a/crRNA expression vector (gifted from John Doench) by BsmBI-v2 Golden Gate Assembly (New England Biolabs, E1602S) following the manufacturer's instructions.

Cell culture

M059K (CRL-2365), LN-18 (CRL-2610), and U-118 MG (HTB-15) were sourced from the American Type Culture Collection (ATCC). Cell line identities were confirmed by STR fingerprinting by M.D. Anderson Cancer Center's Cytogenetic and Cell Authentication Core (Promega Powerplex 16 High Sensitivity Assay). All cell lines were routinely tested for mycoplasma contamination using cells cultured in non-antibiotic medium (PlasmoTest Mycoplasma Detection Assay, InvivoGen).

All cell lines were grown at 37 °C in humidified incubators at 5.0% CO₂. Cell lines were cultured in HEPES modified Dulbecco's modified Eagle's medium (Sigma D6171) supplemented with 10% FBS (Sigma), 1 mM sodium pyruvate (Gibco), 2 mM L-alanyl-L-glutamine dipeptide (Gibco), 1× penicillin-streptomycin (Gibco), and 100 µg/mL Normocin (InvivoGen).

Lentivirus was produced by the University of Michigan Vector Core. Virus stocks were not titered in advance. All transductions were performed in a 6-well plate format across multiple plates with a range of virus volumes and 8 µg/mL polybrene (EMD Millipore). The pool with the most optimal transduction efficiency was expanded and collected for western blotting.

Non-transduced cells were eliminated via selection with 2 µg/mL puromycin dihydrochloride (Gibco). Selection was maintained until all non-transduced control cells reached 0% viability. Once selection with puromycin was complete, three replicates were seeded in a 6-well plate format using the following amounts of viable cells: 4 × 10⁵ LN-18 per well, 6 × 10⁵ U-118 MG per well, 6 × 10⁵ M059K per well.

Western blots

Cells were washed with PBS and harvested in RIPA lysis buffer (Thermo Fisher Scientific, 89900) supplemented with Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific, 87786). Twenty micrograms of cell lysate was denatured in Laemmli SDS-Sample Buffer (Boston BioProducts, BP-110R) and loaded on a precast 4–15% Mini-Protean TGX Stain-Free gel (Bio-Rad). After electrophoresis and transfer, the PVDF membrane (Bio-Rad) was blocked in 5% nonfat milk with TBST (Bio-Rad) at room temperature for 1 h and incubated with primary antibody overnight. MET (3127, 1:1000) antibody

Table 1 Sequences of crRNAs

Genes	crRNA1	crRNA2
CPD	5'-CCAAGGAATAAAGCCGGGTAAATG-3'	5'-GGAACAGAATCGAAGATCACTAA-3'
ADH7	5'-GCCACAGGAATCTGTCGCACAGA-3'	5'-TCTCCACAGGTCAAACCTGGTTC-3'

was purchased from cell signaling technology. CPD (A305-514A-M, 1:1000) and GAPDH (MA5-15738, 1:5000) antibodies were purchased from Thermo Fisher Scientific. The membranes were then washed in TBST three times, followed by incubation with goat anti-rabbit/Mouse secondary antibodies (Cell Signaling Technology, 1:2000–1:5000) for 1 h at room temperature. After washing with TBST, the membranes were developed by films. All western blots were repeated twice. Image quantification was performed in ImageJ.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13059-022-02712-z>.

Additional file 1: Figure S1. Regression model. **Figure S2.** Screenshot of Diffnet web service at <https://diffnet.hart-lab.org>. **Figure S3.** Selected regression model features. **Figure S4.** Raw Western blots from Fig. 5 in main text.

Additional file 2. Tab-delimited table of regression coefficients, with 2918 genomic and lineage features used to predict background-specific essentiality of 2987 genes.

Additional file 3. Tab-delimited table of top 2644 hits from differential PCC analysis.

Additional file 4. Quantitative data from Western blot experiments.

Additional file 5. Review history.

Acknowledgements

EK, MC, LB, CL, and TH were supported by NIGMS grant R35GM130119 and CL and TH were supported by CPRIT grant RP210173. EK was supported by a grant from the Prostate Cancer Foundation. TH is a CPRIT Scholar in Cancer Research (RR160032). This work was supported by the Andrew Sabin Family Foundation Fellowship and by NCI Cancer Center Support Grant P30CA16672.

Peer review information

Kevin Pang was the primary editor of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Review history

The review history is available as Additional file 5.

Authors' contributions

EK developed the hypothesis. EK, TH, and VG performed bioinformatic analysis, with contributions from LN and CB. LN and MC developed the *diffnet* web application under the supervision of CB and TH. LB and CL performed all wet lab experiments. EK and TH drafted the manuscript and all authors edited it. The author(s) read and approved the final manuscript.

Availability of data and materials

Code for the data analysis pipeline and for the *diffnet* website can be accessed on demand at the repository at Github [57] and Zenodo [58].

The *diffnet* website is available at <https://diffnet.hart-lab.org/>.

Declarations

Ethics approval and consent to participate

Ethics approval was not required for this study.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. ²Present Address: Novartis Institutes for BioMedical Research (NIBR), San Diego, CA, USA. ³TRACtion, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. ⁴UTHealth Graduate School of Biomedical Sciences, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ⁵Department of Cancer Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.

Received: 20 September 2021 Accepted: 17 June 2022

Published online: 29 June 2022

References

1. Hart T, Chandrashekhar M, Aregger M, Steinhart Z, Brown KR, MacLeod G, et al. High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities. *Cell*. 2015;163(6):1515–26. <https://doi.org/10.1016/j.cell.2015.11.015>.

2. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen T, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*. 2014;343(6166):84–7. <https://doi.org/10.1126/science.1247005>.
3. Wang T, Birsoy K, Hughes NW, Krupczak KM, Post Y, Wei JJ, et al. Identification and characterization of essential genes in the human genome. *Science*. 2015;350(6264):1096–101. <https://doi.org/10.1126/science.aac7041>.
4. Aguirre AJ, Meyers RM, Weir BA, Vazquez F, Zhang CZ, Ben-David U, et al. Genomic copy number dictates a gene-independent cell response to CRISPR/Cas9 Targeting. *Cancer Discov*. 2016;6(8):914–29. <https://doi.org/10.1158/2159-8290.CD-16-0154>.
5. Behan FM, Iorio F, Picco G, Gonçalves E, Beaver CM, Migliardi G, et al. Prioritization of cancer therapeutic targets using CRISPR–Cas9 screens. *Nature*. 2019;568(7753):511–6. <https://doi.org/10.1038/s41586-019-1103-9>.
6. Meyers RM, Bryan JG, McFarland JM, Weir BA, Sizemore AE, Xu H, et al. Computational correction of copy number effect improves specificity of CRISPR–Cas9 essentiality screens in cancer cells. *Nat Genet*. 2017;49(12):1779–84. <https://doi.org/10.1038/ng.3984>.
7. Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kryukov G, Cowley GS, et al. Defining a cancer dependency map. *Cell*. 2017;170(3):564–576.e16.
8. Boyle EA, Pritchard JK, Greenleaf WJ. High-resolution mapping of cancer cell networks using co-functional interactions. *Mol Syst Biol*. 2018;14(12):e8594.
9. Kim E, Dede M, Lenoir WF, Wang G, Srinivasan S, Colic M, et al. A network of human functional gene interactions from knockout fitness screens in cancer cells. *Life Sci Alliance*. 2019;2(2):e201800278. <https://doi.org/10.26508/lsa.201800278>.
10. Pan J, Meyers RM, Michel BC, Mashtalir N, Sizemore AE, Wells JN, et al. Interrogation of mammalian protein complex structure, function, and membership using genome-scale fitness screens. *Cell Systems*. [cited 2018 May 18]; Available from: <https://www.sciencedirect.com/science/article/pii/S240547121830187X>
11. Rauscher B, Heigwer F, Henkel L, Hielscher T, Voloshanenko O, Boutros M. Toward an integrated map of genetic interactions in cancer cells. *Mol Syst Biol*. 2018;14(2):e7656. <https://doi.org/10.15252/msb.20177656>.
12. Wainberg M, Kamber RA, Balsubramani A, Meyers RM, Sinnott-Armstrong N, Hornburg D, et al. A genome-wide atlas of co-essential modules assigns function to uncharacterized genes. *Nat Genet*. 2021;53(5):638–49. <https://doi.org/10.1038/s41588-021-00840-z>.
13. Wang T, Yu H, Hughes NW, Liu B, Kendirli A, Klein K, et al. Gene essentiality profiling reveals gene networks and synthetic lethal interactions with oncogenic Ras. *Cell*. 2017;168(5):890–903 e15.
14. Aregger M, Lawson KA, Billmann M, Costanzo M, Tong AHY, Chan K, et al. Systematic mapping of genetic interactions for de novo fatty acid synthesis identifies C12orf49 as a regulator of lipid metabolism. *Nature Metabolism*. 2020;2(6):499–513. <https://doi.org/10.1038/s42255-020-0211-z>.
15. Noordermeer SM, Adam S, Setiawati D, Barazas M, Pettitt SJ, Ling AK, et al. The shieldin complex mediates 53BP1-dependent DNA repair. *Nature*. 2018;560(7716):117–21. <https://doi.org/10.1038/s41586-018-0340-7>.
16. Zimmermann M, Murina O, Reijns MAM, Agathangelou A, Challis R, Tarnauskaitė Ž, et al. CRISPR screens identify genomic ribonucleotides as a source of PARP-trapping lesions. *Nature*. 2018;559(7713):285–9. <https://doi.org/10.1038/s41586-018-0291-z>.
17. Amici DR, Jackson JM, Truica MI, Smith RS, Abdulkadir SA, Mendillo ML. FIREWORKS: a bottom-up approach to integrative coessentiality network analysis. *Life Sci Alliance*. 2020;4(2):e202000882. <https://doi.org/10.26508/lsa.202000882>.
18. Sharma S, Dincer C, Weidemüller P, Wright GJ, Petsalaki E. CEN-tools: an integrative platform to identify the contexts of essential genes. *Mol Syst Biol*. 2020;16(10) [cited 2021 Jun 24]. Available from: <https://onlinelibrary.wiley.com/doi/10.15252/msb.20209698>.
19. Ding KF, Finlay D, Yin H, Hendricks WPD, Sereduk C, Kiefer J, et al. Network rewiring in cancer: applications to melanoma cell lines and the Cancer Genome Atlas patients. *Front Genet*. 2018;9:228. <https://doi.org/10.3389/fgene.2018.00228>.
20. Ideker T, Krogan NJ. Differential network biology. *Mol Syst Biol*. 2012;8(1):565. <https://doi.org/10.1038/msb.2011.99>.
21. Kim J, Kim I, Han SK, Bowie JU, Kim S. Network rewiring is an important mechanism of gene essentiality change. *Sci Rep*. 2012;2(1):900. <https://doi.org/10.1038/srep00900>.
22. Cho SB, Kim J, Kim JH. Identifying set-wise differential co-expression in gene expression microarray data. *BMC Bioinformatics*. 2009;10(1):109. <https://doi.org/10.1186/1471-2105-10-109>.
23. Hsu CL, Juan HF, Huang HC. Functional analysis and characterization of differential coexpression networks. *Sci Rep*. 2015;5(1):13295. <https://doi.org/10.1038/srep13295>.
24. Lui TWH, Tsui NBY, Chan LWC, Wong CSC, Siu PMF, Yung BYM. DECODE: an integrated differential co-expression and differential expression analysis of gene expression data. *BMC Bioinformatics*. 2015;16(1):182. <https://doi.org/10.1186/s12859-015-0582-4>.
25. Kim WK, Krumpelman C, Marcotte EM. Inferring mouse gene functions from genomic-scale data using a combined functional network/classification strategy. *Genome Biology*. 2008;9(1):S5. <https://doi.org/10.1186/gb-2008-9-s1-s5>.
26. Lee I, Date SV, Adai AT, Marcotte EM. A probabilistic functional network of yeast genes. *Science*. 2004;306(5701):1555–8. <https://doi.org/10.1126/science.1099511>.
27. Singh-Blom UM, Natarajan N, Tewari A, Woods JO, Dhillon IS, Marcotte EM. Prediction and validation of gene-disease associations using methods inspired by social network analyses. *PLOS ONE*. 2013;8(5):e58977. <https://doi.org/10.1371/journal.pone.0058977>.
28. Greene CS, Krishnan A, Wong AK, Ricciotti E, Zelaya RA, Himmelstein DS, et al. Understanding multicellular function and disease with human tissue-specific networks. *Nat Genet*. 2015;47(6):569–76. <https://doi.org/10.1038/ng.3259>.
29. Guan Y, Gorenshsteyn D, Burmeister M, Wong AK, Schimenti JC, Handel MA, et al. Tissue-specific functional networks for prioritizing phenotype and disease genes. *PLoS Comput Biol*. 2012;8(9):e1002694. <https://doi.org/10.1371/journal.pcbi.1002694>.
30. Rancati G, Moffat J, Typas A, Pavelka N. Emerging and evolving concepts in gene essentiality. *Nat Rev Genet*. 2018;19(1):34–49. <https://doi.org/10.1038/nrg.2017.74>.
31. Janes MR, Zhang J, Li LS, Hansen R, Peters U, Guo X, et al. Targeting KRAS mutant cancers with a covalent G12C-specific inhibitor. *Cell*. 2018;172(3):578–89 e17.

32. O'Leary K. Tracing the origins of KRAS oncogene addiction. *Nat Rev Cancer*. 2021;21(2):69. <https://doi.org/10.1038/s41568-020-00325-x>.
33. Waters AM, Der CJ. KRAS: The critical driver and therapeutic target for pancreatic cancer. *Cold Spring Harb Perspect Med*. 2018;8(9):a031435. <https://doi.org/10.1101/cshperspect.a031435>.
34. Helming KC, Wang X, Wilson BG, Vazquez F, Haswell JR, Manchester HE, et al. ARID1B is a specific vulnerability in *ARID1A*-mutant cancers. *Nature Medicine*. 2014;20(3):251–4. <https://doi.org/10.1038/nm.3480>.
35. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*. 2012;483(7391):603–7. <https://doi.org/10.1038/nature11003>.
36. Kim E, Hart T. Improved analysis of CRISPR fitness screens and reduced off-target effects with the BAGEL2 gene essentiality classifier. *Genome Med*. 2021;13(1):2. <https://doi.org/10.1186/s13073-020-00809-3>.
37. Lord CJ, Quinn N, Ryan CJ. Integrative analysis of large-scale loss-of-function screens identifies robust cancer-associated genetic interactions. *eLife*. 2020;9:e58925.
38. Asati V, Mahapatra DK, Bharti SK. PI3K/Akt/mTOR and Ras/Raf/MEK/ERK signaling pathways inhibitors as anticancer agents: structural and pharmacological perspectives. *Eur J Med Chem*. 2016;109:314–41. <https://doi.org/10.1016/j.ejmech.2016.01.012>.
39. Lavoie H, Therrien M. Regulation of RAF protein kinases in ERK signalling. *Nat Rev Mol Cell Biol*. 2015;16(5):281–98. <https://doi.org/10.1038/nrm3979>.
40. Santarpia L, Lippman SM, El-Naggar AK. Targeting the MAPK-RAS-RAF signaling pathway in cancer therapy. *Expert Opin Ther Targets*. 2012;16(1):103–19. <https://doi.org/10.1517/14728222.2011.645805>.
41. Terrell EM, Durrant DE, Ritt DA, Sealover NE, Sheffels E, Spencer-Smith R, et al. Distinct binding preferences between Ras and Raf family members and the impact on oncogenic Ras signaling. *Mol Cell*. 2019;76(6):872–84 e5.
42. Terrell EM, Morrison DK. Ras-mediated activation of the Raf family kinases. *Cold Spring Harb Perspect Med*. 2019;9(1):a033746. <https://doi.org/10.1101/cshperspect.a033746>.
43. Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, Tamborero D, Schroeder MP, Jene-Sanz A, et al. IntOGen-mutations identifies cancer drivers across tumor types. *Nat Methods*. 2013;10(11):1081–2. <https://doi.org/10.1038/nmeth.2642>.
44. Martínez-Jiménez F, Muiños F, Sentís I, Deu-Pons J, Reyes-Salazar I, Arnedo-Pac C, et al. A compendium of mutational cancer driver genes. *Nat Rev Cancer*. 2020;20(10):555–72. <https://doi.org/10.1038/s41568-020-0290-x>.
45. Zha J, Lackner MR. Targeting the insulin-like growth factor receptor-1R pathway for cancer therapy. *Clin Cancer Res*. 2010;16(9):2512–7. <https://doi.org/10.1158/1078-0432.CCR-09-2232>.
46. Gkeka P, Evangelidis T, Pavlaki M, Lazani V, Christoforidis S, Agianian B, et al. Investigating the structure and dynamics of the PIK3CA wild-type and H1047R oncogenic mutant. *PLoS Comput Biol*. 2014;10(10):e1003895. <https://doi.org/10.1371/journal.pcbi.1003895>.
47. Alarcón C, Cheatham B, Lincoln B, Kahn CR, Siddle K, Rhodes CJ. A Kex2-related endopeptidase activity present in rat liver specifically processes the insulin proreceptor. *Biochem J*. 1994;301(Pt 1):257–65. <https://doi.org/10.1042/bj3010257>.
48. Yu C, Mannan AM, Yvone GM, Ross KN, Zhang YL, Marton MA, et al. High-throughput identification of genotype-specific cancer vulnerabilities in mixtures of barcoded tumor cell lines. *Nat Biotechnol*. 2016;34(4):419–23. <https://doi.org/10.1038/nbt.3460>.
49. Han K, Pierce SE, Li A, Spees K, Anderson GR, Seoane JA, et al. CRISPR screens in cancer spheroids identify 3D growth-specific vulnerabilities. *Nature*. 2020;580(7801):136–41. <https://doi.org/10.1038/s41586-020-2099-x>.
50. Komada M, Hatsuzawa K, Shibamoto S, Ito F, Nakayama K, Kitamura N. Proteolytic processing of the hepatocyte growth factor/scatter factor receptor by furin. *FEBS Lett*. 1993;328(1–2):25–9. [https://doi.org/10.1016/0014-5793\(93\)80958-W](https://doi.org/10.1016/0014-5793(93)80958-W).
51. Lenoir WF, Lim TL, Hart T. PICKLES: the database of pooled in-vitro CRISPR knockout library essentiality screens. *Nucleic Acids Res*. 2018;46(D1):D776–80. <https://doi.org/10.1093/nar/gkx993>.
52. Bandyopadhyay S, Mehta M, Kuo D, Sung MK, Chuang R, Jaehning EJ, et al. Rewiring of genetic networks in response to DNA damage. *Science*. 2010;330(6009):1385–9. <https://doi.org/10.1126/science.1195618>.
53. Lenoir WF, Morgado M, DeWeirdt PC, McLaughlin M, Griffith AL, Sangree AK, et al. Discovery of putative tumor suppressors from CRISPR screens reveals rewired lipid metabolism in acute myeloid leukemia cells. *Nat Commun*. 2021;12(1):6506. <https://doi.org/10.1038/s41467-021-26867-8>.
54. Wang R, Lenoir WF, Wang C, Su D, McLaughlin M, Hu Q, et al. DNA polymerase ϵ compensates for Fanconi anemia pathway deficiency by countering DNA replication stress. *Proc Natl Acad Sci*. 2020;117(52):33436–45. <https://doi.org/10.1073/pnas.2008821117>.
55. Kim HK, Min S, Song M, Jung S, Choi JW, Kim Y, et al. Deep learning improves prediction of CRISPR–Cpf1 guide RNA activity. *Nature Biotechnology*. 2018;36(3):239–41. <https://doi.org/10.1038/nbt.4061>.
56. DeWeirdt PC, Sanson KR, Sangree AK, et al. Optimization of AsCas12a for combinatorial genetic screens in human cells. *Nat Biotechnol*. 2021;39:94–104. <https://doi.org/10.1038/s41587-020-0600-6>.
57. Eiru Kim, Lance C. Novak, Medina Colic, Veronica Gheorghe, Christopher A. Bristow, Traver Hart. Code and datasets. Github. 2022. Available from: https://github.com/hart-lab/dynamic_coessentiality_rewiring/tree/v1.0.1
58. Kim E, Novak LC, Colic M, Gheorghe V, Bristow CA, Hart T. Code and datasets. Zenodo; 2022. Available from: <https://doi.org/10.5281/zenodo.6653887>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.