

RESEARCH ARTICLE

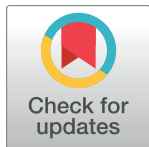
# MEDICI: Mining Essentiality Data to Identify Critical Interactions for Cancer Drug Target Discovery and Development

Sahar Harati<sup>1,2</sup>, Lee A. D. Cooper<sup>1,3,4</sup>, Josue D. Moran<sup>5,6</sup>, Felipe O. Giuste<sup>7</sup>, Yuhong Du<sup>3,8</sup>, Andrei A. Ivanov<sup>8</sup>, Margaret A. Johns<sup>8</sup>, Fadlo R. Khuri<sup>3,9</sup>, Haian Fu<sup>3,8</sup>, Carlos S. Moreno<sup>1,3,6\*</sup>

**1** Department of Biomedical Informatics, Emory University, Atlanta, Georgia, United States of America, **2** Graduate Program in Biomedical Informatics, Emory University, Atlanta, Georgia, United States of America, **3** Winship Cancer Institute, Emory University, Atlanta, Georgia, United States of America, **4** Department of Biomedical Engineering, Emory University, Atlanta, Georgia, United States of America, **5** Graduate Program in Cancer Biology, Emory University, Atlanta, Georgia, United States of America, **6** Department of Pathology and Laboratory Medicine, Emory University, Atlanta, Georgia, United States of America, **7** Medical Scientist Training Program, Emory University, Atlanta, Georgia, United States of America, **8** Department of Pharmacology, Emory University, Atlanta, Georgia, United States of America, **9** Department of Hematology & Medical Oncology, Emory University, Atlanta, Georgia, United States of America

☞ These authors contributed equally to this work.

\* [cmoreno@emory.edu](mailto:cmoreno@emory.edu)



**OPEN ACCESS**

**Citation:** Harati S, Cooper LAD, Moran JD, Giuste FO, Du Y, Ivanov AA, et al. (2017) MEDICI: Mining Essentiality Data to Identify Critical Interactions for Cancer Drug Target Discovery and Development. PLoS ONE 12(1): e0170339. doi:10.1371/journal.pone.0170339

**Editor:** Esmaeil Ebrahimie, Flinders University, AUSTRALIA

**Received:** July 25, 2016

**Accepted:** January 3, 2017

**Published:** January 24, 2017

**Copyright:** © 2017 Harati et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Software is freely available at <https://github.com/cooperlab/MEDICI>. Datasets are available at <https://ctd2.nci.nih.gov/dataPortal>.

**Funding:** This work was supported by National Cancer Institute grant U01 CA168449 to HF. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

Protein-protein interactions (PPIs) mediate the transmission and regulation of oncogenic signals that are essential to cellular proliferation and survival, and thus represent potential targets for anti-cancer therapeutic discovery. Despite their significance, there is no method to experimentally disrupt and interrogate the essentiality of individual endogenous PPIs. The ability to computationally predict or infer *PPI essentiality* would help prioritize PPIs for drug discovery and help advance understanding of cancer biology. Here we introduce a computational method (MEDICI) to predict *PPI essentiality* by combining gene knockdown studies with network models of protein interaction pathways in an analytic framework. Our method uses network topology to model how gene silencing can disrupt PPIs, relating the unknown essentialities of individual PPIs to experimentally observed protein essentialities. This model is then deconvolved to recover the unknown essentialities of individual PPIs. We demonstrate the validity of our approach via prediction of sensitivities to compounds based on PPI essentiality and differences in essentiality based on genetic mutations. We further show that lung cancer patients have improved overall survival when specific PPIs are no longer present, suggesting that these PPIs may be potentially new targets for therapeutic development. Software is freely available at <https://github.com/cooperlab/MEDICI>. Datasets are available at <https://ctd2.nci.nih.gov/dataPortal>.

## Introduction

Advances in high-throughput screening technology have enabled broad investigations of genome-wide gene/protein essentiality in cancer. High-throughput single-gene shRNA/siRNA silencing [1–4] and CRISPR-Cas9 inactivation [5] are well-established experimental approaches to study protein essentiality in genome-wide screens. Observing the proliferative effects of silencing each gene/node in a PPI network can provide insights into tumor biology and help identify promising therapeutic targets, especially when combined with genomic characterizations. Whole-genome siRNA screens have been combined with genomic profiles and drug screens in lung adenocarcinoma to identify context-specific drug sensitivities and their genetic biomarkers [6]. Project Achilles currently provides a pooled shRNA screening database with more than 11,000 genes in 216 cell lines [7]. Systematic analyses of these data have been able to identify specific gene vulnerabilities within genetic contexts in several studies [7–11].

The PPI interface has become increasingly recognized as a tractable target for small molecule therapeutics, as evidenced by recent clinical development of p53/MDM2 and BET bromodomain small molecule inhibitors [2, 12, 13]. Despite the therapeutic potential of protein-protein interactions (PPIs) as drug targets [14], specific analysis of protein-interaction essentiality or the essentiality of *interactions* in biological networks (‘edgetics’) is in its infancy [15, 16]. Current technologies focus on silencing of single genes in large-scale shRNA screens; however, shRNA silencing of a single gene effectively disrupts multiple PPIs and masks the contributions of individual PPIs to the overall protein essentiality. High-throughput technology for interrupting specific PPIs on a whole-interactome scale does not exist, and methods for experimentally measuring the essentiality of individual endogenous PPIs at the genome scale will likely remain an unsolved problem for the foreseeable future. While large-scale PPI screens have measured the effects of disease mutations on specific PPIs [15, 16], they do not provide data on the essentiality of endogenous interactions for the survival of a cell. Thus, we were motivated to develop a computational approach to estimate the essentiality of PPIs by integrating PPI network topology with whole-genome shRNA screens. By measuring the essentiality of every gene (node) in a network, and understanding how proteins are connected through protein interactions (edges), we aim to estimate the essentiality of individual PPIs that are silenced in aggregate as a gene is knocked down by shRNA.

The integration of functional screens with PPI networks has been previously explored with an emphasis on mitigating screening noise to improve the robustness of functional measurements. PPI networks have been integrated with RNAi screens using a diffusion kernel-based method [17] to successfully reduce false-positive and false-negative results in *Drosophila* screens. The IMPACT method used protein interactions as a method for reducing off-target effects and improving the biological interpretation of screened phenotypes [18]. In addition, KEGG networks have been integrated with siRNA screens to refine the insulin-signaling network using a network seeding/pruning approach [18]. A shortest path approach for analysis of PPI networks has been developed and applied to pancreatic cancer [19]. Furthermore, the NEST approach improves on CRISPR data for analysis of gene or node essentiality [20]. However, to our knowledge, no available method leverages genome-scale functional screening resources to compute the importance of individual PPIs within biological networks. Here we describe MEDICI (Mining Essentiality Data to Identify Critical Interactions), a new computational approach that leverages high-throughput gene knockdown data and protein interaction network topology to rank PPIs based on their criticality for the survival and proliferation of

cancer cells. This approach does not predict new PPIs, but rather ranks the relative importance of known PPIs within a cancer cell for cell survival.

## Methods

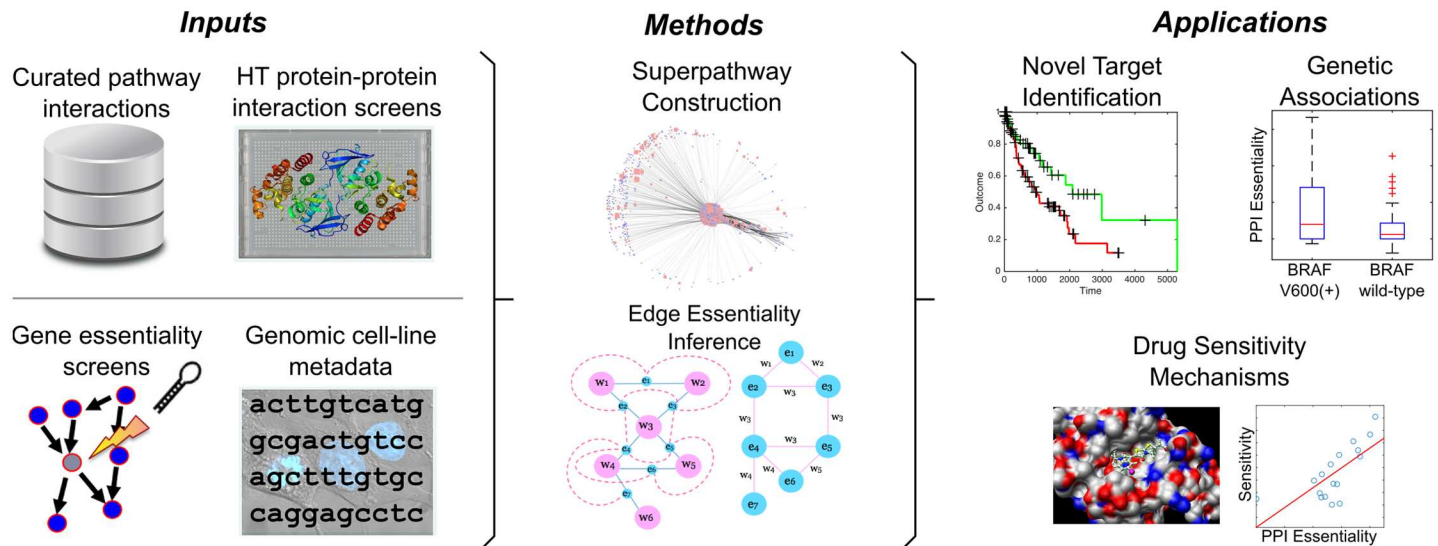
We obtained shRNA knockdown measurements from Project Achilles [7] v2.4.3 (<http://www.broadinstitute.org/achilles>). The .gct file containing multiple shRNAs for each gene (Achilles\_QC\_v2.4.3.rnai.Gs.gct) was analyzed to produce gene-level measurements. We used the value of the second most lethal shRNA for each gene to mitigate spurious off-target effects as previously described [7]. Moreover, the gene solutions for the shRNAs in this dataset have also been corrected for off-target effects using ATARiS software [21]. Cell line metadata was obtained from both Achilles and the Cancer Cell Line Encyclopedia [22] (CCLE, <http://www.broadinstitute.org/ccle>), including histology and organ site, mutations, copy number variation and sensitivities to compounds in the CCLE library.

The superpathway was constructed by combining curated pathway models from the NCI Pathway Interaction Database [23] (PID, <http://pid.nci.nih.gov/>) which can be accessed via MSigDB (<http://software.broadinstitute.org/gsea/msigdb>) [24] with protein-interaction screening results obtained via TR-FRET [25]. The gene symbols for the PID pathway were first mapped to the official HUGO symbols to resolve ambiguity and aid in integration with the Achilles dataset. Pathway entities that could not be mapped to HUGO were excised from the pathway model. All 196 PID pathways were aggregated to form a superpathway with 1548 nodes and 7906 interactions. We added 208 novel interactions identified in PPI screening [25] to this superpathway and annotated them for future reference. We used mutations and copy number alterations from CCLE datasets [22] to remove lost nodes and build 206 cell line context specific networks. Achilles shRNA data was then integrated with each network to produce a gene-essentiality layered superpathway for each cell line. Pathway nodes that did not have corresponding shRNA measurements were excised from the network model. The final set of 7906 PPI essentialities across 206 cell lines is available in [S1 Table](#). All datasets used in the analysis are available for download at the Cancer Target Discovery and Development (CTD2) Data Portal (<https://ctd2.nci.nih.gov/dataPortal/>). We have also conveniently packaged the data necessary for reproducing the analysis on [Synapse](#). We are currently developing a PPI Portal website where PPI essentiality data, as well as data from our OncoPPI screening study [25], can be mined and viewed (A. Ivanov, B. Revennaugh, et al, manuscript in preparation). This website will allow users to export data for analysis in excel or in Cytoscape, or to capture network images as png files.

## Modeling PPI network topologies

Biological networks consist of proteins (as nodes) and interactions (as edges) between them. For this research, we analyze the interactions between proteins with consideration of all other interactions and proteins in the network. To begin, we first consider a PPI between two proteins. These two proteins are typically also involved in interactions with other nodes. If a PPI is essential, then it's more likely that other interactions of those two proteins are essential. Furthermore, one would expect the interactions happening in proximity of that interaction are also essential. In other words, there is a locality of essentiality for PPIs.

To proceed we make the locality of PPIs more formal by inverting the pathway network and representing each of the PPIs as nodes, and each of the proteins as edges. Two PPIs are connected in the network if they share a protein, i.e., each edge in the new graph corresponds to a protein. We thus generate a new graph encoding the proximity for PPIs ([Fig 1](#)).



**Fig 1. Details of the computational framework of MEDICI.** Curated pathway descriptions are integrated with novel interactions discovered by PPI screening to generate an interaction superpathway. Gene essentiality measurements are layered onto the nodes of the superpathway, and the network topology is transformed to the dual graph where the genes become network edges and the gene-interactions become network nodes. Gene essentialities are then diffused over their interactions to infer interaction essentiality weights.

doi:10.1371/journal.pone.0170339.g001

Because we expect that PPIs that are connected with more essential proteins would have higher correlation in essentiality, we choose the edge weights in the dual network to be proportional to the essentiality of the corresponding protein derived from the shRNA screening data.

Once we have constructed our model with the dual graph, we infer the essentiality of PPIs in an unsupervised manner. Following the well-known PageRank algorithm [26], in order to promote the locality of essentiality, we employ an iterative update of essentialities. At each step, the essentiality of each node (or PPI) is updated to be the average of the localities of its neighbors. This is consistent with our intuition of the locality of the essentiality score. If an interaction is essential, it is likely that the interactions nearby are also vital. Furthermore, a less essential interaction is usually close to less important ones. In other words nearby interactions are correlated and the amount of this correlation is proportional to the essentiality of nodes in between and the structure of the network. More interestingly, these updates are structurally motivated, i.e., the scores derived reflect the structural importance of the PPIs in the whole interaction network.

### Calculate PPI essentialities and ranking

A dual network is constructed from the given network, i.e., the network is transformed so that each PPI/edge is represented by a node. Two nodes in the dual graph are connected if and only if their corresponding edges (in the original graph) have a shared end-point. Inspired by the locality of the essentiality, an iterative method is proposed in which the essentiality of a node in the dual graph is averaged over the essentiality of its direct neighbors. Specifically, we define the essentiality of node  $i$  (in the dual graph) at  $t$ th step by  $e_i^{(t)}$ . Also let  $e_i^{(0)}$  be a random initialization between 0 and 1. At step  $(t + 1)$  the score of each node is the weighted average of the score of its neighbors smoothed by its value in previous step with  $0 < \alpha \leq 1.0$ . Intuitively, the significance of each node propagates to its neighbors.

For, example in the network given by Fig 1 the essentiality for node 1 and 2 of the dual graph (edges  $e_1$  and  $e_2$  in the original network) are updated as follows:

$$e_1^{(t+1)} = (1 - \alpha) e_1^{(t)} + \alpha \left( \frac{p_1 e_2^{(t)} + p_2 e_3^{(t)}}{p_1 + p_2} \right)$$

$$e_2^{(t+1)} = (1 - \alpha) e_2^{(t)} + \alpha \left( \frac{p_1 e_1^{(t)} + p_3 e_3^{(t)} + p_3 e_4^{(t)}}{p_1 + p_2 + p_3} \right)$$

The weight of each edge is influenced by the essentiality of the protein. i.e, we have an exponential kernel  $\exp(w * p_i)$ , where  $p_i$  is the essentiality of protein  $i$  and  $w$  is the coefficient. By setting different  $w$  we can incorporate our prior belief on the importance of the essentiality of the protein in the essentiality of the edges.

$$e_1^{(t+1)} = (1 - \alpha) e_1^{(t)} + \alpha \left( \frac{\exp(w * p_1) e_2^{(t)} + \exp(w * p_2) e_3^{(t)}}{\exp(w * p_1) + \exp(w * p_2)} \right)$$

$$e_2^{(t+1)} = (1 - \alpha) e_2^{(t)} + \alpha \left( \frac{\exp(w * p_1) e_1^{(t)} + \exp(w * p_3) e_3^{(t)} + \exp(w * p_3) e_4^{(t)}}{\exp(w * p_1) + \exp(w * p_2) + \exp(w * p_3)} \right)$$

After sufficiently large number of steps the scores converge and the final score is independent of the initialization. It can be easily seen that the stationary values (when the essentiality computation converges) can be simply computed by Eigenvalue decomposition of a variant of the adjacency matrix of the dual graph.

## Smoothing

The above updates are subject to overfitting and are sensitive to noise. Consequently, at each step the new essentiality value may change dramatically because of the changes in the neighboring PPIs. To control this undesirable behavior, we introduce a smoothing step to the model in order to regularize the solution. At each step, the new value is a weighted sum of the previous value and the value inspired by the neighbors. The smoothing parameter,  $0 < \alpha \leq 1$ , controls this tradeoff between the locality and the regularization. The new value is  $(1 - \alpha)$  times the previous value plus  $\alpha$  times the values computed from the neighbors.

We analyzed  $\alpha = 0.9, 0.5, 0.1$  and we observed that the smoothing parameter doesn't have a significant effect on the ranking of essentialities. It only changes the values but won't affect the ordering. A lower  $\alpha$  will damp out the essentialities and make them close to each other. For example  $\alpha = 0$  will result in a uniform essentialities (all equal). The low values of  $\alpha$  make the update rule reluctant to change the essentialities in iterations and the values change smoothly hence, the final values will be close to each other. On the other hand, as we increase  $\alpha$  the effect of neighbors will increase, the changes have higher magnitudes, and the difference is more apparent. The smoothing will make the computed values numerically more stable. The eigenvector computation is pruned to degenerate results when the smallest eigenvalue is close to 0. The smoothing will increase the lowest eigenvalue and therefore the convergence computation will not be affected by degenerate cases. Also, the smoothing can be seen as a denoising approach. If the data is noisy the smoothing will remove the fine-tuning due to the noise. Therefore, variations due to the noise will not affect the model.

We determined that a value of  $\alpha = 0.5$  was optimal for the analyses in this study. A second parameter  $w$  represents how much the essentiality of proteins affects the essentiality of

interactions in between. In other words, larger values of  $w$  show that an interaction is essential if the corresponding proteins are essential, and smaller values of  $w$  show that no matter how essential the corresponding proteins are the interaction is important, e.g., because of its centrality role in the PPI network. Examining different values we set  $w$  to 0.5. See [S1 Methods](#), where we demonstrate that this iterative update is guaranteed to converge and has a closed form answer. Furthermore the final value is independent of the initialization. The final essentiality score ranges from 0 (not essential) to 1 (completely essential to cell survival). All MEDICI software is available for public download at <https://github.com/cooperlab/MEDICI>.

## Data mining and statistical tests

We applied MEDICI to shRNA gene-essentiality profiles from 206 cell lines in the Achilles database [8]. A context specific PPI network was created for each cell line by using mutational and CNV events to re-wire the topology of a prior-knowledge superpathway [23] containing 2,186 proteins and 11,488 PPIs. This superpathway was a subset of the NCI Pathway Interaction Database [23] that is now hosted by the public network data exchange [27] (<http://www.ndexbio.org/>). Gene essentiality measurements were integrated with the re-wired networks, and then deconvolved by constructing a dual-graph to predict unknown PPI essentialities from the measured gene essentialities and the PPI network topology (Fig 1). To evaluate these PPIs in patient samples, we used mutational, CNV and mRNA expression profiles to infer which PPIs are absent in each patient. A Cox proportional hazards analysis was then used to measure the prognostic significance of these groups. PPIs from lung adenocarcinoma derived lines were ranked to identify the most essential PPIs, KS statistics were computed for ranks across these lines, and a positive KS statistic  $> 0.5$  was used as a threshold for significant PPIs in LUAD cell lines.

## Results

### PPI essentiality rankings

We used the context-specific networks to compute the essentialities of 7906 PPIs in 206 cell lines from the project Achilles dataset and have reported these in [S1 Table](#). Context specific networks were based on a superpathway that was constructed by combining 196 curated pathway models from the NCI Pathway Interaction Database with 208 novel interactions identified in PPI screening [25] and mutations and copy number alterations from CCLE datasets [22]. Examination of PPIs with the highest average essentialities determined that the top 20 were dominated by interactions of CTNNB1 and AKT1. Ten of the top 20 PPIs involved CTNNB1, while 8 of 20 involved AKT1. The other two top 20 PPIs were SOS1-SRC and MYC-SMAD4. Highly essential interactions of CTNNB1 in descending order were those with SRC, RAC1, JUN, PTPN11, HDAC1, EGFR, RHOA, CTNNA1, CDC42, and AR. The top interactions of AKT1 were with SRC, RAC1, SOS1, PIK3CA, PTPN11, PIK3R1, JAK1, and MTOR. SRC was a component of 3 of the top 20 PPIs, and the top two PPIs across all cell lines were CTNNB1-SRC and SRC-AKT1, suggesting that the SRC-AKT1-CTNNB1 pathway is of potentially general interest for targeting PPIs. Previous studies have shown synergistic effects using combinations of SRC and AKT inhibitors [28] in small cell lung cancers.

### Clustering analysis

Data were filtered to remove cell lines with large numbers of missing nodes ( $> 20\%$  missing due to gene mutation or deletion) and PPIs with uniform essentiality across cell lines, resulting in a final matrix of 165 cell lines by 5798 PPIs for PPI essentiality clustering analysis. Unsupervised

hierarchical clustering revealed strong clusters of some subsets of PPIs (S1 Fig). In general, the PPIs within these clusters shared a single dominant oncogene (e.g. AKT1 or CTNNB1) or tumor suppressor (e.g. TP53) as one of the interacting partners.

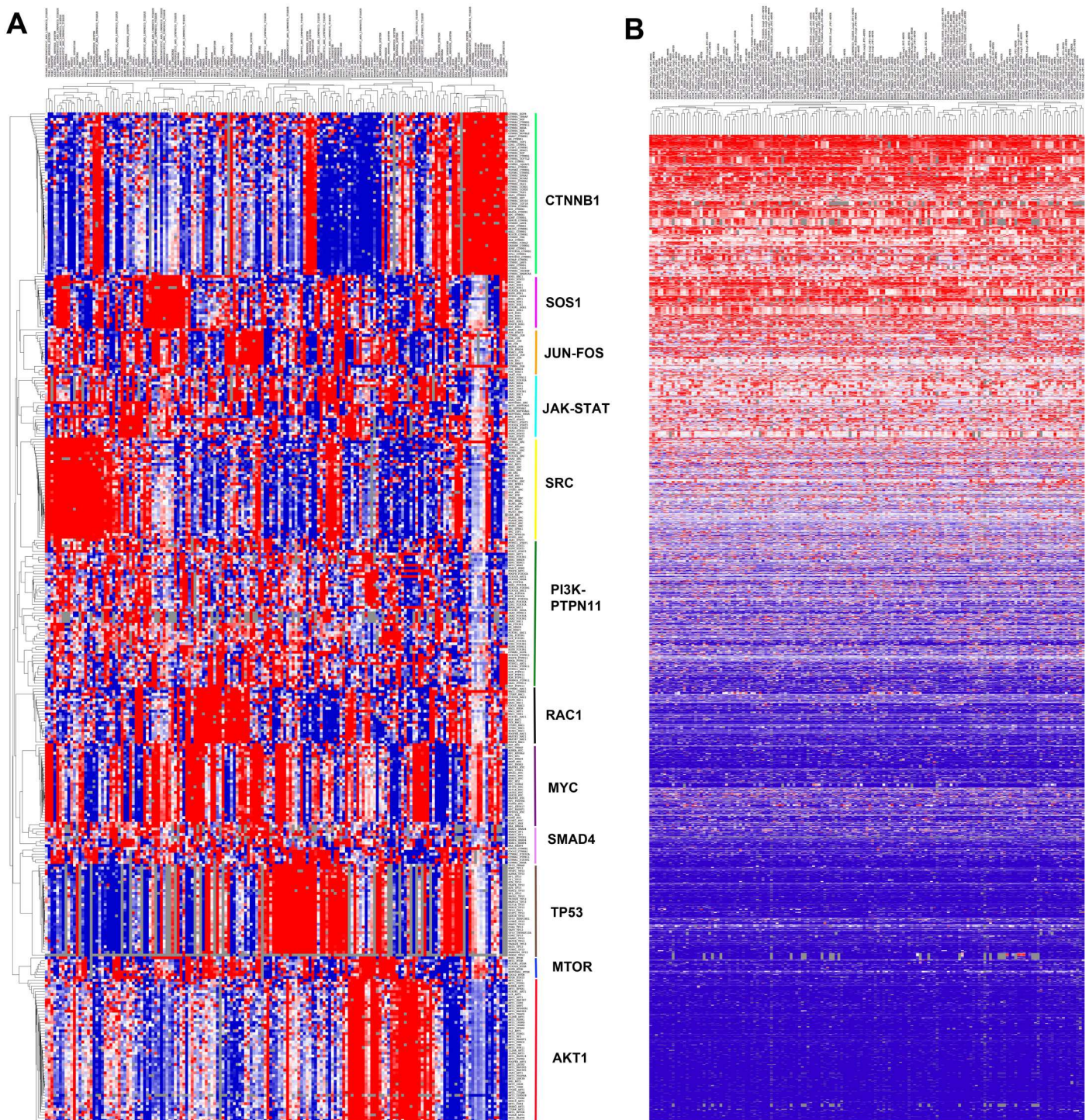
To delve further into those PPIs with therapeutic potential, we analyzed the subset of the 360 most essential PPIs with an average essentiality  $> 0.5$  across all cell lines. An essentiality score of 0 is a non-essential PPI, whereas a score of 1 is a completely essential PPI for cell survival. Clustering of these 360 PPIs across the 165 cell lines identified 12 major PPI clusters. Each of these clusters was strongly associated with either a single protein or a single pathway (Fig 2A). Of particular interest, the PPIs among the top 20 most essential often those linked proteins that dominated particular PPI clusters such as AKT1-SRC, AKT1-MTOR, and CTNNB1-SRC. Interestingly, the cell lines did not cluster primarily by tissue of origin, with the exception of colon-derived cell lines that were driven by the large CTNNB1 cluster. Instead, cell line clustering was likely driven by shared activation of oncogenic pathways.

In contrast, when we clustered the 165 cell lines by protein nodes using the gene-centric Achilles shRNA sensitivity data across 5711 genes, the cell lines clustered primarily by tissue of origin (S2 Fig). A few strong clusters were apparent in the unsupervised clustering of Achilles shRNA data, which included sets of transcription factors, ribosomal proteins, or genes of unknown function. To compare the supervised clustering of edges from Fig 2 using shRNA node data, we extracted the 343 proteins with the top average shRNA sensitivity across all cell lines ( $< -0.05$ ), and performed identical hierarchical clustering (S3 Fig). Cell lines continued to be clustered by tissue of origin, and there were few distinct clusters, with the exception of one small cluster containing MHC class I receptors.

To compare protein and PPI essentiality, we ranked each interaction by its PPI essentiality, and then compared this to the maximum protein essentiality of the interaction source and target. Interactions with a high PPI essentiality but a low source/target essentiality are especially interesting, since these represent vulnerabilities that will not be revealed in single-gene knock-downs. We further computed the ratios of the PPI essentialities to the maximum essentiality of either interacting protein and performed hierarchical clustering of the PPI/maxProteinEssentiality ratio (PPI-MPER) data (Fig 2B). We observed that 2294 (29%) PPIs had a PPI-MPER that was  $> 1$  in over half of the 165 cell lines (S3 Table). Of the 7814 PPIs examined, 6503 (83%) had a PPI-MPER  $> 1$  in at least one cell line. Importantly, 2562 (33%) PPIs had a median PPI-MPER  $< 0.5$ , indicating that the max Protein essentiality was 2-fold or more higher than the PPI essentiality in at least half of the cell lines. These data suggest that the MEDICI software is differentiating crucial PPIs from unimportant PPIs, even when they occur between essential proteins. Among the PPIs with consistently high PPI-MPER values were CDK2\_RB1 [30], PIK3R1\_YWHAZ [31], JAK3\_MAPK14, CREBBP\_CTNNB1 [32], and HIF1A\_SMAD4 [33]. Some of these interactions are well documented as key interactions in critical signal transduction pathways, while others less so, suggesting that they may warrant further investigation.

## Crosstalk of signal transduction pathways

Other PPIs such as YAP1\_CTNNB1 [34], AR\_CTNNB1 [35, 36], and SMAD7\_CTNNB1 [37] were much more essential than either protein (S4 Fig), but in only a subset of cell lines. For example, the AR\_CTNNB1 PPI essentiality rank percentile was very high in a number of cell lines including COLO704 ovarian (99%ile), and BT20 (98%ile) and MDA-MB-453 (97%ile) triple-negative breast cancer cell lines. Both Wnt [38] and androgen signaling [39] have been shown to be important in subsets of ovarian and triple-negative breast cancers [40–42]. Another example of signaling crosstalk identified by this analysis is that of YAP1\_CTNNB1 interactions in L363 myeloma cells (94%ile), CADO-ES1 Ewing's sarcoma cells (96%ile), and



**Fig 2. Clustering of most essential PPIs in the superpathway.** (A) Unsupervised Hierarchical Clustering of the 360 most essential PPIs across the 165 cell lines identifies 12 major clusters. The 360 PPIs with an average essentiality score  $> 0.5$  were used to cluster 165 cell lines used in the Achilles shRNA screening study using Cluster and Java Treeview software [29]. PPI essentiality data was median centered and clustered by average correlation. Red indicates higher essentiality and blue indicates lower essentiality. Major hubs for each cluster are indicated on the right. (B) Clustering of 5798 PPI-MPER values across 165 cell lines. Red indicates PPI essentiality is greater than the max protein essentiality, and blue indicates the PPI essentiality is less than the max protein essentiality.

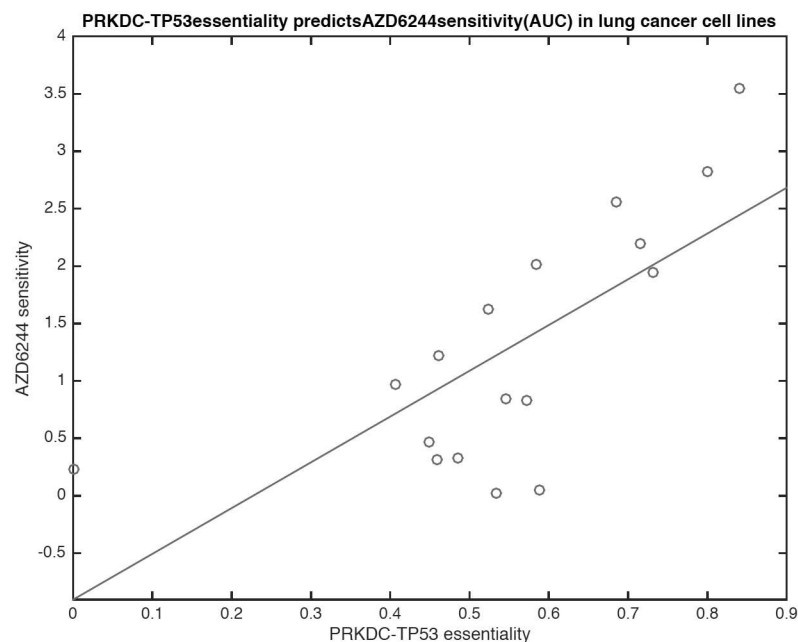
doi:10.1371/journal.pone.0170339.g002



A204 rhabdomyosarcoma cells (96%ile). Wnt and Hippo signaling are important in myelomas, bone, and soft tissue cancers [43–47]. We also observed potential crosstalk of TGF $\beta$  and Wnt signals in colorectal RKO cells (95%ile), KP4 pancreatic cells (97%ile), and 22RV1 prostate cancer cells (99%ile). SMAD7 has been implicated in colon cancer development [48, 49], and in interactions with  $\beta$ -catenin in pancreatic cancer [50] and prostate cancer [37].

## Drug sensitivity

The lack of even low-throughput technology to produce PPI essentiality measurements makes validation of targets difficult. This fact motivated us to evaluate our PPI essentiality predictions through joint analysis of PPI essentialities and drug sensitivity data. We leveraged drug sensitivity data on over 200 cell lines from the Cancer Cell Line Encyclopedia (CCLE) [22] and examined the correlations of PPI essentialities and sensitivity to 19 different compounds. We thus performed a context specific PPI analysis, removing any PPI from a cell-line specific network if either the source or target protein was deleted, had an inactivating mutation, or was not expressed at the RNA level. Because of our context-specific approach that uses genetic alterations and gene-regulation to alter PPI network topology for each cell line, the specific number of PPIs varied from cell line to cell line, but included on average 104 cell lines per PPI. We then computed the statistical significance of correlations of PPI essentiality to drug sensitivity (as defined by AUC in CCLE analysis). Interestingly, the most significant correlation ( $p = 4.63E-07$ ) was that of the BRAF-KRAS interaction to PLX4720, a compound that targets BRAF. The highest absolute correlation ( $r = 0.71$ ,  $p = 0.0013$ ) was between the PRKDC-TP53 PPI and sensitivity to AZD6244, an inhibitor of MAPK1 (Fig 3). This strong correlation is supported by multiple studies that indicate that MAPK signaling is critical for cell cycle arrest in response to DNA damage [51, 52] and that inhibition of p38-MAPK sensitizes lung cancer cells to DNA damaging agents [53].



**Fig 3. Correlating interaction essentialities with drug sensitivity measures provides insights into mechanisms of action.** We correlated drug sensitivity measures from CCLE with interaction essentiality scores to identify critical interactions that predict therapeutic sensitivity. Sensitivity to the MAPK inhibitor AZD6244 is highly correlated with PRKDC-TP53 interaction essentiality, which is consistent with the well established role of p38-MAPK in cell cycle arrest in response to DNA damage [51–53].

doi:10.1371/journal.pone.0170339.g003

**Table 1. Correlations of PPI essentiality with drug sensitivities for 19 CCLE compounds with PPI essentiality data for their respective targets.**

| Compound   | Targets                     | WilcoxonRankSum p-val | FDR    |
|------------|-----------------------------|-----------------------|--------|
| PF2341066  | MET;ALK                     | 9.86E-18              | <0.001 |
| Lapatinib  | EGFR;ERBB2                  | 2.16E-17              | <0.001 |
| Erlotinib  | EGFR;ERBB2                  | 3.60E-12              | <0.001 |
| ZD-6474    | VEGFR;EGFR                  | 2.19E-11              | <0.001 |
| L-685458   | APH1A;NCSTN;PSEN1;PSENEN    | 2.90E-11              | <0.001 |
| Sorafenib  | BRAF;FLT3;KDR;RAF1          | 3.47E-11              | <0.001 |
| PLX4720    | BRAF                        | 7.05E-07              | <0.001 |
| PD-0332991 | CDK4;CDK6                   | 1.01E-05              | <0.001 |
| AZD6244    | MEK                         | 7.65E-04              | 0.002  |
| PD-0325901 | MEK                         | 4.54E-04              | 0.003  |
| Nutlin-3   | MDM2                        | 9.59E-03              | 0.006  |
| RAF265     | BRAF;KDR                    | 8.53E-03              | 0.007  |
| LBW242     | XIAP                        | 1.46E-02              | 0.009  |
| 17-AAG     | HSP90                       | 3.20E-02              | 0.030  |
| AEW541     | IGF1R                       | n.s.                  | 0.161  |
| AZD0530    | SRC;YES;FYN;LYN;BLK;FGR;LCK | n.s.                  | 0.234  |
| Nilotinib  | ABL1;BCR;KIT                | n.s.                  | 0.286  |
| TKI258     | FLT3;FGFR1/3;VEGFR1-4       | n.s.                  | 0.540  |
| PHA-665752 | MET                         | n.s.                  | 0.718  |

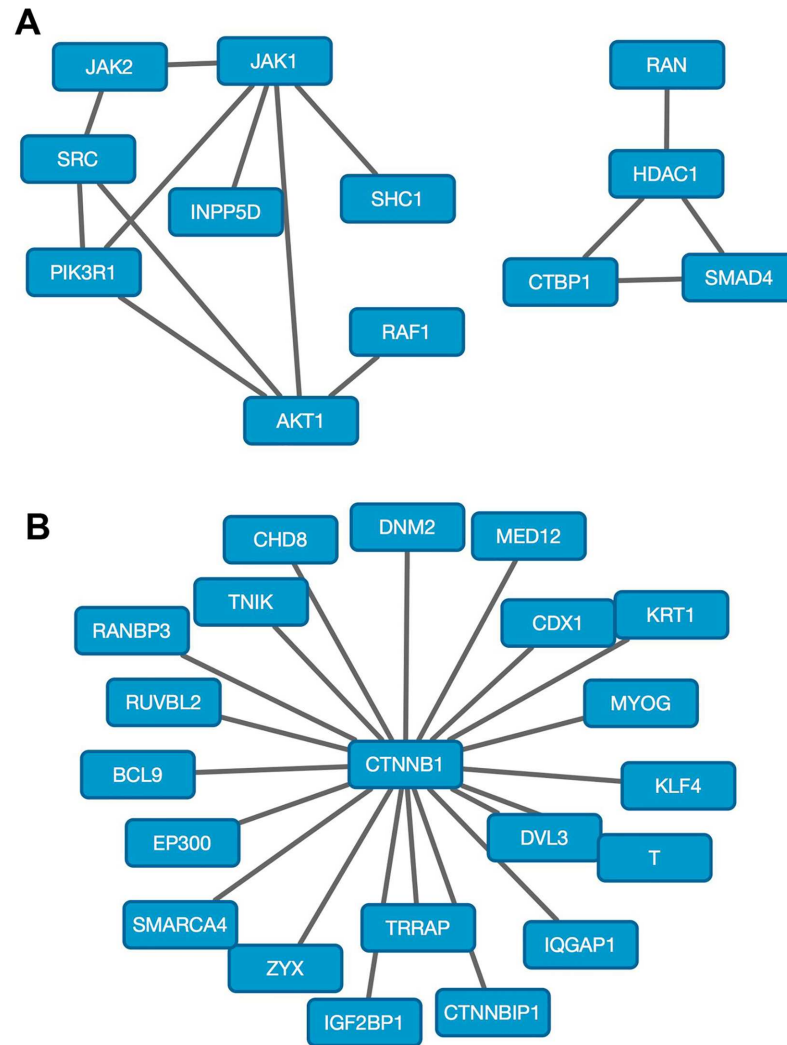
Drug sensitivities were derived from CCLE AUC data. Significance of enrichment for drug target PPI essentialities vs. non-target PPI essentialities was computed by the Wilcoxon Rank Sum test, and was significant for 14 of 19 compounds. Target and non-target gene sets were randomly permuted 1000 times to compute FDR and correct for multiple hypothesis testing.

doi:10.1371/journal.pone.0170339.t001

Comprehensive analysis of drug sensitivity correlations was performed by comparing the strength of correlations for two sets of drug/PPI pairs: 1) drug/PPIs pairs where the PPI involves a protein targeted by that drug and 2) PPIs involving no known target protein. Examining 19 CCLE compounds for which PPI essentiality data was available, we observed that target PPIs were significantly more essential than non-target PPIs by Wilcoxon Rank Sum analysis for 14 of the 19 compounds (Table 1). Random permutation analysis was performed for each set of target vs non-target protein sets to compute false discovery rates and correct for multiple hypothesis testing. Correlations of PPI essentialities for each of the 14 significant drug/PPI pairs were significant by FDR < 0.05 (Table 1).

### Genetic alterations

To further verify MEDICI essentiality predictions, we investigated the associations between known genetic alterations and PPI essentialities. We divided the PPI essentiality data from the 165 cell lines analyzed in Fig 2 based on the presence or absence of wild type PTEN or APC tumor suppressor genes. We hypothesized that cell lines with mutant PTEN would have increased dependence on PPIs in the PI3K-AKT pathway, while those with mutant APC would exhibit increased dependence on  $\beta$ -catenin (CTNNB1). Indeed, when we analyzed PPI essentiality data for cell lines with PTEN mutations (n = 25) relative to those with wild-type PTEN (n = 140), we observed that the most essential PPIs were related to PI3K-AKT signaling (Fig 4A). Top PPIs included JAK1-AKT, JAK1-PIK3R1, SRC-AKT1, and PIK3R1-AKT1 (FDR (BH) = 2.13e-4). Examining cell lines with mutant APC (n = 14) compared to those with



**Fig 4. PPI networks associated with genetic mutations.** (A) Networks of PPIs most increased in essentiality in cells with mutation or loss of the PTEN tumor suppressor gene. The 14 most significant PPIs are shown. Significant differences in PPI essentiality were computed in GenePattern [54]. Networks were visualized with Cytoscape [55]. (B) Networks of PPIs most increased in essentiality in cell lines with mutation or loss of the APC tumor suppressor gene. The 20 most significant PPIs are shown.

doi:10.1371/journal.pone.0170339.g004

wild-type APC (n = 151), the 95 most significant PPIs all involved interactions with  $\beta$ -catenin (FDR(BH) = 2.13e-4). The top 20 PPIs from this analysis are shown in Fig 4B.

### Clinical outcome analysis

We next wanted to address whether loss of a particular PPI might confer a survival advantage to lung cancer patients. To test whether this was the case, we analyzed genomic data from the lung adenocarcinoma (LUAD) [56] datasets from The Cancer Genome Atlas (TCGA) Network. We separated patients into those with or without each of the PPIs in our superpathway network based on the presence or absence of nonsense mutations, frameshift mutations, or homozygous deletions for each of the partner proteins. We then computed log rank p-values for survival based on whether each of the PPIs was present or absent. QQ plots of observed vs

expected log-rank p-values for LUAD patients (Fig 5A) indicate a strong trend of more significant p-values than those expected by chance. However, log rank p-values corrected for multiple tests using Benjamini-Yekutieli correction [57] did not achieve statistical significance, and should therefore be interpreted with caution. We then examined the set of PPIs with nominal log rank p-value  $< 0.05$  that were also essential in lung adenocarcinoma cell lines based on our analysis of Achilles data. For LUAD patients, 15 PPIs were met this criterion (S2 Table, Fig 5B). Of those 15 PPIs, six included JAK1 as a binding partner, two included JAK2, and one was the JAK1-JAK2 interaction. It has recently been shown that inhibition of JAKs attenuates growth of small cell lung cancers *in vitro* and *in vivo* [58], and that activation of JAK signaling induces resistance to EGFR mutations in non-small cell lung cancers [59]. Also of potential interest is an ATM-NBN interaction that is involved in double-strand DNA break repair and PI3KR1 interactions involved in activation of the PI3K pathway.

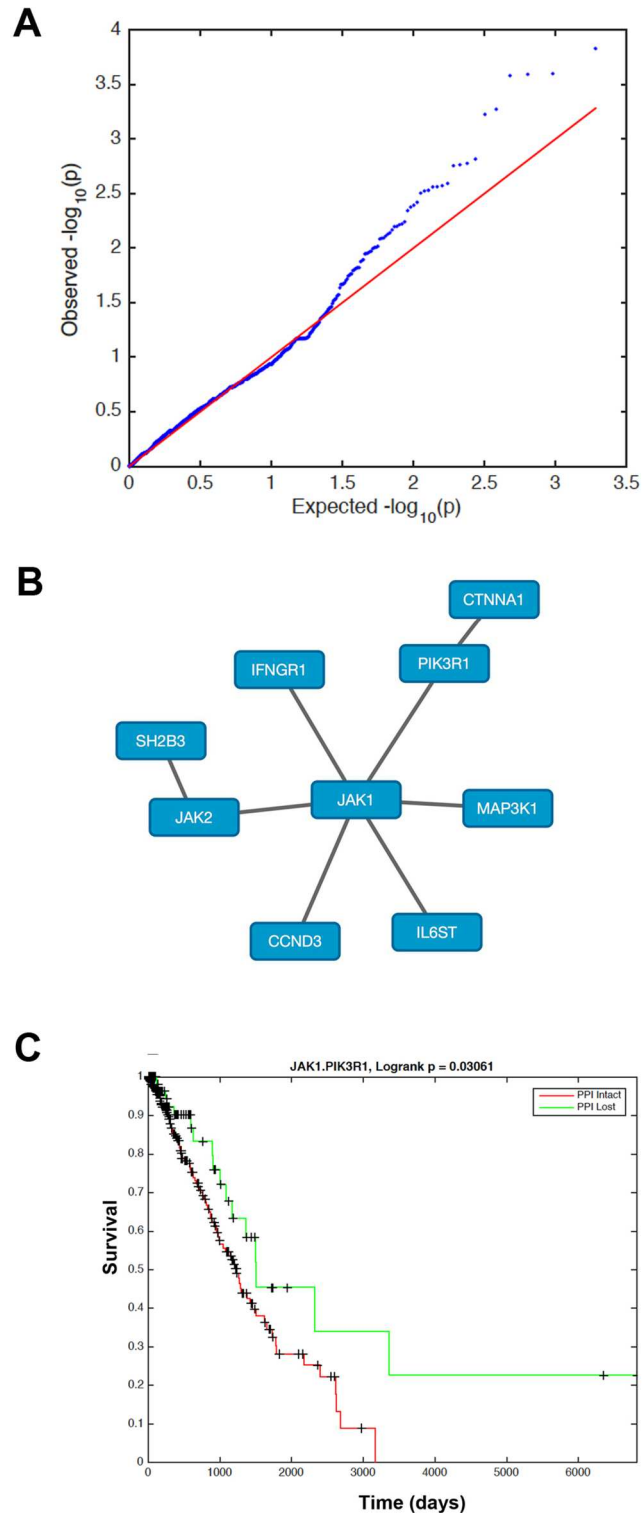
## Discussion

While most cancer drugs either cause DNA damage or target the active sites of enzymes such as protein kinases, a growing number of new compounds interfere with critical protein-protein interactions (PPIs). Although the pharmaceutical industry has been reluctant to develop compounds that target PPIs, it is becoming clear that these types of drugs can provide lethal and specific targeting of cancer cells. Examples of these compounds include Navitoclax (ABT-263) [60], which inhibits BCL2-BAX interactions; Nutlin-3 [61], which blocks MDM2-TP53 interactions; PRI-724 [62], which interferes with  $\beta$ -catenin-CBP interactions; and JQ-1 [13] and I-BET726 [63], which prevent BRD4 binding to acetylated histones. These new compounds are showing promise for treatment of previously difficult to target pathways, including the p53 tumor suppressor [12], Wnt pathway [62], and Myc oncogene [64].

Nevertheless, while there may be on average approximately 10,000–15,000 proteins expressed in any given cell type, there are likely an order of magnitude more PPIs, or over 600,000 PPIs in mammalian cells [65]. Thus, prioritizing which PPIs may be important for drug development is an even greater challenge than identifying individual proteins as potential drug targets. The network topology of cancer drug targets has been shown to be more hub-like than non-drug targets, suggesting that PPIs can be useful in identification of drug target proteins for cancer [66] and neurological diseases [67]. Thus, combinations of computational and experimental methods will be critical to enable identification of the most promising PPIs that are essential to the survival of cancer cells.

In this study we have described a novel approach to computationally infer the essentiality of PPIs based on a combination of network topology and high-throughput screening gene knock-out studies. This study represents our first attempt to provide a framework to predict and prioritize PPIs as potential targets for therapeutic discovery, which is expected to mature with the incorporation of additional datasets and biological insights. While this approach is quite general, here we have applied it to only a single large dataset from Project Achilles [8], which by definition, when integrated with our superpathway, limited our analyses to include only 7906 PPIs. Future studies integrating additional high throughput datasets from siRNA, shRNA, and/or CRISPR/Cas9 screens would integrate multiple complementary datasets that may provide more robust PPI essentiality predictions. Another limitation of our current study is that shRNA screens can have off-target effects. For this reason, we limited our analysis to only those Achilles shRNA targets that were validated by multiple shRNAs that targeted the same genes.

Our analysis is further limited by our understanding and knowledge of the complete PPI network within cancer cells. A variety of approaches including mass spectrometry, yeast two-



**Fig 5. PPI Essentiality association with patient survival.** (A) QQ plot of observed vs. expected log-rank p-values for LUAD patients split based on 5798 PPIs. (B) Network of PPIs with significant log-rank p-value for discriminating survival for TCGA LUAD patients that is centered on JAK1. (C) Kaplan-Meier curve of TCGA LUAD patients separated based on the presence or absence of the JAK1-PIK3R1 PPI. Patients without the JAK1-PIK3R1 PPI have improved survival compared to patients who retain this PPI.

doi:10.1371/journal.pone.0170339.g005

hybrid, and TR-FRET methodologies are continually expanding and improving our understanding of PPI networks. As our grasp on the topology of these PPI networks improves, the potential accuracy, breadth, and utility of inferences regarding PPI essentiality using our methods described here will naturally improve as a result.

Here, we have developed cell-line specific network topologies by removing PPIs in which one partner gene is deleted, mutated, or not expressed. However, other than this, we do not account for differences in gene expression or abundance of various PPI partner proteins. Additional ways to potentially refine our approach would involve analysis of the transcriptional profile of cell lines to deduce which subset of PPIs that are possible within a cell are actually taking place. Currently, our approach assumes that two proteins will interact within a cell line if they are known to interact, are expressed and are not mutated. However, signal transduction pathways can greatly modify the constellation of potential PPIs within a cell, and thus layering pathway activation data onto our current network topologies could improve our approach.

Even so, the analyses presented here indicate that the sets of PPI essentialities that we have computed have strong biological relevance. The correlations of essentialities to drug sensitivities and genetic mutations validate our overall approach and provide new insights into potentially important PPIs that could be used for guiding cancer drug target development.

## Supporting Information

**S1 Fig. Unsupervised hierarchical clustering of all non-zero 5798 PPIs across 165 cell lines.** PPI essentiality data was median centered and clustered by average correlation. Red indicates higher essentiality and blue indicates lower essentiality.  
(TIF)

**S2 Fig. Unsupervised hierarchical clustering of 5711 genes across 165 cell lines based on Achilles direct shRNA sensitivity data.** Cell lines cluster primarily by tissue of origin. The main strong gene clusters were driven by sets of transcription factors, ribosomal proteins, or genes of unknown function.  
(TIF)

**S3 Fig. Unsupervised hierarchical clustering of the 343 proteins with the top average shRNA sensitivity across all cell lines ( $< -0.05$ ).** Cell lines cluster primarily by tissue of origin. Few distinct gene-based clusters are present, with the exception of one small cluster containing MHC class I receptors.  
(TIF)

**S4 Fig. Comparison of protein and protein-interaction essentialities.** To evaluate the utility of MEDICI, we compared estimated PPI essentiality values to the experimentally measured essentialities for individual proteins. Each point in the scatter plots of panels (A-D) represents the essentiality rank of a single PPI versus and the max protein essentiality rank of the constituent proteins in that interaction. Color indicates density of points. Plots for four cell lines (A) A549 lung cancer, (B) BXP3 pancreatic cancer, (C) MKN7 stomach cancer, and (D) U87MG glioma cell lines are shown. The EGFR-ERBB3 PPI is indicated in each plot with an arrow. A large number of PPIs appear in the upper-left of each panel in which PPI essentiality significantly exceeds max constituent protein essentiality. These entries provide insights into interaction-specific sensitivity that cannot be readily observed in the measure protein essentiality data. Ranks/percentiles were used to generate these plots to avoid any consistent biases in the magnitude of PPI versus protein essentialities.  
(TIF)

**S1 Table. Essentiality values for 7906 PPIs across 206 cancer cell lines.** NaN indicates that one of the proteins in the interaction is deleted, mutated, or not expressed in that cell line. (XLS)

**S2 Table. PPIs with nominally significant log-rank p-values for LUAD patient survival PPI and KS.** (XLSX)

**S3 Table. PPI-MPER are provided for 7906 PPIs and 206 cell lines.** Also shown are PPI rank percentiles and maximum protein essentiality rank percentiles. (TXT)

**S1 Methods. Detailed supplementary methods and convergence proof.** (DOCX)

## Acknowledgments

This work was supported by the National Institutes of Health, [grant number U01 CA168449].

## Author Contributions

**Conceptualization:** LADC SH CSM.

**Data curation:** LADC SH JDM FOG CSM.

**Formal analysis:** LADC SH JDM FOG CSM.

**Funding acquisition:** FRK HF CSM.

**Investigation:** JDM YD FOG AAI MAJ FRK HF SH LADC CSM.

**Methodology:** LADC SH CSM.

**Project administration:** LADC MAJ FRK HF CSM.

**Resources:** YD AAI MAJ FRK HF LADC CSM.

**Software:** LADC SH.

**Supervision:** LADC FRK HF CSM.

**Validation:** LADC SH CSM.

**Visualization:** LADC SH CSM.

**Writing – original draft:** LADC SH CSM.

**Writing – review & editing:** JDM YD FOG AAI MAJ FRK HF SH LADC CSM.

## References

1. Berns K, Hijmans EM, Mullenders J, Brummelkamp TR, Velds A, Heimerikx M, et al. A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature*. 2004; 428(6981):431–7. doi: [10.1038/nature02371](https://doi.org/10.1038/nature02371) PMID: [15042092](https://pubmed.ncbi.nlm.nih.gov/15042092/)
2. Brummelkamp TR, Fabius AW, Mullenders J, Madiredjo M, Velds A, Kerkhoven RM, et al. An shRNA barcode screen provides insight into cancer cell vulnerability to MDM2 inhibitors. *Nat Chem Biol*. 2006; 2(4):202–6. doi: [10.1038/nchembio774](https://doi.org/10.1038/nchembio774) PMID: [16474381](https://pubmed.ncbi.nlm.nih.gov/16474381/)
3. Whitehurst AW, Bodemann BO, Cardenas J, Ferguson D, Girard L, Peyton M, et al. Synthetic lethal screen identification of chemosensitizer loci in cancer cells. *Nature*. 2007; 446(7137):815–9. doi: [10.1038/nature05697](https://doi.org/10.1038/nature05697) PMID: [17429401](https://pubmed.ncbi.nlm.nih.gov/17429401/)

4. Colombi M, Molle KD, Benjamin D, Rattenbacher-Kiser K, Schaefer C, Betz C, et al. Genome-wide shRNA screen reveals increased mitochondrial dependence upon mTORC2 addiction. *Oncogene*. 2011; 30(13):1551–65. doi: [10.1038/onc.2010.539](https://doi.org/10.1038/onc.2010.539) PMID: [21170086](https://pubmed.ncbi.nlm.nih.gov/21170086/)
5. Chen S, Sanjana NE, Zheng K, Shalem O, Lee K, Shi X, et al. Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell*. 2015; 160(6):1246–60. doi: [10.1016/j.cell.2015.02.038](https://doi.org/10.1016/j.cell.2015.02.038) PMID: [25748654](https://pubmed.ncbi.nlm.nih.gov/25748654/)
6. Kim HS, Mendiratta S, Kim J, Pecot CV, Larsen JE, Zubovych I, et al. Systematic identification of molecular subtype-selective vulnerabilities in non-small-cell lung cancer. *Cell*. 2013; 155(3):552–66. doi: [10.1016/j.cell.2013.09.041](https://doi.org/10.1016/j.cell.2013.09.041) PMID: [24243015](https://pubmed.ncbi.nlm.nih.gov/24243015/)
7. Cheung HW, Cowley GS, Weir BA, Boehm JS, Rusin S, Scott JA, et al. Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineage-specific dependencies in ovarian cancer. *Proc Natl Acad Sci U S A*. 2011; 108(30):12372–7. doi: [10.1073/pnas.1109363108](https://doi.org/10.1073/pnas.1109363108) PMID: [21746896](https://pubmed.ncbi.nlm.nih.gov/21746896/)
8. Cowley GS, Weir BA, Vazquez F, Tamayo P, Scott JA, Rusin S, et al. Parallel genome-scale loss of function screens in 216 cancer cell lines for the identification of context-specific genetic dependencies. *Scientific Data*. 2014; 1:140035. <http://www.nature.com/articles/sdata201435> - supplementary-information. doi: [10.1038/sdata.2014.35](https://doi.org/10.1038/sdata.2014.35) PMID: [25984343](https://pubmed.ncbi.nlm.nih.gov/25984343/)
9. Helming KC, Wang X, Roberts CW. Vulnerabilities of mutant SWI/SNF complexes in cancer. *Cancer Cell*. 2014; 26(3):309–17. doi: [10.1016/j.ccr.2014.07.018](https://doi.org/10.1016/j.ccr.2014.07.018) PMID: [25203320](https://pubmed.ncbi.nlm.nih.gov/25203320/)
10. Helming KC, Wang X, Wilson BG, Vazquez F, Haswell JR, Manchester HE, et al. ARID1B is a specific vulnerability in ARID1A-mutant cancers. *Nat Med*. 2014; 20(3):251–4. doi: [10.1038/nm.3480](https://doi.org/10.1038/nm.3480) PMID: [24562383](https://pubmed.ncbi.nlm.nih.gov/24562383/)
11. Zhang B, Wang J, Wang X, Zhu J, Liu Q, Shi Z, et al. Proteogenomic characterization of human colon and rectal cancer. *Nature*. 2014; 513(7518):382–7. doi: [10.1038/nature13438](https://doi.org/10.1038/nature13438) PMID: [25043054](https://pubmed.ncbi.nlm.nih.gov/25043054/)
12. Stuhmer T, Chatterjee M, Hildebrandt M, Herrmann P, Gollasch H, Gerecke C, et al. Nongenotoxic activation of the p53 pathway as a therapeutic strategy for multiple myeloma. *Blood*. 2005; 106(10):3609–17. doi: [10.1182/blood-2005-04-1489](https://doi.org/10.1182/blood-2005-04-1489) PMID: [16081689](https://pubmed.ncbi.nlm.nih.gov/16081689/)
13. Filippakopoulos P, Qi J, Picaud S, Shen Y, Smith WB, Fedorov O, et al. Selective inhibition of BET bromodomains. *Nature*. 2010; 468(7327):1067–73. doi: [10.1038/nature09504](https://doi.org/10.1038/nature09504) PMID: [20871596](https://pubmed.ncbi.nlm.nih.gov/20871596/)
14. Ivanov AA, Khuri FR, Fu H. Targeting protein-protein interactions as an anticancer strategy. *Trends Pharmacol Sci*. 2013; 34(7):393–400. doi: [10.1016/j.tips.2013.04.007](https://doi.org/10.1016/j.tips.2013.04.007) PMID: [23725674](https://pubmed.ncbi.nlm.nih.gov/23725674/)
15. Sahni N, Yi S, Taipale M, Fuxman Bass JI, Coulombe-Huntington J, Yang F, et al. Widespread macromolecular interaction perturbations in human genetic disorders. *Cell*. 2015; 161(3):647–60. doi: [10.1016/j.cell.2015.04.013](https://doi.org/10.1016/j.cell.2015.04.013) PMID: [25910212](https://pubmed.ncbi.nlm.nih.gov/25910212/)
16. Zhong Q, Simonis N, Li QR, Charlotheaux B, Heuze F, Klitgord N, et al. Edgetic perturbation models of human inherited disorders. *Mol Syst Biol*. 2009; 5:321. doi: [10.1038/msb.2009.80](https://doi.org/10.1038/msb.2009.80) PMID: [19888216](https://pubmed.ncbi.nlm.nih.gov/19888216/)
17. Wang L, Tu Z, Sun F. A network-based integrative approach to prioritize reliable hits from multiple genome-wide RNAi screens in *Drosophila*. *BMC Genomics*. 2009; 10:220. doi: [10.1186/1471-2164-10-220](https://doi.org/10.1186/1471-2164-10-220) PMID: [19435510](https://pubmed.ncbi.nlm.nih.gov/19435510/)
18. Tu Z, Argmann C, Wong KK, Mitnaul LJ, Edwards S, Sach IC, et al. Integrating siRNA and protein-protein interaction data to identify an expanded insulin signaling network. *Genome Res*. 2009; 19(6):1057–67. doi: [10.1101/gr.087890.108](https://doi.org/10.1101/gr.087890.108) PMID: [19261841](https://pubmed.ncbi.nlm.nih.gov/19261841/)
19. Yuan F, Zhang YH, Wan S, Wang S, Kong XY. Mining for Candidate Genes Related to Pancreatic Cancer Using Protein-Protein Interactions and a Shortest Path Approach. *Biomed Res Int*. 2015; 2015:623121. doi: [10.1155/2015/623121](https://doi.org/10.1155/2015/623121) PMID: [26613085](https://pubmed.ncbi.nlm.nih.gov/26613085/)
20. Jiang P, Wang H, Li W, Zang C, Li B, Wong YJ, et al. Network analysis of gene essentiality in functional genomics experiments. *Genome Biol*. 2015; 16:239. doi: [10.1186/s13059-015-0808-9](https://doi.org/10.1186/s13059-015-0808-9) PMID: [26518695](https://pubmed.ncbi.nlm.nih.gov/26518695/)
21. Shao DD, Tsherniak A, Gopal S, Weir BA, Tamayo P, Stransky N, et al. ATARIS: computational quantification of gene suppression phenotypes from multisample RNAi screens. *Genome Res*. 2013; 23(4):665–78. doi: [10.1101/gr.143586.112](https://doi.org/10.1101/gr.143586.112) PMID: [23269662](https://pubmed.ncbi.nlm.nih.gov/23269662/)
22. 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. doi: [10.1038/nature11003](https://doi.org/10.1038/nature11003) PMID: [22460905](https://pubmed.ncbi.nlm.nih.gov/22460905/)
23. Schaefer CF, Anthony K, Krupa S, Buchoff J, Day M, Hannay T, et al. PID: the Pathway Interaction Database. *Nucleic Acids Res*. 2009; 37(Database issue):D674–9. doi: [10.1093/nar/gkn653](https://doi.org/10.1093/nar/gkn653) PMID: [18832364](https://pubmed.ncbi.nlm.nih.gov/18832364/)
24. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005; 102(43):15545–50. doi: [10.1073/pnas.0506580102](https://doi.org/10.1073/pnas.0506580102) PMID: [16199517](https://pubmed.ncbi.nlm.nih.gov/16199517/)



25. Li Z, Ivanov A, Su R, Qi Q, Webber P, McMillan E, et al. OncoPPI: A Cancer Protein Interaction Network Resource for new Therapeutic Strategies. submitted. 2016.
26. Brin S, Page L. Reprint of: The anatomy of a large-scale hypertextual web search engine. *Computer networks* 2012; 56(18):3825–33.
27. Pratt D, Chen J, Welker D, Rivas R, Pillich R, Rynkov V, et al. NDEX, the Network Data Exchange. *Cell Systems*. 1(4):302–5. doi: [10.1016/j.cels.2015.10.001](https://doi.org/10.1016/j.cels.2015.10.001) PMID: [26594663](https://pubmed.ncbi.nlm.nih.gov/26594663/)
28. Ueda Y, Igishi T, Hashimoto K, Suyama H, Araki K, Sumikawa T, et al. Synergistic cell growth inhibition by the combination of amrubicin and Akt-suppressing tyrosine kinase inhibitors in small cell lung cancer cells: implication of c-Src and its inhibitor. *Int J Oncol*. 2009; 34(3):689–96. PMID: [19212674](https://pubmed.ncbi.nlm.nih.gov/19212674/)
29. Saldanha AJ. Java Treeview—extensible visualization of microarray data. *Bioinformatics*. 2004; 20(17):3246–8. doi: [10.1093/bioinformatics/bth349](https://doi.org/10.1093/bioinformatics/bth349) PMID: [15180930](https://pubmed.ncbi.nlm.nih.gov/15180930/)
30. Akiyama T, Ohuchi T, Sumida S, Matsumoto K, Toyoshima K. Phosphorylation of the retinoblastoma protein by cdk2. *Proc Natl Acad Sci U S A*. 1992; 89(17):7900–4. PMID: [1518810](https://pubmed.ncbi.nlm.nih.gov/1518810/)
31. Neal CL, Xu J, Li P, Mori S, Yang J, Neal NN, et al. Overexpression of 14-3-3zeta in cancer cells activates PI3K via binding the p85 regulatory subunit. *Oncogene*. 2012; 31(7):897–906. doi: [10.1038/onc.2011.284](https://doi.org/10.1038/onc.2011.284) PMID: [21743495](https://pubmed.ncbi.nlm.nih.gov/21743495/)
32. Takemaru KI, Moon RT. The transcriptional coactivator CBP interacts with beta-catenin to activate gene expression. *J Cell Biol*. 2000; 149(2):249–54. PMID: [10769018](https://pubmed.ncbi.nlm.nih.gov/10769018/)
33. Jeon SH, Chae BC, Kim HA, Seo GY, Seo DW, Chun GT, et al. Mechanisms underlying TGF-beta1-induced expression of VEGF and Flk-1 in mouse macrophages and their implications for angiogenesis. *J Leukoc Biol*. 2007; 81(2):557–66. doi: [10.1189/jlb.0806517](https://doi.org/10.1189/jlb.0806517) PMID: [17053163](https://pubmed.ncbi.nlm.nih.gov/17053163/)
34. Rosenbluh J, Nijhawan D, Cox AG, Li X, Neal JT, Schafer EJ, et al. beta-Catenin-driven cancers require a YAP1 transcriptional complex for survival and tumorigenesis. *Cell*. 2012; 151(7):1457–73. doi: [10.1016/j.cell.2012.11.026](https://doi.org/10.1016/j.cell.2012.11.026) PMID: [23245941](https://pubmed.ncbi.nlm.nih.gov/23245941/)
35. Yang F, Li X, Sharma M, Sasaki CY, Longo DL, Lim B, et al. Linking beta-catenin to androgen-signaling pathway. *J Biol Chem*. 2002; 277(13):11336–44. doi: [10.1074/jbc.M111962200](https://doi.org/10.1074/jbc.M111962200) PMID: [11792709](https://pubmed.ncbi.nlm.nih.gov/11792709/)
36. Chesire DR, Ewing CM, Gage WR, Isaacs WB. In vitro evidence for complex modes of nuclear beta-catenin signaling during prostate growth and tumorigenesis. *Oncogene*. 2002; 21(17):2679–94. doi: [10.1038/sj.onc.1205352](https://doi.org/10.1038/sj.onc.1205352) PMID: [11965541](https://pubmed.ncbi.nlm.nih.gov/11965541/)
37. Edlund S, Lee SY, Grimsby S, Zhang S, Aspenstrom P, Heldin CH, et al. Interaction between Smad7 and beta-catenin: importance for transforming growth factor beta-induced apoptosis. *Mol Cell Biol*. 2005; 25(4):1475–88. doi: [10.1128/MCB.25.4.1475-1488.2005](https://doi.org/10.1128/MCB.25.4.1475-1488.2005) PMID: [15684397](https://pubmed.ncbi.nlm.nih.gov/15684397/)
38. Ford CE, Henry C, Llamosas E, Djordjevic A, Hacker N. Wnt signalling in gynaecological cancers: A future target for personalised medicine? *Gynecol Oncol*. 2016; 140(2):345–51. doi: [10.1016/j.ygyno.2015.09.085](https://doi.org/10.1016/j.ygyno.2015.09.085) PMID: [26432042](https://pubmed.ncbi.nlm.nih.gov/26432042/)
39. Rampurwala M, Wisinski KB, O'Regan R. Role of the androgen receptor in triple-negative breast cancer. *Clin Adv Hematol Oncol*. 2016; 14(3):186–93. PMID: [27058032](https://pubmed.ncbi.nlm.nih.gov/27058032/)
40. Dey N, Barwick BG, Moreno CS, Ordanic-Kodani M, Chen Z, Oprea-Ilie G, et al. Wnt signaling in triple negative breast cancer is associated with metastasis. *BMC Cancer*. 2013; 13(1):537.
41. Bilir B, Kucuk O, Moreno CS. Wnt signaling blockage inhibits cell proliferation and migration, and induces apoptosis in triple-negative breast cancer cells. *J Transl Med*. 2013; 11(1):280. Epub 2013/11/06.
42. De P, Carlson JH, Wu H, Marcus A, Leyland-Jones B, Dey N. Wnt-beta-catenin pathway signals metastasis-associated tumor cell phenotypes in triple negative breast cancers. *Oncotarget*. 2016.
43. Deel MD, Li JJ, Crose LE, Lincard CM. A Review: Molecular Aberrations within Hippo Signaling in Bone and Soft-Tissue Sarcomas. *Front Oncol*. 2015; 5:190. doi: [10.3389/fonc.2015.00190](https://doi.org/10.3389/fonc.2015.00190) PMID: [26389076](https://pubmed.ncbi.nlm.nih.gov/26389076/)
44. Simon M, Grandage VL, Linch DC, Khwaja A. Constitutive activation of the Wnt/beta-catenin signalling pathway in acute myeloid leukaemia. *Oncogene*. 2005; 24(14):2410–20. doi: [10.1038/sj.onc.1208431](https://doi.org/10.1038/sj.onc.1208431) PMID: [15735743](https://pubmed.ncbi.nlm.nih.gov/15735743/)
45. Wang DY, Wu YN, Huang JQ, Wang W, Xu M, Jia JP, et al. Hippo/YAP signaling pathway is involved in osteosarcoma chemoresistance. *Chin J Cancer*. 2016; 35(1):47.
46. Derksen PW, Tjin E, Meijer HP, Klook MD, MacGillivray HD, van Oers MH, et al. Illegitimate WNT signaling promotes proliferation of multiple myeloma cells. *Proc Natl Acad Sci U S A*. 2004; 101(16):6122–7. doi: [10.1073/pnas.0305855101](https://doi.org/10.1073/pnas.0305855101) PMID: [15067127](https://pubmed.ncbi.nlm.nih.gov/15067127/)
47. Tiemessen MM, Staal FJ. Wnt signaling in leukemias and myeloma: T-cell factors are in control. *Future Oncol*. 2013; 9(11):1757–72. doi: [10.2217/fon.13.122](https://doi.org/10.2217/fon.13.122) PMID: [24156335](https://pubmed.ncbi.nlm.nih.gov/24156335/)

48. Jiang X, Castelao JE, Vandenberg D, Carracedo A, Redondo CM, Conti DV, et al. Genetic variations in SMAD7 are associated with colorectal cancer risk in the colon cancer family registry. *PLoS One*. 2013; 8(4):e60464. doi: [10.1371/journal.pone.0060464](https://doi.org/10.1371/journal.pone.0060464) PMID: [23560096](https://pubmed.ncbi.nlm.nih.gov/23560096/)
49. Stolfi C, De Simone V, Colantoni A, Franze E, Ribichini E, Fantini MC, et al. A functional role for Smad7 in sustaining colon cancer cell growth and survival. *Cell Death Dis*. 2014; 5:e1073. doi: [10.1038/cddis.2014.49](https://doi.org/10.1038/cddis.2014.49) PMID: [24556688](https://pubmed.ncbi.nlm.nih.gov/24556688/)
50. Zhang H, Liu C, Kong Y, Huang H, Wang C, Zhang H. TGFbeta signaling in pancreatic ductal adenocarcinoma. *Tumour Biol*. 2015; 36(3):1613–8. doi: [10.1007/s13277-014-2757-4](https://doi.org/10.1007/s13277-014-2757-4) PMID: [25355597](https://pubmed.ncbi.nlm.nih.gov/25355597/)
51. Pedraza-Alva G, Koulis M, Charland C, Thornton T, Clements JL, Schlissel MS, et al. Activation of p38 MAP kinase by DNA double-strand breaks in V(D)J recombination induces a G2/M cell cycle checkpoint. *EMBO J*. 2006; 25(4):763–73. doi: [10.1038/sj.emboj.7600972](https://doi.org/10.1038/sj.emboj.7600972) PMID: [16456545](https://pubmed.ncbi.nlm.nih.gov/16456545/)
52. Wu CC, Wu X, Han J, Sun P. p38gamma regulates UV-induced checkpoint signaling and repair of UV-induced DNA damage. *Protein Cell*. 2010; 1(6):573–83. doi: [10.1007/s13238-010-0075-1](https://doi.org/10.1007/s13238-010-0075-1) PMID: [21170151](https://pubmed.ncbi.nlm.nih.gov/21170151/)
53. Tsai MS, Weng SH, Chen HJ, Chiu YF, Huang YC, Tseng SC, et al. Inhibition of p38 MAPK-dependent excision repair cross-complementing 1 expression decreases the DNA repair capacity to sensitize lung cancer cells to etoposide. *Mol Cancer Ther*. 2012; 11(3):561–71. doi: [10.1158/1535-7163.MCT-11-0684](https://doi.org/10.1158/1535-7163.MCT-11-0684) PMID: [22053010](https://pubmed.ncbi.nlm.nih.gov/22053010/)
54. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. GenePattern 2.0. *Nat Genet*. 2006; 38(5):500–1. doi: [10.1038/ng0506-500](https://doi.org/10.1038/ng0506-500) PMID: [16642009](https://pubmed.ncbi.nlm.nih.gov/16642009/)
55. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003; 13(11):2498–504. doi: [10.1101/gr.1239303](https://doi.org/10.1101/gr.1239303) PMID: [14597658](https://pubmed.ncbi.nlm.nih.gov/14597658/)
56. The Cancer Genome Atlas Research N. Comprehensive molecular profiling of lung adenocarcinoma. *Nature*. 2014; 511(7511):543–50. <http://www.nature.com/nature/journal/v511/n7511/abs/nature13385.html> - supplementary-information. doi: [10.1038/nature13385](https://doi.org/10.1038/nature13385) PMID: [25079552](https://pubmed.ncbi.nlm.nih.gov/25079552/)
57. Benjamini Y, Yekutieli D. The Control of the False Discovery Rate in Multiple Testing Under Dependency. *The Annals of Statistics*. 2001; 29(4):1165–88.
58. Lee JH, Park KS, Alberobello AT, Kallakury B, Weng MT, Wang Y, et al. The Janus kinases inhibitor AZD1480 attenuates growth of small cell lung cancers in vitro and in vivo. *Clin Cancer Res*. 2013; 19(24):6777–86. doi: [10.1158/1078-0432.CCR-13-1110](https://doi.org/10.1158/1078-0432.CCR-13-1110) PMID: [24158701](https://pubmed.ncbi.nlm.nih.gov/24158701/)
59. Kim SM, Kwon OJ, Hong YK, Kim JH, Solca F, Ha SJ, et al. Activation of IL-6R/JAK1/STAT3 signaling induces de novo resistance to irreversible EGFR inhibitors in non-small cell lung cancer with T790M resistance mutation. *Mol Cancer Ther*. 2012; 11(10):2254–64. doi: [10.1158/1535-7163.MCT-12-0311](https://doi.org/10.1158/1535-7163.MCT-12-0311) PMID: [22891040](https://pubmed.ncbi.nlm.nih.gov/22891040/)
60. Tse C, Shoemaker AR, Adickes J, Anderson MG, Chen J, Jin S, et al. ABT-263: a potent and orally bioavailable Bcl-2 family inhibitor. *Cancer Res*. 2008; 68(9):3421–8. doi: [10.1158/0008-5472.CAN-07-5836](https://doi.org/10.1158/0008-5472.CAN-07-5836) PMID: [18451170](https://pubmed.ncbi.nlm.nih.gov/18451170/)
61. Vassilev LT, Vu BT, Graves B, Carvajal D, Podlaski F, Filipovic Z, et al. In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science*. 2004; 303(5659):844–8. doi: [10.1126/science.1092472](https://doi.org/10.1126/science.1092472) PMID: [14704432](https://pubmed.ncbi.nlm.nih.gov/14704432/)
62. Lenz HJ, Kahn M. Safely targeting cancer stem cells via selective catenin coactivator antagonism. *Cancer Sci*. 2014; 105(9):1087–92. doi: [10.1111/cas.12471](https://doi.org/10.1111/cas.12471) PMID: [24975284](https://pubmed.ncbi.nlm.nih.gov/24975284/)
63. Wyce A, Ganji G, Smitheman KN, Chung CW, Korenchuk S, Bai Y, et al. BET inhibition silences expression of MYCN and BCL2 and induces cytotoxicity in neuroblastoma tumor models. *PLoS One*. 2013; 8(8):e72967. Epub 2013/09/07. doi: [10.1371/journal.pone.0072967](https://doi.org/10.1371/journal.pone.0072967) PMID: [24009722](https://pubmed.ncbi.nlm.nih.gov/24009722/)
64. Bandopadhyay P, Bergthold G, Nguyen B, Schubert S, Gholamin S, Tang Y, et al. BET bromodomain inhibition of MYC-amplified medulloblastoma. *Clin Cancer Res*. 2014; 20(4):912–25. doi: [10.1158/1078-0432.CCR-13-2281](https://doi.org/10.1158/1078-0432.CCR-13-2281) PMID: [24297863](https://pubmed.ncbi.nlm.nih.gov/24297863/)
65. Stumpf MPH, Thorne T, de Silva E, Stewart R, An HJ, Lappe M, et al. Estimating the size of the human interactome. *Proceedings of the National Academy of Sciences*. 2008; 105(19):6959–64.
66. Mitsopoulos C, Schierz AC, Workman P, Al-Lazikani B. Distinctive Behaviors of Druggable Proteins in Cellular Networks. *PLoS Comput Biol*. 2015; 11(12):e1004597. doi: [10.1371/journal.pcbi.1004597](https://doi.org/10.1371/journal.pcbi.1004597) PMID: [26699810](https://pubmed.ncbi.nlm.nih.gov/26699810/)
67. Newaz K, Sriram K, Bera D. Identification of Major Signaling Pathways in Prion Disease Progression Using Network Analysis. *PLoS One*. 2015; 10(12):e0144389. doi: [10.1371/journal.pone.0144389](https://doi.org/10.1371/journal.pone.0144389) PMID: [26646948](https://pubmed.ncbi.nlm.nih.gov/26646948/)