Network potential identifies therapeutic *miRNA* cocktails in Ewing sarcoma

Davis T. Weaver[1, 2, *], **Kathleen I. Pishas**[3,*], **Drew Williamson**[4,*], **Jessica Scarborough**[1, 2], **Stephen L. Lessnick**[5], **Andrew Dhawan**[2, 6, †], **and Jacob G. Scott**[1, 2, 7, †]

[1]Case Western Reserve University School of Medicine, Cleveland, OH, 44106, USA
[2]Translational Hematology Oncology Research, Cleveland Clinic, Cleveland OH, 44106, USA
[3]Peter MacCallum Cancer Centre, Melbourne, Australia, 3000
[4]Department of Pathology, Brigham & Women's Hospital, Boston, MA, 02144
[5]Nationwide Children's Hospital, Columbus, Ohio, 43205
[6]Division of Neurology, Cleveland Clinic, Cleveland Ohio, 44195
[7]Department of Physics, Case Western Reserve University, Cleveland, OH, 44106, USA
[*]contributed equally
[†]dhawana@ccf.org, scottj10@ccf.org

ABSTRACT

Introduction: Micro-RNA (miRNA)-based therapies are an emerging class of cancer therapies with many potential applications in the field. Because of the broad effects miRNAs can have on different cells and tissues, a network science-based approach is well-equipped to evaluate and identify miRNA candidates and combinations of candidates for the repression of key oncogenic targets.
Methods: We first characterized 6 Ewing sarcoma cell lines using paired mRNA and miRNA sequencing. We then estimated a measure of tumor state, which we term network potential, based on both the mRNA gene expression and the underlying protein-protein interaction network in the tumor. Next, we ranked mRNA targets based on their contribution to network potential. After identifying these mRNA targets, we identified miRNAs and combinations of miRNAs that preferentially act to repress these targets.
Results: We identified TRIM25, APP, ELAV1, RNF4, XPO1 as ideal protein targets for Ewing sarcoma therapy. Using miRNA-mRNA target mappings, we identified miR-3613-3p, let-7a-3p, miR-300, miR-424-5p, and let-7b-3p as the optimal miRNAs for preferential repression of these targets.
Discussion: In this work, we applied a novel pipeline for identification of miRNA candidates for cancer therapy, using Ewing sarcoma as a model system.

Introduction

MicroRNA (miRNA)-based treatments, including anti-sense oligonucleotides, are an emerging class of cancer therapy[1]. Recent work has highlighted the critical importance of miRNAs in the development and maintenance of the cancer phenotype[1–3]. MiRNA deregulation has been implicated in the development of each of the hallmark features of cancer, and restoration of expression of some of these critical downregulated miRNAs has been studied as a potential treatment for several different cancers[3–5]. In particular, in the past decade, anti-sense oligonucleotide inhibitors of the STAT3 transcription factor have shown promise in the settings of lymphoma[6,7] and neuroblastoma[8]. MiR-34 has shown to be effective in pre-clinical studies for treatment of both lung cancer[9–11] and prostate cancer[12]. Finally, mir-34 and let-7 combination therapy has been shown to be effective in pre-clinical studies of lung cancer [11].

MiRNAs have been recognized as potential high-value therapeutics in part due to their ability to cause widespread changes in a cell-signaling network[1]. A single miRNA can bind to and repress multiple mRNA transcripts[3,13–15], a property that can be exploited when designing therapy to maximally disrupt a cancer cell signaling network. This promiscuity of miRNA binding may also increase the risk of off-target effects and toxicity (Figure 1). For example, mir-34 was effective in pre-clinical studies for the treatment of a variety of solid tumors[9–12], only to fail in a phase I clinical trial due to "immune-related serious adverse events"[16]. To capitalize on the promise of miRNA-based cancer therapy while limiting potential toxicity, a systematic, network-based approach is needed to evaluate miRNA candidates and combinations of candidates.

In this work, we build on previous studies applying thermodynamic measures to cell signaling networks in the field of cancer biology[17–19], as well as works that describe a method to use gene homology to map miRNAs to the mRNA transcripts they likely repress[3,13,14]. Reitman et al. previously described a metric of cell state analogous to Gibbs free energy that can be

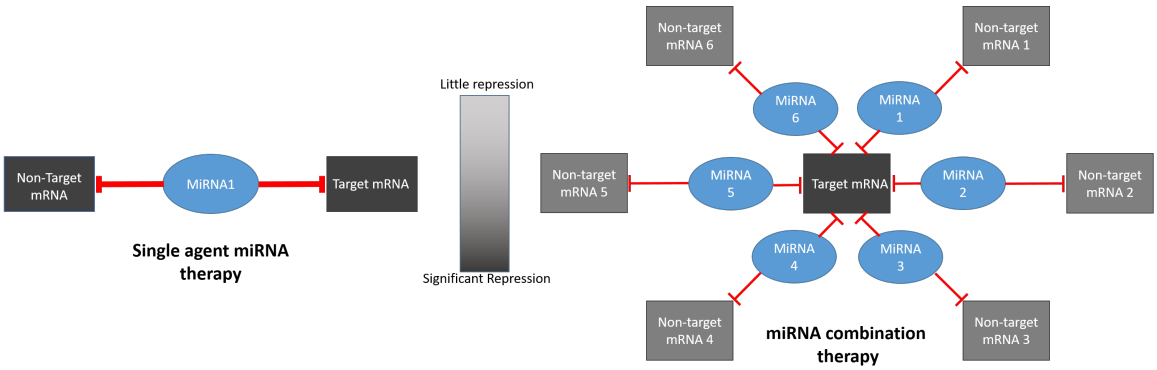

Figure 1. Cartoon describing rationale for focusing on miR combination therapy. With single-agent therapy, both target mRNA and non-target mRNA are inhibited an equal amount, potentially resulting in toxicity due to off-target effects. With miRNA combination therapy, the common target mRNA is inhibited to a greater degree than any individual non-target miRNA.

calculated using the protein interaction network of human cells and relevant transcriptomics data[17]. Gibbs free energy has been correlated with a number of cancer-specific outcomes, including cancer grade and patient survival[18]. Additionally, they leveraged Gibbs and other network measures to identify personalized protein targets for therapy in a dataset of low-grade glioma patients from The Cancer Genome Atlas (TCGA)[17]. Previous work has also highlighted the critical importance of miRNAs to maintenance and development of the oncogenic phenotype, and demonstrated the utility of applying miRNA-mRNA mappings [3].

Building on these works, we developed a novel computational pipeline for the identification of miRNA-based therapies for cancer using Ewing sarcoma as a model system. Ewing sarcoma is a rare malignancy arising from a gene fusion secondary to rearrangements involving the EWS gene[20]. There are 200-300 reported cases each year in the United States, disproportionately affecting children[21]. Ewing sarcoma is extremely prone to developing resistance to available therapies [2], making it an ideal system on which to develop novel therapies to treat resistant tumors or avoid the development of resistance altogether.

1 Methods

1.1 Overview

We characterized six previously described Ewing sarcoma cell lines[22]–A673, ES2, EWS502, TC252, TC32, and TC71–using matched mRNA and miRNA sequencing. We then defined a measure of tumor state, which we term network potential (Equation 1), based on both mRNA gene expression and the underlying protein-protein interaction network (PPI). Next, we ranked mRNA targets based on their contribution to network potential of each cell line, aiming to approximate the relative importance of each mRNA to network stability. Relative importance of each mRNA to network stability was determined by calculating the change in network potential of each network before and after *in silico* repression of each mRNA (ΔG, described in Section 1.5). After identifying these mRNA targets, we then identified miRNA that preferentially acted to repress the most influential of the ranked mRNA targets, with the aim of defining synthetic miRNA-based therapy for down-regulation of these targets. Our computational pipeline is schematized in Figure S1.

1.2 Data sources

We utilized two data sources to develop our Ewing sarcoma cell signaling networks: the BioGRID interaction database[23] and mRNA and miRNA expression data from 6 Ewing sarcoma cell lines, which will be available soon on the sequence read archive.

BioGRID The BioGRID interaction database contains curated data detailing known interactions between proteins for a variety of different species, including Homo sapiens. The data were generated by manual curation of the biomedical literature to identify documented interactions between proteins[23]. To assist in manual curation, the BioGRID project uses a natural language processing algorithm that analyzes the scientific literature to identify manuscripts likely to contain information about novel PPI. The dataset is therefore limited to protein interactions that are reliably reported in the scientific literature. As new research accumulates, substantial changes to the PPI network may occur. For example, between 2016 and 2018, the number of documented PPI in *Homo sapiens* grew from 365,547 to 449,842. The 449,842 documented interactions in 2018 were identified through curation of 27,631 publications[23]. Importantly, the PPI network is designed to represent normal human tissue.

Ewing sarcoma transcriptomics Second, we utilized mRNA expression data from *in vitro* experiments conducted on six Ewing sarcoma cell lines. RNA/miRNA extraction was performed with a Qiagen kit with on-column DNase digestion. These mRNA and miRNA expression data were then normalized to account for between sample differences in data processing and further adjusted using a regularized log (Rlog) transformation[24,25]. Notably, methods for calculating network potential from this type of data require protein concentrations rather than mRNA transcript concentrations. For the purposes of this analysis, we assumed that concentration of protein in an Ewing sarcoma tumor was equivalent to the concentration of the relevant mRNA transcript. A large body of work suggests that mRNA levels are the primary driver of protein levels in a cell under steady state conditions (i.e. not undergoing proliferation, response to stress, differentiation etc)[26–29], a condition that is generally satisfied by transcriptomic data of the type collected in this study[26].

1.3 Network development

We first developed a generic network to represent human cell signaling networks using the BioGRID interaction database[23]. The BioGRID protein-protein interaction network can be downloaded as a non-linear data structure containing ordered pairs of proteins and all the other proteins with which they interact. This data structure can be represented as an undirected graph, with vertex set $\mathscr{V}$, with each vertex representing to a protein, and edge set ($\mathscr{E}$) describing the interactions between proteins.

Using RNA sequencing data from 6 Ewing sarcoma cell lines in triplicate, we then ascribed mRNA transcript concentration for each gene as an attribute to represent the protein concentration for each node in the graph. Through this process, we developed networks specific to each cell line and replicate.

1.4 Network potential calculation

Using the cell signaling network with attached cell line and replicate number specific normalized mRNA expression data, we defined a measure of tumor state following Reitman et al.[17], which we term network potential. We first calculate the network potential of the i-th node in the graph:

$$G_i = C_i \ln\left[\frac{C_i}{\sum C_j + C_i}\right]. \tag{1}$$

where G_i is equal to the network potential of an individual node of the graph, C_i is equal to the concentration of protein corresponding to node G_i, and C_j is the concentration of protein of the j-th neighbor of G_i. Total network potential (G) of the network can then be calculated as the sum over all nodes:

$$G = \sum_i G_i. \tag{2}$$

where G is equal to the total network network potential for each biological replicate of a given cell line. We then compared total network potential across cell lines and biological replicates.

1.5 Ranking of protein targets

After calculation of network potential for each node and the full network, we ranked each protein in the network by network potential and selected the top 5% of proteins for further analysis. For this 5% of all proteins, we simulated "repression" by reducing their expression (computationally) to zero one at a time[30]. Clinically, this would be akin to the application of a drug that perfectly inhibited the protein/mRNA of interest. We then re-calculated network potential for the full network and calculated the change in network potential (δG) by subtracting the new network potential value for the network potential value of the "unrepressed" network. We chose the top 5% of proteins according to the G_i associated with each node for this analysis in order to limit subsequent analysis to proteins with likely biological significance in Ewing sarcoma. We then ranked this top 5% in descending order of the calculated network potential difference. This 5% constituted the set of protein targets on which we conducted further analysis.

1.6 Identification of miRNA candidates

To generate miRNA-mRNA mappings, we implemented a protocol described previously[31]. Briefly, we identified all predicted mRNA targets for each miRNA in our data-set using the miRNAtap database in R, version 1.18.0, as implemented through the Bioconductor targetscan org.Hs.eg.db package, version 3.8.2[13]. We used all five possible databases (default settings): DIANA version 5.061[15], Miranda 2010 release62[32], PicTar 2005 release63[33], TargetScan 7.164[34] and miRDB 5.065[14], with a minimum source number of 2, and the union of all targets found was taken as the set of targets for a given miRNA. Through this mapping, we identified a list of mRNA transcripts that are predicted to be repressed by a given miRNA. Our code is available on Github ($https://github.com/DavisWeaver/MiR_{Combo}Targeting/$).

Using this mapping as well as our ranked list of promising gene candidates for repression from our network analysis, we were able to identify a list of miRNA that we predict would maximally disrupt the Ewing sarcoma cell signaling network when introduced synthetically. To rank miRNA targets, we first identified all the genes on the full target list that a given miRNA was predicted to repress (described in Section 1.5). Next, we summed the predicted ΔG when each of these genes was repressed *in silico* to generate the maximum potential disruption that could be achieved if a given miRNA were introduced synthetically into an Ewing sarcoma tumor. We then ranked miRNA candidates in descending order of the maximum predicted network disruption.

In addition, we identified combinations of miRNAs with overlapping mRNA targets in order to identify promising miRNA cocktails. miRNA cocktails may effectively destabilize cancer cell signaling networks while avoiding systemic toxicities. To do so, we first identified a subset of key mRNA targets to focus on. For this analysis, we limited our list of targets to those that were previously causally implicated in cancer[35]. We then used our miRNA-mRNA mappings to identify a cocktail of miRNAs in which every miRNA was predicted to repress > 2 of the top 5 mRNA transcripts ranked by predicted ΔG when each of these genes was repressed *in silico*.

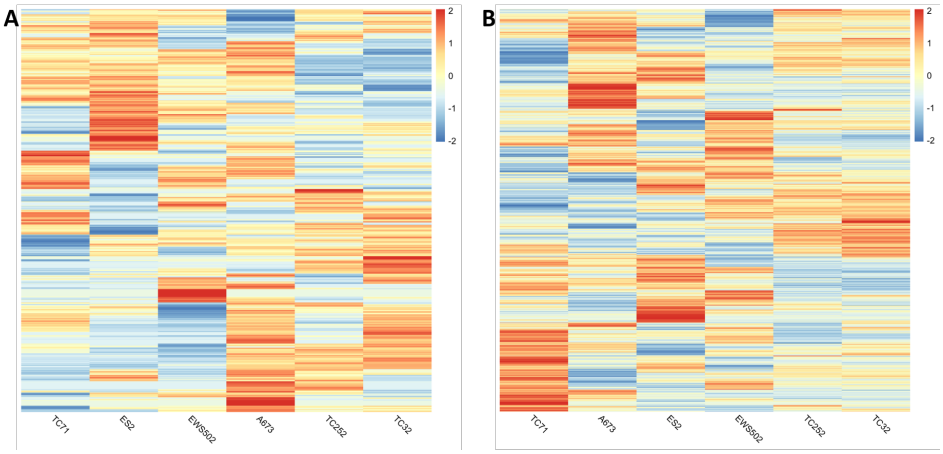

Figure 2. We performed matched mRNA and miRNA sequencing for 6 Ewing sarcoma cell lines. A: Clustered miRNA expression. **B**: Clustered mRNA expression. Rows and columns were clustered using the complete linkage method with euclidean distance as the distance metric. Ewing sarcoma cell line is represented on the x axis while mRNA/miRNA is represented on the y axis.

2 Results

We performed matched mRNA and miRNA sequencing on six Ewing sarcoma cell lines: A673, ES2, EW502, TC252, TC32, and TC71 (Figure 2). Clustered miRNA and MRNA expression are shown in Figure 2. There were distinct differences in mRNA and miRNA sequencing across cell lines.

2.1 Network Overview

We calculated the network potential, a unitless measure of cell state, for each protein in the cell signaling networks for each of the six Ewing sarcoma cell lines in our experiment. An overview of the total network potential for each cell line and evolutionary replicate compared to total mRNA expression is presented in Figure 3. The density plots of network potential and mRNA expression look markedly different (Figure 3A and Figure 3B), indicating that network potential describes different features of a cell signaling network compared to mRNA expression alone. Notably, network potential and mRNA expression for these cell lines is stable across different biological replicates, as demonstrated by the low interquartile range (Figure 3C and 3D). There were modest differences in both mean expression and network potential across cell lines (Figure 3C and 3D). The global average network potential across all samples was -3.45×10^5 with a standard deviation of 1725.

2.2 Identification of Protein Targets

We identified TRIM25, APP, ELAV1, RNF4, and XPO1 as top 5 targets for therapy for each of the 6 cell lines based on the degree of network disruption induced when each gene was modeled as repressed. There was a high degree of concordance between cell lines among the top predicted targets, indicating that these targets may be conserved across Ewing sarcoma (Table S1). Of the top ten predicted targets, 9 are among the top 10 targets for all 6 cell lines.

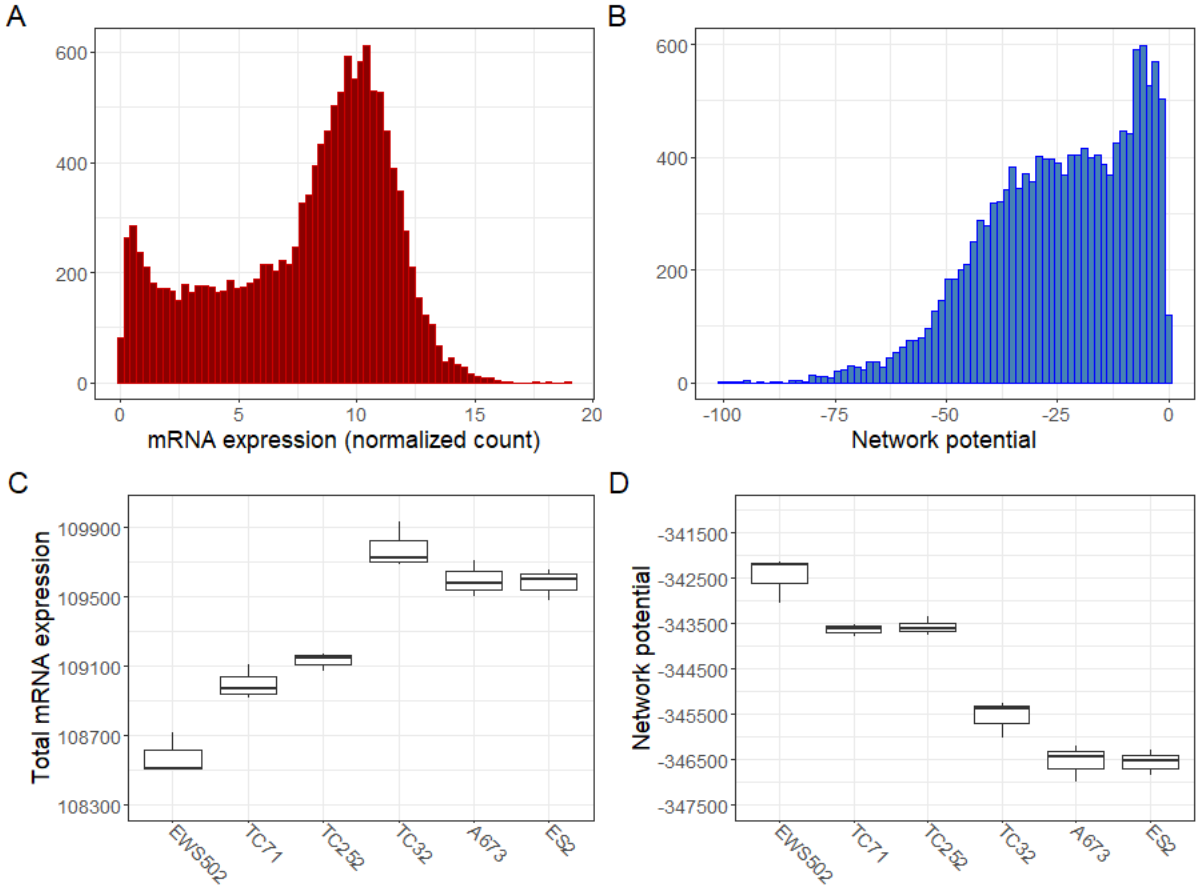

Figure 3. Network potential describes different features of a cell signaling network compared to mRNA expression alone. Panel A and B show violin plots of mRNA expression and network potential, respectively, for each protein in our cell signaling network. mRNA transcripts with an expression level of zero were excluded to better visualize the distribution of genes that are expressed. Panel C and D show box-plots of total mRNA expression and total network potential, respectively, for each cell-line specific cell signaling network.

Some of these identified genes may be essential housekeeping genes highly expressed in all or most cells in the body, making them inappropriate drug targets. TRIM25, and ELAV1, for example, are involved in protein modification and RNA binding, respectively[36]. We therefore repeated this analysis, limiting our search to gene targets that have been causally implicated in cancer[35]. With this limitation in place, we identified XPO1, LMNA, EWSR1, HSP90AA1, and CUL3 as the top 5 targets for therapy when ΔG was averaged for all cell lines. The top 10 cancer-related targets for each cell line can be found in (Table 1). The top 50 protein targets (limited to those causally implicated in cancer) are presented in Figure 4. The top 50 protein targets (not limited to those causally implicated in cancer) can be found in Figure S2. We also conducted gene set enrichment analysis for the all the genes represented in our network. We ranked genes by network potential and compared our gene set to the "hallmarks" pathways set, downloaded from the Molecular Signatures Database (MSigDB)[37,38]. This analysis was conducted using the fGSEA package in R, which uses the Benjamini - Hochberg procedure to correct the false discovery rate[39,40]. Our gene set was enriched (p-value < 0.05) in 22 of the 50 pathways included in the hallmarks set; including apoptosis, DNA repair, mTOR signaling, MYC signaling, and WNT β-catenin signaling. The full results are presented in Table S2.

2.3 Identification of MiRNA Candidates

Candidates for single agent therapy We identified several miRNAs that were predicted to dramatically disrupt the Ewing sarcoma cell signaling network (Figure 5). When averaging all cell lines, we identified miR-3613-3p, let-7a-3p, miR-300, miR-424-5p, and let-7b-3p as the ideal miRs for preferential repression of proteins predicted to be important for Ewing sarcoma signaling network stability. miR-3613-3p, let-7a-3p, miR-300, miR-424-5p, and let-7b-3p were predicted to cause a network network potential increase (driving the system less negative) of 17508, 13180, 12813, 12621 and 12417, respectively (see

	TC252	ES2	A673	TC32	EWS502	TC71
1	**XPO1**	**XPO1**	**XPO1**	**XPO1**	**XPO1**	**XPO1**
2	**LMNA**	**LMNA**	**LMNA**	**LMNA**	NTRK1	EWSR1
3	EWSR1	**HSP90AA1**	**HSP90AA1**	EWSR1	**HSP90AA1**	LMNA
4	**HSP90AA1**	**EWSR1**	**EWSR1**	**HSP90AA1**	**EWSR1**	**HSP90AA1**
5	NTRK1	**CUL3**	**CUL3**	**CUL3**	LMNA	**CUL3**
6	CUL3	KRAS	NTRK1	NTRK1	CUL3	RECQL4
7	**TP53**	**TP53**	RECQL4	**TP53**	RECQL4	KRAS
8	RECQL4	RECQL4	KRAS	KRAS	TP53	BRCA1
9	KRAS	BRCA1	TP53	RECQL4	KRAS	TP53
10	**BRCA1**	EZH2	**BRCA1**	**BRCA1**	**BRCA1**	MYC

Table 1. Top cancer-associated protein targets for each cell line. We ranked potential targets by predicted change in network potential when each protein was modeled as repressed, limited to proteins causally associated in cancer according to the Cosmic database. Proteins that appear in the same position for ≥ 3 cell lines are **bolded**.

Figure 5). The top miR candidate, miR-3613-3p, was predicted to repress 144 distinct mRNA transcripts in the full target set (Table S3), including 33 genes causally implicated in oncogenesis (according to the Cosmic Database[35]). These genes include classic mediators of oncogenesis KRAS, MDM2, SMAD2, SMAD4, and NOTCH1[35]. As with the protein targets, there was a high degree of concordance in the top predicted miR candidates between cell lines (Table 2). It should also be noted that we were able to identify a substantial number of miRNA with potential activity against the Ewing sarcoma cell signaling network. We identified 27 miRNAs with a predicted network potential disruption of greater than $10,000$. For comparison, the largest network change in network potential that could be achieved with a single gene repression across all cell lines was just 2064 (TRIM25).

	TC252	ES2	A673	TC32	EWS502	TC71
1	miR-3613-3p	miR-3613-3p	miR-3613-3p	miR-3613-3p	miR-3613-3p	miR-3613-3p
2	let-7a-3p	let-7a-3p	let-7a-3p	let-7a-3p	let-7a-3p	let-7a-3p
3	miR-300	miR-300	miR-300	miR-300	miR-300	miR-300
4	let-7b-3p	let-7b-3p	let-7b-3p	let-7b-3p	let-7b-3p	miR-4282
5	miR-4282	miR-424-5p	miR-424-5p	miR-424-5p	miR-424-5p	let-7b-3p
6	miR-424-5p	miR-4282	miR-4282	miR-4282	miR-4282	miR-424-5p
7	miR-15a-5p	miR-15a-5p	miR-15a-5p	miR-15a-5p	miR-15a-5p	miR-548a-3p
8	miR-548a-3p	miR-182-5p	miR-548a-3p	miR-195-5p	miR-548a-3p	miR-15a-5p
9	miR-195-5p	miR-590-3p	miR-15b-5p	miR-590-3p	miR-548e-5p	miR-548e-5p
10	miR-15b-5p	miR-195-5p	miR-195-5p	miR-15b-5p	miR-195-5p	miR-15b-5p

Table 2. Top miRNA candidates for single-agent therapy of Ewing sarcoma for each cell line.

miRNA combination therapy Given the documented cases of systemic toxicities associated with miRNA-based therapies, the miRNA that inhibits the most targets might not necessarily be the best drug target. We therefore sought to identify combinations of miRNAs that individually repressed key drug targets. Our hypothesis is that by giving a cocktail of miRNAs with predicted activity against one or multiple identified drug targets, each individual miRNA could be given at a low dose such that only the mRNA transcripts that are targeted by multiple miRNAs in the cocktail are affected. For each cell line, we identified the miRNAs that targeted the largest number of the top 5 mRNA targets. We selected the top five such miRNAs, ranked by number of top mRNA targets that they were predicted to repress, to make up our ideal "miRNA cocktail". For all six cell lines, the top miRNA cocktail for preferential repression of the top 5 mRNA targets was let-7a-3p, let-7b-3p, let-7f-1-3p, miR-302a-5p, and miR-497-3p. We repeated this analysis to identify cocktails that target larger groups of mRNA (the top 10 and 15 mRNA targets) in order to assess the stability of the predicted cocktail to changing conditions. When targeting the top 10 mRNAs identified by predicted change in network potential, the the top miRNA cocktail was miR-4692, let-7a-3p, let-7b-3p, let-7f-1-3p, and miR-15a-5p (6).

When targeting the top 15 mRNA targets, let-7a-3p, let-7b-3p, let-7f-1-3p were again identified as key components of the ideal miRNA cocktail for disruption of the Ewing sarcoma cell signaling network. The full predicted cocktail for each cell line can be found in Table 3.

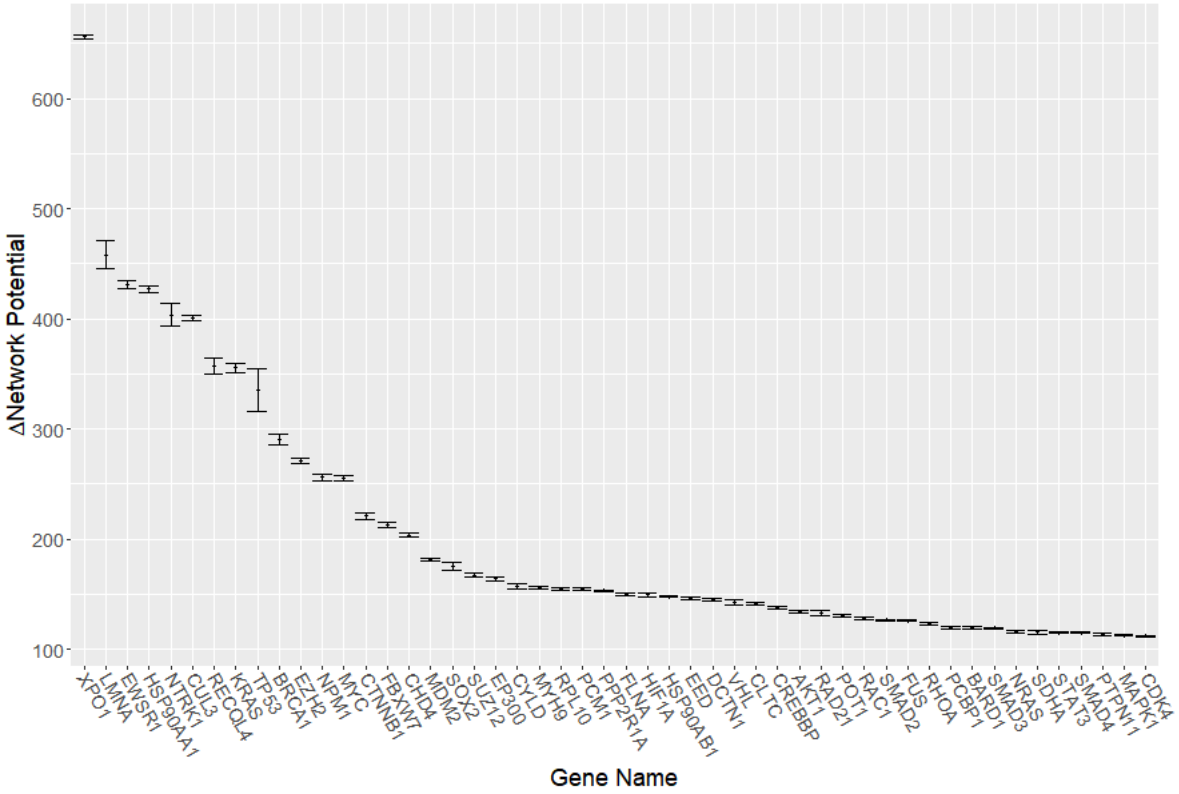

Figure 4. Protein targets ranked by contribution to network stability. When averaging across cell lines, XPO1, LMNA, EWSR1, HSP90AA1, and NTRK1 were identified as the most important proteins in the Ewing sarcoma cell signaling network (when limiting our analysis to proteins causally implicated in cancer[35]). When each protein was simulated as completely repressed, network potential was increased by 656, 458, 431, 427, and 403, respectively. It is reassuring that EWSR1, the kinase associated with Ewing sarcoma development, is identified as highly influential in the cell signaling network by this method.

The optimal number of targets for designing an effective miRNA cocktail is unclear. We present these results to highlight the flexibility of our method for empirically designing miRNA-based therapies for the treatment of cancer. Validation of individual regimens *in vitro* and *in vivo* is necessary to establish clinical utility for any of the regimens identified with this method.

3 Discussion

In this work, we described a novel pipeline for the identification of miRNA candidates for cancer therapy, using Ewing sarcoma as a model system. Our pipeline can also be used to identify targetable proteins that are important to the stability of a cell signaling network. First, we performed matched mRNA and miRNA sequencing on six Ewing sarcoma cell lines (available soon on the sequence read archive). We then defined a metric of cell state, network potential. Using network potential, we identified the most important proteins in the cell signaling network for each of the 6 cell lines. Notably, this set of proteins was strongly enriched in genes involved in the canonical miRNA biogenesis pathway[3], which supports the crucial role miRNA play in modulation of cell signaling.

To our knowledge, ours is the first generic pipeline that analyzes RNA sequencing data to generate protein targets and miRNA therapeutic candidates using network potential as an organizing principle. In previous work, researchers have explored "perturbation biology", which is a combined experimental-computational approach to predict the response of cells to perturbations[41,42]. Perturbation biology relies on systemic perturbations of cells with drug combinations followed by measurement of response profiles. Network models are developed from these data and used to nominate novel drug combinations with predicted efficacy[41,42]. In addition, previous work has used network potential to retrospectively identify individualized drug targets for patients with low-grade glioma[17]. We extend this work by simplifying the network analysis and building the capability to identify miRNA therapeutic candidates given a set of protein targets.

Prior to application of these findings for treatment of Ewing sarcoma or any other disease, specific *in vitro* and *in vivo*

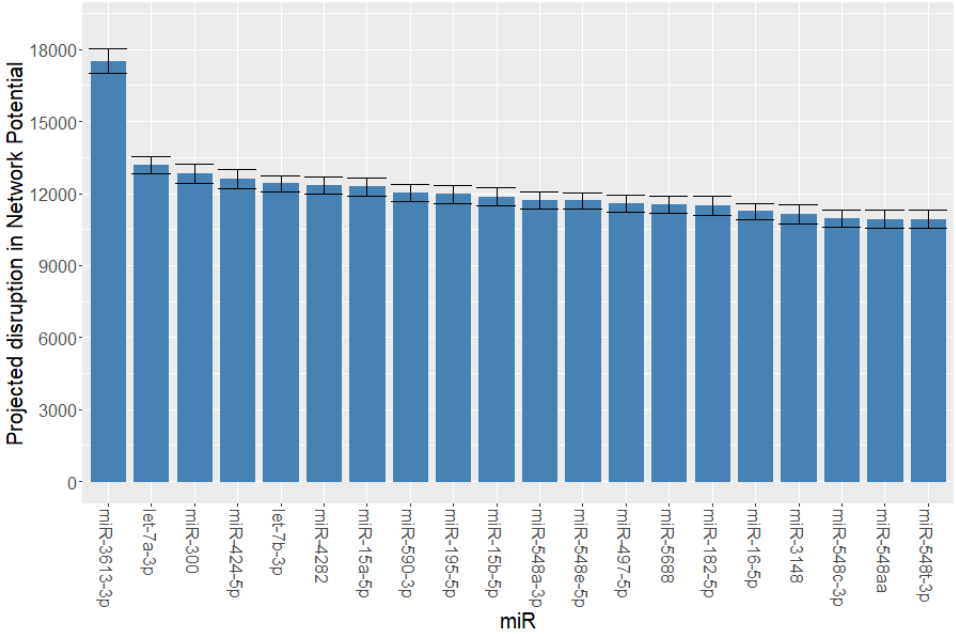

Figure 5. Top miRNA candidates for repression of the Ewing sarcoma cell signaling network. We identified the top 20 miRNA for treatment of Ewing sarcoma, ranked by their predicted disruption of the Ewing sarcoma cell signaling network. mir-3613-3p was predicted to induce the largest network disruption, with a predicted change in network potential of 17,509.

	TC252	ES2	A673	TC32	EWS502	TC71
1	let-7a-3p	let-7a-3p	let-7a-3p	let-7a-3p	miR-4692	let-7a-3p
2	miR-4692	miR-4692	miR-4692	miR-4692	let-7a-3p	miR-4692
3	let-7b-3p	let-7b-3p	let-7b-3p	let-7b-3p	let-7b-3p	let-7b-3p
4	let-7f-1-3p	let-7f-1-3p	let-7f-1-3p	let-7f-1-3p	let-7f-1-3p	let-7f-1-3p
5	miR-9-5p	miR-483-3p	miR-9-5p	miR-9-5p	miR-483-3p	miR-483-3p

Table 3. Predicted miRNA cocktail for each cell line for repression of the top 15 protein targets.

validation is needed. In addition, the protein targets and miRNA candidates we identified in our model dataset of 6 Ewing sarcoma cell lines are consistent with the literature on Ewing sarcoma and cancer cell signaling. Of the top 50 protein targets that we identified, 15 have previously been causally implicated in cancer[35], including EWSR1, the proposed driver of Ewing sarcoma development. Canonical drivers of oncogenesis KRAS, TP53, and MYC[35] were also identified as key protein targets in the Ewing sarcoma cell signaling network for these cell lines, providing further validation for our computational pipeline.

In addition, many of the miRNA we identified as potential therapeutic candidates have been previously studied due to their association with cancer outcomes, including members of the let-7 family, miR-300, miR-424-5p, miR-4282, miR-15a-5p, and miR-590-3p. Loss of expression of the let-7 family of miRNA has been widely implicated in cancer development[43–46]. In Ewing sarcoma specifically, low levels of let-7 family miRNA have been correlated with disease progression or recurrence[43]. The let-7 family of miRNA have also been studied as treatment for non-small cell lung cancer in the pre-clinical setting[11]. Loss of miR-300 has been previously correlated with development and aggressiveness of hepatocellular carcinoma[47]. Loss of miR-300 has also been implicated in oncogenesis of pituitary tumors[48]. Reduced expression of miR-424-5p and miR-4282 have each been implicated in the development of basal-like breast cancer[49,50]. miR-15a-5p has been shown to have anti-melanoma activity[51]. In addition, miR-590-3p has been show to suppress proliferation of both breast cancer[52], and hepatocellular carcinoma[53]. The broad literature linking many of our proposed miRNA candidates for Ewing sarcoma treatment to the development and maintenance of cancer highlights the ability of our computational pipeline to identify potentially promising therapeutic candidates.

In addition, we used mRNA concentration as a surrogate for protein concentration in designing our cell signaling network. While this is not true in all cases, it is likely a reasonable approximation under steady state conditions[26–29] (see Section 1.2 for more detail).

Despite these limitations, our pipeline can serve as a valuable tool for the identification of key proteins in a cancer cell

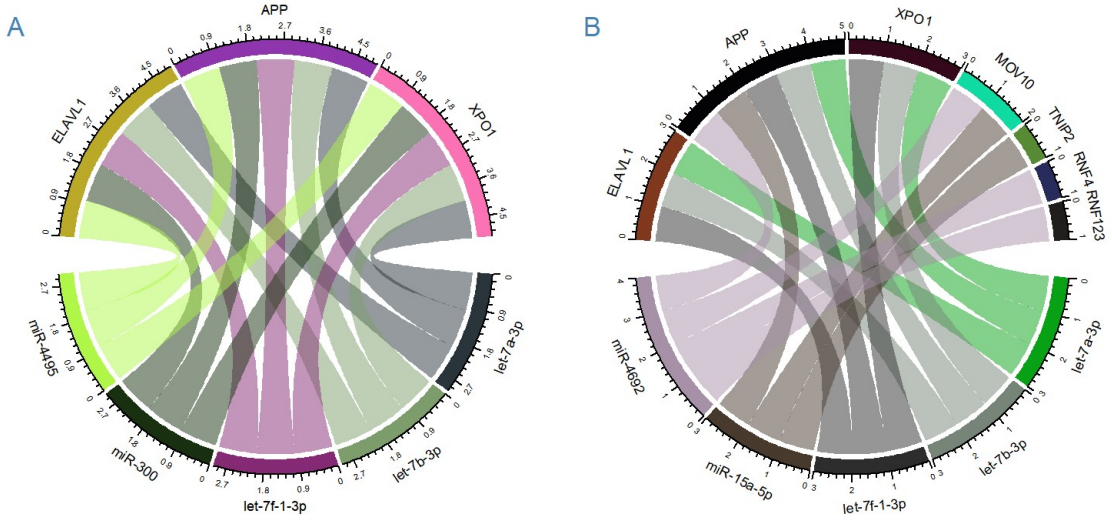

Figure 6. Sample miRNA Combination Therapy. We identified multiple miRNA that share important protein targets. In sample combination A, 5 miRNA (miR-4495, mir-300, let-7f-1-3p, let-7b-3p and let-7a-3p) each independently target 3 of the 5 most important protein targets (ELAV1, APP, and XPO1). In sample combination B, 5 miRNA (miR-4692, miR-15a-5p, let-7f-1-3p, let-7b-3p and let-7a-3p) each repress at least 3 of the 10 most important protein targets.

signaling network and for the identification of miRNA that target these proteins. We also present a rudimentary approach for identification of optimal miRNA combinations that target key proteins in the cancer cell signaling network. Future work will focus on the development of programs that can identify miRNA combinations that both target key cancer cell signaling proteins and limit off-target toxicity. This novel method can facilitate the rapid identification of key proteins in any cancer cell signaling network for which mRNA sequencing data is available. This may facilitate more rapid drug discovery and assist in the discovery of proteins and miRNA that are significant in cancer.

Acknowledgments

We acknowledge experimental support from Julia Selich-Anderson. JGS would like to thank the NIH Loan Repayment Program for their generous support and the Paul Calabresi Career Development Award for Clinical Oncology (NIH K12CA076917). Kathleen I. Pishas acknowledges financial support from the NHMRC CJ Martin Overseas Biomedical Fellowship (APP1111032) and Alex's Lemonade Stand Young Investigator Grant (APP37138). Davis T. Weaver acknowledges financial support from the Cancer Center Trainee Award for Cancer Research from the Case Comprehensive Cancer Center

Author contributions

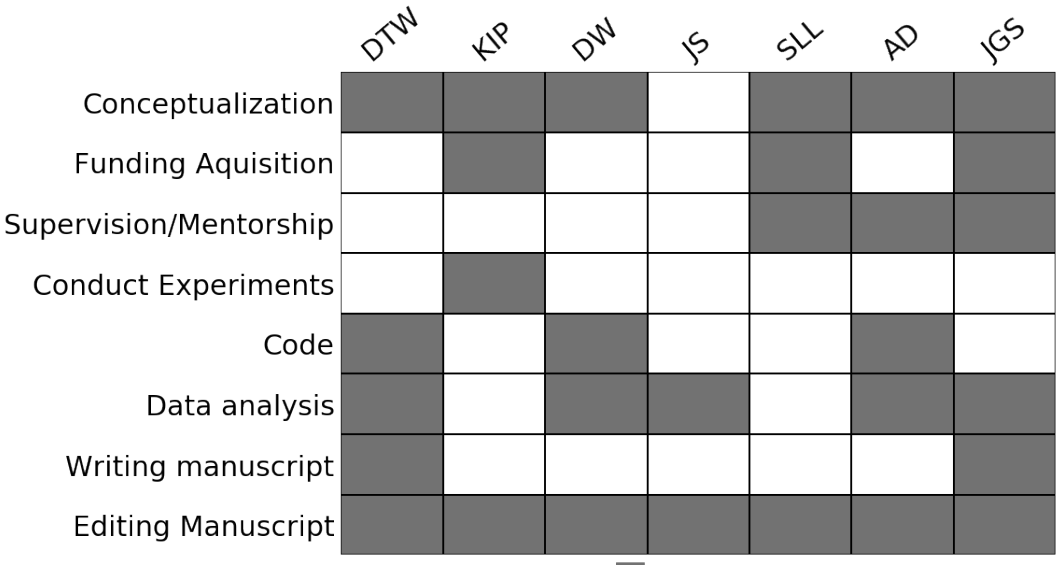

Figure 7. **Grey denotes contribution**

References

1. Rupaimoole, R. & Slack, F. J. MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. *Nat. Rev. Drug Discov.* **16**, 203–222, DOI: 10.1038/nrd.2016.246 (2017).

2. Riggi, N. *et al.* EWS-FLI-1 modulates miRNA145 and SOX2 expression to initiate mesenchymal stem cell reprogramming toward Ewing sarcoma cancer stem cells. *Genes & development* **24**, 916–32, DOI: 10.1101/gad.1899710 (2010).

3. Dhawan, A., Scott, J. G., Harris, A. L. & Buffa, F. M. Pan-cancer characterisation of microRNA across cancer hallmarks reveals microRNA-mediated downregulation of tumour suppressors. *Nat. Commun.* **9**, 5228, DOI: 10.1038/s41467-018-07657-1 (2018).

4. LIN, X., YANG, Z., ZHANG, P. & SHAO, G. miR-154 suppresses non-small cell lung cancer growth in vitro and in vivo. *Oncol. Reports* **33**, 3053–3060, DOI: 10.3892/or.2015.3895 (2015).

5. Lin, K. *et al.* Loss of MIR15A and MIR16-1 at 13q14 is associated with increased TP53 mRNA, de-repression of BCL2 and adverse outcome in chronic lymphocytic leukaemia. *Br. J. Haematol.* **167**, 346–355, DOI: 10.1111/bjh.13043 (2014).

6. Reilley, M. J. *et al.* STAT3 antisense oligonucleotide AZD9150 in a subset of patients with heavily pretreated lymphoma: results of a phase 1b trial. *J. for ImmunoTherapy Cancer* **6**, 119, DOI: 10.1186/s40425-018-0436-5 (2018).

7. Hong, D. *et al.* AZD9150, a next-generation antisense oligonucleotide inhibitor of STAT3 with early evidence of clinical activity in lymphoma and lung cancer. *Sci. Transl. Medicine* **7**, 185–314, DOI: 10.1126/scitranslmed.aac5272 (2015).

8. Odate, S. *et al.* Inhibition of STAT3 with the Generation 2.5 Antisense Oligonucleotide, AZD9150, Decreases Neuroblastoma Tumorigenicity and Increases Chemosensitivity. *Clin. Cancer Res.* **23**, 1771–1784, DOI: 10.1158/1078-0432.CCR-16-1317 (2017).

9. Kasinski, A. L. & Slack, F. J. miRNA-34 prevents cancer initiation and progression in a therapeutically resistant K-ras and p53-induced mouse model of lung adenocarcinoma. *Cancer research* **72**, 5576–87, DOI: 10.1158/0008-5472.CAN-12-2001 (2012).

10. Wiggins, J. F. *et al.* Development of a Lung Cancer Therapeutic Based on the Tumor Suppressor MicroRNA-34. *Cancer Res.* **70**, 5923–5930, DOI: 10.1158/0008-5472.CAN-10-0655 (2010).

11. Stahlhut, C. & Slack, F. J. Combinatorial Action of MicroRNAs let-7 and miR-34 Effectively Synergizes with Erlotinib to Suppress Non-small Cell Lung Cancer Cell Proliferation. *Cell Cycle* **14**, 2171–2180, DOI: 10.1080/15384101.2014.1003008 (2015).

12. Liu, C. *et al.* The microRNA miR-34a inhibits prostate cancer stem cells and metastasis by directly repressing CD44. *Nat. medicine* **17**, 211–5, DOI: 10.1038/nm.2284 (2011).

13. Maciej Pajak, T. I. S. miRNAtap: miRNAtap: microRNA Targets— Aggregated Predictions., DOI: 10.18129 (2019).

14. Wong, N. & Wang, X. miRDB: an online resource for microRNA target prediction and functional annotations. *Nucleic Acids Res.* **43**, D146–D152, DOI: 10.1093/nar/gku1104 (2015).

15. Maragkakis, M. *et al.* DIANA-microT web server: elucidating microRNA functions through target prediction. *Nucleic Acids Res.* **37**, W273–W276, DOI: 10.1093/nar/gkp292 (2009).

16. O'Neill, V. A Multicenter Phase I Study of MRX34, MicroRNA miR-RX34 Liposomal Injection - Full Text View - ClinicalTrials.gov, DOI: NCT01829971.

17. Rietman, E. A., Scott, J. G., Tuszynski, J. A. & Klement, G. L. Personalized anticancer therapy selection using molecular landscape topology and thermodynamics. *Oncotarget* **8**, 18735–18745, DOI: 10.18632/oncotarget.12932 (2017).

18. Rietman, E. A., Platig, J., Tuszynski, J. A. & Lakka Klement, G. Thermodynamic measures of cancer: Gibbs free energy and entropy of protein-protein interactions. *J. biological physics* **42**, 339–50, DOI: 10.1007/s10867-016-9410-y (2016).

19. Rietman, E. & Tuszynski, J. A. Using Thermodynamic Functions as an Organizing Principle in Cancer Biology. 139–157, DOI: 10.1007/978-3-319-74974-7{_}8 (Springer, Cham, 2018).

20. Mackintosh, C., Madoz-Gúrpide, J., Ordóñez, J. L., Osuna, D. & Herrero-Martín, D. The molecular pathogenesis of Ewing's sarcoma. *Cancer biology & therapy* **9**, 655–67 (2010).

21. Esiashvili, N., Goodman, M. & Marcus, R. B. Changes in Incidence and Survival of Ewing Sarcoma Patients Over the Past 3 Decades. *J. Pediatr. Hematol.* **30**, 425–430, DOI: 10.1097/MPH.0b013e31816e22f3 (2008).

22. Pishas, K. I. *et al.* Therapeutic targeting of KDM1A/LSD1 in ewing sarcoma with SP-2509 engages the endoplasmic reticulum stress response. *Mol. Cancer Ther.* **17**, 1902–1916, DOI: 10.1158/1535-7163.MCT-18-0373 (2018).

23. Oughtred, R. *et al.* The BioGRID interaction database: 2019 update. *Nucleic Acids Res.* **47**, D529–D541, DOI: 10.1093/nar/gky1079 (2019).

24. Zwiener, I., Frisch, B. & Binder, H. Transforming RNA-Seq Data to Improve the Performance of Prognostic Gene Signatures. *PLoS ONE* **9**, e85150, DOI: 10.1371/journal.pone.0085150 (2014).

25. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550, DOI: 10.1186/s13059-014-0550-8 (2014).

26. Liu, Y., Beyer, A. & Aebersold, R. On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell* **165**, 535–50, DOI: 10.1016/j.cell.2016.03.014 (2016).

27. Vogel, C. & Marcotte, E. M. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* **13**, 227–232, DOI: 10.1038/nrg3185 (2012).

28. Li, J. J., Bickel, P. J. & Biggin, M. D. System wide analyses have underestimated protein abundances and the importance of transcription in mammals. *PeerJ* **2**, e270, DOI: 10.7717/peerj.270 (2014).

29. Jovanovic, M. *et al.* Dynamic profiling of the protein life cycle in response to pathogens. *Science* **347**, 1259038–1259038, DOI: 10.1126/science.1259038 (2015).

30. Albert, R., Jeong, H. & Barabási, A. L. Error and attack tolerance of complex networks. *Nature* **406**, 378–382, DOI: 10.1038/35019019 (2000).

31. Dhawan, A. *et al.* Collateral sensitivity networks reveal evolutionary instability and novel treatment strategies in ALK mutated non-small cell lung cancer. *Sci. Reports* **7**, 1232, DOI: 10.1038/s41598-017-00791-8 (2017).

32. Enright, A. J. *et al.* MicroRNA targets in Drosophila. *Genome Biol.* **5**, R1, DOI: 10.1186/gb-2003-5-1-r1 (2003).

33. Lall, S. *et al.* A Genome-Wide Map of Conserved MicroRNA Targets in C. elegans. *Curr. Biol.* **16**, 460–471, DOI: 10.1016/j.cub.2006.01.050 (2006).

34. Friedman, R. C., Farh, K. K.-H., Burge, C. B. & Bartel, D. P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **19**, 92–105, DOI: 10.1101/gr.082701.108 (2008).

35. Tate, J. G. *et al.* COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* **47**, D941–D947, DOI: 10.1093/nar/gky1015 (2019).

36. Stelzer, G. *et al.* The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. In *Current Protocols in Bioinformatics*, vol. 54, 1–1, DOI: 10.1002/cpbi.5 (John Wiley & Sons, Inc., Hoboken, NJ, USA, 2016).

37. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740, DOI: 10.1093/bioinformatics/btr260 (2011).

38. Liberzon, A. *et al.* The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Syst.* **1**, 417–425, DOI: 10.1016/j.cels.2015.12.004 (2015).

39. Korotkevich, G., Sukhov, V. & Sergushichev, A. A. An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. *bioRxiv* 060012, DOI: 10.1101/060012 (2016).

40. Benjamini, Y., Yoav ; Hochberg. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B-Methodological 1995.pdf. *J. Royal Stat. Soc. Ser. B (Methological)* **57**, 289–300, DOI: 10.2307/2346101 (1995).

41. Molinelli, E. J. *et al.* Perturbation Biology: Inferring Signaling Networks in Cellular Systems. *PLoS Comput. Biol.* **9**, DOI: 10.1371/journal.pcbi.1003290 (2013).

42. Korkut, A. *et al.* Perturbation biology nominates upstream–downstream drug combinations in RAF inhibitor resistant melanoma cells. *eLife* **4**, DOI: 10.7554/eLife.04640 (2015).

43. Hameiri-Grossman, M. *et al.* The association between let-7, RAS and HIF-1 in Ewing Sarcoma tumor growth. *Oncotarget* **6**, 33834–48, DOI: 10.18632/oncotarget.5616 (2015).

44. Guo, M., Zhao, X., Yuan, X., Jiang, J. & Li, P. MiR-let-7a inhibits cell proliferation, migration, and invasion by down-regulating PKM2 in cervical cancer. *Oncotarget* **8**, 28226–28236, DOI: 10.18632/oncotarget.15999 (2017).

45. Tang, R. *et al.* MiR-let-7a inhibits cell proliferation, migration, and invasion by down-regulating PKM2 in gastric cancer. *Oncotarget* **7**, 5972–84, DOI: 10.18632/oncotarget.6821 (2016).

46. Pan, X.-M. *et al.* A let-7 KRAS rs712 polymorphism increases colorectal cancer risk. *Tumor Biol.* **35**, 831–835, DOI: 10.1007/s13277-013-1114-3 (2014).

47. Wang, R. *et al.* miR-300 regulates the epithelial-mesenchymal transition and invasion of hepatocellular carcinoma by targeting the FAK/PI3K/AKT signaling pathway. *Biomed. & Pharmacother.* **103**, 1632–1642, DOI: 10.1016/J.BIOPHA.2018.03.005 (2018).

48. Liang, H.-q. *et al.* The PTTG1-targeting miRNAs miR-329, miR-300, miR-381, and miR-655 inhibit pituitary tumor cell tumorigenesis and are involved in a p53/PTTG1 regulation feedback loop. *Oncotarget* **6**, 29413–27, DOI: 10.18632/oncotarget.5003 (2015).

49. Wang, J., Wang, S., Zhou, J. & Qian, Q. miR-424-5p regulates cell proliferation, migration and invasion by targeting doublecortin-like kinase 1 in basal-like breast cancer. *Biomed. & Pharmacother.* **102**, 147–152, DOI: 10.1016/J.BIOPHA.2018.03.018 (2018).

50. J. Zhao, G.-Q. J. MiR-4282 inhibits proliferation, invasion and metastasis of human breast cancer by targeting Myc. *Eur. review for medical pharmacological sciences* DOI: 10.26355/eurrev{_}201812{_}16643 (2018).

51. Alderman, C. & Yang, Y. The anti-melanoma activity and oncogenic targets of hsa-miR-15a-5p. *RNA & disease (Houston, Tex.)* **3** (2016).

52. Rohini, M., Gokulnath, M., Miranda, P. & Selvamurugan, N. miR-590–3p inhibits proliferation and promotes apoptosis by targeting activating transcription factor 3 in human breast cancer cells. *Biochimie* **154**, 10–18, DOI: 10.1016/J.BIOCHI.2018.07.023 (2018).

53. Ge, X. & Gong, L. MiR-590-3p suppresses hepatocellular carcinoma growth by targeting TEAD1. *Tumor Biol.* **39**, 101042831769594, DOI: 10.1177/1010428317695947 (2017).

Supplemental Materials

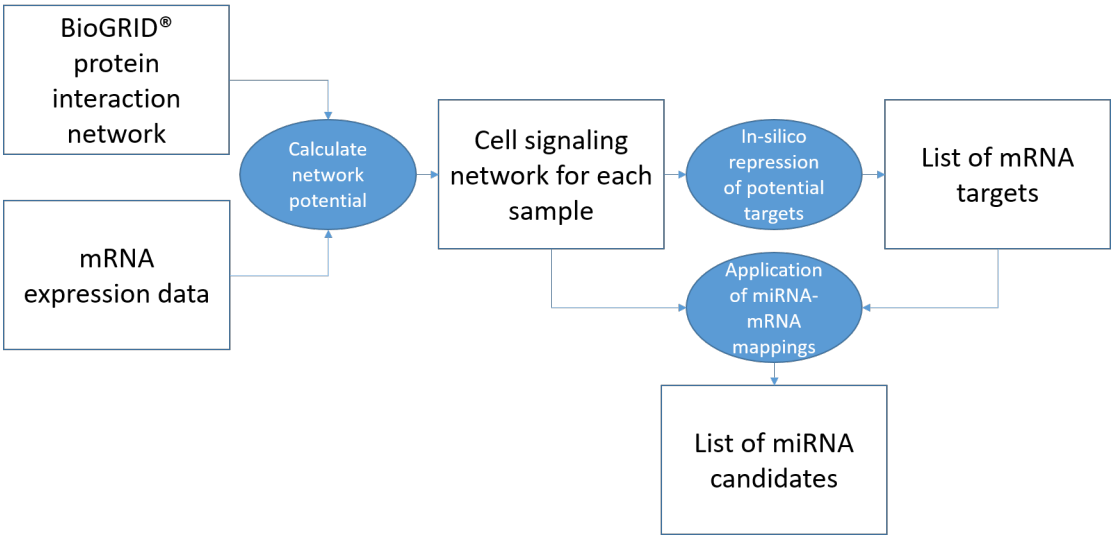

Figure S1. Simplified schematic of our computational pipeline. Squares represent data, ellipses represent computational steps, https://github.com/DavisWeaver/MiR$_{Combo_Targeting}$/

3.1 Additional Analyses

As described in the main text, we ranked proteins according to their contribution to network stability by calculating the change in network potential following complete *in silico* repression of each protein. In the main text, we limited our analysis to proteins that had been causally implicated in cancer according to the cosmic database[35]. Here, we present the top 50 proteins (when network potential for all 6 cell lines was averaged) ranked by contribution to network stability, not limited to proteins that were causally implicated in cancer (Figure S2).

We also analyzed each cell line individually to identify the top protein targets for each cell line. In the main text, we limited this analysis to proteins that had been causally implicated in cancer[35]. Here, we present the top protein targets for each cell line, not limited to those proteins that had previously been causally implicated in cancer (Table S1). 5t

	TC252	ES2	A673	TC32	EWS502	TC71
1	TRIM25	TRIM25	TRIM25	TRIM25	TRIM25	TRIM25
2	APP	APP	APP	APP	APP	APP
3	ELAVL1	ELAVL1	ELAVL1	ELAVL1	ELAVL1	ELAVL1
4	RNF4	RNF4	RNF4	RNF4	RNF4	RNF4
5	XPO1	XPO1	XPO1	XPO1	XPO1	XPO1
6	NXF1	NXF1	NXF1	NXF1	NXF1	NXF1
7	UBC	TNIP2	UBC	UBC	UBC	UBC
8	TNIP2	UBC	TNIP2	TNIP2	TNIP2	TNIP2
9	MOV10	MOV10	MOV10	MOV10	MOV10	MOV10
10	RNF123	RNF123	RNF123	LMNA	RNF123	RNF123

Table S1. Top protein targets for each cell line. We ranked potential targets by predicted change in network potential when each protein was modeled as repressed.

Gene set enrichment analysis We conducted gene set enrichment analysis, using all genes in our Ewing sarcoma cell signaling network as the gene set. We ranked this set of genes by network potential and used the "hallmarks" pathways set from the Molecular Signatures Database as the genomic background[37,38]. We used the "fgsea" R package version 1.8.0 to conduct the analysis using the following settings: nperm $= 100$, minSize $= 1$, maxSize $= \infty$, nproc $= 0$, gseaParam $= 1$, BPPARAM =NULL[39]. We found our gene set to be significantly enriched in several pathways related to oncogenesis, including DNA repair, apoptosis, and MTOR signaling (Table S2). These results indicate that network potential can identify a cancer-specific signal from mRNA expression data.

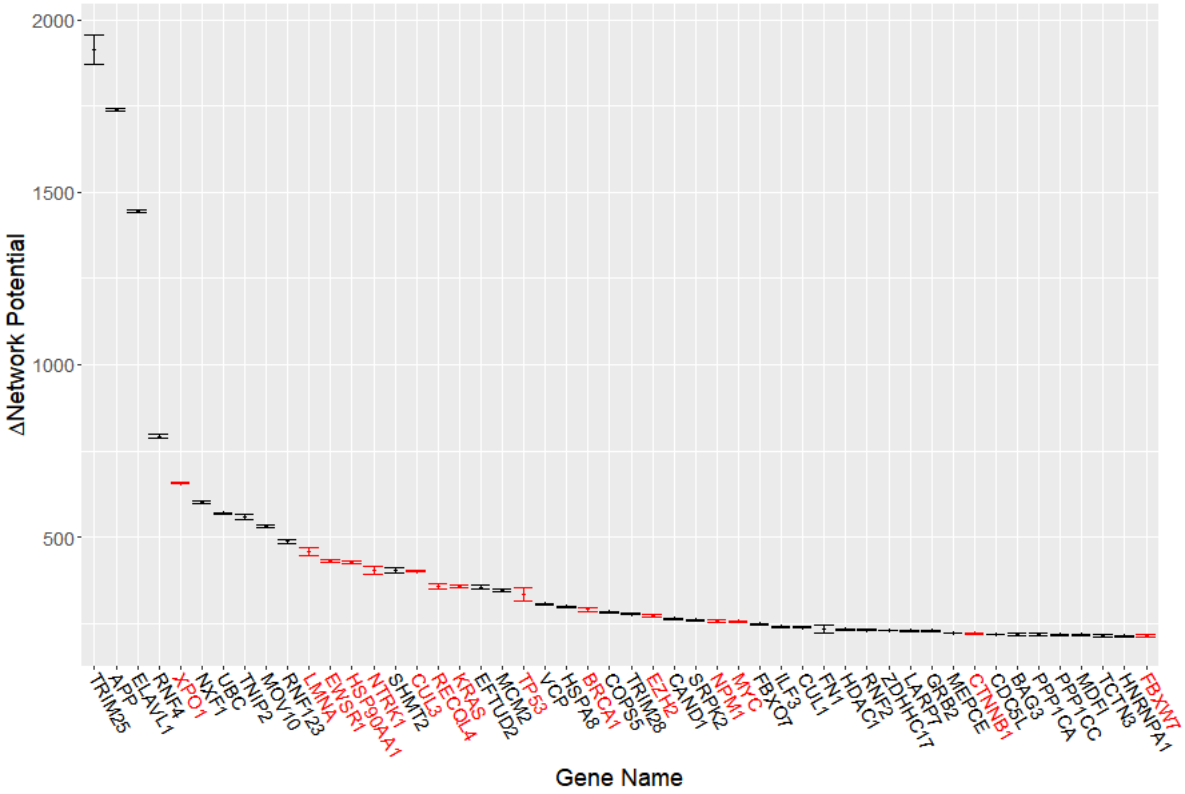

Figure S2. Protein targets ranked by contribution to network stability. When averaging across cell lines, TRIM25, APP, ELAV1, RNF4, and XPO1 were identified as the most important proteins in the Ewing sarcoma cell signaling network. When each protein was simulated as completely repressed, network potential was increased by 1,914, 1,739, 1,445, 792, and 656, respectively. It is notable that EWSR1, the kinase associated with Ewing sarcoma development, is considered highly influential in the cell signaling network by this method. Genes that have previously been causally implicated in cancer according to the Cosmic database are highlighted in red [35].

As described in the main text, miR-3613-3p was the top miR for predicted repression of the Ewing sarcoma cell signaling network across all 6 cell lines. Here we present the full list of the 144 distinct mRNA miR-3613-3p is predicted to repress among our full list of protein targets (Table S3). The full list of protein targets was generated by ranking all proteins in the cell signaling network according to network potential and taking the top 5% of all proteins.

Table S3. List of targets for miR-3613-3p

	miR-3613-3p targets
1	ABCE1
2	ABL1
3	ACTB
4	AHNAK
5	ATP6V1A
6	BARD1
7	CALCOCO2
8	CALM1
9	CAPRIN1
10	CAPZA1
11	CBX3

Continued on next page

Table S3 – continued from previous page

	miR-3613-3p targets
12	CBX5
13	CCAR2
14	CDC27
15	CDC42
16	CDC5L
17	CDC73
18	CEP44
19	CPSF6
20	CREBBP
21	CSE1L
22	CSNK1A1
23	CTNNB1
24	CTTN
25	CUL3
26	DAZAP1
27	DCUN1D1
28	DDX6
29	EIF3A
30	ELAVL1
31	EP300
32	EZR
33	FBXW7
34	FGFR1
35	FUBP1
36	G3BP1
37	G3BP2
38	GFPT1
39	GOLT1B
40	GSK3B
41	HDAC3
42	HIF1A
43	HNRNPA3
44	HNRNPF
45	HNRNPH1
46	HNRNPU
47	HP1BP3
48	HUWE1
49	IPO7
50	KIF5B
51	KRAS
52	LARP1
53	LMNB1
54	MAP3K7
55	MCM4
56	MDM2
57	MTCH2
58	NAP1L1
59	NCBP1
60	NDUFA4
61	NEDD1
62	NF2
63	NOTCH1
Continued on next page	

Table S3 – continued from previous page

	miR-3613-3p targets
64	NR3C1
65	NUDT21
66	PDIA3
67	PIK3R1
68	POLR2A
69	PPP1CB
70	PRKAA1
71	PRKAR1A
72	PRPF40A
73	PTGES3
74	PTK2
75	RAD23B
76	RB1
77	RBMX
78	RBPJ
79	RC3H1
80	RDX
81	REST
82	RNF2
83	RPL15
84	SERBP1
85	SET
86	SIN3A
87	SKP2
88	SMAD2
89	SMAD4
90	SMAD9
91	SMARCAD1
92	SOAT1
93	SP1
94	SPAG9
95	SPTBN1
96	SREK1
97	SRPK1
98	SRPK2
99	SRRM1
100	SRSF2
101	SRSF5
102	STAU1
103	SUMO1
104	SUMO2
105	SYNCRIP
106	TBL1XR1
107	TCEA1
108	TCF3
109	TCF4
110	TES
111	TOP1
112	TUBB
113	UBA52
114	UBE2D3
115	USP47
Continued on next page	

Table S3 – continued from previous page

	miR-3613-3p targets
116	VAPA
117	XPO1
118	YAP1
119	YWHAH
120	YWHAZ
121	ZDHHC17
122	ZWINT
123	TERF2
124	CCND2
125	IGF2BP3
126	KHSRP
127	RICTOR
128	DPYSL2
129	GNB1
130	KMT2A
131	LGR4
132	SSR3
133	TXLNA
134	SMC2
135	TSNAX
136	IGF1R
137	LAMP2
138	CCNT1
139	ATXN1
140	CIRBP
141	CLINT1
142	KPNA1
143	XRN1
144	GOPC

361

	pathway	pval	padj	ES	NES	size
1	MITOTIC_SPINDLE	0.01	0.03	0.61	1.59	197
2	TGF_BETA_SIGNALING	0.01	0.03	0.61	1.49	53
3	DNA_REPAIR	0.01	0.03	0.58	1.47	146
4	G2M_CHECKPOINT	0.01	0.03	0.70	1.81	191
5	APOPTOSIS	0.01	0.03	0.51	1.31	159
6	ADIPOGENESIS	0.01	0.03	0.47	1.22	191
7	ANDROGEN_RESPONSE	0.01	0.03	0.53	1.34	96
8	PROTEIN_SECRETION	0.01	0.03	0.61	1.55	93
9	UNFOLDED_PROTEIN_RESPONSE	0.01	0.03	0.61	1.56	109
10	PI3K_AKT_MTOR_SIGNALING	0.01	0.03	0.59	1.50	104
11	MTORC1_SIGNALING	0.01	0.03	0.58	1.52	196
12	E2F_TARGETS	0.01	0.03	0.70	1.83	197
13	MYC_TARGETS_V1	0.01	0.03	0.81	2.09	197
14	MYC_TARGETS_V2	0.01	0.03	0.63	1.56	58
15	FATTY_ACID_METABOLISM	0.01	0.03	0.47	1.21	155
16	OXIDATIVE_PHOSPHORYLATION	0.01	0.03	0.59	1.54	184
17	ALLOGRAFT_REJECTION	0.01	0.03	0.48	1.25	195
18	WNT_BETA_CATENIN_SIGNALING	0.02	0.05	0.55	1.31	42
19	GLYCOLYSIS	0.02	0.05	0.46	1.20	193
20	UV_RESPONSE_UP	0.02	0.05	0.47	1.20	155
21	HEME_METABOLISM	0.02	0.05	0.45	1.15	190
22	SPERMATOGENESIS	0.02	0.05	0.48	1.21	126

Table S2. Genes ranked by network potential are enriched for several biological pathways related to cancer Pathways with an adjusted p-value < 0.05 are shown above. "ES" refers to enrichment score and "NES" refers to the normalized enrichment score. Size refers to the number of genes in the pathway that were also present in our mRNA expression dataset.