In-silico identification of Prognostically Inversely Correlated miRNAs and mRNAs (PIC's) in multiple cancers

Chirayu Pankaj Goswami & Zi-Xuan Wang

Molecular and Genomic Pathology, Thomas Jefferson University Hospital, Philadelphia, USA

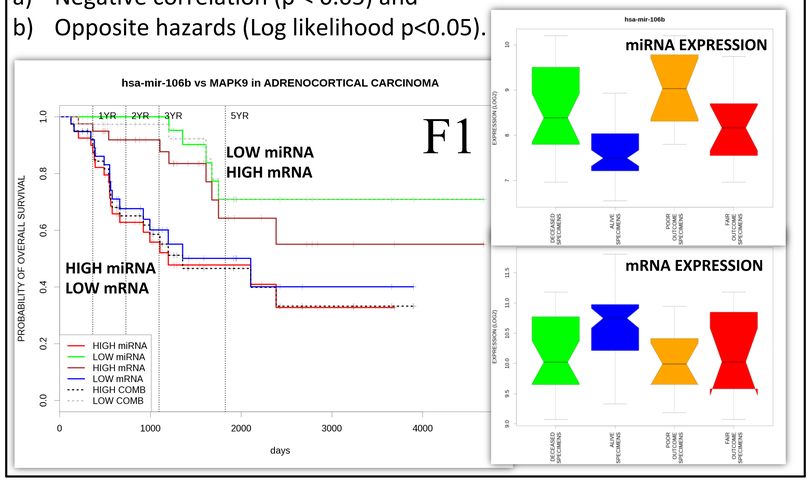
ABSTRACT

Despite numerous methods available to identify potential mRNA targets for miRNAs, prognostic relationship of these molecules in diseases like cancers where deregulation of gene expression is a major pathogenic factor, has not yet been emphasized. We performed in-silico identification of prognostically inversely correlated miRNA - mRNA pairs (PIC's) in multiple cancers using expression data from The Cancer Genome Atlas. Partners in a PIC show inverse correlation of expression and opposite hazard implication. Using a three step approach, we identified a total of 1,253,443 PIC's from 23 cancer types, several of which have previously been shown to have a predicted or experimentally validated relationship. A maximum 375,621 PICs were identified in Lower Grade Gliomas, while a minimum 300 PICs were identified in Prostate adenocarcinoma. Four miRNA-mRNA pairs were identified as PICs in 7 different cancer types. Two miRNA-mRNA pairs were identified as PICs in 5 different cancer types where the mRNA is also a validated target of miRNA. Organ specific analysis was performed to identify PICs common to cancers from same or related tissue of origin. We have also developed a database PROGTar for hosting our analysis results. PROGTar is available freely for non-commercial use at www.xvm145.jefferson.edu/progtar. We believe our method and analysis results will provide a novel prognostically relevant, pan-cancer perspective to study of miRNA-mRNA interactions and miRNA target validation.

METHODOLOGY

Detailed methodology for identification of PICs is provided in Poster 18 "PROGTar: A database of prognostically inversely correlated miRNA and Genes in multiple cancers". Briefly, For each cancer type, pair-wise correlations of miRNA and mRNA expressions were performed using sequencing data. Cox proportional hazard analysis was performed separately for miRNAs and mRNAs using clinical survival related data and expression data. Correlation and Proportional hazard results were then merged to identify miRNA-mRNA pairs which showed

a) Negative correlation ($p < 0.05$) and
b) Opposite hazards (Log likelihood $p<0.05$).



Pairs were then ranked and annotated. We used TargetScan and miRWALK databases to annotate the pair with predicted/validated target information.

RESULTS

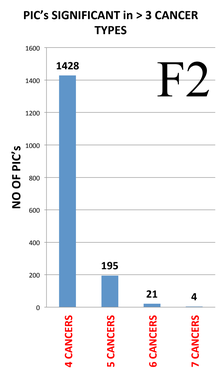
Table 1 shows the cancer types that were included in our analysis along with # of specimens available, # of PICs identified, # of unique miRNA and mRNA identifies, and # of PICs where partners have a predicted or/and validated relationship per cancer type. PICs show inverse correlation of expression between partners and opposite hazards. An example PIC with corresponding expression boxplots is shown in Figure F1. Several PICs identified in our analysis were prognostically relevant in more than one cancer type. Figure F2 shows no of PICs that were identified as significant in more than 3 cancer types.

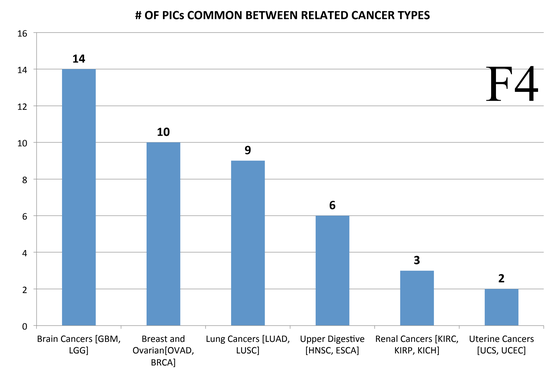
We used TargetScan and MirWALK databases to identify how many PICs identified in our analysis have previously been shown to have predicted or/ and validated relationship between partner miRNA and mRNA. Figure F3 shows for each cancer, percentage of total PICs identified for the cancer type where partners have a predicted and/or validated relationship. A maximum of 5.38% of total PICs in COAD had mRNA as predicted target for miRNA in the pair, while a maximum of 1.23% PICs in CESC had mRNA as validated target for mRNA in the pair.

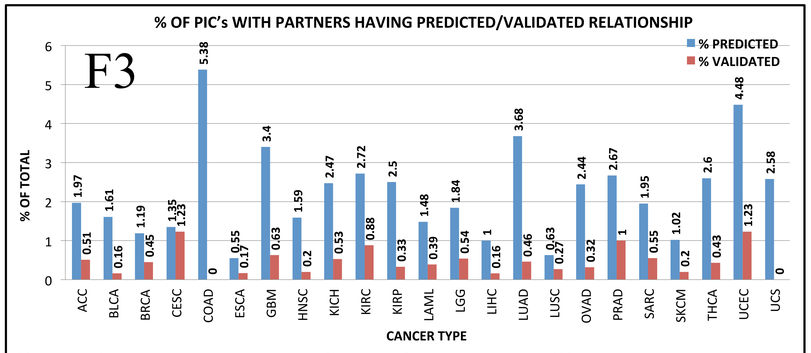
Table 1: Cancer types analyzed

	CANCER	# SAMPLES	# PIC's	# miRNA	# mRNA	P*	V*	P+V
1	Adrenocortical carcinoma [ACC]	79	136463	143	5069	2686	695	140
2	Bladder Urothelial Carcinoma [BLCA]	317	31466	77	2336	506	49	6
3	Breast invasive carcinoma [BRCA]	656	137788	100	5168	1636	619	64
4	Cervical squamous cell carcinoma and endocervical adenocarcinoma[CESC]	281	17595	92	2329	238	217	15
5	Colon adenocarcinoma [COAD]	224	502	11	259	27	0	0
6	Esophageal carcinoma [ESCA]	180	2378	42	494	13	4	1
7	Glioblastoma multiforme [GBM]	534	46068	57	3482	1569	287	86
8	Head and Neck squamous cell carcinoma [HNSC]	448	24935	76	2269	396	49	5
9	Kidney Chromophobe [KICH]	66	37525	93	2494	925	198	35
10	Kidney renal clear cell carcinoma [KIRC]	234	140965	108	6888	3830	1240	219
11	Kidney renal papillary cell carcinoma [KIRP]	220	65222	87	4028	1630	213	66
12	Acute Myeloid Leukemia [LAML]	84	18774	46	2860	278	74	15
13	Brain Lower Grade Glioma [LGG]	416	375621	182	8169	6903	2010	208
14	Liver hepatocellular carcinoma [LIHC]	293	7528	56	1399	75	12	3
15	Lung adenocarcinoma [LUAD]	371	50828	68	3675	1869	234	63
16	Lung squamous cell carcinoma [LUSC]	283	2557	45	630	16	7	0
17	Ovarian serous cystadenocarcinoma [OVAD]	300	3805	27	1138	93	12	3
18	Prostate Adenocarcinoma [PRAD]	413	300	8	218	8	3	1
19	Sarcoma [SARC]	242	54862	98	2639	1072	301	57
20	Skin Cutaneous Melanoma [SKCM]	373	26203	59	3981	268	53	10
21	Thyroid carcinoma [THCA]	476	16669	71	1577	434	72	20
22	Uterine Corpus Endometrial Carcinoma [UCEC]	143	2278	21	561	102	28	6
23	Uterine Carcinosarcoma [UCS]	55	969	27	409	25	0	0

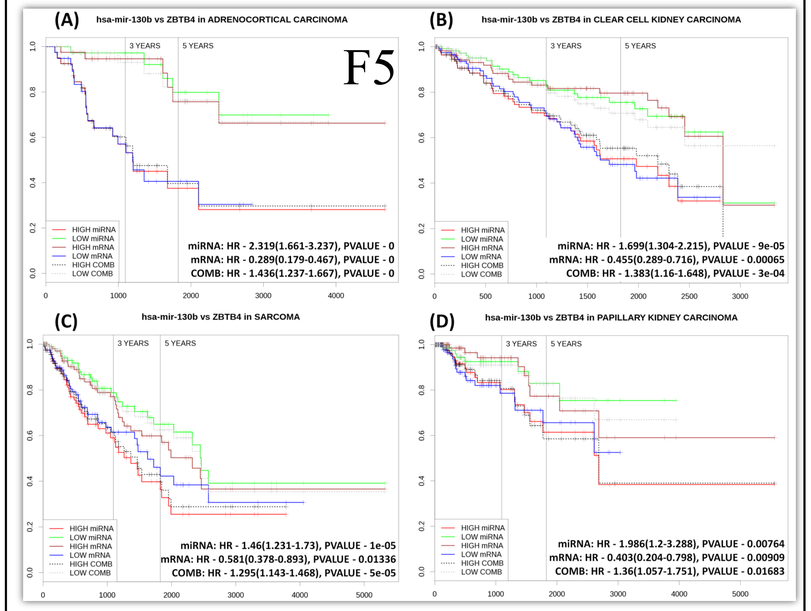
* P = PREDICTED, V = VALIDATED







Since miRNAs are expressed in heavily organ specific manner, we looked into PICs common between cancers of same/related tissues of origin. We found that several PICs were common in cancers from same organ/related cancers (See Figure F4). Figure F5 shows an example PIC which is significant in 4 different cancer types where mRNA (ZBTB4) in the partner is a predicted as well as validated target of the miRNA (mir-130b).



CONCLUSIONS

We have performed in-silico identification of Prognostically relavant interplay of miRNAs and mRNAs in multiple cancers. We have also created a web application to host results generated from our analysis. We believe the web application will be useful for the scientific community in pursuing molecules of their interest, or for identifying novel targets for hypothesis generation studies.

We also believe that our analysis results will provide a new direction to validation of miRNA target information. Results generated from our analysis may facilitate generation of novel prognostic and predictive signatures and miRNA directed therapeutics.