JCB*i* JOURNAL OF CLINICAL BIOINFORMATICS

DATABASE   **Open Access**

PROGmiR: a tool for identifying prognostic miRNA biomarkers in multiple cancers using publicly available data

Chirayu Pankaj Goswami[1*] and Harikrishna Nakshatri[2]

Abstract

Background: Identification of prognostic biomarkers is hallmark of cancer genomics. Since miRNAs regulate expression of multiple genes, they act as potent biomarkers in several cancers. Identification of miRNAs that are prognostically important has been done sporadically, but no resource is available till date that allows users to study prognostics of miRNAs of interest, utilizing the wealth of available data, in major cancer types.

Description: In this paper, we present a web based tool that allows users to study prognostic properties of miRNAs in several cancer types, using publicly available data. We have compiled data from Gene Expression Omnibus (GEO), and recently developed "The Cancer Genome Atlas (TCGA)", to create this tool. The tool is called "PROGmiR" and it is available at www.compbio.iupui.edu/progmir. Currently, our tool can be used to study overall survival implications for approximately 1050 human miRNAs in 16 major cancer types.

Conclusions: We believe this resource, as a hypothesis generation tool, will be helpful for researchers to link miRNA expression with cancer outcome and to design mechanistic studies. We studied performance of our tool using identified miRNA biomarkers from published studies. The prognostic plots created using our tool for specific miRNAs in specific cancer types corroborated with the findings in the studies.

Keywords: miRNA, Prognostics, Cancer, Pan-cancer, Database, Signature, Biomarker

Background

MiRNAs are small non coding RNA molecules, 18-25 bases at maturity, and exert a profound effect on regulation of genetic machinery of the cell. MiRNAs are found in both plant and animal cells and act as key regulators of gene expression. The first miRNA, a small transcript of lin-4 gene, was discovered as an antisense molecule to lin-14 mRNA in C. elegans [1]. Lin-14 gene synthesizes lin-14 protein, lower levels of which are essential for normal embryonic development of C. elegans larvae. The first miRNA to be categorically studied was 21 nucleotide long let-7 miRNA [2]. MiRNA were subsequently also discovered in plants [3]. MiRNA are found in almost every species ranging from unicellular organisms such as yeast to primates, but the number of miRNAs in different species ranges from a few dozens to thousands.

Since miRNAs act as key regulators of gene expression, their role is evident in several diseases where gene expression is altered. For this reason, despite their recent discovery, miRNAs have been studied to a great extent in cancers, where pathological changes are bought about primarily due to altered gene expression. Recent studies have suggested a role of miRNAs in many cancers. MiRNAs have so far been reported to be up or down regulated in several types of cancers [4]. In cancer, miRNAs may act as tumor suppressors or tumor promoters (oncogenes). MiR17-92 cluster is a prominent oncogenic miRNA cluster [5], whereas miR-34 family, which is induced by the tumor suppressor gene p53 [6], displays tumor suppressor activity in multiple cancers. MiRNAs have been reported to play these roles in both cancers of hematopoietic origin as well as in solid tumors. Also, several miRNAs have been found to be directly associated with tumor progression and metastasis [7-10].

The first miRNA to be associated with cancer was discovered in Chronic Lymphocytic Leukemia (CLL). Croce

* Correspondence: cgoswami@iupui.edu
[1]Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, 410 W 10th Street, Indianapolis, IN 46202, USA
Full list of author information is available at the end of the article

et al discovered a novel miRNA signature associated with prognosis and disease progression in CLL [11]. By systematic profiling of miRNAs in several cancers, Golub et al [12] observed a general downregulation of miRNAs in cancer samples. Subsequently, miRNAs were discovered in several cancers to be of prognostic importance. For instance, Monzo et al recently discovered a 25 miRNA signature for Hodgkin's Lymphoma (HL) and also identified miR-135a as a key player in prognostic outcome of HL [13]. In breast cancer, miR-155 [14] and miR-21 [15] are upregulated, while miR-10b has been implicated in invasiveness and metastasis of breast cancer [7]. MiR17-5p, miR-20a, miR-335, and miR-126 along with other miRNAs have also been found to be of importance in breast cancer [8,16]. In two separate studies, miR-143 and miR-145 were found to be downregulated in colorectal cancer [17,18]. Murakami et al analyzed Hepatocellular Carcinoma (HCC), chronic hepatitis and normal liver tissue for miRNAs and found an eight miRNA signature predictive of HCC [19]. MiR-224, miR-18, and pre-mir-18 were upregulated and miR-199a*, miR-200a, miR-199a, miR-125a, and miR-195 were downregulated in HCC samples compared to normal liver samples. In recent studies, miR-21, miR-10b, and miR-222 were found to be upregulated and miR-200c and miR-203 were downregulated in HCC. In Non Small Cell Lung Cancer (NSCLC), miR-128b functions as a tumor suppressor miRNA by controlling the production of EGFR [20]. MiR-128b along with a five miRNA signature constituting of miR-221, miR-137, miR-372, miR-182, and let-7 has been shown to be predictive of outcome in NSCLC patients [21]. In pancreatic cancer, separate studies identified miR-126 as a biomarker for pancreatic cancer [22-24].

MiRNA regulate gene expression in several ways, especially in pathological situations. Altered expression of some genes, particularly transcription factors, may result in altered expression of other miRNAs, which as a feed forward action causes altered expression of yet other genes resulting in heavy deregulation of normal molecular machinery of the cell. For example, we reported recently specific down-regulation of miR-22 in metastatic breast cancer cells compared to primary tumor cells, which results in elevated expression of the transcription factor EVI-1 in metastatic cancer cells [25]. EVI-1 being an epigenetic modulator of gene expression can profoundly alter gene expression pattern in metastatic cells.

Altered expression of miRNAs in cancers may occur due to genetic abnormalities, altered transcription, altered post-transcriptional events, or altered epigenetic factors. Consequently, dysregulated miRNAs alter cellular machinery at genomic/epigenomic level. For instance, miR-29 inhibits the expression of DNMT3A and DNMT3B, which are involved in DNA methylation in lung cancer

[26], whereas miR-101 has been shown to regulate histone methyltransferase EZH2 in prostate cancer [27]. MiR-15a and miR-16-1 were one of the first miRNAs to be discovered whose expression was altered due to genetic abnormality in chronic lymphocytic leukemia [28]. In a global analysis of mouse genome, several miRNAs were found to be associated with sites of frequent genomic abnormalities [29]. A similar trend was observed in the human genome while studying cancers [30]. Altered transcriptional regulation of miRNAs is also a major cause of deregulation of miRNA expression in certain cancers. It has been noted that transcription factors (TFs) may induce miRNAs by activating transcription of pri-miRNAs. Several miRNAs have also been experimentally shown to be directly regulated by TFs in cancers and quite often these microRNAs target the same transcription factors that induce them as a feedback loop [31-39]. miRNA expression can also be altered during post transcriptional events. For instance, levels of miRNA processing enzymes DROSHA and DICER were found to be altered in many cancers [40-44]. Also, in silico analysis of breast cancer tissue specimens has identified CpG islands near dozens of miRNAs genomic locations [45]. Furthermore, DNA hypomethylation induced release of miRNA silencing in colorectal cancer [46] suggests an epigenetic regulation of miRNA expression in cancers.

Altered miRNA expression can bring about pathophysiological changes in cellular machinery in several ways. Important mechanisms are induction of apoptosis, alteration of cell cycle, increased invasive and metastatic characteristics in the cells. For example, miR-29b [47], miR-34s [31], miR-15a and miR-16 [48] participate in tumorigenesis by targeting anti-apoptotic genes. MiR-221 and miR-222 in Glioblastoma (GBM) [49] and prostate cancer [50] have been shown to target p27, which restricts the cell cycle to G1 stage by preventing G1-to-S transition. As mentioned before, miR-10b [6] is suggested to impart metastatic characteristics to the cell, while miR-18 and miR-19 have been reported to repress TSP-1 and CTGF [37], both control angiogenesis.

All the above mentioned studies emphasize the remarkable multifaceted role miRNAs play as potent biomarkers in cancer. The biomarker capabilities of miRNAs have been studied sporadically (in specific studies for specific cancers), and the data for many such studies are available in public repositories. Global profiling of miRNAs in cancers has been done using the traditional PCR techniques, but more frequently using array platforms and sequencing. Recent advances in sequencing technologies have enhanced our capacity to study the global transcriptome of cancer populations, including miRNAs and other non-coding RNAs. Currently, there is no resource available which enables users to study biomarker capabilities of miRNAs in different cancers. Researchers have to go

through searching and processing publically available datasets for identifying and published information on prognostic implications of miRNA in cancer of interest. The wealth of data available publically thus remains underutilized, and this motivated us to create a platform where miRNAs can be studied for several types of cancers as prognostic biomarkers. In this paper, we present a first of a kind web based tool for studying prognostic importance of miRNAs in several types of cancers. Our tool is called PROGmiR, and it is available online freely for academic and non commercial purposes. PROGmiR allows users to study overall survival in form of prognostic plots using miRNA expression data from several publically available patient series. The data in our tool comes from Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) [51]. We have compiled the data on 16 different cancer types from both these sources. A list of cancer types available in our database along with other statistics on the data sets is available in Table 1. Our tool uses miRNA expression data from these datasets to create overall survival Kaplan-Meier (K-M) plots. Plots can be created for individual miRNAs as well as for average expression of a group (signature) of miRNAs in any of the cancer mentioned in Table 1.

Currently, our tool is restricted to studying only overall survival, but in the future, as more data become publically available, we aim to extend our tool to enable users to study additional prognostic measures such as metastasis free survival, recurrence free survival, and effects of specific therapies using our tool. Currently we

present this tool as a hypothesis generation tool only. In future versions, we would also like to add demographic and clinical covariates such as age, race and hormonal and therapy statuses to be included in the analysis to produce more meaningful and accurate prognostic plots. Since miRNAs have significant importance as biomarkers, and since our tool covers almost all major cancers from more than 10 tissues of origin, we believe this tool will help researchers in identifying novel prognostic markers as well as in formulating hypothesis for mechanistic studies in the future. In particular, our tool will allow investigators to obtain preliminary evidence on whether microRNA functions as a tumor suppressor or oncogene across all cancer types or has dual role depending on cancer type.

Construction and content
Workflow
To create prognostic plots for miRNAs, our application uses miRNA expression and overall survival data. MiRNA expression data are in form of array expression data or sequencing data based on the platform. Our application is a web based PHP [56] script which uses R [57] in the backend to compute survival plots for miRNAs of interest. The R script uses library 'Survival' for creating survival plots. Survival plots can be created for 3-year, 5-year or full follow up survival time. Figure 1 describes the workflow of the application. Prognostic plots in our application can be created for single or multiple miRNAs. When multiple miRNAs are entered, our application creates prognostic plots for miRNAs

Table 1 Description of various data sources included in progmir database

	Cancer type	Dataset	Source	Platform	No of samples	No of miRNAs*	Reference
1	Adrenocortial Carcinoma	GSE22816	GEO	Agilent-025987 Human miRNA Microarray Release 14.0	22	200	[52]
2	Acute Myeloid Leukemia	LAML	TCGA	Illumina GA miRNA Seq	164	704	[51]
3	Brain Lower Grade Glioma	LGG	TCGA	Illumina HISeq miRNA Seq	29	1046	[51]
4	Glioblastoma multiforme	GBM	TCGA	UNC miRNA 8X15K	487	470	[51]
5	Breast Invasive Carcinoma	BRCA	TCGA	Illumina GA and HISeq	727	1046	[51]
6	Non-small-cell Lung Cancer	GSE16025	GEO	mirVANA miRNA Bioarray V2	60	328	[53]
7	Small Cell Lung Cancer	GSE27435	GEO	Capitalbio mammal microRNA V3.0	42	1638	[54]
8	Lung Adenocarcinoma	LUAD	TCGA	Illumina GA and HISeq	79	1046	[51]
9	Lung Squamous Cell Carcinoma	LUSC	TCGA	Illumina GA and HISeq	185	1046	[51]
10	Hepatocellular Carcinoma	GSE31384	GEO	CapitalBio custom Human microRNA array	166	682	[55]
11	Head and Neck Squamous Cell Carcinoma	HNSC	TCGA	Illumina GA and HISeq	89	1046	[51]
12	Ovarian Cystadenocarcinoma	OV	TCGA	UNC miRNA 8X15K	46	705	[51]
13	Rectal Adenocarcinoma	READ	TCGA	Illumina GA miRNA Seq	38	705	[51]
14	Renal Clear Cell Carcinoma	KIRC	TCGA	Illumina GA and HISeq	546	1046	[51]
15	Stomach Ademocarcinoma	STAD	TCGA	Illumina GA and HISeq	79	1046	[51]
16	Uterine Corpus Endometroid Carcinoma	UCEC	TCGA	Illumina GA and HISeq	358	1046	[51]

Figure 1 Line diagram delineating workflow of PROGmiR.
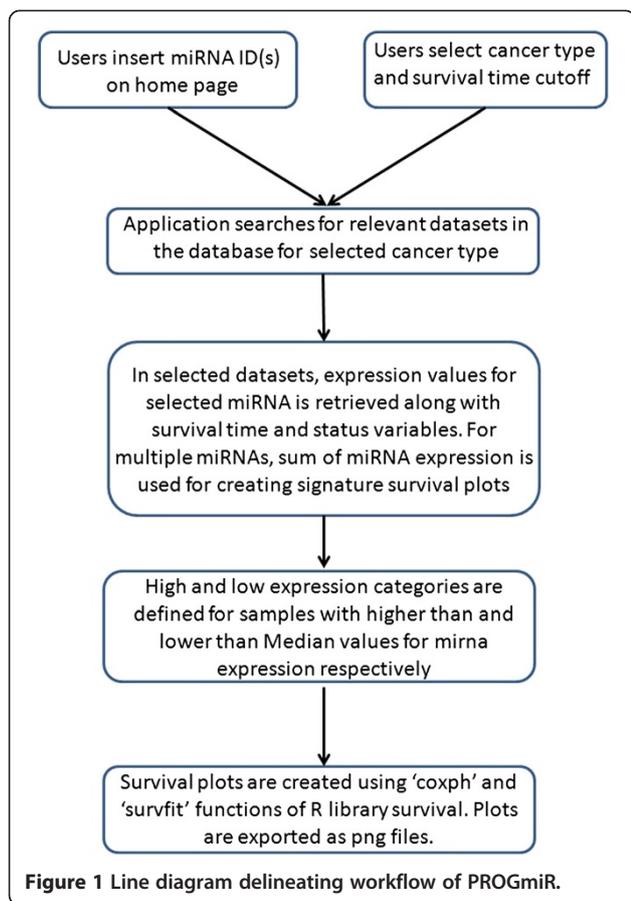
individually as well as a combined prognostic plot for all miRNAs entered. This makes it possible to study prognostic implications of a miRNA signature in different types of cancer. For combined plots, sum of miRNA expressions of all miRNAs entered is computed and used in creation of prognostic plots.

Users enter miRNA ID(s) on the home page of web application and select the cancer type in which prognostic plots have to be visualized. For some cancer types, there are more than one expression dataset present. For each dataset, one prognostic plot is created for each miRNA and one plot is created for sum of expressions of all miRNAs entered combined. The program retrieves expression data for each entered miRNA along with corresponding survival annotations. Hazard ratio and p value are then calculated using function 'coxph' and prognostic plots are visualized using function 'survfit', both functions in library 'survival' for R. To estimate hazard ratio and corresponding p value, continuous miRNA expression is used. For creating prognostic plots, samples are categorized into 'HIGH' and 'LOW' miRNA expression categories by bifurcating at median miRNA expression. The plots are exported as .png files and these files are visualized on the results page as individual

miRNA plots and sum of expression plots divided by datasets. For each plot, 'exp(Coef)' as estimate of hazard ratio, and Log likelihood p value generated by the 'coxph' model are also provided.

Data
We downloaded miRNA expression data from GEO and TCGA. GEO data were downloaded for separate platforms as described in Table 1. The TCGA data were in form of RNA sequencing data except for Glioblastoma, where it was on array platform. Sequencing data from TCGA were available in form of 'reads per million (Level 3)' for each miRNA. GEO data were downloaded in form of series matrices from the GEO website. Survival data associated with clinical samples were also downloaded from the respective data sources along with the expression data. The time to death data were converted into days to death for data sources where months or years to death were reported. Batch effect arising due to processing of samples at different times was removed statistically, if present. In most cases, unlike mRNA expression data, the data available were in already processed format and did not have to be normalized further. Since we do not combine expression data from all datasets for a particular cancer type, any other normalization was again not relevant. The data were stored as R data sets and computation is done directly using these datasets in our application.

Web application
We have created a web application for implementation of our tool. As mentioned previously, the web application was written in PHP5 and uses R scripts in backend to create survival plots. The web application consists of a home page and results page. Users can input individual miRNAs or a comma delimited list of miRNAs on the home page. Users also select one cancer type in which prognostic plots have to be created. Also available on the home page is options to truncate plots to a 3 Yr or 5 Yr follow up or to create plots for full follow up time. Since different studies have been performed on different platforms, the number of miRNAs profiled in each study differs. Prognostic plots are created for each dataset for available miRNAs only. Upon submitting the information on home page, results in form of KM plots are provided on the results page.

Utility and discussion
We have compiled miRNA expression data from 16 cancer types in our data base. Median survival and follow-up times as well as number of events for each dataset have been summarized in Table 2. We used only overall survival data in our database, as other survival functions such as metastasis free survival and relapse free survival,

Table 2 Global statistics on survival related variables for the datasets available in PROGmiR

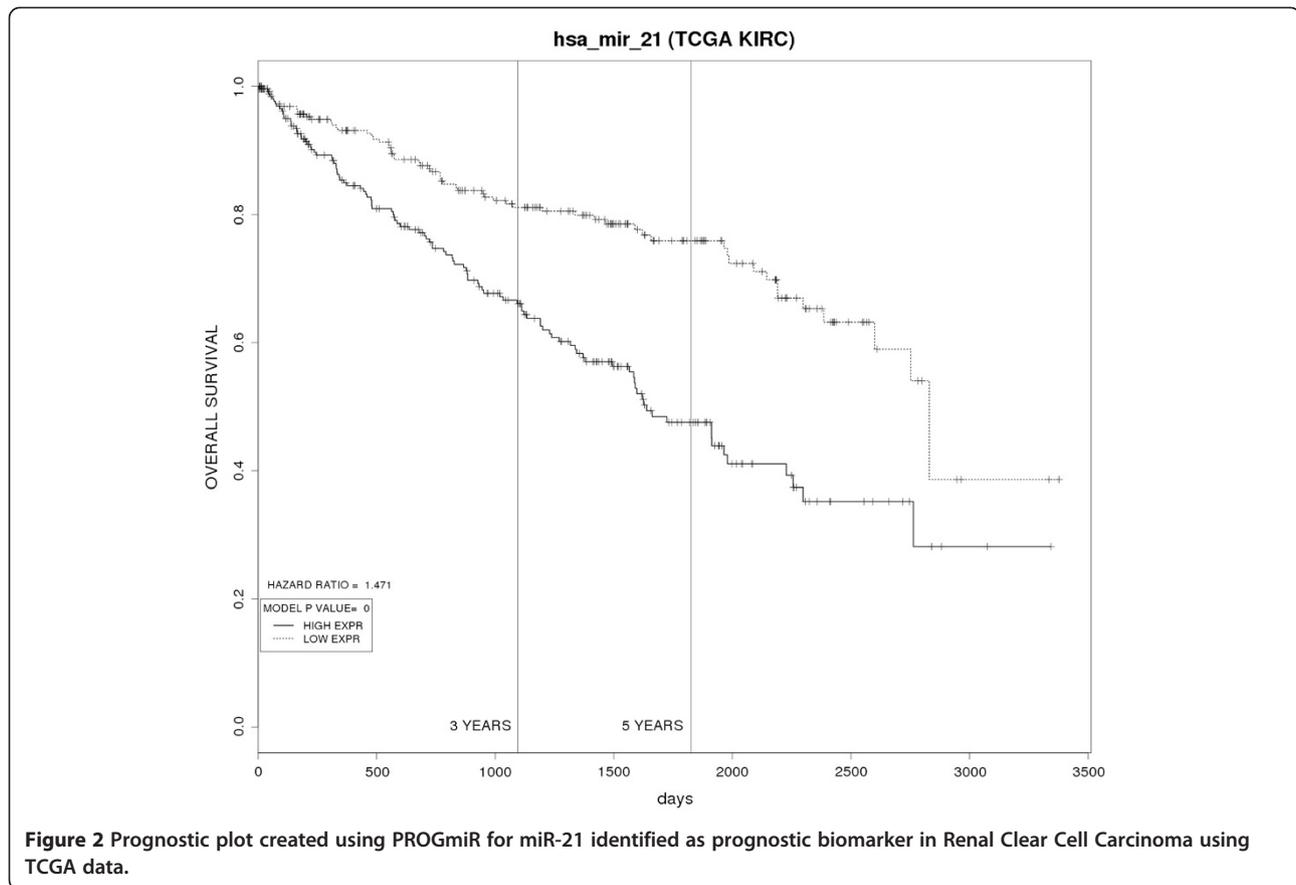|	Cancer type	Dataset	Total no of samples	Survival event		No event	
				No of samples	Median survival[§] (range)	No of samples	Median survival[§*] (range)
1	Adrenocortial Carcinoma	GSE22816	22	6	600 (120-4860)	16	885 (60-4980)
2	Acute Myeloid Leukemia	LAML	164	100	303 (28-1706)	64	699 (28-2861)
3	Brain Lower Grade Glioma	LGG	29	23	788 (96-1915)	6	1317 (242-6423)
4	Glioblastoma multiforme	GBM	487	383	377 (3-3880)	104	266 (3-2817)
5	Breast Invasive Carcinoma	BRCA	727	91	1563 (157-4456)	636	466 (1-6795)
6	Non-small-cell Lung Cancer	GSE16025	60	26	1980 (210-2490)	34	1500 (120-2580)
7	Small Cell Lung Cancer	GSE27435	42	18	743 (234-2400)	24	1314 (426 -2490)
8	Lung Adenocarcinoma	LUAD	79	27	701 (22-1318)	52	400 (1-2161)
9	Lung Squamous Cell Carcinoma	LUSC	185	77	544 (12-5296)	108	640 (3-4299)
10	Hepatocellular Carcinoma	GSE31384	166	73	450 (30 – 2280)	93	1320 (420-2430)
11	Head and Neck Squamous Cell Carcinoma	HNSC	89	27	395 (128 – 2318)	62	359 (45 – 4115)
12	Ovarian Cystadenocarcinoma	OV	46	21	887 (9-1756)	25	912 (141 – 2099)
13	Rectal Adenocarcinoma	READ	38	3	316 (59 – 1184)	35	183 (28 – 2192)
14	Renal Clear Cell Carcinoma	KIRC	546	173	722 (2-2830)	373	1307 (4 – 3377)
15	Stomach Adenocarcinoma	STAD	79	14	258 (19 – 881)	65	109 (2 – 2131)
16	Uterine Corpus Endometroid Carcinoma	UCEC	358	27	413 (58 – 3251)	331	527 (1 – 5690)

§ in days.
* Follow up time.

were not available for all studies. Studies in which numbers of survival events were less than 5 were not included in the database. In TCGA data, there is an innate problem of several samples been annotated as having a zero follow-up time. Such samples were removed from the final datasets. Our final data consists of a total of 3117 samples in 16 cancer types and appximately 1050 miRNAs profiled. Unlike related applications available for studying survival implications of mRNAs in breast and ovarian cancers [58,59], we did not merge all samples for the same cancer type into one large dataset owing to the fact that it may lead to over-fitting of data due to normalization of samples coming from different studies. For this reason, we kept different datasets belonging to the same cancer type separate.

For assessing performance of our tool, we created survival plots for miRNAs that have been implicated as having prognostic importance in published studies. Dahiya et al [60], have shown that a high expression of miR-21 correlates with poor overall survival in renal clear cell carcinoma. We used our tool to create prognostic plot for miR-21 in renal cell carcinoma in TCGA data. Figure 2 shows a poor overall survival for patients with high miR-21 expression compared to patients having low miR-21 expression levels. The hazard ratio and p value for the proportional hazards model is also given in the figure.

In another study, Chen et al [61], have demonstrated down-regulation of several isoforms of miR-181 (miR-181-a, miR-181-b, miR-181-c and miR-181-d) being correlated with poor overall survival in acute myeloid leukemia (AML). We used our tool to create prognostic plots for the above mentioned miRNA isoforms using TCGA data. All the isoforms and combined sum of expression of isoforms showed the same pattern in plots created using our tool as shown in the paper, with poor overall survival observed in the group having low expression of miR-181 isoforms. Additional file 1: Figures S1-S4 provided in supplementary data show prognostic plots for miR-181 isoforms in AML data. Additional file 1: Figure S5 shows a prognostic plot for sum of the 4 miRNAs showing correlation of overall survival with sum of miRNA expression. Similarly, in another study, Annilo et al [62], have demonstrated low expression of miR-374a to be associated with poor outcome in non small cell lung cancer. We used our tool to create overall survival prognostic plots for miR-374a in squamous cell carcinoma of lung (LUSC) using TCGA data. The plot (Additional file 1: Figure S6) showed the miRNA low expression arm demonstrating a poor outcome in LUSC samples.

In each of the case studies we performed, prognostic plots created by our tool corroborated with the findings in the published studies. Since this tool is basically a pipeline, this fact makes the tool validated. Somasundaram et al [63] recently used the TCGA data itself to identify 10 miRNAs having prognostic implications in glioblastoma. Our tool used the same data and showed similar results (Additional file 1: Figures S7-S16) for the selected 10 miRNAs.

Figure 2 Prognostic plot created using PROGmiR for miR-21 identified as prognostic biomarker in Renal Clear Cell Carcinoma using TCGA data.

In future updates to our tool, we plan to expand the repository of our tool by adding new datasets as and when they become available. We also plan to include more survival functions, such as 'metastasis free survival' and 'relapse free survival' to our tool in future versions.

Conclusions

We believe this tool will prove useful for hypothesis generation and testing as well as for mechanistic studies. Considering the impact of miRNAs as prognostic biomarkers in several cancer types, such preliminary findings will also benefit researchers formulate research plans. Since our tool covers data from all major cancer types, researchers working on wide array of cancers will be able to formulate their study hypotheses based on findings from this tool.

Availability and requirements

The database is available freely for non commercial and academic usage at www.compbio.iupui.edu/progmir.

Additional file

Additional file 1: Figures S1-S4. Prognostic plot created using PROGmiR for isoforms a, b, c and d of miRNA hsa-miR-181 identified as prognostically important biomarker in Acute Myeloid Leukemia (AML) by

Chen et al, using TCGA data. **Figure S5**. Prognostic plot for sum of expression of hsa-mir-181 isoforms a,b,c and d in TCGA AML data. **Figure S6**. Prognostic plot created using PROGmiR for miRNA hsa-miR-374 identified as prognostically important biomarker in Lung Squamous cell carcinoma (LUSC) by Annilo et al, using TCGA data. **Figures S7-S16**. Prognostic plot created using PROGmiR for 10 miRNAs identified as prognostically important biomarker in Glioblastoma (GBM) by Somasundaram et al, using TCGA data.

Author details
[1]Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, 410 W 10th Street, Indianapolis, IN 46202, USA. [2]Departments of Surgery, Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN 46202, USA.

References

1. Lee RC, Feinbaum RL, Ambros V: **The C. Elegans heterochromic gene lin-4 encodes small RNAs with antisense complementarity to lin-14.** *Cell* 1993, **75**(5):843–854.
2. Pasquinelli AE, Reinhart BJ, Slack F, Martindale MQ, Kuroda MI, Maller B, Hayward DC, Ball EE, Degnan B, Müller P, Spring J, Srinivasan A, Fishman M, Finnerty J, Corbo J, Levine M, Leahy P, Davidson E, Ruvkun G: **Conservation of the sequence and temporal expression of let-7 heterochromic regulatory RNA.** *Nature* 2000, **408**(6808):86–89.
3. Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP: **MicroRNAs in plants.** *Genes Dev* 2002, **16**:1616–1626.
4. Lee YS, Dutta A: **MicroRNAs in cancer.** *Annu Rev Pathol* 2009, **4**:199–227.
5. Mavrakis KJ, Wolfe AL, Oricchio E, Palomero T, de Keersmaecker K, McJunkin K, Zuber J, James T, Khan AA, Leslie CS, Parker JS, Paddison PJ, Tam W, Ferrando A, Wendel HG: **Genome-wide RNA-mediated interference screen identifies miR-19 targets in notch-induced T-cell acute lymphoblastic leukaemia.** *Nat Cell Biol* 2010, **12**(4):372–379.
6. Raver-Shapira N, Marciano E, Meiri E, Spector Y, Rosenfeld N, Moskovits N, Bentwich Z, Oren M: **Transcriptional activation of miR-34a contributes to p53-mediated apoptosis.** *Mol Cell* 2007, **26**(5):731–743.
7. Ma L, Teruya-Feldstein J, Weinberg RA: **Tumour invasion and metastasis initiated by microRNA-10b in breast cancer.** *Nature* 2007, **449**(7163):682–688.
8. Tavazoie SF, Alarcón C, Oskarsson T, Padua D, Wang Q, Bos PD, Gerald WL, Massagué J: **Endogenous human microRNAs that suppress breast cancer metastasis.** *Nature* 2008, **451**(7175):147–152.
9. Ma L, Young J, Prabhala H, Pan E, Mestdagh P, Muth D, Teruya-Feldstein J, Reinhardt F, Onder TT, Valastyan S, Westermann F, Speleman F, Vandesompele J, Weinberg RA: **miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis.** *Nat Cell Biol* 2010, **12**(3):247–256.
10. Valastyan S, Reinhardt F, Benaich N, Calogrias D, Szász AM, Wang ZC, Brock JE, Richardson AL, Weinberg RA: **A pleiotropically acting microRNA, miR-31, inhibits breast cancer metastasis.** *Cell* 2009, **137**(6):1032–1046.
11. Calin GA, Ferracin M, Cimmino A, Di Leva G, Shimizu M, Wojcik SE, Iorio MV, Visone R, Sever NI, Fabbri M, Iuliano R, Palumbo T, Pichiorri F, Roldo C, Garzon R, Sevignani C, Rassenti L, Alder H, Volinia S, Liu CG, Kipps TJ, Negrini M, Croce CM: **MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia.** *N Engl J Med* 2005, **353**:1793–1801.
12. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA, Downing JR, Jacks T, Horvitz HR, Golub TR: **MicroRNA expression profiles classify human cancers.** *Nature* 2005, **435**(7043):834–838.
13. Navarro A, Diaz T, Martinez A, Gaya A, Pons A, Gel B, Codony C, Ferrer G, Martinez C, Montserrat E, Monzo M: **Regulation of JAK2 by miR-135a: prognostic impact in classic Hodgkin lymphoma.** *Blood* 2009, **114**:2945–2951.
14. Iorio MV, Ferracin M, Liu CG, Veronese A, Spizzo R, Sabbioni S, Magri E, Pedriali M, Fabbri M, Campiglio M, Ménard S, Palazzo JP, Rosenberg A, Musiani P, Volinia S, Nenci I, Calin GA, Querzoli P, Negrini M, Croce CM: **MicroRNA gene expression deregulation in human breast cancer.** *Cancer Res* 2005, **65**:7065–7070.
15. Zhu S, Si ML, Wu H, Mo YY: **MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1).** *J Biol Chem* 2007, **282**:14328–14336.
16. Yu Z, Wang C, Wang M, Li Z, Casimiro MC, Liu M, Wu K, Whittle J, Ju X, Hyslop T, McCue P, Pestell RG: **A cyclinD1/microRNA 17/20 regulatory feedback loop in control of breast cancer cell proliferation.** *J Cell Bio* 2008, **182**:509–517.
17. Akao Y, Nakagawa Y, Naoe T: **MicroRNA-143 and -145 in colon cancer.** *DNA Cell Biol* 2007, **26**:311–320.
18. Michael MZ, O'Connor SM, van Holst Pellekaan NG, Young GP, James RJ: **Reduced accumulation of specific microRNAs in colorectal neoplasia.** *Mol Cancer Res* 2003, **1**:882–891.
19. Murakami Y, Yasuda T, Saigo K, Urashima T, Toyoda H, Okanoue T, Shimotohno K: **Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues.** *Oncogene* 2006, **25**:2537–2545.
20. Weiss GJ, Bemis LT, Nakajima E, Sugita M, Birks DK, Robinson WA, Varella-Garcia M, Bunn PA: **Jr, Haney J, helfrich BA, Kato H, Hirsch FR, Franklin WA: EGFR regulation by microRNA in lung cancer: correlation with clinical response and survival to Gefitinib and EGFR expression in cell lines.** *Ann Oncol* 2008, **19**:1053–1059.
21. Yu SL, Chen HY, Chang GC, Chen CY, Chen HW, Singh S, Cheng CL, Yu CJ, Lee YC, Chen HS, Su TJ, Chiang CC, Li HN, Hong QS, Su HY, Chen CC, Chen WJ, Liu CC, Chan WK, Chen WJ, Li KC, Chen JJ, Yang PC: **MicroRNA signature predicts survival and relapse in lung cancer.** *Cancer Cell* 2008, **13**:48–57.
22. Baskerville S, Bartel DP: **Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes.** *RNA* 2005, **11**:241–247.
23. Sood P, Krek A, Zavolan M, Macino G, Rajewsky N: **Cell-type-specific signatures of microRNAs on target mRNA expression.** *Proc Natl Acad Sci* 2006, **103**:2746–2751. USA.
24. Shingara J, Keiger K, Shelton J, Laosinchai-Wolf W, Powers P, Conrad R, Brown D, Labourier E: **An optimized isolation and labeling platform for accurate microRNA expression profiling.** *RNA* 2005, **11**:1461–1470.
25. Patel JB, Appaiah HN, Burnett RM, Bhat-Nakshatri P, Wang G, Mehta R, Badve S, Thomson MJ, Hammond S, Steeg P, Liu Y, Nakshatri H: **Control of EVI-1 oncogene expression in metastatic breast cancer cells through microRNA miR-22.** *Oncogene* 2011, **30**(11):1290–1301.
26. Fabbri M, Garzon R, Cimmino A, Liu Z, Zanesi N, Callegari E, Liu S, Alder H, Costinean S, Fernandez-Cymering C, Volinia S, Guler G, Morrison CD, Chan KK, Marcucci G, Calin GA, Huebner K, Croce CM: **MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B.** *Proc Natl Acad Sci USA* 2007, **104**(40):15805–15810.
27. Varambally S, Cao Q, Mani RS, Shankar S, Wang X, Ateeq B, Laxman B, Cao X, Jing X, Ramnarayanan K, Brenner JC, Yu J, Kim JH, Han B, Tan P, Kumar-Sinha C, Lonigro RJ, Palanisamy N, Maher CA, Chinnaiyan AM: **Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer.** *Science* 2008, **322**(5908):1695–1699.
28. Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch E, Aldler H, Rattan S, Keating M, Rai K, Rassenti L, Kipps T, Negrini M, Bullrich F, Croce CM: **Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia.** *Proc Natl Acad Sci USA* 2002, **99**(24):15524–15529.
29. Sevignani C, Calin GA, Nnadi SC, Shimizu M, Davuluri RV, Hyslop T, Demant P, Croce CM, Siracusa LD: **MicroRNA genes are frequently located near mouse cancer susceptibility loci.** *Proc Natl Acad Sci USA* 2007, **104**(19):8017–8022.
30. Zhang L, Huang J, Yang N, Greshock J, Megraw MS, Giannakakis A, Liang S, Naylor TL, Barchetti A, Ward MR, Yao G, Medina A, O'Brien-Jenkins A, Katsaros D, Hatzigeorgiou A, Gimotty PA, Weber BL, Coukos G: **MicroRNAs exhibit high frequency genomic alterations in human cancer.** *Proc Natl Acad Sci USA* 2006, **103**(24):9136–9141.
31. Bommer GT, Gerin I, Feng Y, Kaczorowski AJ, Kuick R, Love RE, Zhai Y, Giordano TJ, Qin ZS, Moore BB, MacDougald OA, Cho KR, Fearon ER: **p53-Mediated activation of miRNA34 candidate tumor-suppressor genes.** *Curr Biol* 2007, **17**(15):1298–1307.
32. Löffler D, Brocke-Heidrich K, Pfeifer G, Stocsits C, Hackermüller J, Kretzschmar AK, Burger R, Gramatzki M, Blumert C, Bauer K, Cvijic H, Ullmann AK, Stadler PF, Horn F: **Interleukin-6 dependent survival of multiple myeloma cells involves the Stat3-mediated induction of microRNA-21 through a highly conserved enhancer.** *Blood* 2007, **110**(4):1330–1333.
33. Schulte JH, Horn S, Otto T, Samans B, Heukamp LC, Eilers UC, Krause M, Astrahantseff K, Klein-Hitpass L, Buettner R, Schramm A, Christiansen H, Eilers M, Eggert A, Berwanger B: **MYCN regulates oncogenic MicroRNAs in Neuroblastoma.** *Int J Cancer* 2008, **122**(3):699–704.
34. Woods K, Thomson JM, Hammond SM: **Direct regulation of an oncogenic micro-RNA cluster by E2F transcription factors.** *J Biol Chem* 2007, **282**(4):2130–2134.
35. Sylvestre Y, De Guire V, Querido E, Mukhopadhyay UK, Bourdeau V, Major F, Ferbeyre G, Chartrand P: **An E2F/miR-20a autoregulatory feedback loop.** *J Biol Chem* 2007, **282**(4):2135–2143.
36. Tarasov V, Jung P, Verdoodt B, Lodygin D, Epanchintsev A, Menssen A, Meister G, Hermeking H: **Differential regulation of microRNAs by p53 revealed by massively parallel sequencing: miR-34a is a p53 target that induces apoptosis and G1-arrest.** *Cell Cycle* 2007, **6**(13):1586–1593.
37. Dews M, Homayouni A, Yu D, Murphy D, Sevignani C, Wentzel E, Furth EE, Lee WM, Enders GH, Mendell JT, Thomas-Tikhonenko A: **Augmentation of tumor angiogenesis by a Myc-activated microRNA cluster.** *Nat Genet* 2006, **38**(9):1060–1065.

38. Kulshreshtha R, Ferracin M, Wojcik SE, Garzon R, Alder H, Agosto-Perez FJ, Davuluri R, Liu CG, Croce CM, Negrini M, Calin GA, Ivan M: A microRNA signature of hypoxia. *Mol Cell Biol* 2007, **27**(5):1859–1867.

39. Shi XB, Xue L, Yang J, Ma AH, Zhao J, Xu M, Tepper CG, Evans CP, Kung HJ, deVere White RW: An androgen-regulated miRNA suppresses Bak1 expression and induces androgen-independent growth of prostate cancer cells. *Proc Natl Acad Sci USA* 2007, **104**(50):19983–19988.

40. Kumar MS, Lu J, Mercer KL, Golub TR, Jacks T: Impaired microRNA processing enhances cellular transformation and tumorigenesis. *Nat Genet* 2007, **39**(5):673–677.

41. Blenkiron C, Goldstein LD, Thorne NP, Spiteri I, Chin SF, Dunning MJ, Barbosa-Morais NL, Teschendorff AE, Green AR, Ellis IO, Tavaré S, Caldas C, Miska EA: MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype. *Genome Biol* 2007, **8**(10):R214.

42. Chiosea S, Jelezcova E, Chandran U, Luo J, Mantha G, Sobol RW, Dacic S: Overexpression of dicer in precursor lesions of lung adenocarcinoma. *Cancer Res* 2007, **67**(5):2345–2350.

43. Muralidhar B, Goldstein LD, Ng G, Winder DM, Palmer RD, Gooding EL, Barbosa-Morais NL, Mukherjee G, Thorne NP, Roberts I, Pett MR, Coleman N: Global microRNA profiles in cervical squamous cell carcinoma depend on drosha expression levels. *J Pathol* 2007, **212**(4):368–377.

44. Karube Y, Tanaka H, Osada H, Tomida S, Tatematsu Y, Yanagisawa K, Yatabe Y, Takamizawa J, Miyoshi S, Mitsudomi T, Takahashi T: Reduced expression of dicer associated with poor prognosis in lung cancer patients. *Cancer Sci* 2005, **96**(2):111–115.

45. Lehmann U, Hasemeier B, Christgen M, Müller M, Römermann D, Länger F, Kreipe H: Epigenetic inactivation of microRNA gene hsa-mir-9-1 in human breast cancer. *J Pathol* 2008, **214**(1):17–24.

46. Lujambio A, Ropero S, Ballestar E, Fraga MF, Cerrato C, Setién F, Casado S, Suarez-Gauthier A, Sanchez-Cespedes M, Git A, Spiteri I, Das PP, Caldas C, Miska E, Esteller M: Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. *Cancer Res* 2007, **67**(4):1424–1429.

47. Mott JL, Kobayashi S, Bronk SF, Gores GJ: mir-29 regulates Mcl-1 protein expression and apoptosis. *Oncogene* 2007, **26**(42):6133–6140.

48. Cimmino A, Calin GA, Fabbri M, Iorio MV, Ferracin M, Shimizu M, Wojcik SE, Aqeilan RI, Zupo S, Dono M, Rassenti L, Alder H, Volinia S, Liu CG, Kipps TJ, Negrini M, Croce CM: miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proc Natl Acad Sci USA* 2005, **102**(39):13944–13949.

49. Gillies JK, Lorimer IA: Regulation of p27Kip1 by miRNA 221/222 in glioblastoma. *Cell Cycle* 2007, **6**(16):2005–2009.

50. Galardi S, Mercatelli N, Giorda E, Massalini S, Frajese GV, Ciafrè SA, Farace MG: miR-221 and miR-222 expression affects the proliferation potential of human prostate carcinoma cell lines by targeting p27Kip1. *J Biol Chem* 2007, **282**(32):23716–23724.

51. https://tcga-data.nci.nih.gov/tcga.

52. Özata DM, Caramuta S, Velázquez-Fernández D, Akçakaya P, Xie H, Höög A, Zedenius J, Bäckdahl M, Larsson C, Lui WO: The role of microRNA deregulation in the pathogenesis of adrenocortical carcinoma. *Endocr Relat Cancer* 2011, **18**(6):643–655.

53. Raponi M, Dossey L, Jatkoe T, Wu X, Chen G, Fan H, Beer DG: MicroRNA classifiers for predicting prognosis of squamous cell lung cancer. *Cancer Res* 2009, **69**(14):5776–5783.

54. http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE27435.

55. http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE31384.

56. www.php.net.

57. http://www.r-project.org.

58. Gyorffy B, Lanczky A, Szallasi Z: Implementing an online tool for genome-wide validation of survival-associated biomarkers in ovarian-cancer using microarray data from 1287 patients. *EndocrRelat Cancer* 2012, **19**(2):197–208.

59. Gyorffy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q, Szallasi Z: An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1809 patients. *Breast Cancer Res Treatment* 2010, **123**(3):725–731.

60. Zaman MS, Shahryari V, Deng G, Thamminana S, Saini S, Majid S, Chang I, Hirata H, Ueno K, Yamamura S, Singh K, Tanaka Y, Tabatabai ZL, Dahiya R: Up-regulation of microRNA-21 correlates with lower kidney cancer survival. *PLoS One* 2012, **7**(2):e31060.

61. Li Z, Huang H, Li Y, Jiang X, Chen P, Arnovitz S, Radmacher MD, Maharry K, Elkahloun A, Yang X, He C, He M, Zhang Z, Dohner K, Neilly MB, Price C, Lussier YA, Zhang Y, Larson RA, Le Beau MM, Caligiuri MA, Bullinger L, Valk PJ, Delwel R, Lowenberg B, Liu PP, Marcucci G, Bloomfield CD, Rowley JD, Chen J: Up-regulation of a HOXA-PBX3 homeobox-gene signature following down-regulation of miR-181 is associated with adverse prognosis in patients with cytogenetically abnormal AML. *Blood* 2012, **119**(10):2314–2324.

62. Vōsa U, Vooder T, Kolde R, Fischer K, Välk K, Tōnisson N, Roosipuu R, Vilo J, Metspalu A, Annilo T: Identification of miR-374a as a prognostic marker for survival in patients with early-stage nonsmall cell lung cancer. *Genes Chromosomes Cancer* 2011, **50**(10):812–822.

63. Srinivasan S, Patric IR, Somasundaram K: A ten-microRNA expression signature predicts survival in glioblastoma. *PLoS One* 2011, **6**(3):e17438.