Integrating Protein Networks for Identifying Cooperative miRNA Activity in Disease Gene Signatures

Mohammed Alshalalfa
Computer Science Department
University of Calgary
Calgary, Canada
Email: msalshal@ucalgary.ca

Reda Alhajj
Computer Science Department
University of Calgary
Calgary, Canada
Email: alhajj@ucalgary.ca

***Abstract*—MicroRNAs are a major class of regulatory molecules involved in a broad range of biological processes and complex diseases. A useful step for understanding their functional role is detecting their influence on genome-wide expression profiles. In this work we use elastic-net regression model that incorporates direct and indirect effects of miRNAs on protein networks for identifying regulatory effect of miRNAs on a list of genes from expression experiment (disease gene signature). Elastic-net regression is used to identify miRNAs whose targets are enriched in disease gene signature. Integrating direct and indirect effects of miRNAs on protein network revealed more significant miRNA enrichment in prostate gene signatures compared to using direct effects alone. Integrating protein networks into regression model revealed significant enrichment of miR-16-1 in upregulated genes in prostate cancer which indicates its putative tumor suppression activity.**

***Index Terms*—miRNA; gene list; regression model; prostate cancer; miRNA enrichment**

I. Introduction

MicroRNAs(miRNAs) are small (18-24) nt long non-coding RNAs that are a major class of regulatory molecules that added another dimension to gene expression regulation. Since the discovery of microRNAs [1], they emerged as a new layer of gene regulation that dramatically influence genes by binding to their 3'UTR and inactivate them by promoting their degradation or translational repression. Each miRNA is estimated to regulate hundreds of genes, and over 50% of the human protein-coding are regulated by miRNAs [2]. MiRNAs play a key role in regulating diverse cellular functions, such as development, proliferation, apoptosis, and metabolism [3] and are associated with a growing list of diseases including cancer [4], [5]. A growing body of evidence suggests that miRNAs impact gene expression in many cancer types including prostate cancer [6]–[8]. Several studies have investigated the role of miRNAs in cancer using mRNA and miRNA expression profiling [9], [10]. Their findings suggested that most diseases are attributed to more than one miRNA that affect hundreds of genes. Better understanding the regulatory role of miRNAs in cancer development and progression requires exploring their cooperative influence on target proteins and their protein partners. High-throughput genomics technologies are witnessing a revolution and becoming a standard routine in many experimental laboratories. As a result, gene sets are delivered on a regular basis. Gene sets can be differentially expressed genes that bear important knowledge to understand the change of status of the cell. It is also expected that the differentially expressed genes bear a signature of regulatory activity from miRNAs as it is anticipated that the collective effect of miRNAs may lead to a dramatic change in the expression of their targets that might lead to diseases.

A common step to reveal the biological knowledge in gene signatures and infer overrepresentation of particular gene sets is enrichment analysis. Analysis of overrepresented features in lists of genes is a powerful tool for associating function with biological effect. A number of tools employing available gene annotation and pathways have been developed like DAVID (http://david.abcc.ncifcrf.gov/), GeneMANIA (genemania.org), Geneset2miRNA [11], Expression2kinases [12], and Enrichnet and network based enrichment. Several tools have also been developed to identify miRNAs that target significant subset of the given gene signature. Gene set enrichment analysis can be grouped into two classes. The first takes ordered list of all genes and then analyze the overrepresentation in the top set compared with the rest. GSEA [13] and SylArray [14] are two methods that require ordered gene signature. The other class that constitute a big portion of the existing methods assess the statistical significance of the overlap between gene signature and miRNA targets. A recent survey by Huang et al [15] list 68 distinct bioinformatics enrichment tools. Most of them share a similar workflow that assess the significance of the overlap between two sets. Mainly statistical tests such as fisher test, hypergeometric tests are used for enrichment analysis. The drawback of the first one is that it requires a big list of genes that might be not available or might be not of interest for biologists as they are concerned with one particular gene signature that might be less than 100 genes. The disadvantage of the second method is that it requires a cut-off value to select the gene signature.

Characterizing the downstream effect of miRNA on protein

978-1-4673-2560-8/12/$31.00 ©2012 IEEE

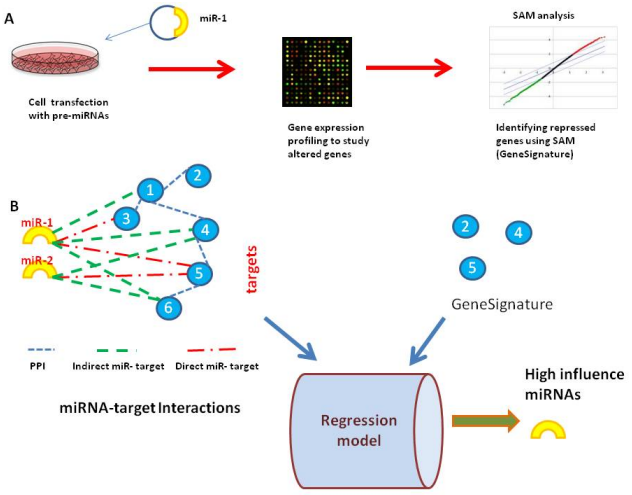

Fig. 1. An overview of Identifying miRNA gene signature and constructing influential miRNA-target interactions for regression analysis

interactions gained considerable body of attention in the past few years. Protein degree in PPI networks showed to be correlated with the number of targeting miRNAs; hub proteins are targeted by large number of distinct miRNAs [16]. Topological features of proteins in PPI showed to be useful to eliminate false discoveries in miRNA-target prediction algorithms [17]. Other studies demonstrated that PPI context of miRNA targets provides more representative information about miRNA function compared to using direct targets alone [16]. These observations shed light on the influence of miRNAs on the PPI subnetwork involving the target, and highlights the importance of considering target protein partners when searching for functional miRNA-target interactions. To our knowledge, there is no method that takes into account the downstream partners of miRNA target to assess the enriched functional miRNAs.

The primary goal of this work is assess if a single or multiple miRNAs are directly or indirectly affecting the disease gene signature. Given a gene signature and a miRNA-gene influence matrix, elastic-net regression model is used to predict the influence coefficient of each miRNA on the gene signature. The resulted coefficents are used to reveal miRNA enrichemnt in the gene signature.

II. Materials and Methods

A. miRNA target interactions

Human miRNA target computational predictions for miRNA with conserved 3'UTR were taken from TargetScan 5.1 [18][PredNet], and experimentally validated miRNA and their targets were taken from mirTarbase [19] and miRecord [20]. We used the union of mirTarbase and miRecord as a source of experimentally validated miRNA-target interactions [ExpNet].

B. Protein networks

We combined undirected functional interactions from Reactome [21], which includes proteins physically interacting, proteins sharing biological function and regulatory interactions, and physical protein interactions from the HPRD database [22].

C. Integrating direct and indirect association between miRNAs and targets

Here we refer to mRNAs (m) that have miRNA (miR) binding site in their 3'UTR as direct interaction. These interactions are extracted from PredNet. mRNAs that do not have miRNA binding site but its protein partners in the PPI network have miRNA binding site, are indirect interactions. Here we constructed a miRNA-target network ($miRNet$) as follows:

$$miRNet(m, miR) = \begin{cases} 1 & \text{if PredNet(m,miR)=1,} \\ & \text{OR} \\ & \text{if PredNet(m,miR)=0 and PredNet(k,miR)=1} \\ 0 & \text{other cases} \end{cases} \quad (1)$$

where k is set of m partners. We only considered indirect interactions between m and miR that have k of size greater than 10. This means we considered miRNAs that affect proteins indirectly through at least 10 protein partners.

D. Gene lists to assess model performance

To assess the performance of our proposed regression model, we used affymetrix gene expression data from LNCaP cell lines treated with pre-miR-1, pre-miR-27b and pre-miR-206 that was retrieved from [23] under the access number GSE31620. Significant analysis of micorarray (SAM) [24] was used to identify differentially expressed genes. 80 genes were identified to be upregulated after pre-miR-1 treatment and 88 were down regulated. 62 were upregulated after pre-miR-206 treatment and 83 were down regulated. 157 genes were upregulated after pre-miR-27b treatment and 51 were down regulated. We used another set of downregulated genes after treating HeLa cells with miR-1, miR-124 or miR-373. We also constructed a test dataset by extracting the targets of 11 prostate miRNA that have important role in prostate cancer development [2] from both ExpNet (460) targets and PredNEt (1854). We also used a publicly available mRNA expression from Taylor paper deposited in NCBI (GSE21032) to identify top downregulated genes in prostate cancer; 480 genes were identified as down regulated and 51 as upregulated using SAM.

E. Regression modeling for miRNA enrichment

We used PredNet and ExpNet miRNA-target interaction networks to determine the initial variables (miRNAs) in the regression model. Each variable represents the influence of miRNA on all targets in miRNA-target interactions. Let $GeneSignature$ represents the gene signature of a disease, $PredNet_j$ the miRNA-target influence profile of miRNA (j) on all target genes. β_j the strength of the impact of miRNA

(j) on gene signature.β_j is used to assess the enrichemtn of miRNA (j) targets in the gene signature. The proposed regression model can be written as follows:

$$GeneSignature = \sum_{j=1}^{miR} PredNet_j * \beta_j + \lambda P_\alpha(\beta) \quad (2)$$

where

$$P_\alpha(\beta) = \sum_{j=1}^{miR} [\frac{1}{2}(1-\alpha)\beta_j^2 + \alpha|\beta_j|] \quad (3)$$

is the elastic-net penalty. $P\alpha$ is a compromise between the ridge regression penalty (α =0), and the lasso penalty (α =1). This penalty is particularly useful when there are many correlated predictor variables as in the case of miRNAs. β is the regression coefficient of each variable, which indicates how each miRNA explains the gene signature. λ is a factor that determines the sparsity of the solution, as λ increases, the number of nonzero components of β decreases.

We next integrated protein protein interactions into the regression model to take into account the indirect influence of miRNA on targets. We constructed the indirect influence of miRNA on targets by counting the number of paths of length 2 between a miRNA and each protein target. This represents a miRNA that targets a modulator protein that interacts with the target proteins. We only considered paths that have 10 protein modulators. Interactions of paths length equals 1 between miRNA and protein target are considered as direct target. We rewrite the model as:

$$GeneSignature = \sum_{j=1}^{miR} \beta_j * miRNet_j + \lambda P_\alpha(\beta) \quad (4)$$

where $miRNet_j$ is the joint direct and indirect influence of $miRNA_j$ on GeneSignature, p is the total number of proteins in the PPI network. In this model, β represent enrichment of each miRNA targets in the disease gene signature. An overall description of the framework followed in this work to identify influential miRNAs is displayed in Figure 1.

III. Results

A. Parameter optimization

To optimize α and λ, we first selected 20 alpha values ranging from 0 (ridge regression) to 1 (lasso regression) and for each α we evaluated 100 values of λ. To select optimal λ for each α and fit the elastic-net regression, we conducted 10-fold cross validation and chose λ with the minimum mean cross-validated error (λ-min). In case, different λ values have equal mean cross-validated error, we select the maximum λ. We selected α=0.6 as λ-min values started to get steady. After selecting α as 0.6, we evaluated 100 values of λ to select the optimal one. We used glmnet implementation in matlab from http://www-stat.stanford.edu/ tibs/glmnet-matlab/ to find miRNA coefficients.

B. Proposed regression model is able to identify influential miRNA after pre-miRNA treatment

We first assessed the performance of the proposed regression method using several gene lists reported by recently published studies that used microarray analysis to reveal genes whose expression is affected by miRNA treatment. For example, in [23] LNCaP cell lines were treated with pre-miRNA (pre-miR-1, pre-miR206, and pre-miR27b) and downregulated genes were identified using differential gene expression analysis. In another study [25], miR-1, miR-124, and miR-373 RNA duplexes were transfected into HeLa cells and gene expression profiling was conducted to identify downregulated genes as a result of miRNA treatment. The downregulated gene lists are used to evaluate the performance of enrichment tools. Full summary of the gene lists used in this study is shown in Table I

We used PredNet miRNA-gene interactions as variable set to train the regression model and the down regulated gene set as response in the regression model. miRNA coefficients from the regression model are used to assess the enrichment of miRNAs in the gene set. In the pre-miR-1 downregulated genes, the regression model ranked miRNA-1 first with the highest coefficient value. In the pre-miR-206 downregulated genes, the regression model showed that miR-1 and miRNA-206 have the highest coefficient that explains 25% of the downregulated genes. In the downregulated genes after miR-27b treatment, the model showed that miRNA-9 has the highest coefficient and miRNA-27b ranked second. We compared the results of the proposed model with Geneset2miRNA and Expression2Kinases tools that find miRNA enrichment in gene sets. We then used the miRNA-mRNA network that combined both direct and indirect interactions (miRNet) to predict miRNA influence on gene signatures. Here we anticipate that gene signature includes both direct targets of miRNAs and secondary targets (indirect) that are influenced by the direct target. PPI-based regression model showed to perform as good as regression model that does not consider PPI, but it showed to have stronger coefficients for the targeting miRNAs. Table II summarizes the comparison among the four methods. The four methods showed to rank the overexpressed miRNA first.

We then used another set of downregulated genes after treating HeLa cells with miR-1, miR-124 and miR-373. Regression models accurately identified the targeting miRNAs as high influential miRNAs with very strong coefficients (0.9). On the other hand, The other two methods ranked the targeting miRNAs second (Table II).

We next analyzed miRNA enrichment in the whole set of differentially expressed genes in LNCaP cells. Expression2Kinase and regression models ranked miR-1 and miR-206 first, but miR-27b was ranked 10th in Expression2Kinase, 3rd by regression model, 15th using GeneSet2miRNA and 2nd by PPI based regression model.

The results of our method demonstrated that it is able to infer correct miRNAs from gene lists downregulated after pre-miRNA treatment.

TABLE I

SUMMARY OF GENE LISTS USED IN THIS STUDY TO VALIDATE THE PERFORMANCE OF PROPOSED METHOD AND COMPARE WITH EXISTING ALGORITHMS

Experiment description	Enriched miRNAs	Number of reported genes	Reference
HeLa cells transfected with miR-1	*miR-1*	*96 repressed genes*	*[25]*
HeLa cells transfected with miR-124	*miR-124*	*174 repressed genes*	*[25]*
HeLa cells transfected with miR-373	*miR-373*	*65 repressed genes*	*[25]*
LNCaP cells treatedwith pre-miR-1	*miR-1*	*88 repressed genes and 80 up-regulated genes*	*[23]*
LNCaP cells treatedwith pre-miR-206	*miR-206*	*83 repressed genes and 62 up-regulated genes*	*[23]*
LNCaP cells treatedwith pre-miR-27b	*miR-27b*	*51 repressed genes and 157 upregulated genes*	*[23]*
Genes predicted to be targets of 11 prostate miRNAs extracted from PRedNet	*miR(1, 204, 205, 143, 145, 221, 222, 27b, 133b, 31, let-7*	*1854*	*[2], [26], [27]*
Altered genes in prostate cancer using Taylor data		*480 downregulated in prostate and 51 upregulated in prostate*	*[28]*

TABLE II

RANK OF IDENTIFIED ENRICHED MIRNAS IN GENE LISTS DOWNREGULATED AND DIFFERENTIALLY EXPRESSED GENES AFTER MIRNA TREATMENT

	PPI-based regression model	Regression model	Expression2Kinase	GeneSet2miRNA
Downregulated gene set in LNCaP cells				
pre-miRNA-1	1^{st}	1^{st}	1^{st}	1^{st}
pre-miRNA-206	1^{st}	1^{st}	1^{st}	2^{nd}
pre-miRNA-27b	1^{st}	2^{nd}	1^{st}	1^{st}
Differentially expressed gene set in LNCaP				
pre-miRNA-1	1^{st}	2^{nd}	1^{st}	2^{nd}
pre-miRNA-206	1^{st}	1^{st}	1^{st}	2^{nd}
pre-miRNA-27b	2^{nd}	3^{rd}	10^{th}	15^{th}
Downregulated gene set in HeLa cells				
miRNA-1	1^{st}	1^{st}	1^{st}	2^{nd}
miRNA-124	1^{st}	1^{st}	2^{nd}	2^{nd}
miRNA-373	1^{st}	1^{st}	2^{nd}	2^{nd}

TABLE III

ENRICHMENT OF THE 11 MIRNAS IN GENE TARGET SET USING THREE METHODS. $\surd$ INDICATES THAT THE MIRNA IS IDENTIFIED IN THE TOP 11 ENRICHED MIRNAS

	PPI-based regression model	regression model	Expression2Kinase	GeneSet2miRNA
miR-1	$\surd$	$\surd$	$\surd$	$\surd$
miR-204	$\surd$	$\surd$	$\surd$	$\surd$
miR-143	$\surd$	$\surd$	$\surd$	$\surd$
miR-145	$\surd$	$\surd$	$\surd$	$\surd$
miR-205	$\surd$	$\surd$	$\surd$	$\surd$
miR-221	$\surd$	$\surd$	$\surd$	
miR-31	$\surd$	$\surd$		
miR-27b	$\surd$	$\surd$		$\surd$
Let-7a	$\surd$	$\surd$		
miR-133b	$\surd$	$\surd$		
miR-222	$\surd$			$\surd$

C. Proposed regression model correctly identified the 11 miRNA prostate signature

We identified a set of prostate miRNAs that showed to play a role in prostate cancer (Table III). We extracted their targets (1854 targets) from PredNet and then applied the four methods to predict enriched miRNAs in the target set. Table III shows the methods that were able to rank the targeting miRNA in the top 11. Results showed that regression models correctly infer miRNAs in gene lists enriched with multiple miRNAs.

D. Proposed regression model predicted miRNAs implicated prostate cancer

After evaluating the performance of the regression models on gene sets repressed after miRNA treatment, and accurately inferring the treating miRNA, we applied it to 480 down-regulated genes in prostate cancer that were extracted from Taylor data using SAM. We used both regression models and Expression2kinase to find miRNA enrichment. There was only one common miRNA (miR-30a-5p) between the three sets of miRNAs with no significant enrichment. However, regression models identified several miRNAs that have potential role in prostate cancer. miR-146b, miR-206, miR-212, miR-223, miR-29b and miR-409-3p showed to be enriched in downregulated genes. miR-146b and miR-206 have already shown a role in prostate cancer by targeting ROCK1 [29] and HDAC4 [23] respectively.

We analyzed miRNA enrichment in the upregulated gene set (51 genes) and miR-16-1 was found by the proposed regression model as the most significant enriched miRNA. PPI-based regression model identified miRNA-16-1 as the most significant and miR-222, miR-338 and miR-34c as less significant miRNAs. Using Expression2Kinase to identify the

enriched miRNAs in the 51 gene set, miR-143 and miR-217 were found significant (p=0.02) but miR-16-1 was not identified as significantly enriched.

E. Validating predicted miRNAs using independent miRNA expression data

After identifying the top 10 enriched miRNAs in genes downregulated prostate cancer, we extracted their expression from Taylor miRNA expression sets to assess their diagnostic role in prostate cancer. Using SVM to assess the ability of the miRNAs expression to classify cancer vs normal samples, we found miRNAs extracted using regression model classify samples with 90% accuracy and 85% for miRNAs extracted by Expression2Kinase.

IV. Discussion

Gene expression profiling emerged as a routine protocol to characterize dysregulated pathways. Characterizing the functional role of miRNAs from gene expression profiling lead to developing several tools employing miRNA gene functional annotations. Most of these methods are based on assessing the statistical overlap between two gene sets like fisher test and hypergeometric test. In this work we proposed an elastic-net regression model to predict miRNAs that have high influence on gene list with potential to be enriched in gene lists. We first used miRNA-mRNA network that only includes direct targets of miRNAs and then combined it with miRNA-mRNA network that includes indirect association between miRNAs and targets. Here we aim to assess the power of including the downstream partners of miRNA targets on miRNA enrichment analysis.

Assessing the performance of miRNA enrichment tools and algorithms is challenging due to lack of benchmark data. In here we used gene expression data from LNCaP cell lines treated with three miRNAs (miR-1, miR-206, and miR-27b) and HeLa cell lines treated with miR-1, miR-124 and miR-373. We anticipate that downregulated genes after miRNA treatment are regulated directly or indirectly by treating miRNAs. Thus we evaluated the performance of the methods based on the rank of the miRNAs, the higher the rank of miRNA, the more influential it is. Based on the results in Table II, we showed that regression models demonstrated a proof of concept for miRNA enrichment.

The four methods (two existing methods, two proposed) ranked miR-1 and miR-206 first when using the downregulated genes after miRNA treatment, however , some methods disagreed on miR-27b rank as shown in Table II. This could be due to the size of targets of miRNAs. We showed that the bigger the set of miRNA targets, the better the performance of the methods. However, regression based model showed to be not sensitive to the size of miRNA targets. When we analyzed the enriched miRNAs in repressed genes in HeLa cells we found that regression models accurately predicted the miRNA transfected, however, Expression2Kinase ranked miR-124 second in HeLa cells treated with miR-124 and ranked miR-506 first. Same with HeLa cells treated with miR-373, Expression2Kinase ranked miR-373 second after miR-520. Geneset2miRNA also ranked the transfected miRNAs second. When we looked at miRNA families, we found that miR-1 and miR-206, miR-124 and miR-506, and miR-373 and miR-520 are from the same family and target the same targets. This suggest that methods based on statistical overlap significance are sensitive to redundant miRNA sets, however, regression based methods are not due to the elastic net penalty that gives high penalty for redundant sets.

To assess the performance of the methods on noisy gene lists, we run the methods on differentially expressed genes (both downregulated and upregulated genes). We found that all methods still rank miR-1 and miR-206 high but non-regression methods were unable to rank miR-27b high. This is because miRNA-1 and miR-206 are from the same family and they target large portion of genes (200 from ExpNet and 500 from PredNet) and miR-27b target smaller number of genes. These results demonstrated that all methods perform accurately on clean gene lists that only includes either direct targets or secondary targets. On the other hand, the non-regression method showed to be very sensitive to noisy gene lists that contain off targets. Results also showed that non-regression methods are sensitive to the number of miRNA targets, and they performed poorly on miR-27b as it targets small number of genes. Since we do not always have clean gene lists, we prepared another test gene list of length 1854 that contains targets of 11 miRNAs that have been studied in the literature to play a role in prostate cancer [2], [29]. miRNA-1 target almost 500 genes, miR-204 and miR-203 target about 200 genes. We aimed to assess the ability of the methods to identify enriched miRNAs in gene sets regulated by multiple miRNAs. This test gene list demonstrated two main results. First, regression models simultaneously identify enriched miRNAs, unlike the other two methods that test against each miRNA set individually. Second, regression based models are able to handle enrichment analysis of genes enriched with multiple miRNAs.

We finally tested the methods against genes that are differentially expressed in prostate cancer. This gene set is both noisy as it contains secondary targets of miRNAs and it is anticipated to be enriched in multiple miRNAs as it is argued that several miRNAs do play a regulatory role in prostate cancer progression. Both regular regression and PPI-based regression method identified miR-16-1 as a tumor suppressor that target the upregulated genes in prostate cancer. PPI-based regression showed to generate stronger coefficient of miR-16-1 compared to regular regression. miR-16-1 showed to be downregulated in advanced prostate cancer and a prognostic marker that is in clinical trials [2]. Several studies have shown that miR-16-1 is downregulated in advanced prostate cancer and it targets BCL2,CCND1, and WNT3A genes [29]. Down regulation of miR-16-1 lead to overexpression of their target BCL2,CCND1 and WNT3A that leads to increased survival and invasion. This suggests that miR-16-1 is tumor suppressor that have significant therapeutic potential. Delivery of miR-16-1 to prostate cancer xenografts caused tumor regression [2]. Unfortunately miR-16-1 expression was not found in Taylor

data nor in GSE23022 data. miR-143 also showed to be downregulated in prostate cancer and acts as anti-metastatic miRNAs by targeting ERK5 gene [29].

These findings showed a promise of using regression models for miRNA enrichment analysis in particular and gene set enrichment analysis in general. This could give us more insight on the functional characterization of gene lists and this model can be applied to infer disease miRNA functional association as future direction.

V. Conclusion

Interpreting gene lists resulted from new experiments is becoming a crucial task to understand dysregulation in biological processes. We proposed a regression model with elastic-net penalty to predict miRNAs with high influence on gene lists. The proposed regression model succeeded in proof-of-concept experiments and showed to be not sensitive to miRNA target set nor to noisiness in the gene set. It also provided evidence that it is not affected by redundant gene sets and it uses elastic net penalty to eliminate redundant sets. The proposed model demonstrated that modeling gene enrichment problem as regression model is effective and promising to identify miRNAs associated with diseases. The proposed method predicted miR-16-1 as a tumor suppressor and miR-146b as an oncogene in prostate cancer.

Acknowledgment

Authors would like to thank NSERC, Alberta Innovates and University of Calgary for funding.

References

[1] G. Ruvkun, "Molecular biology: Glimpses of a tiny rna world," *Science*, vol. 294, pp. 797–799, 2001.

[2] A. Gordanpour, R. Nam, L. Sugar, and A. Seth, "Micrornas in prostate cancer: from biomarkers to molecularly-based therapeutics," *Prostate Cancer Prostatic Dis*, vol. 1, p. 6, 2012.

[3] V. Ambros, "The function of animal mirnas," *Nature*, vol. 431, pp. 350–355, 2004.

[4] Y. Pang, C. Young, and H. Yuan, "Micrornas and prostate cancer," *Acta Biochim Biophys Sin*, vol. 42, pp. 363–369, 2010.

[5] A. Watahiki, Y. Wang, J. Morris, K. Dennis, H. ODwyer, M. Gleave, P. Gout, and Y. Wang, "Micrornas associated with metastatic prostate cancer," *PLoS One*, vol. 6, p. e24950, 2011.

[6] L. He and G. Hannon, "Micrornas: small rnas with a big role in gene regulation," *Nature Reviews Genetics*, vol. 5, p. 631, 2004.

[7] G. Calin and C. Croce, "Microrna signatures in human cancers," *Nature Reviews Cancer*, vol. 6, pp. 857–866, 2006.

[8] A. Esquela-Kerscher and F. Slack, "Oncomirs microrna with a role in cancer," *Nature Reviews Cancer*, vol. 6, pp. 259–269, 2006.

[9] J. Cho, R. Gelinas, K. Wang, A. Etheridge, M. Piper, K. Batte, D. Dakhallah, J. Price, D. Bornman, S. Zhang, C. Marsh, and D. Galas, "Systems biology of interstitial lung diseases: integration of mrna and microrna expression changes," *BMC Medical Genomics*, vol. 4, p. 8, 2011.

[10] S. Sevli, A. Uzumcu, M. Solak, M. Ittmann, and M. Ozen, "The function of micrornas, small but potent molecules, in human prostate cancer," *Prostate Cancer Prostatic Dis*, vol. 13, pp. 208–217, 2010.

[11] A. Antonov, S. Dietmann, P. Wong, D. Lutter, and H. Mewes, "Geneset2mirna: finding the signature of cooperative mirna activities in the gene lists," *Nucleic Acids Res*, vol. 37, pp. W323–W328, 2009.

[12] E. Chen, H. Xu, S. Gordonov, M. Lim, M. Perkins, and A. Ma'ayan, "Expression2kinases: mrna profiling linked to multiple upstream regulatory layers." *Bioinformatics*, vol. 28, pp. 105–111, 2012.

[13] A. Subramanian, P. Tamayo, V. Mootha, S. Mukherjee, B. Ebert, M. Gillette, A. Paulovich, S. Pomeroy, T. Golub, E. Lander, and J. Mesirov, "Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles," *PNAS*, vol. 102, pp. 15 545–15 550, 2005.

[14] N. Bartonicek and A. Enright, "Sylarray: a web server for automated detection of mirna effects from expression data." *Bioinformatics*, vol. 26, pp. 2900–2901, 2010.

[15] D. Huang, B. Sherman, and R. Lempicki, "Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists," *Nucleic Acids Res*, vol. 37, pp. 1–13, 2009.

[16] C. Hsu, H. Juan, and H. Huang, "Characterization of microrna-regulated protein-protein interaction network," *Proteomics*, vol. 8, pp. 1975–1979, 2008.

[17] M. Sualp and T. Can, "Using network context as a filter for mirna target prediction," *Biosystems*, vol. 105, pp. 201–209, 2011.

[18] A. Grimson, K. Farh, W. Johnston, P. Garrett-Engele, L. Lim, and D. Bartel, "Microrna targeting specificity in mammals: Determinants beyond seed pairing," *Molecular Cell*, vol. 27, pp. 91–105, 2007.

[19] S. Hsu, F. Lin, W. Wu, C. Liang, W. Huang, W. Chan, W. Tsai, G. Chen, C. Lee, C. Chiu, C. Chien, M. Wu, C. Huang, A. Tsou, and H. Huang, "mirtarbase: a database curates experimentally validated microrna target interactions," *Nucleic Acids Res*, vol. 39, pp. D163–D169, 2011.

[20] F. Xiao, Z. Zuo, G. Cai, S. Kang, X. Gao, and T. Li, "mirecords: an integrated resource for microrna target interactions," *Nucleic Acids Res*, vol. 37, pp. D105–D110, 2009.

[21] G. Wu, X. Feng, and L. Stein, "A human functional protein interaction network and its application to cancer data analysis," *Genome Biology*, vol. 11, p. R53, 2010.

[22] S. Peri and a. et, "Human protein reference database as a discovery resource for proteomics," *Nucl. Acids Res*, vol. 32, pp. D497–D501, 2004.

[23] R. Hudson, M. Yi, D. Esposito, S. Watkins, A. Hurwitz, H. Yfantis, D. Lee, J. Borin, M. Naslund, R. Alexander, T. Dorsey, R. Stephens, C. Croce, and S. Ambs, "Microrna-1 is a candidate tumor suppressor and prognostic marker in human prostate cancer," *Nucleic Acids Res*, pp. 1–15, 2011.

[24] V. Tusher, R. Tibshirani, and G. Chu, "Significance analysis of microarrays applied to the ionizing radiation response," *Proc Natl Acad Sci*, vol. 98, pp. 5116–5121, 2001.

[25] L. Lim, N. Lau, P. Garrett-Engele, A. Grimson, J. Schelter, J. Castle, D. Bartel, P. Linsley, and J. Johnson, "Microarray analysis shows that some micrornas downregulate large numbers of target mrnas," *Nature*, vol. 433, pp. 769–773, 2005.

[26] M. Ozen, C. Creighton, M. Ozdemir, and M. Ittmann, "Widespread deregulation of microrna expression in human prostate cancer," *Oncogene*, vol. 27, pp. 1788–1793, 2008.

[27] K. Porkka, M. Pfeiffer, K. Waltering, R. Vessella, T. Tammela, and T. Visakorpi, "Microrna expression profiling in prostate cancer," *Cancer Research*, vol. 67, p. 6130, 2007.

[28] B. Taylor, N. Schultz, H. Hieronymus, A. Gopalan, Y. Xiao, B. Carver, V. Arora, P. Kaushik, E. Cerami, B. Reva, Y. Antipin, N. Mitsiades, T. Landers, I. Dolgalev, J. Major, M. Wilson, N. Socci, A. Lash, A. Heguy, J. Eastham, H. Scher, V. Reuter, P. Scardino, C. Sander, C. Sawyers, and W. Gerald, "Integrative genomic profiling of human prostate cancer," *Cancer Cell*, vol. 18, pp. 11–22, 2010.

[29] S. Saini, S. Majid, and R. Dahiya, "Diet, micrornas and prostate cancer," *Pharm Res*, vol. 27, pp. 1014–1026, 2010.