In Silico Approach to Construct miRNA mRNA Module

Roma Chandra*[1], Sumit Kumar[*]

*Shri Venkateshwara University, Gajraula, Amroha, Uttar Pradesh

[1]IILM College of Engineering & Technology, Greater Noida

Abstract: **miRNAs are regulators as they cause mRNA cleavage playing a major role in RNA interference pathway. They have roles in various biological processes. The aim is to study miRNA and their associated genes which further can be associated to drugs in next module to study pharmacogenomics. miRNA mRNA modules were obtained which are combinations of miRNAs and mRNAs that represents similar biological functions. In silico study using GSEA was done.**

Keywords: **miRNA, miRNA mRNA module, drugs.**

I. INTRODUCTION

The idea about miRNA came in 1993 while the term was coined in 2001. [1,2]They are about a length of 20-25 nucleotides and are non coding RNA with involvement in various biological processes[3,4]. They have major roles in transcription and translation as they cause regulation of genes[5,6]. miRNA play roles both as promoter and inhibitor and are related to many plant and animal diseases. Gene regulation is a common area of research which involves study of miRNAs targeting multiple genes. These miRNAs are responsible for the regulation of target genes along with functions like development and diseases. Approaches involving development of modules using expression data included construction of bipartite graph, probabilistic optimization technique, rule induction method, etc. The study included a new approach using GSEA software for construction of relationship between miRNA and their target genes to attain the objective of miRNA mRNA module construction.

II. MATERIALS AND METHODS

2.1.) Developing a Problem

As the aim was to construct module representing relationship between miRNA and their target genes, thus primary goal was to construct dataset of both miRNA and mRNA respectively. miRNA and mRNA expression data set files were constructed which represented the expression score values of both miRNA and mRNA during normal and cancerous condition. A new file representing miRNA target binding score [7,8] was also constructed.

2.2.) Preparing Dataset

miRNA expression data set file represented expression of 217 genes while mRNA expression data set file represented expression of 16,063 genes. The samples used included 68 tumor and 13 normal types including 11 tissues (colon, kidney, ovary, uterus, pancreas, bladder, breast, lung, mesoderm, prostate).mRNA dataset file was represented by their probe sets.Figure 1 represents part of mRNA expression dataset file and Figure 2 represents part of miRNA expression dataset file.

i.) mRNA expression dataset file:

Probe set	Symbol	N_COLON	T_COLON	T_PAN	T_KID	T_BLDR	N_PROST	T_OVARY
A28102_at	GABRA3	9.32291	5.57077	6.02803	6.96991	8.61192	5.97093	7.81616
AB000114_at	OMD	9.1426	5	5	5	8.62137	5	6.15925
AB000115_at	C1orf29	8.80008	5.65317	6.90311	7.69415	9.12905	6.78375	8.19336
AB000220_at	SEMA3C	9.27447	7.15842	7.66432	6.14831	9.12983	8.30444	7.20462
AB000381_s_at	GML	5	5	5	6.2507	5.3724	5	5

Figure 1: Part of miRNA expression dataset file

ii.) miRNA expression dataset file:

Both the above files were constructed in gene matrix text file format. The third file representing miRNA target site matrix was constructed using miRNA target site scores which were obtained using PICTAR.

miRNA	NC	TC	NP	TP	NK	TK	NB	TB
hsa-let-7d	7.3	9.8	9.5	7.8	9.5	8.1	9.4	9.3
hsa-let-7e	7.2	9.3	8.0	6.8	9.6	8.1	8.2	8.6
hsa-miR-1	11.4	11.3	8.0	5.6	7.6	5.0	7.7	5.7
hsa-miR-101	9.5	8.9	9.4	9.2	11.1	7.5	9.8	10.1
hsa-miR-103	8.8	9.6	8.8	9.0	10.9	6.7	9.6	10.9

Figure 2: Part of miRNA expression profile dataset

iii.) miRNA target binding score matrix:

miRNA	APC	COL1A2	ESR1	FGFR3	GCH1	JAG1	PKD2	PTEN	BLMH
hsa-let-7d	0	3.06	0	0	0	0	0	0	0
hsa-let-7e	0	2.53	0	0	0	0	0	0	0
hsa-miR-1	0	0	0	0	3.18	0	0	0	0
hsa-miR-101	0	0	0	0	0	0	0	0	0
hsa-miR-103	0	0	0	0	0	0	0	0	2.95
hsa-miR-106b	0	0	0	0	0	0	3.03	2.34	0

Figure 3: Part of miRNA target binding score matrix

Correlation was studied to obtain relationship between target genes using data analysis tool XLSTAT.

2.3.) Correlation

Correlation represents relationship among miRNA and mRNA expression datasets. miRNA mRNA correlation matrix was constructed using both miRNA expression profile and mRNA expression profile. Now the miRNA mRNA correlation matrix is combined with miRNA target binding score matrix to develop miRNA mRNA module matrix. This was formed between the miRNA and mRNA datasets with relation showing to those sets which have positive correlation and and the target binding score. Then to filter out the miRNA mRNA combinations the data was undergone GSEA analysis.

2.4.) GSEA Analysis

GSEA or Gene Set Enrichment Analysis is an in silico method to find out whether the given genes are statistically significant for the different biological states. GSEA analysis was done using prepared dataset files. The results represented in a relational representation between miRNA and target genes. Further the fitness function was obtained for the calculated miRNA mRNA combinations.

2.5.) Calculate Fitness Function:

Fitness of a module (x, y) can be measured as following:

$F(A, B') = xBS_{A'B'} + yEC_{A'} + zEC_{B'} + VOLUME$

- $BS_{A'B'}$ is binding score mean of miRNA target binding scoring matrix subset, consisting of A', B'.
- $EC_{A'}$ and $EC_{B'}$ are Pearson correlation mean from all the possible pairs of miRNA and mRNA for the subset taken and is known as expression coherence (EC) score.
- Where $VOLUME = w(w_a(P'_a/P_a) + w_b(P'_b/P_b))$
- P'_a and P'_b are the subset size
- w, w_a and w_b have values 0.1,0.5,0.5 respectively
- x, yand , z have values 0.6,0.3,0.1 respectively

Modules	$BS_{A',B'}$	$EC_{A'}$	$EC_{B'}$	Fitness
hsa-mir-143,hsa-mir-181a,NOVA1,ZFP36l1	2.1925	0.481718	0.438387	0.26
hsa-mir-125b,hsa-mir-145,DAG1,YES1,BMPR2,PTPRF	0.1498	0.070905	-0.0653	0.1053
hsa-mir-27a,hsa-mir-143,NOVA1,CDH5,ADD3	0.1848	0.77441	0.310605	0.375
hsa-mir-149,hsa-mir,29a,BCL2L2,PLAG1,SP1,CBX1	0.1672	0.332532	0.149044	0.216
hsa-mir-17-5p,hsa-mir-25,CIC,EDG1,SSFA2,PCAF,SALL1	0.167	0.783068	0.135963	0.356
hsa-mir-134,hsa-mir-15a,KPNA3,EPHA7	0.1648	0.272878	0.140155	0.202
hsa-mir-198,hsa-mir-30e,MAPRE1,NCOR2,NRIP1	0.2583	0.499113	0.170378	0.322
hsa-mir-101,hsa-mir-190,HAS2,PPP3R1,DAG1	0.1355	0.519051	0.58623	0.244

Figure 4: Some of miRNA mRNA modules and their calculated fitness

The above is the list of miRNA mRNA modules which were obtained after calculation of fitness function.

III. RESULTS AND DISCUSSION

3.1) miRNA mRNA modules:

Information obtained from literature study showed that module consisting of miR-212, miR-132, HIC1, OVCA2, BCL6, miR-212 and miR-132 are present at the same location on the chromosome within 300 bp which may be on a transcript (polycistronic miRNAs)[9,10]. Their presence was seen on a loss of heterozygosity (LOH) region as noticed in hepatocellular carcinomas. Tumor suppressor genes were present in upstream regions of hypermethylated in cancer 1 (HIC1) and downstream regions of ovarian cancer gene 2 (OVCA2). Thus, tumor suppression involves polycistronic miRNAs. Gene miR-127 was down regulated in 75% of human cancer cells. It was treated by chromatin remodelling drugs and then induced which finally resulted in down regulation of B cell CLL/lymphoma 6 (BCL6) thus showing its role as tumor suppressor in combination to miR-212 and miR-132. Literature study and GSEA analysis results were compared from where few of the mentioned modules were obtained.

S.no.	miRNA mRNA module
1	hsa-miR-143,hsa-miR-181a,NOVA1,ST8SIA4,ZFP36L1
2	hsa-miR-125b,hsa-miR-145,DAG1,NEDD9,YES1,BMPR2,PTPRF
3	hsa-miR-126,hsa-miR-181b,NOVA1,PCAF,EIF4A2
4	hsa-miR-212,hsa-miR-132,HIC1,OVCA2,BCL6
5	hsa-miR-27a,hsa-miR-143,NOVA1,CDH5,ADD3
6	hsa-miR-101,hsa-miR-19a,hsa-miR-221,ATXN1,CTCF,RAB1A
7	hsa-let-7e,hsa-miR-26a,ARID3A,TAF5,HAS2,NOVA1,AKAP6,DYRK1A
8	hsa-miR-149,hsa-miR-29a,BCL2L2,PLAG1,SP1,CBX1
9	hsa-miR-17-5p,hsa-miR-25,CIC,EDG1,SSFA2,PCAF,SALL1
10	hsa-miR-134,hsa-miR-15a,KPNA3,RUNX1T1,EPHA7

Figure 5: Some of the miRNA mRNA modules obtained

Study of tumor suppressor miRNAs in a module shows the oncogenic behaviour of target genes. Like for example target genes, EIF4A2, GUSB, ACVR2B were said to behave as oncogenic. EIF4A2 is a known relative in BCL6 translocation. On the other hand when GUSB was over expressed in human it resulted in greater tumor susceptibility. MSI-H colorectal cancer originates due to activin signalling loss caused when ACVR2 mutated.

Genes RPL34, RPL13A and SNRPD3 were found to be involved both in malignant and non malignant tumor. RPL13A have a significant expression in prostate cancer tissue. Genes present in this module were involved in biological functions. They showed involvement in metabolism when gene ontology was studied while on literature study it was found that they were involved in cancer. Thus, by suppressing the tumor causing modules or by using tumor suppressive module cancer can be regulated.

IV. CONCLUSION

miRNAs behave both as tumor suppressors and oncogenes. Oncogenes (ex: miR-21 and miR-17-92) are those genes whose increased expression or improper activation causes oncogenesis. On the other hand tumor suppressive genes (ex: let-7 and miR-34) are those genes that protect cells from being cancerous but their inactivation definitely lead to cancer. miRNAs regulate mRNAs causing gene regulation. Thus, module construction of miRNA and mRNA provide knowledge about how they are involved in various biological processes. miRNA basically act on pathways regulating genome and gene products. Regulated genes are both oncogenic and tumor suppressive which are supposed to be drug targets by various pharmaceutical and biotechnology industry. As miRNAs have multiple target genes thus they have high therapeutic role.. miRNA with their target mRNA could be studied for their integrative function by the formation of miRNA mRNA module from their expression values and target binding scores.Now considering the functionally similar miRNA mRNA modules, miRNAs could be obtained which are related to the genes which are part of these pathways showing the same effect of miRNA as the genes have. miRNAs are so small but they act just in a huge way to obtain such regulatory networks. Gene regulation caused by these miRNAs helps in targeting several tumor causatives. Thus, miRNAs are important in therapeutics helping to avoid various diseases including not only cancers but various heart disease like cardiac hypertrophy, neurological diseases like Alzheimer's syndrome, regulating cholesterol, causing immunodeficiency and regulating viral diseases.

REFERENCES

[1]. Lee RC, et al. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell. 1993;75:843–54.

[2]. Reinhart B, et al. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature. 2000; 403: 901–6.

[3]. Ambros V. The functions of animal microRNAs. *Nature*. 2004; 431:350–5

[4]. Bartel D. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*.2004; 116:281–97.

[5]. Carthew, R.W. Gene regulation by microRNAs. *Curr. Opin. Genet.Dev*. 2006, *16*: 203-8.

[6]. Guarnieri, D.J., DiLeone, R.J. MicroRNAs: a new class of gene regulators. *Ann. Med*. 2008, *40*: 197-208.

[7]. John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS. Human MicroRNA targets. PLoS Biol. (2004) 2:e363.

[8]. Betel D, Wilson M, Gabow A, Marks DS, Sander C. The microRNAorg resource: targets and expression. Nucleic Acids Res. 36: D149-D153 (2008)

[9]. Medina, P.P., Slack, F.J. microRNAs and cancer: An overview.*Cell Cycle*. 2008, *17*: 7-16.

[10]. Zhang, B., Pan, X., Cobb, G.P., Anderson, T.A. microRNAs as oncogenes and tumor suppressors. *Dev. Biol*. 2007, *302*: 1-12.