ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

microRNA mRNA Module Prediction Leading to Microarray Data Analysis to Develop Gene Regulatory Network

Roma Chandra

[*] Department of Biotechnology, College of Engineering & Technology, IILM - AHL Greater Noida, UP, India

Abstract

microRNAs (miRNAs) are those small RNAs that not only recognize but also regulate mRNA target genes. Evidences are relevant enough to show that they are key regulators for various functions including development as well as diseases like cancer. They have a major role as regulator through RNA-interference pathway. The regulation occurs when mRNA is cleaved or has repressing transcription. Aim is to predict miRNA mRNA regulatory modules which consists of miRNAs and target genes that work together in gene regulation influencing biological processes. Primary step includes studying the combined effect of miRNAs and mRNAs on various cellular states in order to get the knowledge of those miRNAs and mRNAs which are closely related to each other. Individual as well as inter relational expression profiles, binding scores and correlations were studied. miRNA mRNA modules constructed were combinations of miRNA and mRNAs that represents similar biological functions. Fitness of these combinations was calculated followed by a probabilistic learning method. Most of the members in the module were reported to be related to cancer. Consequently the method provides source for miRNA and target sets found to be closely related parts in the gene regulatory pathways. Microarray data was analysed to obtain refined clusters which further constructs the gene regulatory networks. Signalling pathways resulted from the target genes for the miRNAs showed both activating and inhibitory roles. Microarrays are used to study the collective behaviour of genes. Genes control cell response to environmental changes through gene regulation. The most important use of gene regulatory network is as we study the relationship between genes a triggering mechanism is studied as well as they help to study genes for the disease treatment.

Keywords: miRNA, Gene regulation, Regulatory modules, Regulatory networks, Microarray analysis

Introduction

In 1993 Lee and his co-workers gave some idea about microRNA but the term was given in 2001. lin-4 and let-7 were the first microRNA discovered in *Caenorhabditis elegans* [1, 2]. MicroRNAs (miRNAs) are transcribed non coding RNA with a limited length of nearly 20 to 25 nucleotides found to be involved in numerous biological processes [3,4]. They affect the transcription and translation processes at gene level as they tend to bind to the 3′ untranslated regions of mRNA targets causing gene regulation [5, 6] due to either mRNA destruction or translational inhibition. Till recent studies miRNAs were considered to have negative effect on gene expression but now they are found to be involved positively[7].They have functions both in plants and animals from leaf development to larval development. They are causative of many diseases affecting nearly every organ of human body. They are tumor causing as well as tumor suppressor [8, 9].And they are found to be involved in various processes at cellular level like apoptosis, hematopoietic differentiation, t cell differentiation. They have an important part in the scenario of cancer, as around 50% are found on cancer related genomic sites thus showing their involvement in the disease [10].Gene regulation by miRNAs is really noticeable as a single miRNA targets many genes. MicroRNAs (miRNAs) recognize and regulate their target genes. They regulate various functions related to development and disease, including cancer. As it is difficult to define position and function of miRNAs in regulatory networks, approaches involving module network from expression data can be helpful. Module network is formed from miRNA mRNA data using a probabilistic optimization technique [11]. Concept of miRNA regulatory modules was first told by Yoon

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

and De Micheli. They are said to be combinations of miRNA and their target mRNAs having similar function and biological process defined with the help of weighted bipartite graph. miRNA targets are identified and relational graph are constructed with miRNA and seed binding to a common target are selected. Thus candidates are constructed for miRNA regulatory modules. miRNA mRNA modules were discovered by applying if then rule induction method, a machine learning technique on miRNA and mRNA expression profiles with target binding information. Rules were obtained from confidence and coverage concepts [12].Confidence defines positive class examples to number of all classes in dataset and coverage defines the positive class alone. Another method for module formation is probabilistic learning method. Module formed from expression profiles of miRNA, mRNA and target binding information. Then for set of combinations fitness was calculated. Best scoring ones were selected and probability was obtained iteratively till the maximum generations to produce modules.

Gene regulatory networks (GRN) contain genes which are connected to each other only when their expression is activated or inhibited by another genes expression. Proteins control cellular processes like reproduction, metabolism and responses. On the other hand every single protein is constructed by the help of a gene. Some of the proteins obtained from genes work as promoter or suppressor to other genes that are responsible for further protein formation. Hence, genes regulate other genes. Thus, we can say that GRN are on-off switches for cell that work at genetic level. A GRN consist of nodes and edges where nodes are genes and edges are directional representation of genes regulation or modulation. It involves mainly three objectives:

- miRNA mRNA modules prediction.
- Study of the microarray datasets and its analysis.
- Predicting gene regulatory network from the analysed microarray data.

miRNA mRNA modules prediction was the main objective as it correlates the miRNAs and mRNAs on biological basis. Microarray analysis was done to obtain cluster of related genes and then these were used to form the gene regulatory network. Thus, we can see a link is formed between miRNA to mRNA to regulatory network showing how each one is connected and how study of each is related biologically.

Materials and methods

Problem Formulation

Goal was to find miRNA–mRNA modules on the basis of miRNA and mRNA expression profiles and the miRNA target binding score matrix.

- $A = \{a_1, a_2, \ldots, a_{Pa}\}$, represents the set of miRNAs.
- $B = \{b_1, b_2, \ldots, b_{Pb}\}$, represents the set of mRNAs , where Pa= number of miRNAs and Pb= number of mRNAs.
- Subset (A', B') was supposed to be a module, where $|A'| \leq |A|$ and $|B'| \leq |B|$.
- Pa × Pas (Pb × Pbs) matrices were formed representing the expression profiles of miRNAs and mRNAs respectively.
- Pa × Pb represents miRNA target binding scoring matrix.

Thus modules were selected which contained miRNA and mRNA subsets.

Dataset Preparation

Gene cluster text files for mRNA expression for 16,063 genes (Affy chips) and microRNA expression for 217 microRNA genes (ad hoc bead-based profiling method) were taken. These expression data sets contained 89 human tissue samples (68 tumor and 13 normal) for 11 tissues (colon, pancreas, kidney, bladder, prostate, ovary, uterus, lung, mesoderm, breast.) Probe set were taken for mRNA data file. miRNA data files and mRNA data files were obtained as matrices as following:

Figure 1 shows the expression profile data of mRNA for different tumor samples and Figure 2 shows the expression profile data for miRNA for different samples (both normal & tumor type) from tissues like colon, pancreas, kidney, ovary , etc.

i.) mRNA expression profile dataset file:

Probe set	Symbol	TC2	TP2	TK1	TBL1	TPR1	TO4	TUT1
A28102_at	GABRA3	5.16662	6.99321	7.04974	6.79776	6.49894	6.46628	6.17345
AB000114_at	OMD	5	6.06779	5	6.10229	6.69284	6.65208	5.6753
AB000115_at	C1orf29	7.69092	7.03985	7.11246	7.5488	5.82759	8.36482	6.0533
AB000220_at	SEMA3C	6.16463	6.65429	5.027	8.8278	9.37305	6.93015	6.71709
AB000381_s_at	GML	5.4516	5	5.9561	5	5	6.379	5

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

Figure 1: Part of miRNA expression profile dataset

ii.) miRNA expression profile dataset file:

miRNA	NC3	TC3	NP1	TP1	NK1	TK1	NB1	TB1
hsa-let-7d	7.33001	9.77485	9.53409	7.75051	9.45428	8.14952	9.43228	9.32095
hsa-let-7e	7.16847	9.32795	7.99561	6.84044	9.55354	8.13731	8.22252	8.57449
hsa-miR-1	11.4035	11.3218	8.03782	5.56835	7.62402	5	7.69383	5.71716
hsa-miR-101	9.47113	8.94853	9.43007	9.23957	11.0572	7.53196	9.77271	10.0803
hsa-miR-103	8.83539	9.60971	8.75641	9.04829	10.8527	6.69919	9.58058	10.8846

Figure 2: Part of miRNA expression profile dataset

miRNA data files and mRNA data files were obtained as matrices.Now the third matrix , miRNA target site matrix was constructed. To form this matrix, miRNA target site scores were obtained by using PICTAR.

iii.) miRNA target binding score matrix:

	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

The process continued with the formation of miRNA and mRNA expression profile correlation matrices by using XLSTAT. XLSTAT is a complete solution for data analysis and visualization on MS Excel.

Correlation

Correlation represents how much a variable is related to another variable. Correlation matrix was developed using miRNA expression profile and mRNA expression profile. From the two matrices miRNA mRNA correlation matrix and miRNA target binding score matrix, the final matrix called the miRNA mRNA module matrix was constructed. 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 further filter out the miRNA mRNA combinations the data was undergone GSEA analysis.

GSEA Analysis

Gene Set Enrichment Analysis is a computational method by which we can know whether a known defined genes set is statistically significant with a definable difference between two biological states (e.g. phenotypes).GSEA analysis was done which resulted in the relational representation between miRNA and target genes which was represented by matrix.The leading edge analysis results shows about the relationship between miRNA and the target genes i.e. they are up regulated for either tumor or normal. So the tumor related genes and miRNAs were filtered. On considering both the miRNA mRNA module matrix and that obtained from the leading edge analysis, the

miRNA mRNA combinations were obtained. Further the fitness function was obtained for the calculated miRNA mRNA combinations.

Calculate Fitness Function:

Fitness of a module (A', B') can be measured as following:

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

- $BS_{A'B'}$ is the mean of the binding score of the subset of the miRNA target binding scoring matrix, consisting of A', B'.
- $EC_{A'}$ and $EC_{B'}$ are the mean for the Pearson correlation between all the possible pairs of miRNA and mRNA for the subset taken and is known as expression coherence (EC) score.
- Where $\text{VOLUME} = w\,(w_a\,(P'_a/P_a) + w_b\,(P'_b/P_b))$
- 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

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

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

After calculating the fitness function, the above miRNA mRNA modules were finally obtained.

Learning Algorithm

In order to obtain miRNA mRNA modules a learning algorithm was used based on co evolutionary learning and EDA. Co evolutionary learning results cooperatively for two populations in relation to each other while EDA i.e. estimation of distribution algorithm produces new forms by the use of probability distribution.From the probabilities for miRNAs and mRNAs new populations were obtained. The process was iterated till maximum number of generations was obtained. Thus, miRNA mRNA regulatory modules were obtained.

The modules were obtained as such but now these modules have to be used to generate regulatory network. For this purpose microarray data was analysed and then from there the miRNA mRNA combinations were selected which were the modules obtained here and from them the regulatory network was obtained. One of the reasons why microarray analysis [13] was done was to obtain the expression level of genes on genomic scale.

Microarray Data Analysis

Human biological genome was taken for analysis [14] after which miRNA profile data has been taken from the gene expression omnibus (GEO) - NCBI database for *homo sapiens* species. Data consists of 32 paired samples (which are both cancerous and normal) from 14 different patients and 8 different cancer types (breast, lymphoma, ovary, testes, liver, lung, colon, prostate). Data was categorized on two basis: cancer type and tissue type (cancerous or normal). Data was preprocessed as well as normalised. Further data was filtered and statistically analysed. Statistical analysis is done to know can the difference between the observations be explained by chance alone or how significant the difference is. For example, the expression of a gene can be statistically signified between a cancer patient and its healthy control.

Clustering

Microarray produces large data which undergoes dimensionality reduction to make the storage and analysis of data easy. Clustering is one of such methods. [15] Microarray data analysis works to cluster genes or samples having similar expression profile into biologically and functionally similar set. Clusters of miRNA were obtained after clustering. Each of the clusters was used to predict the targets. These clusters were studied using Target Scan which predicts miRNA targets from the clusters as it searches miRNA seed regions for presence of conserved 8mer and 7 mer sites. Perl script has been created to obtain the targets [16][17]. The Target Scan algorithm given by Lewis et al is as follows:

Step 1: UTRs are searched as 2-8 nucleotide long miRNA are checked for Watson Crick complementarity for 7 nucleotide match.

Step 2: Matches allowed in both directions as wobble pairs are allowed while it stops at mismatches.

Step 3: Matching is done for rest of the base pairs to miRNAs 3'region using RNA fold program.

Stage 4: UTRs are ranked .Free energy is finally calculated for every miRNA mRNA module.

Results and discussion

miRNA mRNA modules:

From the information obtained after literature study it was found that for 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). 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

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

down regulation of B cell CLL/lymphoma 6 (BCL6) thus showing its role as tumor suppressor in combination to miR-212 and miR-132.

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.

3.2) Gene Ontology:

Gene Ontology maintains database which contains genes and their products in the form of three hierarchical structures which represents molecular functions, biological processes and cellular components. [18] So each gene was studied for its behaviour. Software tools like Go Miner, MAP Finder and Onto Express are used to explore GO relationships.

Pathway Analysis

A biological system contains gene products which have to be graphically represented what is called a pathway. Pathways are thus combinations of genes and proteins constituting a biological task. [19] As for proteins that synthesize metabolites in a cell, metabolic pathway occurs. Similarly, proteins involved in signal transduction from cell membrane to nucleus forms the signal transduction pathways. Pathway formed is based on NLP (Natural Language Processing).The relations for the pathway databases were obtained from published Pub med abstracts by text mining. For the miRNAs, the genes have been obtained and for those genes the pathways have been obtained. The following pathways have been obtained.

Some of the pathways obtained are listed as following:

NO.	NAME	DETAIL	FEW GENES	FUNCTION
1	Alpha6Beta4Integrin	Integrin protein complex with α and β chain heterodimerises to form receptor	ABL1,EGFR,YES,BPAG	Hemidesmosomes formation, cell adhesion, attachment, apoptosis
2	Androgen Receptor	Binds to specific DNA sequences influencing androgen responsive gene transcription	AES,BRCA1,CDC2,CDK9	Regulates gene expression in glycolysis, growth, cell development, apoptosis, lipid metabolism
3	EGFR1	Uses EGF receptor with receptor tyrosine kinases	AKT1,BCAR1,CREB	Activates RAS and MAP kinase pathways; involved in apoptosis and cell cycle.

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

		leading to receptor dimerization, protein kinase activity		
4	Hedgehog	Secretive proteins which includes Sonic, Desert, Indian which causes transmembrane binding on diffusion	CCNB1,DHH,PTCH,RAB23	Important in cellular processes, cell cycle proliferation, cell regulation and embryogenesis.
5	ID	Members of helix loop helix family with 4 members -Id1 to Id4. Consists of transcription factors and stimulation concerned with ligands	ADD1,CDK2,ID1,SMAD3	Regulate cell growth, differentiation, apoptosis, senescence and angiogenesis.
6	Kit Receptor	Involves kinases , adaptor molecules	ABL1,BTK, CRKL,FGR	Regulates gene expression, apoptosis, cell proliferation , chemotaxis
7	NOTCH	Cell surface receptor activated by ligand	APH2,DTX1 ,FHL1,GATA1	Regulates downstream genes regulation, control apoptosis, development ,cell cycle
8	TGFBR	Multifunctional cytokine that activates receptor	SMAD2,SMAD3,TGBR3 ,CDC2	Regulates gene expression, related to growth, migration, myelination
9	WNT	Secretive proteins that binds to cell surface receptors and downstream signalling components	CDC2 AKT1,BCL9 ,CDH1	Cell proliferation, development, adhesion, apoptosis

Figure 6: Cancer Signalling Pathways

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

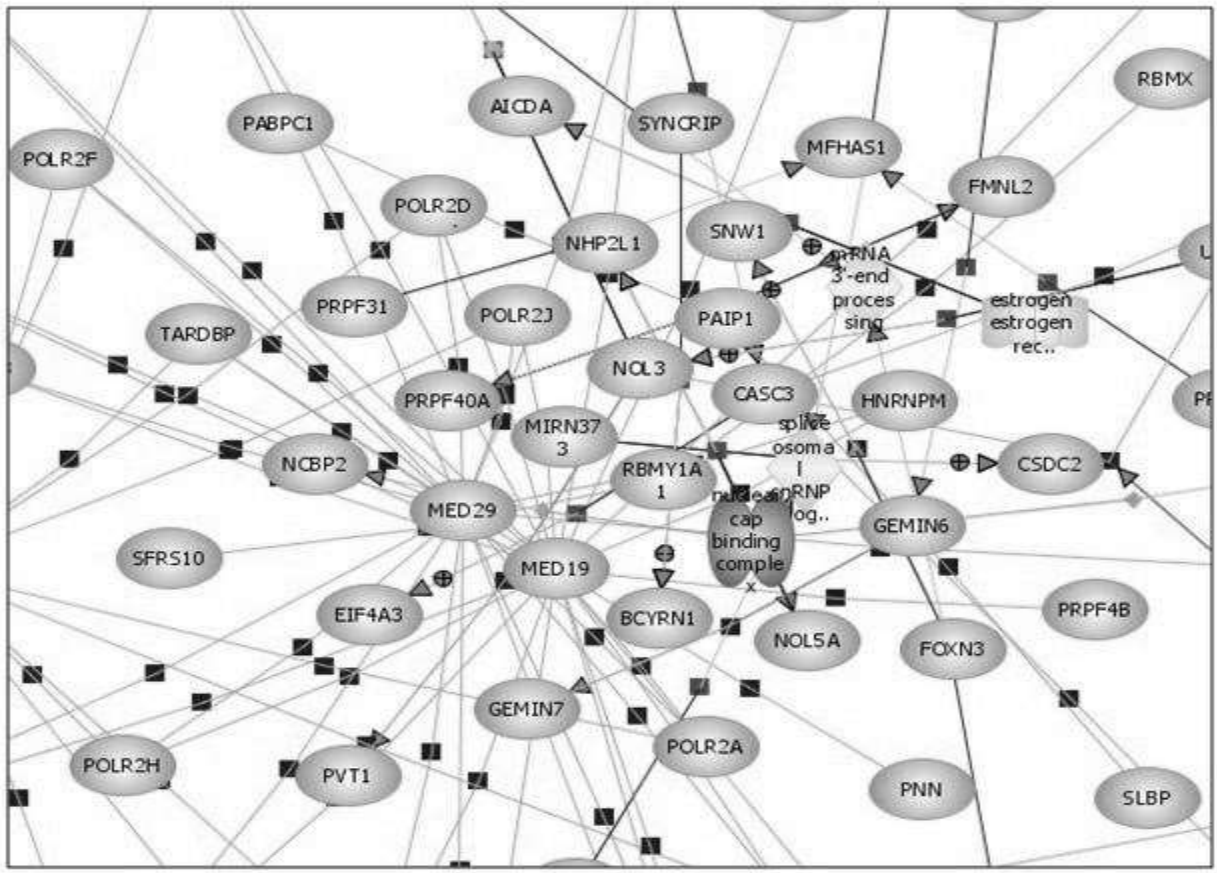

Figure 7: A closer view of a small part of Network targets and regulators

Gene expression analysis in combination with perturbations, treatments, mutations were studied for processes like signal transduction and metabolism which helps to get knowledge about the molecular and functional effects on genes. Gene expression data helps in studying molecular pathways for the regulatory effect of genes. The pathway specifies the role of entities in a molecular reaction which is both participatory and regulatory. Gene interactions which may be activating or inhibitory, direct or indirect were all under pathway study. Reactants and products have participatory roles while catalysts act as regulators which could be effective both positively and negatively. Here, MTF1 up regulates expression of HNF1A while TNF up regulates binding of HNF1A.SPINK7 modulates transport of MT1F. Cd2+ catalysis binding of MT1X. MT1G catalysis binding of MT1A and protein modification of FLNC. FLNC negatively catalysis HNF4A.

Conclusion

miRNA therapy is most frequently seen in cancer as miRNAs behave both as tumor suppressors and oncogenes. miRNAs undergoes misregulation leading to cancer development as they function both as oncogenic or tumor

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

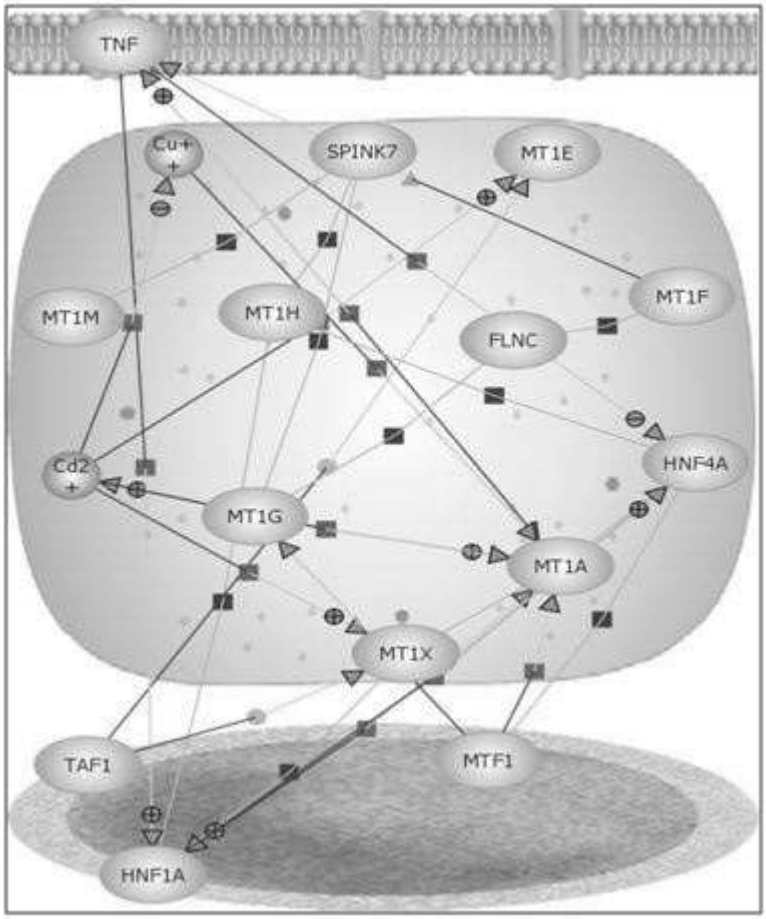

Figure 8: Mt-Heavy Metal Pathway

suppressive. 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. These miRNAs regulate mRNAs causing gene regulation. Thus, combinations of miRNA and mRNA provide knowledge about how they are involved in various biological processes. miRNA controls pathways and act as switches for genome regulating gene products. miRNAs thus act as gene regulators. Regulated genes are both oncogenic and tumor suppressive which are supposed to be drug targets by various pharmaceutical and biotechnology industry. miRNAs have high therapeutic role as they have multiple target genes. miRNAs are natural which makes them have less chance to cause non specific side effects. miRNA with their target mRNA could be studied for their integrative function by the formation of miRNA mRNA module from their expression profiles and target binding scores.

Matrices formed from miRNAs and mRNAs showed correlation not only among themselves but also between genes. Fitness was calculated to obtain miRNA mRNA dependencies over each other. Learning algorithm applied showed probability of miRNA related to mRNA. GSEA and leading edge analysis ranked the genes and again formed matrix for miRNA mRNA. Modules were thus formed considering all the matrices. Fitness function and learning algorithm were useful enough to obtain better correlated miRNA mRNA genes. Where there was GSEA to rank genes, there was microarray analysis to further get some functionally correlated clusters of genes. As for the data was undergone microarray analysis, clustering was done which formed certain clusters each of which showed how genes are related to each other. From these clusters gene regulatory pathways have been obtained. And thus miRNAs develop to be a part of regulatory network in relation with their target mRNAs whose information could be obtained from miRNA mRNA modules. Gene regulatory pathways were formed that

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

showed relations among the genes whether its activation or inhibition. Also pathways related to cancer signalling as well as immune signalling were obtained showing the involvement of genes. 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. Vasudevan, S., Tong, Y., Steitz, J.A. Switching from repression to activation: microRNAs can up-regulate translation. *Science* 2007,*318*: 1931-4.
8. Medina, P.P., Slack, F.J. microRNAs and cancer: An overview.*Cell Cycle.* 2008, *17*: 7-16.
9. Zhang, B., Pan, X., Cobb, G.P., Anderson, T.A. microRNAs as oncogenes and tumor suppressors. *Dev. Biol*. 2007, *302*: 1-12.
10. Calin, G.A., Sevignani, C., Dumitru, C.D., Hyslop, T., Noch, E.,Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M.,Croce, C.M. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc. Natl.Acad. Sci. USA* 2004, *101*: 2999-3004.
11. Eric Bonnet, Marianthi Tatari, Anagha Joshi, Tom Michoel, Kathleen Marchal, Geert Berx, Yves Van de Peer,Module Network Inference from a Cancer Gene Expression Data Set Identifies MicroRNA Regulated Modules. *PLoS ONE*, Vol. 5, No. 4. (14 April 2010), e10162.
12. Dang Hung Tran, Kenji Satou and Tu Bao Ho, Finding microRNA regulatory modules in human genome using rule induction . *BMC Bioinformatics* 2008, **9**(Suppl 12)**:**S5
13. R. Keira Curtis, Matej Ores and Antonio Vidal-Puig Pathways to the analysis of microarray data TRENDS in Biotechnology Vol.23 No.8 August 2005.
14. Hegde, P. et al. (2000) A concise guide to cDNA microarray analysis. *Biotechniques* 29, 548–562.
15. Eisen, M.B. et al. (1998) Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci.* U. S. A. 95, 14863–14868.
16. John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS. Human MicroRNA targets. PLoS Biol. (2004) 2:e363.
17. Betel D, Wilson M, Gabow A, Marks DS, Sander C. The microRNAorg resource: targets and expression. Nucleic Acids Res. 36: D149-D153 (2008)
18. Ashburner, M. et al. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*., 25(1):25–9.
19. Curtis R. Keira, Ores Matej and Vidal-Puig Antonio. Pathways to the analysis of microarray Data, Elsevier 2005 Aug; 23(8):429-35.