Identification of microRNA functional targets based on microRNA and mRNA Co-expression Network Analysis

BY

Bhavisha Chapatwala

B.E., S.V.N.I.T., Surat, 2003

THESIS

Submitted as partial fulfillment of the requirements

for the degree of Master of Science in Bioinformatics

in the Graduate College of the

University of Illinois at Chicago, 2013.

Chicago, Illinois

Defense Committee:

Dr. Yang Dai, Chair and Advisor

Dr. Hui Lu, Bioinformatics

Dr. Bhaskar DasGupta, Computer Science

ABSTRACT

Background:

MicroRNAs are essential key regulators of gene expression. They have significance in essential biological process. MicroRNA expression patterns are promising biomarkers for several tumor types including breast cancer. Many computational approaches are proposed to classify miRNA functions in recent years. Here, we propose an integrative approach to identify miRNA modules and its functional targets through the analysis of global miRNA and mRNA expression data. Our interest is to identify functionally correlated miRNA-mRNA modules that are involved in specific biological processes.

Results:

The Weighted Gene Co-expression Network Analysis (WGCNA) methodology was applied to analyze miRNA and mRNA expression data in order to determine the statistically significant modules of miRNA and the function of their targets. The process can be divided into three categories: (1) identify which mRNAs were targeted by which miRNAs, (2) determination of miRNA regulatory modules, i.e. to identify a group of co-expressed miRNAs and mRNAs. (3) Investigation of the miRNA regulatory modules i.e. to find an involvement in specific biological process for a particular miRNA module.

Conclusion:

We used mRNA and miRNA expression data from Espen Enerly breast cancer study. The proposed framework effectively captured miRNA modules. Through Gene Ontology analysis, several biological processes involving miRNAs and their targeted mRNAs were identified. To determine coherent miRNA-mRNA modules, we demonstrated that mRNAs in one module exhibit higher correlation with the miRNAs in a module. However, due to the fact that only the small numbers of mRNA modules were detected from the WGCNA analysis for this datasets, we were not able to find other miRNA-mRNA modules. For that reason we converted our focus to the other miRNAs which are not related to any modules. Therefore, the effectiveness of this approach has to be further investigated using other datasets.

ACKNOWLEDGEMENT

TABLE OF CONTENTS

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

1.1 WHAT IS microRNAs?

MicroRNAs, also known as miRNAs, were first discovered in 1993 by Victor Ambros, Rosalind and Rhonda Feinbaum [15]. But not until early 2000s, miRNAs were recognized as an individual class of biological regulators with conserved functions. They regulate gene expression through target mRNA degradation or translational gene silencing. They play an important role in many biological processes and in the development of many diseases like cancer.

A miRNA, which plays a role in transcriptional and post-transcriptional regulation of gene expression, is a small non-coding RNA molecule. MiRNAs are ~22 nucleotide RNA sequences that bind to complementary sequences in the 3' UTR of multiple targets mRNAs, resulting in a gene silencing via translational repression or target degradation [3, 6]. miRNAs target ~60% of all genes, are abundantly present in all human cells. They are well conserved in organisms and suggest that they are a vital part of genetic regulation with an ancient origin. The plant miRNAs may bind their targets in both coding and non-coding regions, whereas the animal miRNAs exhibit partial complementarity to their mRNA targets. The majority of miRNAs are transcribed from independent transcription units, but some are transcribed from introns of pre-mRNAs [2]. miRNA genes are found in a cluster of 2-7 genes having highly similar expression

profiles suggesting that transcription of these miRNAs is controlled by common promoters [6]. The miRNA genes found between the introns are not transcribed by their own promoter, but they are processed from the introns [2]. Some miRNAs are expressed at different stages of development; some are expressed in different cells [2].

The miRNA gene is transcribed into primary miRNA (pri-miRNA) by an enzyme, polymerase. Then it is processed by a nuclear RNase type III enzyme (Drosha) to produce a 60 -70 nucleotide long stem loop precursor miRNA (pre-miRNA) [7]. Drosha cleaves both strands of the stem near the base of primary stem loop [2]. The pre-miRNA then exported to the cytoplasm by the nuclear export factor Exportin 5 and the Ran-GTP cofactor and trimmed by dicer into miRNA:miRNA* duplex [7, 8]. Drosha processes one end of miRNA:miRNA* duplex in nucleus and Dicer processes other end in cytoplasm [2]. One strand of miRNA:miRNA* duplex is identified by the RNA-induced silencing complex (RISC) and the other strand is generally degraded [9]. The miRNA targets the specific 3'UTR of mRNA transcript.

Computational approaches have been unparalleled tools in understanding the biology of miRANs. Many web-based miRNA data-bases are available to provide thousands of published miRNA sequences, annotation and potential miRNA target genes. Computations algorithms are developed to pri-miRNAs and to search for homologous conserved miRNA genes in several animal species.

The pathways of miRNA biogenesis in animal cells are shown below:

2
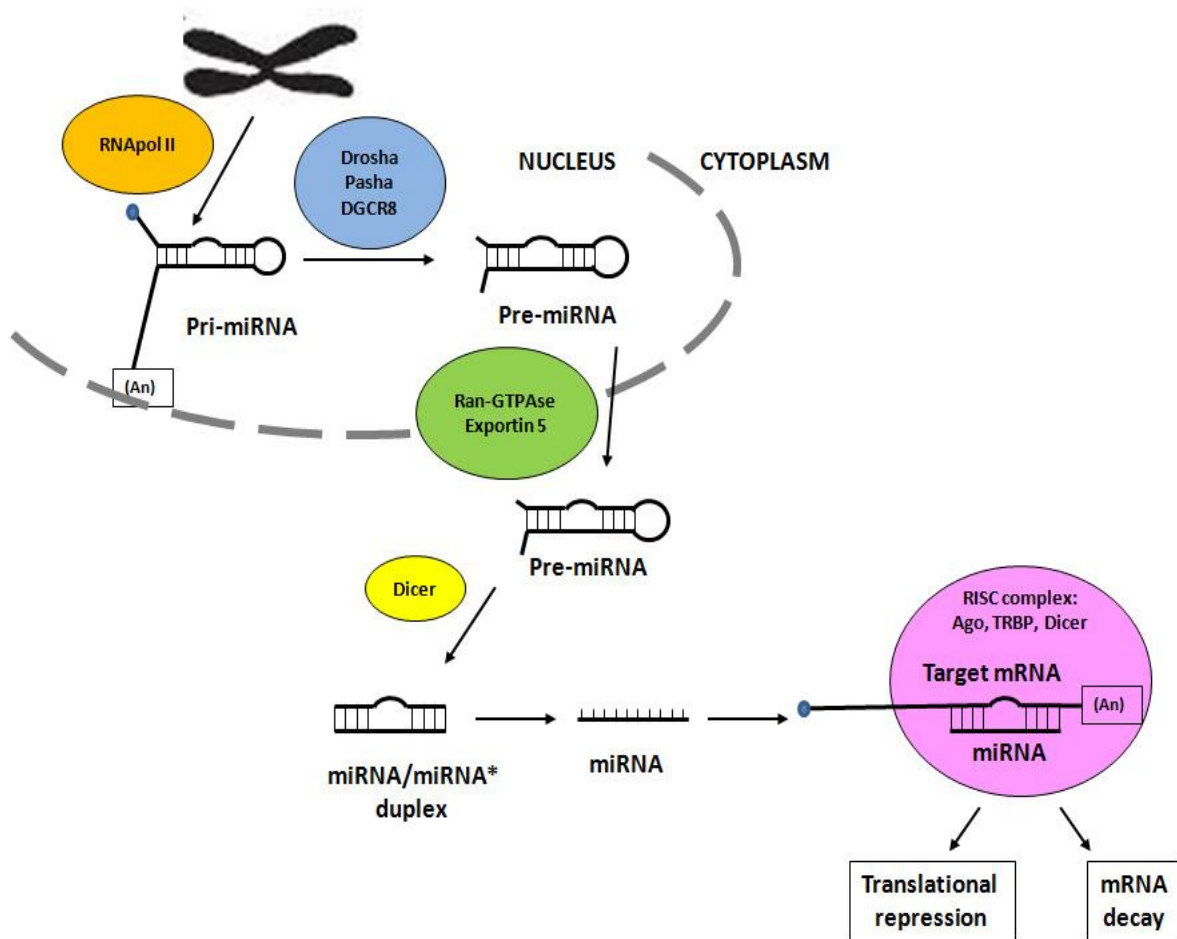
Figure 1 an animal miRNA Biogenesis [1]

The function of miRNA depends on the gene it targets. Experimentally, it is difficult to identify new miRNA targets, even though there are many experimentally validated miRNAs. The miRNA binds to the mRNA and it causes the mRNA cleavage or inhibits the translation. In general, mRNA cleavage occurs in plants and translation repression occurs in animals [4]. A miRNA may have multiple different mRNA targets, and a target might be targeted by multiple miRNAs. The figure below shows the various configurations for miRNA-mRNA duplex.
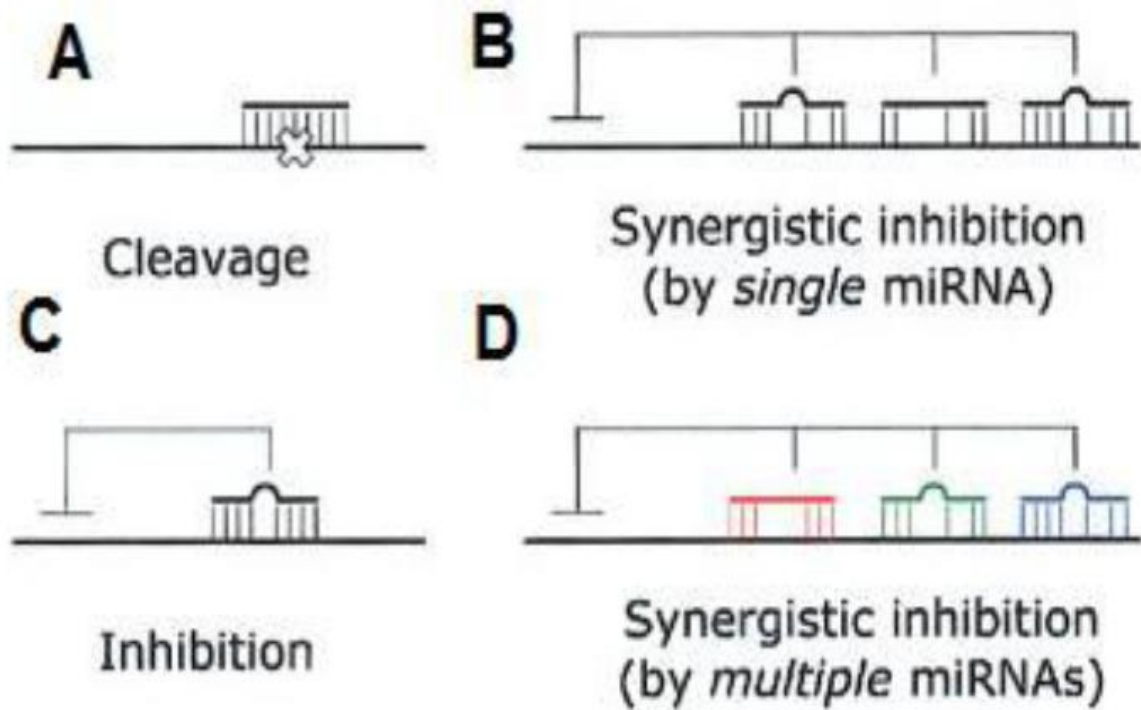
Figure 2 miRNA-mRNA duplex different configurations

In figure 2, A. represents near-perfect binding sites for one miRNA, B. represents multiple target sites for one miRNA, C. represents strong binding sites for one miRNA, and D. represents multiple target sites for multiple miRNAs.

Most miRNA based computational methods comprise of the prediction of miRNA genes and their targets. To fulfill this requirement many web-based resources are being developed. They can be used as computational target prediction tools, which can provide number of targets for experimental validation.

1.2 miRNA TARGET PREDICTION AND ITS REGULATORY ROLE

Many algorithms have been developed to predict miRNA targets. Prediction of miRNA targets in plant is very naïve because of perfect complementarity between miRNA and mRNA. Nonetheless it is tough in animals because of lack of perfect complementarity between miRNA and mRNA interaction. As a result, there are many different computational approaches to predict miRNA targets. Since miRNAs are short, they have limited sequence complementarity to their targets. The miRNA target prediction principles used by most of the approaches are almost similar [4].

Some prediction criterions are described below:

1. The miRNA and 3'UTR region of mRNA have complementarity between them, especially between the seed region of miRNA and mRNA. Complementarity between miRNA and mRNA can be of 3 types: 5'- dominant canonical, 5'- dominant seed and 3' compensatory [13].

2. The thermodynamics of miRNA and mRNA interaction can be computed by currently available RNA folding packages and is used in many prediction algorithms [4].

3. 3'UTR target regions of many miRNAs are highly conserved over many species [3].

The following figure shows Secondary structure of miRNA-mRNA interaction.
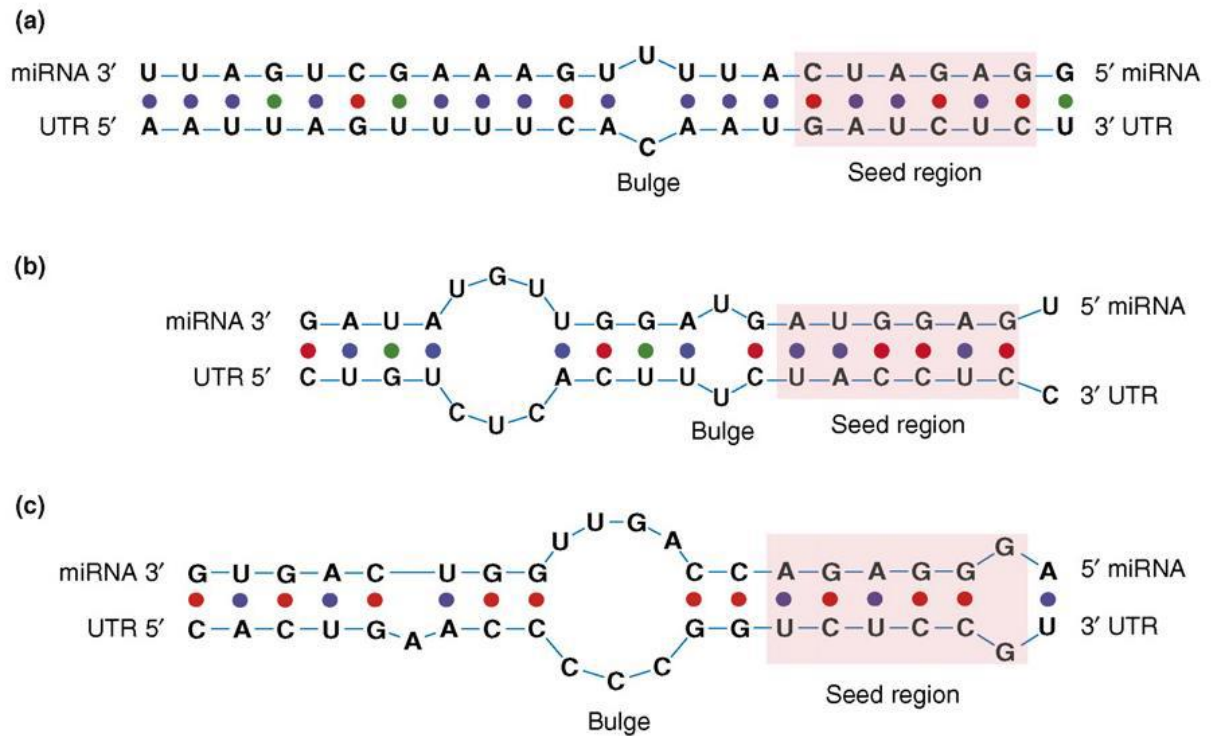
Figure 3: secondary structure of miRNA-mRNA interaction

(a) Good or perfect complementarity at both the 5' and 3' ends of the miRNA.

(b) Perfect seed region complementarity at 5' end of the miRNA, but poor 3' complementarity.

(c) Seed region has a mismatch or wobble but 3" end has an excellent complementarity.

miRNA target prediction approaches can be classified into 3 categories:

1. Complementarity searching based methods;

2. thermodynamics based methods;

3. Other methods.

There are many miRNA target prediction tools: DIANA-microT, EIMMo, miRanda, MirTarget2, miTarget, PicTar, Suport Vector Machine (SVM), rna22, RNahybrid, TargetScan, TargetScanS.

There are many miRNA Target databases available: TarBase, MiRDB, and MiRecords.

As we know there are many target prediction tools and databases available, we used MicroCosm Target Version 5 database that uses miRanda prediction tool [12] to identify potential binding sites for a given miRNA in genomic sequences. Here, prediction is purely sequence-based; we matched the sequence of miRNA and mRNA. We don't consider any condition whether the target is actually regulated in particular (breast cancer) cell type.

miRNAs are important regulators of various biological processes including cell differentiation, cell death, cell adhesion, cell proliferation, immune response, defense response, inflammatory response, signaling pathway, tissue homeostasis and apoptosis. Recent studies showed that differentially expressed miRNAs in different types of cancer, such as, breast cancer, colon cancer, kidney cancer, lung cancer, prostate cancer and ovarian cancer. Recently, great efforts have been made to simplify their regulatory mechanism.

The goal of this study is to predict the module-wise miRNA targets by applying a new approach of Weighted Gene Co-expression Network Analysis (WGCNA) [17] in combination with miRNA target prediction tool [12].

1.3 WEIGHTED GENE CO-EXPRESSION NETWORK ANALYSIS

Correlation network enables network based gene screening methods which can be used in various biological applications such as human genetics (for cancer), mouse genetics, yeast genetics, analysis of brain imaging data, etc. Further, it can be used to find modules of interconnected nodes, and highly connected hub nodes, which is centrally located in the module. It can identify significant modules, annotate all network nodes within identified modules, define network neighbor-hood of a given nodes, screen nodes on module membership information and contrast one network with another network [17].

A Weighted Gene Co-expression Network Analysis (WGCNA) is one of the applications of correlation network. A WGCNA is a method for describing the correlation patterns between genes and miRNAs across samples. A WGCNA is used to find out modules of highly correlated genes. It summarizes such modules using the module eigen-gene or an intra-modular hub gene, relating modules to external information and calculating module membership measures.

A WGCNA is all about letting the data speak for themselves. It does not assume prior pathway information but constructs modules in an unsupervised fashion. It can be interpreted as a biologically motivated data reduction scheme. A WGCNA starts from the level of thousands of genes, identifies clinically interesting gene modules, and finally uses gene significance to identify key genes in the disease pathways for further validation. A WGCNA alleviates the multiple testing problems inherent in microarray

data analysis. It focuses on the relationship between a few (typically less than 10) modules.

In the above mentioned work, the analysis of the interaction is directly focused on targets. In this study, we demonstrated a novel integrative method to analyze miRNA and mRNA expression data in combination with Weighted Gene Co-expression Network Analysis (WGCNA) methodology. We combined all information, which leads us to predict module-wise miRNA targets and their effects on regulation of predicted genes.

In this work, the focus is on the differential expression analysis and WGCNA methodology. Here, we have demonstrated how to construct a co-expression network, how to identify the modules and how these modules are related to Gene information from expression data. We further computed the significance of miRNA and mRNA modules and construct the network using expression data with the ultimate goal to predict module-wise miRNA targeted genes.

CHAPTER 2

METHODS

2.1 FRAMEWORK

A novel structure using mRNA and miRNA expression data from Espen Enerly [16] breast tumor study was demonstrated. The patients in this study were divided into two groups: Estrogen Receptor Positive (ER+) and Estrogen Receptor Negative (ER-).

The flowchart below represents our framework

Data Input and Filtering Method

Differential Expression Analysis

Weighted Gene Co-expression Network Analysis

miRNA – gene Target Prediction and its Significance

Construct module-wise miRNA – gene Target Prediction Network

.
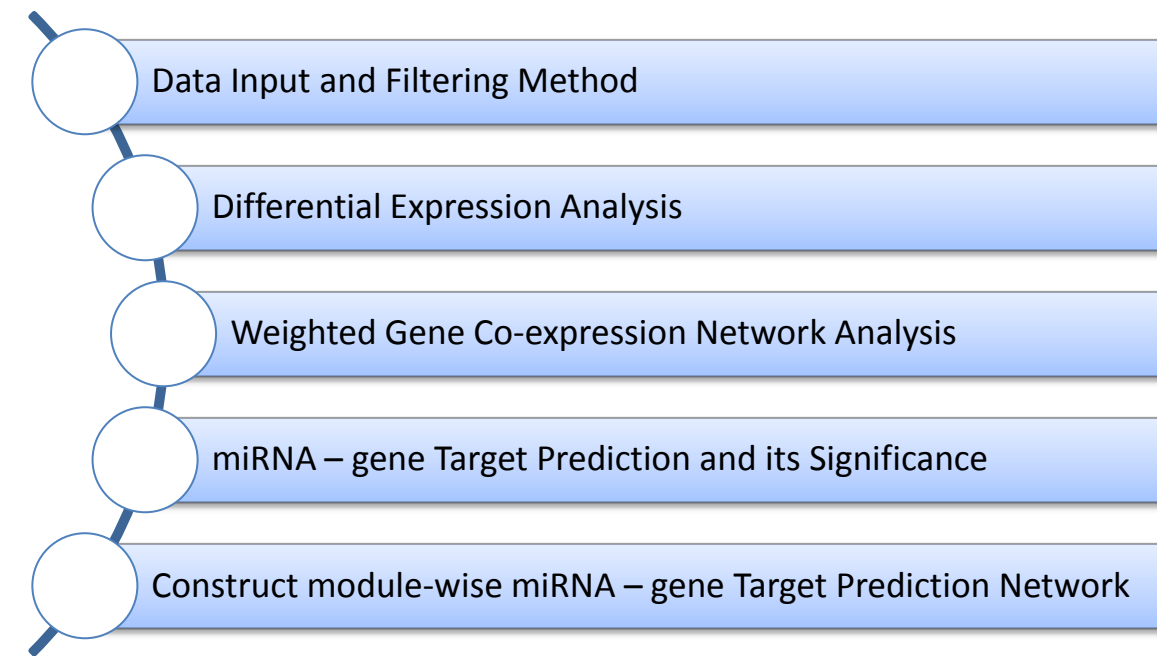
Figure 4: Framework of developed method

2.2 HYPOTHESIS AND FLOWCHART

A. Dataset and Filtering Method

We started with mRNA and miRNA expression data from Espen Enerly breast cancer study [16]. The patients in this study were divided into two groups: Estrogen Receptor Positive (ER+) and Estrogen Receptor Negative (ER-). The dataset consists of 60 ER+ and 35 ER- samples. Here, we used Espen Enerly pre-processed dataset. The expression data for this dataset were normalized.

➔ **Filtering Methods [37]**

1. For miRNAs, discard probes that are not associated with Homo sapiens.
2. For mRNAs, discard probes that are not associated to an Entrez gene IDs.

B. Differential Expression Analysis

Differentially expressed genes between ER+ and ER- samples were obtained. In this section, we provided technical details about how we obtain differentially expressed genes [26].

The differential expression analysis was performed on normalized data. For the mRNA data, the normalized expression data were used and discarded the probes that are not associated to an Entrez gene IDs. For the miRNA data, we discarded the probes which are not associated with Homo sapiens.

➔ **Hypothesis**

H_0: miRNA/mRNA is not differentially expressed in ER- samples

H_a: miRNA/mRNA is differentially expressed in ER- samples

The differential expression analysis was performed as described below:

a) Welch's t-test is an adaptation of student's t-test. We used Welch's t-test for two-sample unequal variances to find out the p-value.

b) Arrange unadjusted p-values in an ascending order.

c) Convert unadjusted p-values into adjusted p-values using Benjamini-Hochberg correction method.

d) The adjusted p-values threshold was set to 0.05.

e) Report only those probes whose adjusted p-value is less than 0.05.

Adjusted p-value is less than 0.05 than reject H_0 i.e. miRNA/mRNA is differentially expressed in ER- samples.

C. Overview of WGCNA Methodology

To construct a network, we began with the calculation of Pearson's correlation for all pairs of genes. We weighted the Pearson correlation by taking their absolute value and raising them to the power of β. This effectively served to emphasize strong correlations and punish weak correlations on an exponential scale. These weighted correlations represented the strengths between genes in the network. By accumulation of these connection strengths for each gene, we produced a single number that describes how strongly that gene is connected to all other genes in the network. The general

12

framework of Weighted Gene Co-expression Network Analysis was used.

The flowchart below presents a brief overview of Weighted Gene Co-expression Network Analysis [17].

Construct a gene co-expression network

•**Rationale:** Make use of interaction patterns among
 genes
•**Tools:** Correlation as a measure of co-expression

Identify Modules

•**Rationale:** Module based analysis
•**Tools:** Hierarchical clustering, Dynamic Tree Cut

Relate modules to external information

•Gene Information: Ontology, functional enrichment
•**Rationale:** Find biologically interesting modules
•**Tools:** DAVID web based tool

Figure 5: overview of WGCNA methodology

D. Predicting miRNA Target

The miRanda algorithm was one of the first miRNA target prediction algorithms and is widely used for target prediction by multiple interfaces including http://microRNA.org and MicroCosm Targets, available at http://www.ebi.ac.uk/enright-srv/microcosm/htdocs/targets/v5/ . All miRNA – gene predictions were downloaded from MicroCosm Targets Version 5.0 that currently uses the miRanda algorithm. The algorithm ranks the probability of each gene to be a miRNA target and the probability of

each miRNA to target a gene. The algorithm uses a weighted scoring system and rewards complementarity at the 5' end of the miRNA and 3' end of mRNA. Currently it demands strict complementarity at this so-called seed region. miRanda is a miRNA target prediction algorithm that searches highly conserved 3'UTR targets matching the seed region of miRNAs.

We used miRanda algorithm to identify potential binding sites for a given miRNA in genomic sequences. miRanda method was originally developed to predict miRNA target genes in Drosophila melanogaster [12], but was also used to predict human miRNA targets. Enright, A. J., B. John et al. (2004) [12] improved the method by implementing a strict model for the binding sites that require almost perfect complementarity in Drosophila. Their analysis also suggested that miRNA genes, which comprise around 1% of the human genome, control the production of protein for 10% or more of all human genes.

The resulting binding sites are then evaluated thermodynamically, using the Vienna RNA folding package. The false positive rate is between 24% and 39%.

E. Fisher's Exact Test

Fisher's test is used to detect group difference. Fisher's test is basically used for categorical data. We used Fisher's exact test [33] to show statistical significance between miRNA modules to mRNA module. This involves 2 × 2 contingency table. The fisher's test calculates an exact probability value for the relationship between two different variables. If there is a small value in one of the cell of the contingency table the

fisher's exact test is preferred. The p-value from the Fisher's exact test decides the significance [34]. Exact p-value tend to be more conservative than most approximate estimates, such as Chai squared test.

F. Construct a module-wise miRNA Target Prediction Network

To construct a module-wise miRNA target prediction network, we used the Biological Network Analyzer (BiNA) version 2.3.1. BiNA provides sophisticated visualization style for biological networks. For this, we used the concept of hierarchical and dynamic graph structures with background imaging. A complex data framework allows mapping of almost any data to the network.

CHAPTER 3

RESULTS

A. Data set after Filtering Method and Differential Expression Analysis

Here, Espen Enerly pre-processed dataset was used. The Espen Enerly dataset contains mRNA and miRNA expression profiles on ER+/ER- breast tumors. The expression data for this dataset were normalized. The Table 1 below displays the information about the dataset [16], number of samples in ER+ and ER- and number of probes before pre-processing, before and after filtering method plus number of probes that are differentially expressed.

Differentially expressed genes are shown in Supplementary Table 1.

Dataset Name	Number of samples		Number of Probes			
	ER+	ER-	Before pre-processing	Before Filtering	After Filtering	Differentially Expressed
Espen Enerly (miRNA)	60	35	729	498	477	49
Espen Enerly (mRNA)	60	35	41094	12837	12605	3030

Table 1: Dataset information after each steps

B. WGCNA Methodology

1) Construct the co-expression network

To construct a network, first, we calculated Pearson correlations [25] for all pairs of genes in the network. We weighted the Pearson correlations by taking their absolute value and raising them to power β, because data can be noisy and the number of samples is often small. We emphasized strong correlations and punished weak correlations on an exponential scale. Weighted correlations represented the connection strengths between genes in the network. For each gene, the connectivity is described as how strongly that gene is connected to all other genes in the network. We used the flowchart to present a brief overview of Weighted Gene Co-expression Network Analysis as described earlier in methods.

Briefly, the absolute value of the Pearson correlation coefficient was calculated for all pairwise comparisons of gene expression values across samples. The Pearson correlation matrix is then transformed into an adjacency matrix A. We considered networks where adjacency matrix A_{ij} calculates the connection strength between node i to node j i.e. connection strength between gene pairs.

Adjacency matrix is defined as,

$$A_{ij} = |corr(x_i, x_j)|^{\beta}$$

We studied networks whose adjacencies satisfy the following conditions:

$$0 \leq A_{ij} \leq 1$$

$$A_{ij} = A_{ji}$$

$$A_{ii} = 1$$

A weighted network adjacency can be defined by raising the co-expression similarity to a power [5, 10]. The function adjacency calculates the adjacency matrix from the expression data. The adjacency implies that the weighted adjacency A_{ij} between two genes is proportional to their similarity on a logarithmic scale.

$$\log(A_{ij}) = \beta \times \log(|corr(x_i, x_j)|)$$

Adjacency functions for weighted networks are required to choose threshold parameter by applying the scale-free topology criterion. The network connectivity k(i) of the ith gene, expression profile x(i) is the sum of the connection strength with all other genes in the network i.e. it shows how i[th] gene is correlated with all other genes in the network.

To choose a power β, we used the scale free topology criterion explained in Zhang and Horvath 2005 [17].

2) Scale-free topology criterion

Many co-expression networks satisfy the scale free property [17]. The network exhibits a scale free topology if the frequency distribution p(k) of the connectivity follows :

$$p(k) \sim k^{-\gamma}$$

Here, the power γ has nothing to do with β that is used to define the co-expression network. To visualize the scale free topology, we plotted log(p(k)) versus log(k). The

model fitting index R^2 of the linear model that regress $\log(p(k))$ on $\log(k)$. If R^2 of the model approaches 1, then there is a straight line relationship between $\log(p(k))$ and $\log(k)$. We only considered those powers that lead to a network satisfying scale free topology i.e. $R^2 > 0.80$. We considered the following points, when choosing the adjacency parameter: (i) the mean connectivity should be high so that network contains enough information, (ii) the slope of the regression line between $\log(p(k))$ and $\log(k)$ should be negative. We found the relationship between R^2 and β is characterized by a saturation curve. We used the lowest power β where saturation is reached. In this case, we chose default value of β for unsigned network, i.e. $\beta = 6$.

3) Identify Modules (Module Detection)

Once the network is constructed, next step is module detection. Modules are clusters of highly interconnected genes. In unsigned co-expression network, modules correspond to clusters of genes with high absolute correlations.

We used average linkage hierarchical clustering coupled with a gene dissimilarity measure to define a cluster tree of the network. The default choice is the Topological Overlap Matrix (TOM) based dissimilarity measure [17, 23, 37]. A pair of genes is said to have high topological overlap if they are both strongly connected to the same group of genes. Topological overlap of two genes reflects their relative interconnectivity. The Topological Overlap dissimilarity is used as an input of hierarchical clustering [24],

$$TOM_{ij} = \frac{\sum_u a_{iu} a_{uj} + a_{ij}}{\min(k_i, k_j) + 1 - a_{ij}}$$

19

$$dissTOM_{ij} = 1 - TOM_{ij}$$

Where, k_i is the number of connections of a node with $k_i = \sum_u a_{iu}$ and $k_j = \sum_u a_{ju}$. The use of topological overlap serves as a filter to exclude isolated connections during the network construction.

Dissimilarity measure can also defined as,

$$dissA_{ij} = 1 - A_{ij}$$

This dissimilarity measure, directly using adjacency matrix, computationally is much faster than the Topological Overlap measure and often leads to approximately similar modules. Here, we used dynamic branch cutting method [38] that offers the following advantages: (i) it is capable of identifying nested clusters, (ii) it is flexible, and (iii) it is suitable for automation. WGCNA implements two types of dynamic branch cutting method. (i) Considers the shape parameters. (ii) Hybrid method that combines the advantages of hierarchical clustering and Partitioning around Medoids (PAM). One drawback is that it can be difficult to determine how many clusters are present in the given data set.

Module-wise genes and miRNAs are shown in the implementation of framework of results section.

4) Functional Enrichment Analysis of Module genes

We selected the genes targeted by miRNA different modules and combined them [30]. The combination of selected genes could be used as input of functional enrichment

analysis software such as EASE, KEGG, Webgestalt, Ingenuity, etc. Here, we used web based software DAVID tool [27, 31]. Module-wise functional enrichment analysis results are shown in the implementation of framework of results section.

5) miRNA – gene Target Prediction

We used MicroCosm Targets Version 5 database that used miRanda algorithm to predict the targets. miRanda algorithm is described in the methods section. Module-wise miRNA - gene target predictions are shown in the implementation of framework of results section.

6) Fisher's Exact Test

It is a statistical significance test, which is used in the analysis of contingency tables. The notion behind Fisher's Exact Test is shown in the table below:

		miRNA j Module		
		Target (Y)	Not Target (Y')	
mRNA	Gene in (X)	A	b	$a + b = \theta_X$
i	Gene not in (X')	C	d	$c + d = \theta_{X'}$
Module		$a + c = \theta_Y$	$b + d = \theta_{Y'}$	$n = a + b + c + d$

Table 2: Notion behind Fisher's Exact Test

Fisher follows the Hyper-geometric distribution:

$$pr(X = a) = \frac{\binom{\theta_X}{a}\binom{\theta_{X'}}{c}}{\binom{n}{\theta_Y}} = \frac{\theta_X! \; \theta_{X'}! \; \theta_Y! \; \theta_{Y'}!}{n! \; a! \; b! \; c! \; d!}$$

The p-value of Fisher's Exact Test given by

$$p - value = \; \Sigma \; pr(X \geq a)$$

We can consider that $p - value$ less than 0.05 is significant.

Here we found out that one module has genes that are associated with breast cancer in ER- samples with their up/down regulation. Module-wise test results are shown in the implementation of framework of results section.

7) Construction of miRNA – gene Prediction Network

The BiNA was described in the methods section. The whole module-wise prediction is shown in the section of implementation of framework of results.

C. IMPLEMENTATION of FRAMWORK

To implement the framework, we used two different programming languages i.e. R script and JAVA script. The framework is divided into smaller scripts. The results are stored in CSV files for each script.

Script 1: Pre-processed data and Normalization

The script was implemented in R language. It displays how one can convert miRNA and

mRNA expression data from .RData files to excel or .csv files. The miRNA data set contains three objects: eset, eset.pos and eset.neg i.e. expression matrix. This expression matrix is for 60 ER+ samples and 35 ER- samples. The rows are probes and columns are samples. The mRNA data set contains only eset object. The first 60 columns are ER+ samples and the rest 35 columns are ER- samples. The row names are Entrez gene IDs. After using filtering methods, the output is stored in eset.1miRNA.csv and eset.1mRNA.csv.

Normalization is done by generic function. The normalized data are stored in normalizedmiRNA.csv and normalizedmRNA.csv. miRNA files store miRNA name and samples, whereas mRNA files store Gene Name, Entrez Gene IDs and samples.

The script generates following output files:

For miRNA,

miRNA	GSM487074	GSM487075	GSM487076	GSM487077	GSM487079
hsa-let-7a	0.999973	0.999789	0.999867	0.99975	0.999655
hsa-let-7a*	0.251205	0.243577	0.241152	0.232222	0.241279
hsa-let-7b	0.999954	0.999893	0.999936	0.999567	0.999844
hsa-let-7b*	0.274006	0.27924	0.27535	0.261129	0.261013
hsa-let-7c	0.998916	0.997075	0.997324	0.993872	0.99274
hsa-let-7c*	0.245459	0.263959	0.251259	0.249377	0.250977
hsa-let-7d	0.97876	0.958962	0.95501	0.937092	0.929989
hsa-let-7d*	0.266529	0.267674	0.257304	0.255911	0.259014
hsa-let-7e	0.982685	0.970821	0.95451	0.984406	0.954175
hsa-let-7e*	0.255339	0.257482	0.246138	0.251169	0.243241

For mRNA,

Gene Name	EntrezGene	GSM496925	GSM496926	GSM496927	GSM496928
GATC	283459	0.403149	0.538238	0.35197	0.509032
EIF4E1B	253314	0.340214	0.254113	0.256225	0.22542
A1BG	1	0.505659	0.485732	0.461343	0.532706
A2M	2	0.993427	0.996312	0.995695	0.99536
A2ML1	144568	0.31756	0.346977	0.26309	0.455636
A4GALT	53947	0.40038	0.536961	0.472546	0.492588
A4GNT	51146	0.123726	0.093063	0.114332	0.159346
AAAS	8086	0.664529	0.581762	0.644784	0.558541
AACS	65985	0.778443	0.847711	0.846013	0.83547

Script 2: P- Value and Differential Expression Analysis

Again we used R language to implement the script. The differential expression analysis is performed as descried in the methods section. In order to find the p-value of normalized data of miRNA and mRNA, Welch's test was used. Then, the adjusted p-value was found by using Benjamini-Hochberg correction method. Further, we considered only those miRNAs and mRNAs whose adjusted p-value is less than 0.05. As a result, we found 49 miRNAs and 3030 mRNAs are differentially expressed. The results stored in pvalmir49.csv and pvalmrna3030.csv.

The output of this script looks like:

For miRNA,

	miRNA	pAdjusted.index	rawp	BH
1	hsa-miR-29c*	204	2.65E-10	1.26E-07
2	hsa-miR-149	85	1.42E-09	3.39E-07
3	hsa-miR-190b	125	7.46E-09	1.15E-06
4	hsa-miR-342-3p	246	9.66E-09	1.15E-06
5	hsa-miR-342-5p	247	4.53E-08	4.32E-06
6	hsa-miR-339-5p	240	5.09E-07	4.04E-05

7	hsa-miR-29b-2*	202	2.63E-06	0.000179
8	hsa-miR-224	178	5.74E-06	0.000342
9	hsa-miR-505	335	7.24E-06	0.000383
10	hsa-miR-146b-5p	81	2.95E-05	0.001293

Table 3: (a) miRNAs differentially expressed between ER+ and ER-

For mRNA,

Gene Name	EntrezGene	pAdjusted.index	rawp	BH
KCNK15	60598	5845	6.20E-17	7.82E-13
ESR1	2099	3757	4.83E-14	3.05E-10
C6orf97	80129	1738	3.25E-13	1.36E-09
AGR3	155465	973	1.85E-12	4.80E-09
PLCD4	84812	8436	1.90E-12	4.80E-09
RAB30	27314	9061	2.43E-12	5.10E-09
TES	26136	10927	4.46E-12	8.04E-09
PARD6B	84612	8006	6.06E-12	8.49E-09
CA12	771	1844	6.06E-12	8.49E-09
GFRA1	2674	4647	1.13E-11	1.42E-08

Table 3: (b) mRNAs differentially expressed between ER+ and ER-

Script 3: Weighted Gene Co-expression Network Methodology

A. Scale-free topology Criterion to choose β for Adjacency matrix

To choose a power β, we used a scale free criterion on differentially expressed 49 miRNAs and 3030 mRNAs. We chose power β = 6, which is large enough to have network exhibits the approximate scale free topology. Here, we focused on the linear

regression model fitting index R^2 that quantifies how well a network satisfies a scale free topology. The result showed network properties for different choices of the power β.

For miRNA: ER+ samples

	Power	SFT.R.sq	slope	truncated.R.sq	mean.k.	median.k.	max.k.
1	1	0.0685	0.742	0.303	9.36	9.38	13.2
2	2	0.00457	0.0768	-0.108	2.93	2.85	5.1
3	3	0.142	-0.22	0.371	1.23	1.13	2.56
4	4	0.573	-0.446	0.517	0.643	0.538	1.53
5	5	0.556	-0.734	0.476	0.391	0.287	1.13
6	6	0.635	-0.711	0.584	0.264	0.167	0.889
7	7	0.102	-1.71	-0.154	0.192	0.103	0.726
8	8	0.178	-2.4	0.0177	0.147	0.0657	0.632
9	9	0.123	-1.92	-0.126	0.117	0.0425	0.595
10	10	0.174	-2.84	-0.0579	0.0955	0.0279	0.561
11	12	0.219	-3.35	0.00713	0.0673	0.0129	0.499
12	14	0.27	-4.14	0.0717	0.05	0.00586	0.444
13	16	0.252	-3.46	0.0484	0.0384	0.0025	0.396
14	18	0.271	-3.56	0.109	0.0303	0.00114	0.352
15	20	0.271	-3.62	0.0641	0.0244	0.000525	0.314

Table 4: (a) choices of power β in miRNA ER+samples

Scale Free	Rsquared	slope
1	0.59	-0.71
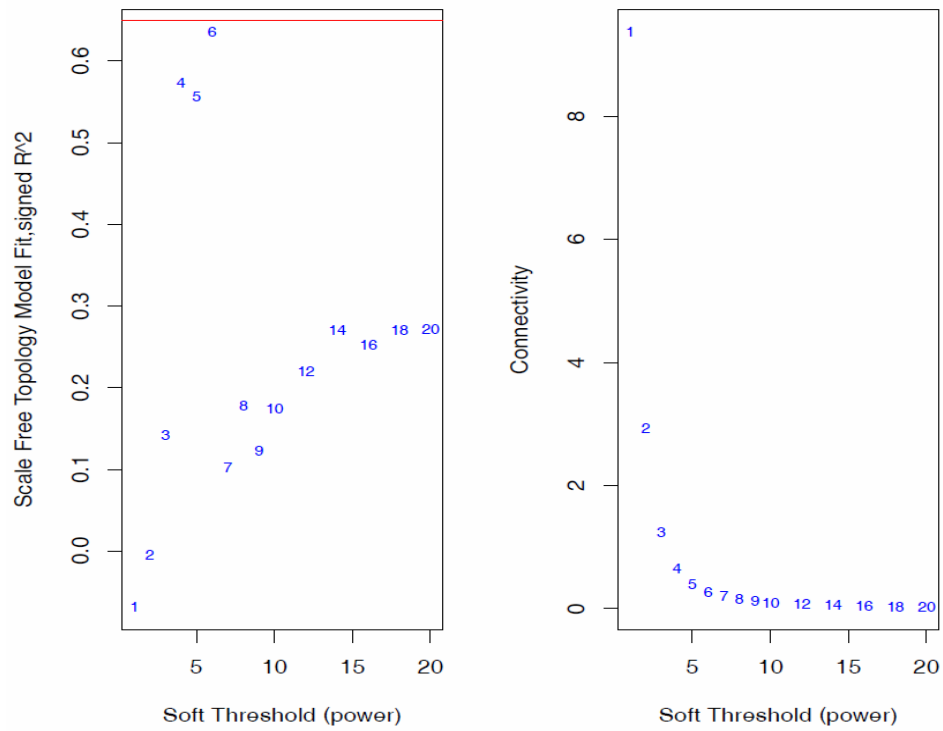
Figure 6: (a) choices of power β for miRNA ER+ samples

ER- samples

	Power	SFT.R.sq	slope	truncated.R.sq	mean.k.	median.k.	max.k.
1	1	0.00538	-0.235	0.481	11	10.7	15.9
2	2	0.0224	-0.254	0.587	3.79	3.7	6.7
3	3	0.0791	-0.294	0.602	1.66	1.65	3.36
4	4	0.235	-0.496	0.213	0.868	0.943	1.94
5	5	0.359	-0.415	0.176	0.522	0.501	1.26
6	6	0.694	-0.51	0.609	0.349	0.306	0.94
7	7	0.0563	-1.38	-0.178	0.253	0.179	0.779
8	8	0.107	-2.37	-0.139	0.194	0.111	0.736
9	9	0.19	-3.11	0.00653	0.155	0.0726	0.702
10	10	0.227	-3.3	0.125	0.128	0.0489	0.672
11	12	0.282	-3.94	0.167	0.0925	0.0224	0.618

12	14	0.327	-4.53	0.168	0.0708	0.0108	0.569
13	16	0.367	-4.78	0.187	0.0562	0.00485	0.525
14	18	0.343	-4.4	0.161	0.0457	0.00223	0.484
15	20	0.361	-4.09	0.186	0.0379	0.00104	0.447

Table 4: (b) choices of power β in miRNA ER- samples

Scale Free	Rsquared	slope
1	0.66	-0.51
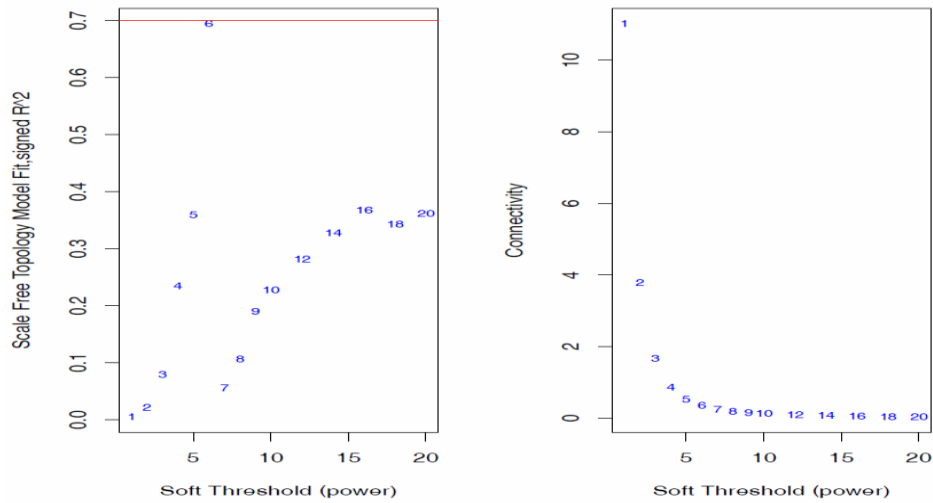
Figure 6: (b) choices of power β for miRNA ER- samples

For mRNA: ER+ samples

	Power	SFT.R.sq	slope	truncated.R.sq	mean.k.	median.k.	max.k.
1	1	0.0852	-1.32	0.979	494	4.90E+02	792
2	2	0.39	-1.94	0.99	124	1.20E+02	288
3	3	0.603	-2.2	0.986	38.9	3.62E+01	122
4	4	0.716	-2.33	0.983	14.2	1.24E+01	57.1
5	5	0.769	-2.43	0.986	5.77	4.76E+00	28.7
6	6	0.792	-2.37	0.969	2.58	1.96E+00	15.2
7	7	0.828	-2.23	0.953	1.25	8.69E-01	8.65
8	8	0.89	-2.25	0.971	0.649	4.06E-01	5.82

9	9	0.894	-2.34	0.946	0.36	1.97E-01	4.42
10	10	0.947	-2.18	0.984	0.213	9.85E-02	3.44
11	12	0.963	-1.79	0.956	0.0886	2.76E-02	2.18
12	14	0.366	-2.42	0.291	0.0458	8.24E-03	2.04
13	16	0.377	-2.18	0.32	0.0283	2.68E-03	1.93
14	18	0.352	-2.45	0.317	0.0201	9.16E-04	1.83
15	20	0.336	-2.21	0.304	0.0156	3.20E-04	1.74

Table 4: (c) choices of power β in mRNA ER+ samples

scaleFree	Rsquared	slope
1	0.77	-2.37
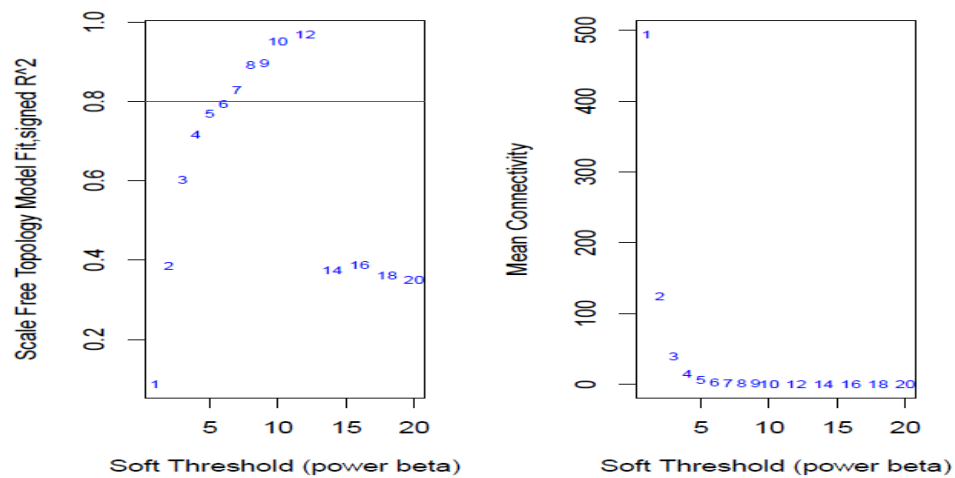
Figure 6: (c) choices of power β for mRNA ER+ samples

ER-samples

	Power	SFT.R.sq	Slope	truncated.R.sq	mean.k.	median.k.	max.k.
1	1	0.252	-1.91	0.947	628	6.06E+02	1060
2	2	0.536	-2.44	0.951	196	1.78E+02	498
3	3	0.711	-2.44	0.982	74.8	6.33E+01	270
4	4	0.795	-2.47	0.982	32.7	2.53E+01	159
5	5	0.839	-2.47	0.992	15.7	1.11E+01	99.6

6	6	0.876	-2.38	0.995	8.16	5.20E+00	65.2
7	7	0.9	-2.29	0.989	4.5	2.59E+00	44.2
8	8	0.916	-2.2	0.985	2.62	1.37E+00	30.8
9	9	0.929	-2.1	0.981	1.59	7.53E-01	22
10	10	0.967	-1.95	0.986	1.01	4.26E-01	16.1
11	12	0.935	-1.82	0.92	0.447	1.46E-01	9
12	14	0.911	-1.74	0.886	0.224	5.52E-02	5.93
13	16	0.957	-1.63	0.944	0.125	2.25E-02	4.37
14	18	0.961	-1.56	0.951	0.076	9.42E-03	3.37
15	20	0.931	-1.51	0.919	0.0503	4.12E-03	2.75

Table 4: (d) choices of power β in mRNA ER- samples

scaleFree	Rsquared	slope
1	0.86	-2.38
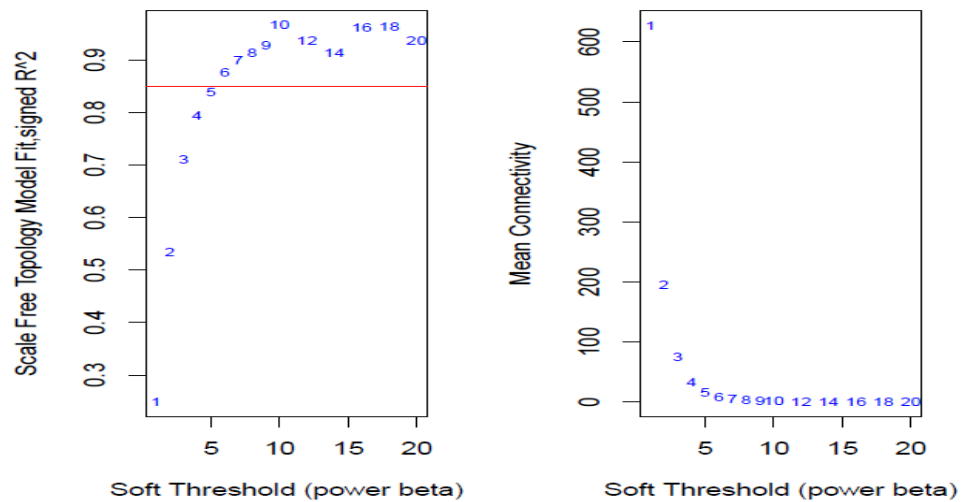
Figure 6: (d) choices of power β for mRNA ER- samples

Network Connectivity of ER+ and ER- samples

For miRNA,

Connectivity is highly preserved in miRNA samples.
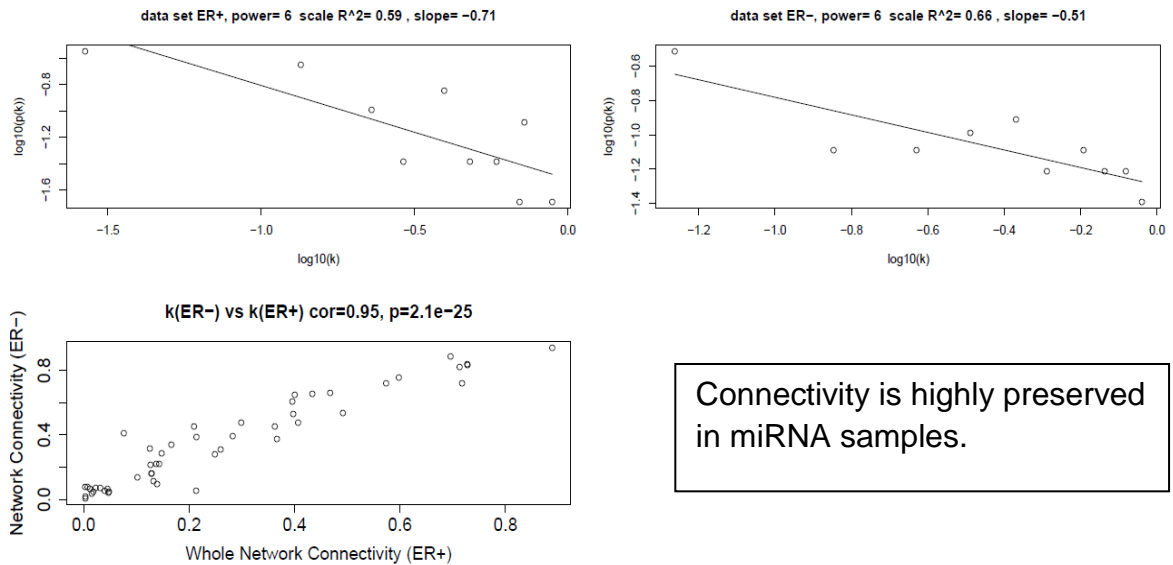
In co-expression networks, the connectivity measures how correlated a miRNA/gene is with all other network miRNAs/genes.

For mRNA,

Connectivity is not highly preserved in mRNA samples.
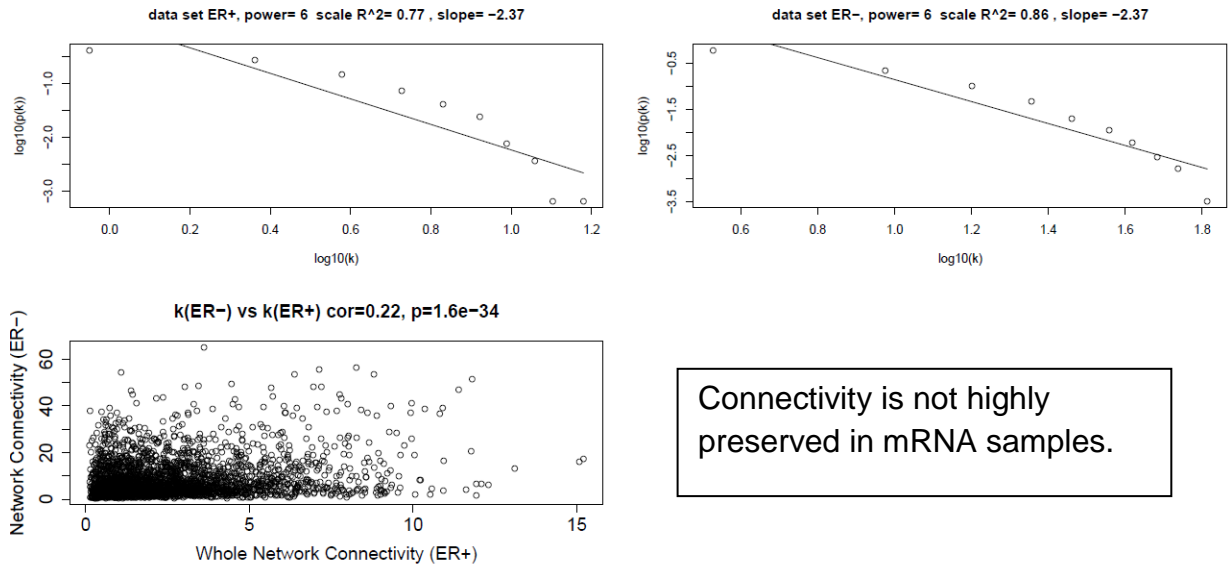
Figure 7: Network connectivity for (a) miRNAs and (b) mRNAs

B. Adjacency matrix, Dissimilarity measure and Module Detection

After choosing β for miRNA and mRNA expression data for two phenotypes ER+ and ER-, the adjacency for each was found. It is useful to find modules in the network. In order to implement this, R script was used. The theory is described in the method section. Some of the conditions should be satisfied to find out the network adjacencies,

$$0 \leq A_{ij} \leq 1$$

$$A_{ij} = A_{ji}$$

$$A_{ii} = 1$$

Adjacency matrix looks like,

For miRNA, ER+ samples: Result stored in ADJER+.txt

	[,1]	[,2]	[,3]	[,4]	[,5]	--	--	[,49]
[1,]	1.00E+00	8.54E-04	8.05E-02	7.83E-09	7.27E-05			1.14E-09
[2,]	8.54E-04	1.00E+00	2.66E-04	1.74E-09	6.10E-07			3.85E-06
[3,]	8.05E-02	2.66E-04	1.00E+00	5.04E-07	2.61E-04			9.81E-11
[4,]	7.83E-09	1.74E-09	5.04E-07	1.00E+00	1.21E-01			1.05E-05
[5,]	7.27E-05	6.10E-07	2.61E-04	1.21E-01	1.00E+00			1.93E-07
[49]	1.14E-09	3.85E-06	9.81E-11	1.05E-05	1.93E-07	--	--	1.00E+00

ER- samples Result stored in ADJER-.txt

	[,1]	[,2]	[,3]	[,4]	[,5]	--	--	[,49]
[1,]	1	0.000942	0.085988	6.43E-05	5.37E-10			3.34E-07
[2,]	0.000942	1	0.000257	4.72E-06	1.32E-08			1.99E-08

[3,]	0.085988	0.000257	1	0.000237	6.78E-08			4.86E-07
[4,]	6.43E-05	4.72E-06	0.000237	1	0.125976			0.000865
[5,]	5.37E-10	1.32E-08	6.78E-08	0.125976	1			2.58E-05
[49,]	3.34E-07	1.99E-08	4.86E-07	0.000865	2.58E-05	--	--	1.00E+00

Table 5: (a) adjacency matrix for miRNA

For mRNA, ER+ samples: Result stored in ADJER+m.txt

	[,1]	[,2]	[,3]	[,4]	[,5]	--	--	[,3030]
[1,]	1	0.003892	7.72E-06	0.000341	7.09E-06			0.000325
[2,]	0.003892	1	0.006697	0.000496	0.000452			0.000249
[3,]	7.72E-06	0.006697	1	0.001686	3.26E-14			0.000135
[4,]	0.000341	0.000496	0.001686	1	0.000144			0.000195
[5,]	7.09E-06	0.000452	3.26E-14	0.000144	1			8.76E-07
[3030,]	0.000325	0.000249	0.000135	0.000195	8.76E-07			1

ER- : Result stored in ADJER-m.txt

	[,1]	[,2]	[,3]	[,4]	[,5]	--	--	[,3030]
[1,]	1	0.001632	0.066333	0.04377	0.002669			0.00076
[2,]	0.001632	1	0.016613	0.00794	0.008784			0.006989
[3,]	0.066333	0.016613	1	0.050861	0.019168			0.001297
[4,]	0.04377	0.00794	0.050861	1	0.08481			0.011573
[5,]	0.002669	0.008784	0.019168	0.08481	1			0.098296
[3030,]	0.00076	0.006989	0.001297	0.011573	0.098296			1

Table 5: (b) adjacency matrix for mRNA

Dissimilarity can be approximately measured by

$$dissA_{ij} = 1 - A_{ij}$$

Its output is in the table below:

For miRNA, ER+ samples: Result stored in disTOMER+.txt

	[,1]	[,2]	[,3]	[,4]	[,5]	--	--	[,49]
[1,]	0	0.998869	0.911965	0.999987	0.999777			0.999997
[2,]	0.998869	0	0.999573	0.999998	0.999997			0.999996
[3,]	0.911965	0.999573	0	0.999964	0.999521			0.999998
[4,]	0.999987	0.999998	0.999964	0	0.879671			0.999999
[5,]	0.999777	0.999997	0.999521	0.879671	0			0.999999
[49]	0.999997	0.999996	0.999998	0.999999	0.999999			0

ER- : Result stored in disTOMER-.txt

	[,1]	[,2]	[,3]	[,4]	[,5]	--	--	[,49]
[1,]	0	0.998869	0.911965	0.999987	0.999777			0.999997
[2,]	0.998869	0	0.999573	0.999998	0.999997			0.999996
[3,]	0.911965	0.999573	0	0.999964	0.999521			0.999998
[4,]	0.999987	0.999998	0.999964	0	0.879671			0.999998
[5,]	0.999777	0.999997	0.999521	0.879671	0			0.999999
[49,]	0.999997	0.999996	0.999998	0.999998	0.999999			0

Table 6: (a) dissimilarity measures for miRNA

For mRNA, ER+ samples: Result stored in disTOMER+m.txt

	[,1]	[,2]	[,3]	[,4]	[,5]	--	--	[,3030]
[1,]	0	0.995585	0.998901	0.998693	0.999615			0.999471
[2,]	0.995585	0	0.993915	0.996741	0.997268			0.998318
[3,]	0.998901	0.993915	0	0.994627	0.999377			0.99801
[4,]	0.998693	0.996741	0.994627	0	0.999338			0.9992
[5,]	0.999615	0.997268	0.999377	0.999338	0			0.999882
|								
|								
[3030,]	0.999471	0.998318	0.99801	0.9992	0.999882			0

ER- samples: Result stored in disTOMER-m.txt

	[,1]	[,2]	[,3]	[,4]	[,5]	--	--	[,3030]
[1,]	0	0.989214	0.971138	0.968537	0.986701			0.989539
[2,]	0.989214	0	0.980514	0.980713	0.982671			0.991045
[3,]	0.971138	0.980514	0	0.953618	0.976725			0.98471
[4,]	0.968537	0.980713	0.953618	0	0.947647			0.968189
[5,]	0.986701	0.982671	0.976725	0.947647	0			0.945542
|								
|								
[3030,]	0.989539	0.991045	0.98471	0.968189	0.945542			0

Table 6: (a) dissimilarity measures for mRNA

Module Detection

For miRNA, ER+ samples
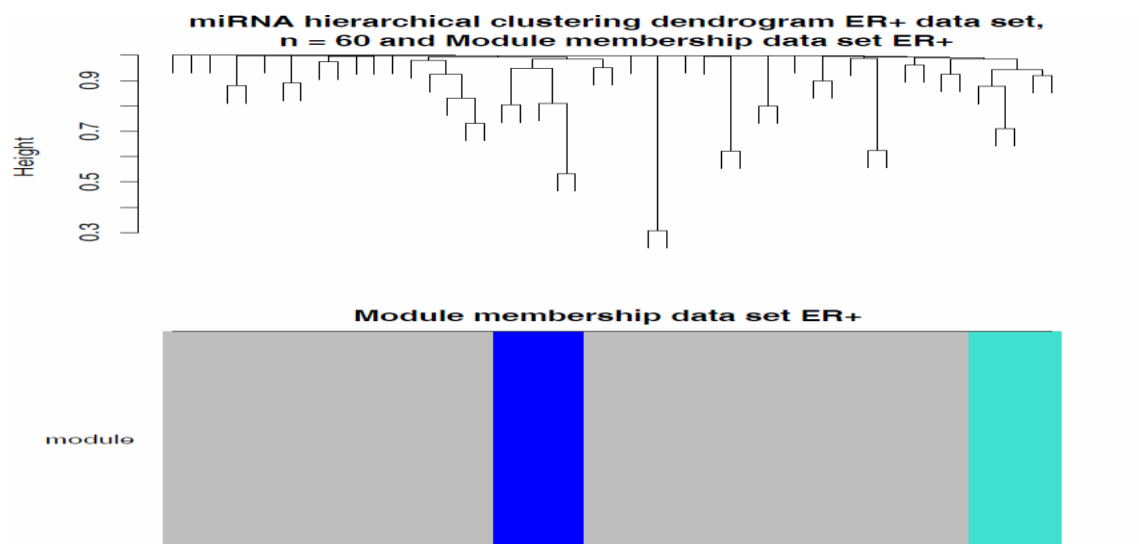
From, clustering we can say that miRNA modules are highly preserved.
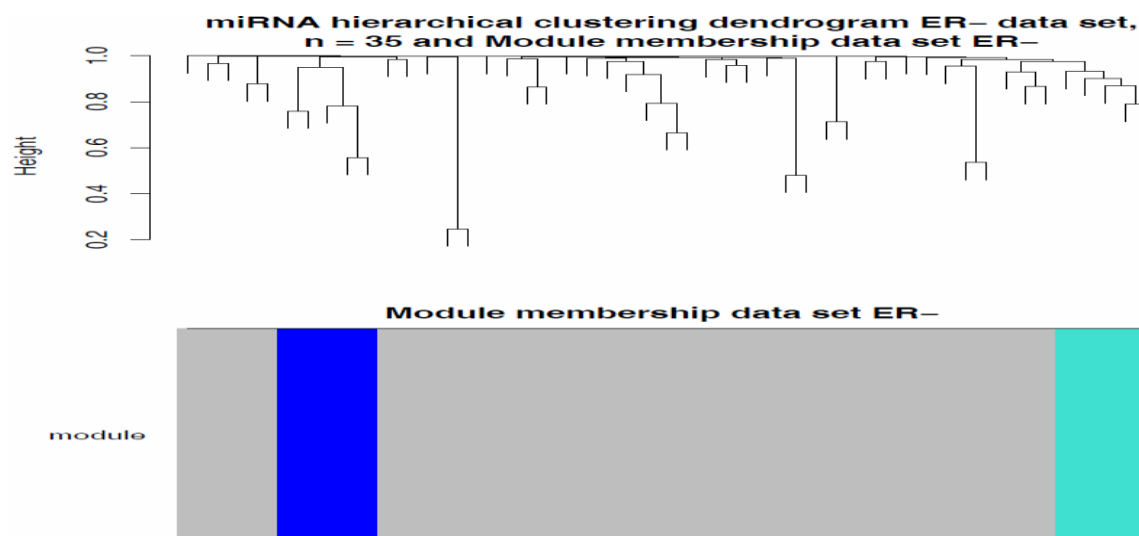
ER- samples

Figure 8: (a) Hierarchical clustering for miRNA ER+ and ER- samples

For mRNA,

ER+ samples

mRNA hierarchical clustering dendrogram ER+ data set,
n = 60 and Module membership data set ER+

Module membership data set ER+
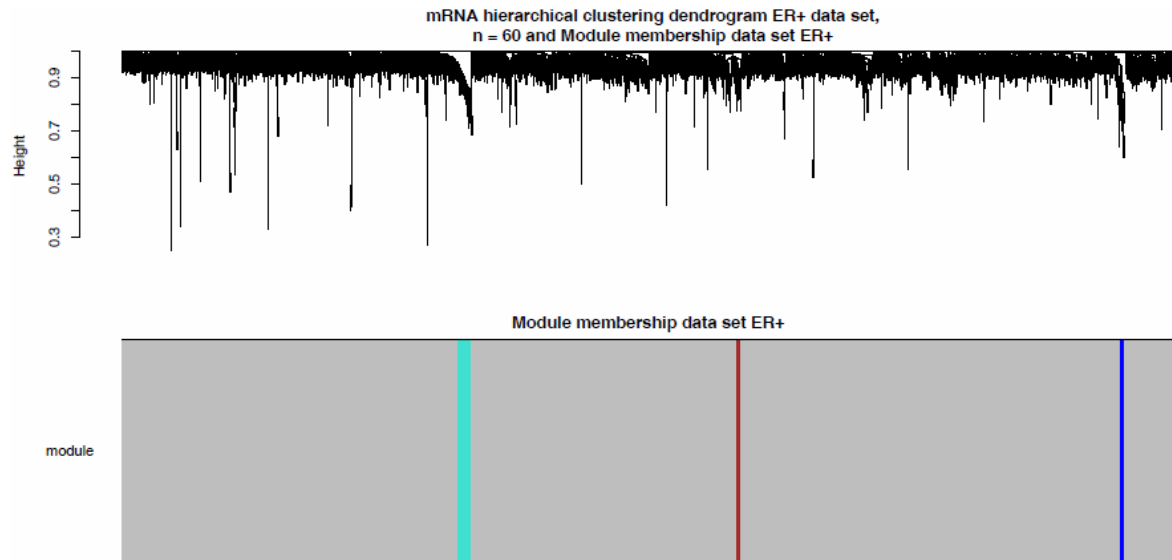
From, clustering we can say that mRNA modules are not highly preserved.

ER- samples

mRNA hierarchical clustering dendrogram ER– data set,
n = 35 and Module membership data set ER–

Module membership data set ER–

Figure 8: (b) Hierarchical clustering for mRNA ER+ and ER- samples

The whole module-wise network is shown in the following table:
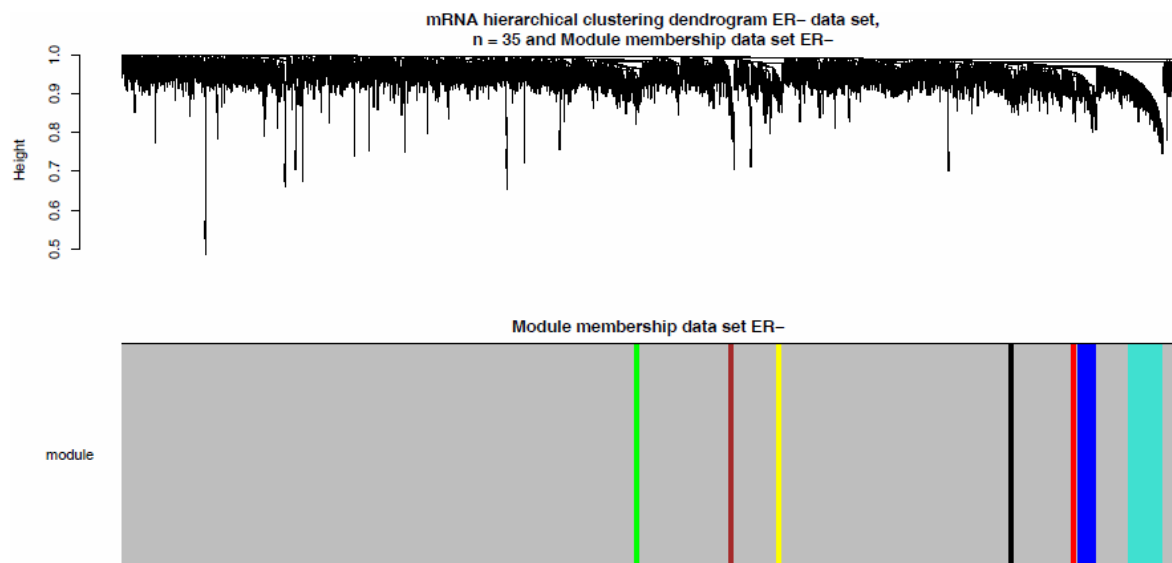
For miRNA ER+ samples is stored in: CancerNetmiRNAER+

colorh1	miRNA	pAdjusted.index	rawp	BH
blue	hsa-miR-505	335	7.24E-06	0.000383
blue	hsa-miR-18a	122	0.000333	0.008364
blue	hsa-miR-505*	336	0.000662	0.012271
blue	hsa-miR-19a	144	0.001451	0.019222
blue	hsa-miR-18b	123	0.001511	0.019484

turquoise	hsa-miR-29c*	204	2.65E-10	1.26E-07
turquoise	hsa-miR-190b	125	7.46E-09	1.15E-06
turquoise	hsa-miR-29b-2*	202	2.63E-06	0.000179
turquoise	hsa-miR-29c	203	3.64E-05	0.001336
turquoise	hsa-miR-148b	84	0.000956	0.015193

grey	hsa-miR-149	85	1.42E-09	3.39E-07
grey	hsa-miR-342-3p	246	9.66E-09	1.15E-06
grey	hsa-miR-342-5p	247	4.53E-08	4.32E-06
grey	hsa-miR-339-5p	240	5.09E-07	4.04E-05
grey	hsa-miR-224	178	5.74E-06	0.000342
grey	hsa-miR-146b-5p	81	2.95E-05	0.001293
grey	hsa-miR-99b	476	3.12E-05	0.001293
grey	hsa-miR-135b	63	3.25E-05	0.001293
grey	hsa-let-7e*	10	4.59E-05	0.001564
grey	hsa-miR-374a	266	5.48E-05	0.001742
grey	hsa-miR-339-3p	239	6.13E-05	0.001826
grey	hsa-miR-628-3p	409	0.000233	0.00653
grey	hsa-miR-499-5p	327	0.000306	0.0081
grey	hsa-miR-125a-5p	42	0.000386	0.009212
grey	hsa-miR-452	302	0.000488	0.010283
grey	hsa-miR-223	176	0.000489	0.010283
grey	hsa-miR-625	406	0.000496	0.010283
grey	hsa-miR-26a	187	0.000617	0.012262
grey	hsa-miR-10b*	29	0.000669	0.012271
grey	hsa-miR-9	456	0.000781	0.013752

grey	hsa-miR-629	411	0.000807	0.013752
grey	hsa-miR-9*	457	0.000935	0.015193
grey	hsa-miR-623	404	0.001177	0.018108
grey	hsa-miR-181d	111	0.001266	0.018434
grey	hsa-miR-378*	277	0.001275	0.018434
grey	hsa-miR-23a	179	0.001318	0.018497
grey	hsa-miR-148a	82	0.001365	0.018599
grey	hsa-miR-423-5p	288	0.001982	0.024122
grey	hsa-miR-326	227	0.001986	0.024122
grey	hsa-miR-101*	19	0.002023	0.024122
grey	hsa-miR-103	20	0.002294	0.026688
grey	hsa-miR-432	296	0.002761	0.031356
grey	hsa-miR-424	289	0.00307	0.03405
grey	hsa-let-7e	9	0.004105	0.044073
grey	hsa-miR-26b*	189	0.004158	0.044073
grey	hsa-miR-146a	80	0.004494	0.046599
grey	hsa-let-7i	14	0.00471	0.047801
grey	hsa-miR-30a*	209	0.004912	0.048653
grey	hsa-miR-375	269	0.004998	0.048653

miRNA ER- samples is stored in : CancerNetmiRNAER-

colorh2	miRNA	pAdjusted.index	rawp	BH
blue	hsa-miR-505	335	7.24E-06	0.000383
blue	hsa-miR-18a	122	0.000333	0.008364
blue	hsa-miR-505*	336	0.000662	0.012271
blue	hsa-miR-19a	144	0.001451	0.019222
blue	hsa-miR-18b	123	0.001511	0.019484

turquoise	hsa-miR-29c*	204	2.65E-10	1.26E-07
turquoise	hsa-miR-190b	125	7.46E-09	1.15E-06
turquoise	hsa-miR-29b-2*	202	2.63E-06	0.000179
turquoise	hsa-miR-29c	203	3.64E-05	0.001336
turquoise	hsa-miR-148b	84	0.000956	0.015193

grey	hsa-miR-149	85	1.42E-09	3.39E-07
grey	hsa-miR-342-3p	246	9.66E-09	1.15E-06
grey	hsa-miR-342-5p	247	4.53E-08	4.32E-06
grey	hsa-miR-339-5p	240	5.09E-07	4.04E-05
grey	hsa-miR-224	178	5.74E-06	0.000342
grey	hsa-miR-146b-5p	81	2.95E-05	0.001293
grey	hsa-miR-99b	476	3.12E-05	0.001293
grey	hsa-miR-135b	63	3.25E-05	0.001293
grey	hsa-let-7e*	10	4.59E-05	0.001564
grey	hsa-miR-374a	266	5.48E-05	0.001742
grey	hsa-miR-339-3p	239	6.13E-05	0.001826
grey	hsa-miR-628-3p	409	0.000233	0.00653
grey	hsa-miR-499-5p	327	0.000306	0.0081
grey	hsa-miR-125a-5p	42	0.000386	0.009212
grey	hsa-miR-452	302	0.000488	0.010283
grey	hsa-miR-223	176	0.000489	0.010283
grey	hsa-miR-625	406	0.000496	0.010283
grey	hsa-miR-26a	187	0.000617	0.012262
grey	hsa-miR-10b*	29	0.000669	0.012271
grey	hsa-miR-9	456	0.000781	0.013752
grey	hsa-miR-629	411	0.000807	0.013752
grey	hsa-miR-9*	457	0.000935	0.015193
grey	hsa-miR-623	404	0.001177	0.018108
grey	hsa-miR-181d	111	0.001266	0.018434
grey	hsa-miR-378*	277	0.001275	0.018434
grey	hsa-miR-23a	179	0.001318	0.018497
grey	hsa-miR-148a	82	0.001365	0.018599
grey	hsa-miR-423-5p	288	0.001982	0.024122
grey	hsa-miR-326	227	0.001986	0.024122
grey	hsa-miR-101*	19	0.002023	0.024122
Grey	hsa-miR-103	20	0.002294	0.026688
Grey	hsa-miR-432	296	0.002761	0.031356
Grey	hsa-miR-424	289	0.00307	0.03405
Grey	hsa-let-7e	9	0.004105	0.044073
Grey	hsa-miR-26b*	189	0.004158	0.044073
Grey	hsa-miR-146a	80	0.004494	0.046599
Grey	hsa-let-7i	14	0.00471	0.047801
Grey	hsa-miR-30a*	209	0.004912	0.048653
Grey	hsa-miR-375	269	0.004998	0.048653

Table 7: module-wise listing for miRNA

Module-wise listing of miRNAs and mRNAs for ER+ and ER- samples are stored in a file CancerNetmRNAER+ and CancerNetmiRNAER- respectively.

Script 4: miRNA Functional Target Prediction

The script was implemented in JAVA. We used MicroCosm Target Version 5 data base to predict the targets, which was implemented on miRanda algorithm. miRanda is described in methods. It reads the miRNA predefined targets and mRNA in modules from the individual file. The miRNA targeted gene is found in the mRNA module using dynamic programming. This represents module-wise target prediction.

The module-wise results are stored in mimTargetERpos and mimTargetERneg

Script 5: Fisher's Exact Test

This was executed in R. The theory is described in Method section. Fisher's exact test shows the miRNA targets are enriched within the mRNA modules or not. Here, we used one-sided p-value that shows positive association between miRNA modules and mRNA modules that means miRNA targets are enriched in mRNA modules.

Here are the results of comparison of differentially expressed Vs non-differentially expressed targets of the miRNA modules in ER+ and ER- samples.

	miRNA Module	ER+ samples ER- samples	P-value (One sided)	Significance P-value < 0.05
1	Blue	ER+	0.3953	No
		ER-	0.3953	No

2	Turquoise	ER+	1	No
		ER-	1	No
3	Grey	ER+	0.9878	No
		ER-	0.9773	No

Table 8: (a) Fisher's test for miRNA modules enriched within differentially expressed genes and non-differentially expressed genes

From above table we can see that, the one sided P-value is not less than 0.05 which is statistically not significant. It means that the same genes that are not targeted by miRNAs modules are found in the mRNA modules.

The module-wise results had shown in the table below:

ER+ samples:

	miRNA Module	mRNA Module	P – value (One sided)	Significance P-value < 0.05
1	Blue	Blue	1	No
		Brown	0.7533	No
		Turquoise	0.7740	No
		Grey	0.1829	No
2	Turquoise	Blue	0.1526	No
		Brown	0.8851	No
		Turquoise	0.1247	No
		Grey	0.9347	No
3	Grey	Blue	0.8767	No
		Brown	0.9828	No
		Turquoise	0.5962	No
		Grey	0.1284	No

Table 8: (b) Fisher's test for miRNA modules enriched within mRNA modules

ER- samples:

	miRNA Module	mRNA Module	P – value (One sided)	Significance P-value < 0.05
1	Blue	Black	6.3*10^(-5)	Yes
		Blue	0.9466	No
		Brown	0.5724	No
		Green	0.5386	No
		Red	0.5030	No
		Turquoise	0.8062	No
		Yellow	1	No
		Grey	0.5673	No
2	Turquoise	Black	0.2233	No
		Blue	0.3939	No
		Brown	0.7956	No
		Green	0.9439	No
		Red	0.7335	No
		Turquoise	0.2247	No
		Yellow	0.7663	No
		Grey	0.7052	No
3	Grey	Black	0.5260	No
		Blue	0.9491	No
		Brown	0.02834	Yes
		Green	0.2473	No
		Red	0.7192	No
		Turquoise	0.3694	No
		Yellow	0.4277	No
		Grey	0.7093	No

Table 8: (c) Fisher's test for miRNA modules enriched within mRNA modules

From above tables, we didn't find one sided P-value is not less than 0.05 which is statistically not significant. It means that the same genes targeted by miRNAs modules are not enriched in the mRNA modules in ER+ samples; but we found that the one sided P-value is less than 0.05 which is statistically significant. It means that the same

genes targeted by miRNA Blue modules are enriched in the mRNA Black modules.

FUNCTIONAL ENRICHMENT ANALYSIS:

Functional Enrichment Analysis is done using web based tool DAVID. As we found different modules for miRNA and mRNA, here we presented module-wise Functional Enrichment is shown in ER+ samples and ER- samples. We combined the targets of miRNA one module and miRNA other module in mRNA module, we showed the enrichment of miRNA module to mRNA module. The result contains Gene Symbol, Gene Name, Chromosome, KEGG Pathways, GO terms: BP (Biological Process), CC (Cellular Component) and MF (Molecular Function). The result stored for modules in appendix.

From Fisher's Exact Test, miRNA Blue and Grey modules and mRNA Black and Brown modules are significant in ER- samples. The significant mRNA modules i.e. Black and Brown were shown.

Functional enrichment analysis for individual modules stored in the following files:

Black mRNA module for ER- samples: annotationtableBL.xlsx

Blue mRNA module for ER- samples: annotationtableBLU.xlsx

Brown mRNA module for ER- samples: annotationtableBR.xlsx

Green mRNA module for ER- samples: annotationtableGR.xlsx

Red mRNA module for ER- samples: annotationtableR.xlsx

Turquoise mRNA module for ER- samples: annotationtableTUR.xlsx

Yellow mRNA module for ER- samples: annotationtableY.xlsx

Clustering information for combined targets of miRNA two modules with their biological processes is stored in the table below; we merely showed those categories whose FDR is less than 10%:

Blue Module of mRNA targeted by miRNA Blue and Turquoise module

Category	Term	Count	%	PValue	Genes	Fold Enrichment	FDR
GOTERM_BP_FAT	GO:0001501~ skeletal system development	3	27.27 273	0.0078 09	BMP1, COL3A1, TRAF6	18.17465	9.5303 12

Table 9: (a) functional enrichment analysis for mRNA

Other mRNA targeted by miRNA Blue and Turquoise module

Category	Term	Count	%	PValue	Genes	Fold Enrichment	FDR
GOTERM_BP_FAT	GO:0006898~ receptor-mediated endocytosis	7	1.343 57	0.0040 91	DAB2, LRP1, DRD3, FOLR1, IGF2R, SORL1, LRP2	4.557951	6.8728 22

Table 9: (b) functional enrichment analysis for mRNA

Interpretation of above tables: miRNA in Blue and Turquoise modules are involved in 2 different biological processes such as skeletal system development and receptor mediated endocytosis. More information stored as feabluembluenturmi.xlsx and feagreymbluenturmi.xlsx

Other mRNA targeted by miRNA Blue module and other miRNAs

Category	Term	Count	%	PValue	Genes	Fold Enrichment	FDR
GOTERM_BP_FAT	GO:0018212~peptidyl-tyrosine modification	11	0.784593	0.004021	OSM, IL12RB2, TPST1, ERBB4, FYN, IFNG, CLK4, RELN, INSR, DDR2, BTK	2.867869	7.097772
GOTERM_BP_FAT	GO:0016477~cell migration	36	2.56776	0.004236	NRTN, CCK, GIPC1, CXCL12, ITGAM, VCAM1, AZU1, CTTNBP2, SBDS, DOCK2, OVOL2, SAA1, IFNG, KRT2, CLASP2, CAP1, NR2F1, PTPRK, LMX1B, BARHL2, NEUROG2, SIX4, COL5A1, SLIT2, CDH13, ID1, FYN, ITGA5, LRP6, ADAM17, HBEGF, RELN, SELE, PLAU, MYH10, LRP5	1.632305	7.462677
GOTERM_BP_FAT	GO:0001501~skeletal system development	40	2.853067	0.004851	MMP9, CYTL1, FHL2, HOXD13, POSTN, EXTL1, GLI3, GLI1, SBDS, HOXC9, CHD7, HOXA4, JUND, ANKRD11, COL12A1, COL11A2, AXIN2, PCSK5, MINPP1, CMKLR1, DLL3, HSPG2, IGF1, SIX4, SMAD1, NPR3, CACNA1S, INHBA, CTSK, CHRDL2, RPS6KA3, TULP3, HOXC11, KAZALD1, COL1A2, TFAP2A, STC1, ATP6V0A4, EIF2AK3, CDH11	1.569196	8.502661

46

GOTERM_BP_FAT	GO:0007155~cell adhesion	76	5.420827	0.004973	DLC1, CADM4, MYBPC2, COL21A1, CLDN6, CLSTN1, CASK, LMO7, L1CAM, POSTN, LY9, CD151, CXCL12, DDR2, CDH22, AZGP1, CD96, DGCR6, COL12A1, DLG5, ZYX, COL11A2, BOC, CIB1, F11R, PTPRK, ICAM4, ICAM2, FLOT2, COL22A1, PCDHB2, ACTN1, MFGE8, CTNNA1, AMBP, HES1, ARVCF, CD36, HAS1, GPR56, NPTN, ADAM17, LAMC2, RELN, TGFB1I1, LIMS1, GNE, CDH3, ITGAM, ALCAM, VCAM1, NPHP4, LGALS3BP, COL7A1, FAT4, ITGB7, BAI1, CD4, SSX2IP, HAPLN3, LRRN2, COL15A1, HSPG2, CD99, ITGA3, NID2, COL5A1, MUC4, CDH13, COL14A1, ITGA5, PKP3, SELE, IL2, FEZ1, CDH11	1.3587	8.707062
GOTERM_BP_FAT	GO:0010324~membrane invagination	30	2.1398	0.005098	DRD3, LDLR, ADORA2A, SORL1, SNX2, EEA1, ITSN2, ASGR1, DAB2, FOLR1, FCN2, CAP1, TRIP10, DBNL, MRC1L1, MFGE8, ELMO3, LMBR1L, CD36, LRP1, GAPVD1, IGF2R, LRP12, LRP6, SH3KBP1, LRP2, BIN1, CD14, DNM1, LRP5	1.706501	8.915501
GOTERM_BP_FAT	GO:0006897~endocytosis	30	2.1398	0.005098	DRD3, LDLR, ADORA2A, SORL1, SNX2, EEA1, ITSN2, ASGR1, DAB2, FOLR1, FCN2, CAP1, TRIP10, DBNL, MRC1L1, MFGE8, ELMO3, LMBR1L, CD36, LRP1, GAPVD1, IGF2R, LRP12, LRP6, SH3KBP1, LRP2, BIN1, CD14, DNM1, LRP5	1.706501	8.915501

47

GOTERM_BP_FAT	GO:0022610~biological adhesion	76	5.420827	0.005125	DLC1, CADM4, MYBPC2, COL21A1, CLDN6, CLSTN1, CASK, LMO7, L1CAM, POSTN, LY9, CD151, CXCL12, DDR2,	1.356761	8.60429

					CDH22, AZGP1, CD96, DGCR6, COL12A1, DLG5, ZYX, COL11A2, BOC, CIB1, F11R, PTPRK, ICAM4, ICAM2, FLOT2, COL22A1, PCDHB2, ACTN1, MFGE8, CTNNA1, AMBP, HES1, ARVCF, CD36, HAS1, GPR56, NPTN, ADAM17, LAMC2, RELN, TGFB1I1, LIMS1, GNE, CDH3, ITGAM, ALCAM, VCAM1, NPHP4, LGALS3BP, COL7A1, FAT4, ITGB7, BAI1, CD4, SSX2IP, HAPLN3, LRRN2, COL15A1, HSPG2, CD99, ITGA3, NID2, COL5A1, MUC4, CDH13, COL14A1, ITGA5, PKP3, SELE, IL2, FEZ1, CDH11		
GOTERM_BP_FAT	GO:0040008~regulation of growth	42	2.99572	0.005275	MYOD1, IL9R, DRD3, MYL2, GDF5, DDR2, DAB2, FAM107A, CHD7, HSF1, PPP2CA, IFNG, CREG1, TAF9, ACTL6A, AGRN, PRL, INSR, BRMS1L, ENO1, ADAM10, CRYAB, SPTBN4, BARHL2, ATRN, PSRC1, IGF1, CAPRIN2, OSM, VEGFB, NTRK3, INHBA, CDH13, CTH, EP300, LRP12, KAZALD1, HBEGF, ADAM17, UTS2R, IGFBP2, IL2	1.541355	9.211902
GOTERM_BP_FAT	GO:0006333~chromatin assembly or disassembly	20	1.426534	0.005665	HIST2H2AA3, HIST1H2BC, HIST1H1C, SUV39H1, H1FX, NAP1L3, CBX6, CHD8, CHD7, SET, CDYL2, HIST1H2BI, HIST1H3A, CHD1, H2AFX, HIST3H2A, CDY2A, HIST1H2AM, ASF1A, HDAC8, HIST1H4H	1.970762	9.859617

Table 9: (c) functional enrichment analysis for mRNA

Interpretation of above tables: miRNA in Blue module and other miRNAs are involved in different biological processes such as cell migration, cell and biological adhesion, regulation of growth, skeletal system development and endocytosis, etc. more information stored in feagreymbluengreymi.xlsx

Other mRNA targeted by miRNA Turquoise module and other miRNAs

Category	Term	Count	%	PValue	Genes	Fold Enrichment	FDR
GOTERM_BP_FAT	GO:0030198~extracellular matrix organization	19	1.371841	0.001169	LMX1B, MMP9, ADAMTSL4, ELN, HSPG2, ANXA2P1, SPINK5, COL5A1, APLP2, COL14A1, KAZALD1, SMOC1, FOXF1, COL1A2, COL12A1, LOX, COL11A2, B4GALT7, COL11A1	2.320621	2.111877
GOTERM_BP_FAT	GO:0001501~skeletal system development	42	3.032491	0.001219	MMP9, CYTL1, FHL2, HOXD13, EXTL1, GLI3, GLI1, SBDS, HOXC9, CHD7, HOXA4, JUND, ANKRD11, COL12A1, PKD1, COL11A2, AXIN2, COL11A1, PCSK5, MINPP1, CMKLR1, DLL3, HSPG2, IGF1, SIX4, SMAD1, NPR3, ANXA2P1, CACNA1S, INHBA, CTSK, CHRDL2, RPS6KA3, TULP3, HOXC11, KAZALD1, COL1A2, TFAP2A, STC1, ATP6V0A4, EIF2AK3, CDH11	1.672409	2.201234
GOTERM_BP_FAT	GO:0030199~collagen fibril organization	9	0.649819	0.001374	COL14A1, LMX1B, COL1A2, COL12A1, LOX, COL11A2, ANXA2P1, COL11A1, COL5A1	3.942108	2.477641

GOTERM_BP_FAT	GO:0016477~cell migration	37	2.671 48	0.0018 18	NRTN, CCK, GIPC1, CXCL12, ITGAM, VCAM1, AZU1, CTTNBP2, SBDS, DOCK2, OVOL2, SAA1, CKLF, IFNG, KRT2, CLASP2, CAP1, NR2F1, PTPRK, LMX1B, BARHL2, NEUROG2, SIX4, COL5A1, SLIT2, CDH13, ID1, FYN, ITGA5, LRP6, ADAM17, HBEGF, RELN, SELE, PLAU, MYH10, LRP5	1.7028 51	3.2653 6
GOTERM_BP_FAT	GO:0018108~peptid yl-tyrosine phosphory lation	11	0.794 224	0.0025 98	OSM, IL12RB2, ERBB4, FYN, IFNG, CLK4, ABI1, RELN, INSR, DDR2, BTK	3.0375 18	4.6358 74
GOTERM_BP_FAT	GO:0048870~cell motility	39	2.815 884	0.0035 01	SLC22A16, NRTN, CCK, GIPC1, CXCL12, ITGAM, VCAM1, AZU1, CTTNBP2, SBDS, DOCK2, OVOL2, SAA1, CKLF, IFNG, KRT2, CLASP2, CAP1, NR2F1, PTPRK, LMX1B, BARHL2, NEUROG2, SIX4, COL5A1, SLIT2, CDH13, ID1, FYN, ITGA5, CATSPER1, LRP6, ADAM17, HBEGF, RELN, SELE, PLAU, MYH10, LRP5	1.6136 53	6.1997 35
GOTERM_BP_FAT	GO:0051674~localiz ation of cell	39	2.815 884	0.0035 01	SLC22A16, NRTN, CCK, GIPC1, CXCL12, ITGAM, VCAM1, AZU1, CTTNBP2, SBDS, DOCK2, OVOL2, SAA1, CKLF, IFNG, KRT2, CLASP2, CAP1, NR2F1, PTPRK, LMX1B, BARHL2, NEUROG2, SIX4, COL5A1, SLIT2, CDH13, ID1, FYN, ITGA5, CATSPER1, LRP6, ADAM17, HBEGF, RELN, SELE, PLAU, MYH10, LRP5	1.6136 53	6.1997 35
GOTERM_BP_FAT	GO:0018212~peptid yl-tyrosine modificati on	11	0.794 224	0.0036 13	OSM, IL12RB2, ERBB4, FYN, IFNG, CLK4, ABI1, RELN, INSR, DDR2, BTK	2.9109 55	6.3910 08

					ERBB4, LMX1B, MMP9, ADAMTSL4, ELN, HSPG2, PCDHB2, CACNB4, ANXA2P1, CACNA1S, SPINK5, COL5A1, APLP2, COL14A1, KAZALD1, SMOC1, FOXF1, COL1A2, COL12A1, AGRN, LOX, COL11A2, B4GALT7, COL11A1		
GOTERM_ BP_FAT	GO:00430 62~extrac ellular structure organizati on	24	1.732 852	0.0043 77		1.8702 84	7.6926 98

Table 9: (d) functional enrichment analysis for mRNA

Interpretation of above tables: miRNA in Turquoise module and other miRNAs are involved in different biological processes such as cell migration, cell motility, localization of cell, skeletal system development, extracellular matrix organization and extracellular structure organization. More information stored as feagreymturngreymi.xlsx

Functional enrichment analysis for individual modules stored in the following files:

Blue mRNA module for ER- samples: annotationBLU.xlsx

Other mRNA module for ER- samples: annotationGREY.xlsx

Results for genes in all modules of ER+ samples with Gene Ontology and KEGG Pathway information with genes location on chromosome were shown in the following files:

Blue mRNA module for ER+ samples: annotationtableBLU.xlsx

Brown mRNA module for ER+ samples: annotationtableBR.xlsx

Turquoise mRNA module for ER+ samples: annotationtableTUR.xlsx

Functional enrichment analysis for individual modules stored for ER+ samples in the following files:

Other mRNA for ER+ samples: annotationGREY.txt

UP/DOWN REGULATION OF miRNAS AND GENES

Finally, key genes targeted by miRNAs within all modules were identified. Then their up regulation and down regulation compare to estrogen receptor negative samples were demonstrated.

miRNA expression with Estrogen Receptor negative factor is shown in the table below.

Modules	Up Regulated in ER- samples	Down Regulated in ER- samples
BLUE	hsa-miR-18a, hsa-miR-18b, hsa-miR-19a, hsa-miR-505, hsa-miR-505*	
TURQUOISE		hsa-miR-29c*, hsa-miR-190b, hsa-miR-29b-2*, hsa-miR-29c, hsa-miR-148b
GREY	hsa-miR-224, hsa-miR-146b-5p, hsa-miR-135b, hsa-miR-374a, hsa-miR-452, hsa-miR-223, hsa-miR-9, hsa-miR-9*, hsa-miR-378*, hsa-miR-23a, hsa-miR-148a, , hsa-miR-424	hsa-miR-149, hsa-miR-342-3p, hsa-miR-342-5p, hsa-miR-339-5p, hsa -miR-99b, hsa-let-7e*, hsa-miR-339-3p, hsa-miR-628-3p, hsa-miR-499-5p, hsa-miR-125a-5p, hsa-miR-625, hsa-miR-26a, hsa-miR-10b*, hsa-miR-629, hsa-miR-623, hsa-miR-181d, hsa-miR-423-5p, hsa-miR-326, hsa-miR-101*, hsa-miR-103, hsa-miR-432, hsa-let-7e, hsa-miR-26b*, hsa-miR-146a, hsa-let-7i, hsa-miR-30a*, hsa-miR-375

Table 10: (a) module-wise up/down regulation for miRNA in ER-

Based on literature, some miRNAs, shown in Red color, are related to breast cancer.

mRNA expression with ER+ vs ER- factor are stored as updownregulation.xslx , it shows up and down regulation of genes in ER- samples.

Modules	Up Regulated in ER- samples	Down Regulated in ER- samples
BLACK	NUP133,TCP11L1,ARHGAP29, FBXO46,PAPSS1,C21orf99, DHRS1,NR1H4,RPL7L1,SPRED2,OLFML3	PEX3,CR2,WNT11,BTNL9
BLUE	NES,C20orf194,FAM86B1,TLE2, MOCS3,HCLS1,CEBPB,ASH2L, LBP, ZNF563,LYK5,ASB13, PPFIBP2,STAG3,PLA2G3, GABRP,ACCN4,COL3A1,BMP1, CTAGEP,CASP8AP2,KCTD21, HES5,FRK,WDR76,MTERF, SULF1,LPIN1	APOBEC3D,PER2,OSBP2,ZNF294,TRAF3IP3, C7orf31,UTRN, FOXC1,SLC39A6,DNAJC5B, ZNF195,ZNF71,REEP3, MIPOL1, PODN,HR,TRAF6, ICT1,FTSJ2,UQCRFS1,SOCS2, INTS4,SNORA70,CERK,CCDC23, VWCE
BROWN		CYB561D1,MR1,PRSS21,LSS, ZNF289,ABCC5, DFNB31, PIGM,CD2, MUC3A,PHACTR2, IBCH,CYC1, C21orf121,RHPN1, GPATCH8
GREEN	TBX4	OMA1,PPFIA1,CCDC102A, NAV1, TMEM39A, PRKDC, FLRT3,SYT6, ANG, JPH3, BCAS3,LGALS3, KCNK16, RNH1, ZC3H18
RED	SAMD14,C13orf30,NANOG, KPNA6,SPATA13,RNF11, SLC35B2,RRP12,RBM15,ATP2A1, ZNF165,MDM4,TOP1MT,ACTR6	NKPD1

TURQUOISE	HIST1H2BM,LENG4,C3orf19, MLLT3,FBXO8,GOSR2,PAGE2, SETDB2,SREBF1,HOXD10, ATP6AP2, PROS1,C15orf48, MSH3,YPEL3,TNFSF11,PTPRZ1,MEGF10, IGSF21,FABP6,PMM2, GDI2,CYP4F11,KL,TANC1, RGS19,ZDHHC21,SPRY2, ENTPD3,MBOAT1,LSAMP, CTNNAL1,PCM1,PRB1,WDFY4, BEX1,TRPM4,CLEC14A,TRIP13, LCN12,AMD1,GPR110,RBPMS2, CDH5,C2orf15,UPF3A,PIP4K2A, COX11,SLC2A6,ZNF652, C10orf84,TM2D2,RPS6KA4, PLCH1,ENG, SMARCC2,HCRTR2, TCL1A,ISG20,OCA2,FGFRL1, TGS1,CD59,CD207,PIM1,ZBTB2, MAPK8IP2,DPH4,SH3D19, ZNF415,PRDM1,TTC16,FAM19A4, GALM,LPHN2,PLA2G12A,ACRV1, MGC24039,LAMB2,FAM29A, GPNMB,ZNF462,AKAP13	LOH3CR2A,LRSAM1, SYNGAP1, C13orf18, CREG1, CILP2,LPO,HLA-DQB1, MAP1A, AMTA1,C10orf107, C17orf68, BVES,BAG3,LOC51149, PBX1, LZIC,GRM8,SUSD3
YELLOW		TMEM58,MVD,CYP2U1,RARA, HPS5,LYPD6,DEPDC6,NBR1, BAHD1,APOL4,FCGR2B, LEPRE1, RIPK2,C21orf71, LILRB3,ROR

Table 10: (b) module-wise up/down regulation for mRNA in ER-

VISUALIZATION:

The visualization of module-wise miRNA - gene target prediction network is done using BiNA software. From the Fisher's Exact Test, we can determine that miRNA Blue and Grey Modules and mRNA Black and Brown modules are important in ER- samples.

miRNA Blue module targets mRNA Black module is shown in the table below:

miRNA Blue Module	Regulation in ER- samples	Target genes mRNA Black Module	Regulation in ER- samples
hsa-miR-505	Up	TCP11L1	Up
hsa-miR-18a	Up	--------	
hsa-miR-505*	Up	BTNL9, PAPSS1,	Down, Up
hsa-miR-19a	Up	NUP133, OLFML3, PEX3, CR2,	Up, Up, Down, Down
hsa-miR-18b	Up	DHRS1	Up

Table 11: (a) inverse correlation of miRNAs with their targets

From above table, we can say that miRNAs hsa-miR-505* and hsa-miR-19a have expression pattern that inversely correlated with targeted genes BTNL9 and PEX3, CR2, respectively. They definitely could be the functional targets.

Other miRNA targets mRNA Brown module is shown in the table below:

miRNA	Regulation in ER- samples	Target genes mRNA Brown Module	Regulation in ER- samples
hsa-miR-149	Down	RHPN1	Down
hsa-miR-342-3p	Down	LSS, CD2	Down, Down
hsa-miR-342-5p	Down	--------	
hsa-miR-339-5p	Down	DFNB31	Down
hsa-miR-224	Up	CHRM3	Down
hsa-miR-146b-5p	Up	--------	
hsa-miR-99b	Down	--------	
hsa-miR-135b	Up	CD2	Down
hsa-let-7e*	Down	--------	
hsa-miR-374a	Up	--------	
hsa-miR-339-3p	Down	CYC1, ZNF289	Down, Down
hsa-miR-628-3p	Down	--------	
hsa-miR-499-5p	Down	--------	
hsa-miR-125a-5p	Down	ZNF289	Down
hsa-miR-452	Up	--------	
hsa-miR-223	Up	--------	

hsa-miR-625	Down	--------	
hsa-miR-26a	Down	MR1	Down
hsa-miR-10b*	Down	--------	
hsa-miR-9	Up	--------	
hsa-miR-629	Down	CHRM3	Down
hsa-miR-9*	Up	HIBCH	Down
hsa-miR-623	Down	--------	
hsa-miR-181d	Down	--------	
hsa-miR-378*	Up	DFNB31	Down
hsa-miR-23a	Up	ABCC5	Down
hsa-miR-148a	Up	--------	
hsa-miR-423-5p	Down	RHPN1	Down
hsa-miR-326	Down	--------	
hsa-miR-101*	Down	--------	
hsa-miR-103	Down	--------	
hsa-miR-432	Down	--------	
hsa-miR-424	Up	PHACTR2	Down
hsa-let-7e	Down	MUC3A, ABCC5	Down, Down
hsa-miR-26b*	Down	--------	
hsa-miR-146a	Down	--------	
hsa-let-7i	Down	ABCC5	Down
hsa-miR-30a*	Down	--------	
hsa-miR-375	Down	--------	

Table 11: (b) inverse correlation of miRNAs with their targets

From above table, we can say that miRNAs hsa-miR-224, -135b, -424, -378* and -23a have expression pattern that inversely correlated with targeted genes CHRM3, CD2, HIBCH, DFNB31 and ABCC5.

Inverse correlation between miRNA and their targets suggests that they definitely could be the functional targets.

The whole module-wise miRNA – gene target prediction results are stored in a file;

ER+ samples: mimTargetERpos and ER- samples: mimTargetERneg

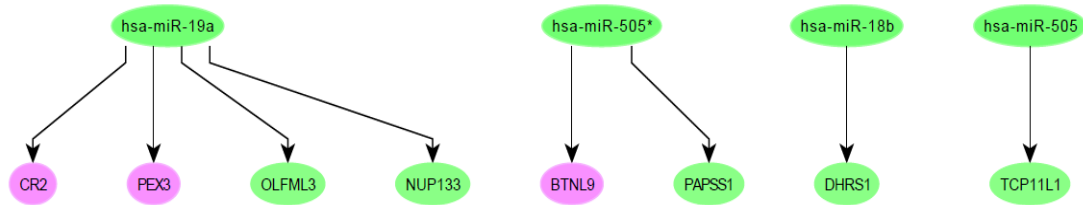miRNA Blue Module target the mRNA Black module:

Figure 9: (a) miRNA targeted genes in blue module of miRNA

Here pink color indicates down regulation of miRNA/gene and pistachio color indicates up regulation in ER- samples.

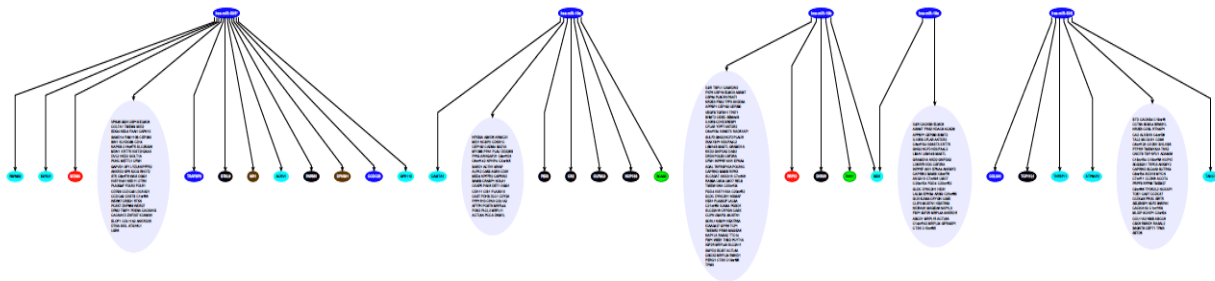miRNA Blue Module target the mRNA all modules including other mRNAs:

Figure 9: (b) miRNA targeted genes in blue module of miRNA

Similarly, miRNA Turquoise Module target the mRNA all module:
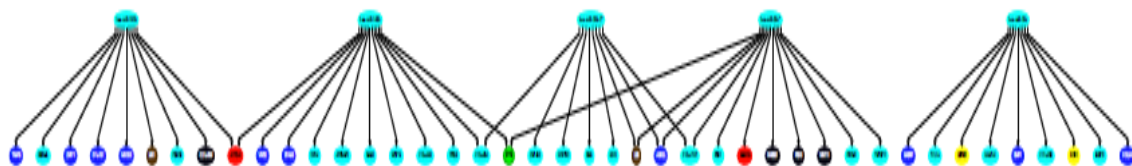

Figure 9: (c) miRNA targeted genes in turquoise module of miRNA

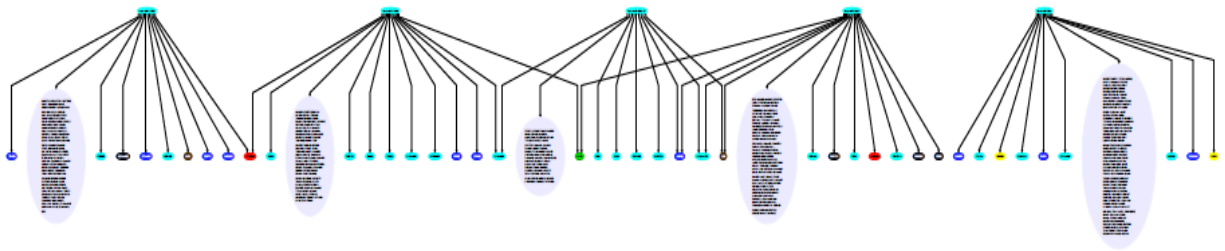miRNA Blue Module target the mRNA all modules including other mRNAs:

Figure 9: (d) miRNA targeted genes in turquoise module of miRNA

Similarly, miRNA other Module targets the mRNA Brown module:

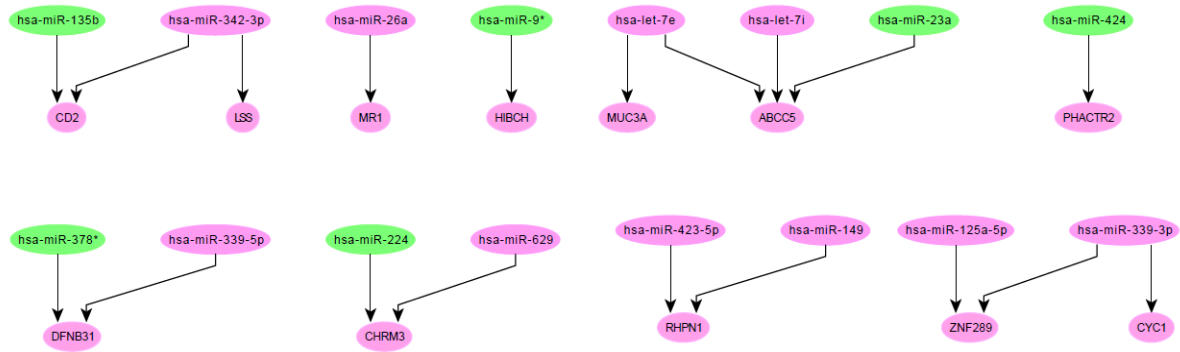Figure 9: (e) miRNA targeted genes in rest of miRNAs

Here pink color indicates down regulation of miRNA/gene and pistachio color indicates up regulation in ER- samples.

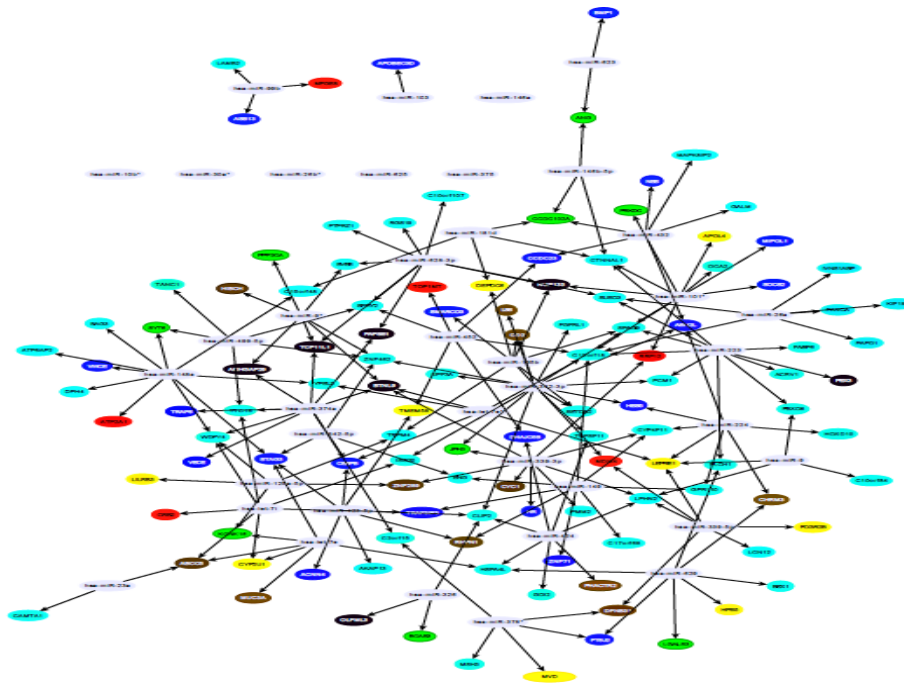miRNA other Module targets all mRNA modules including other mRNAs:

Figure 9: (f) miRNA targeted genes in rest of miRNAs

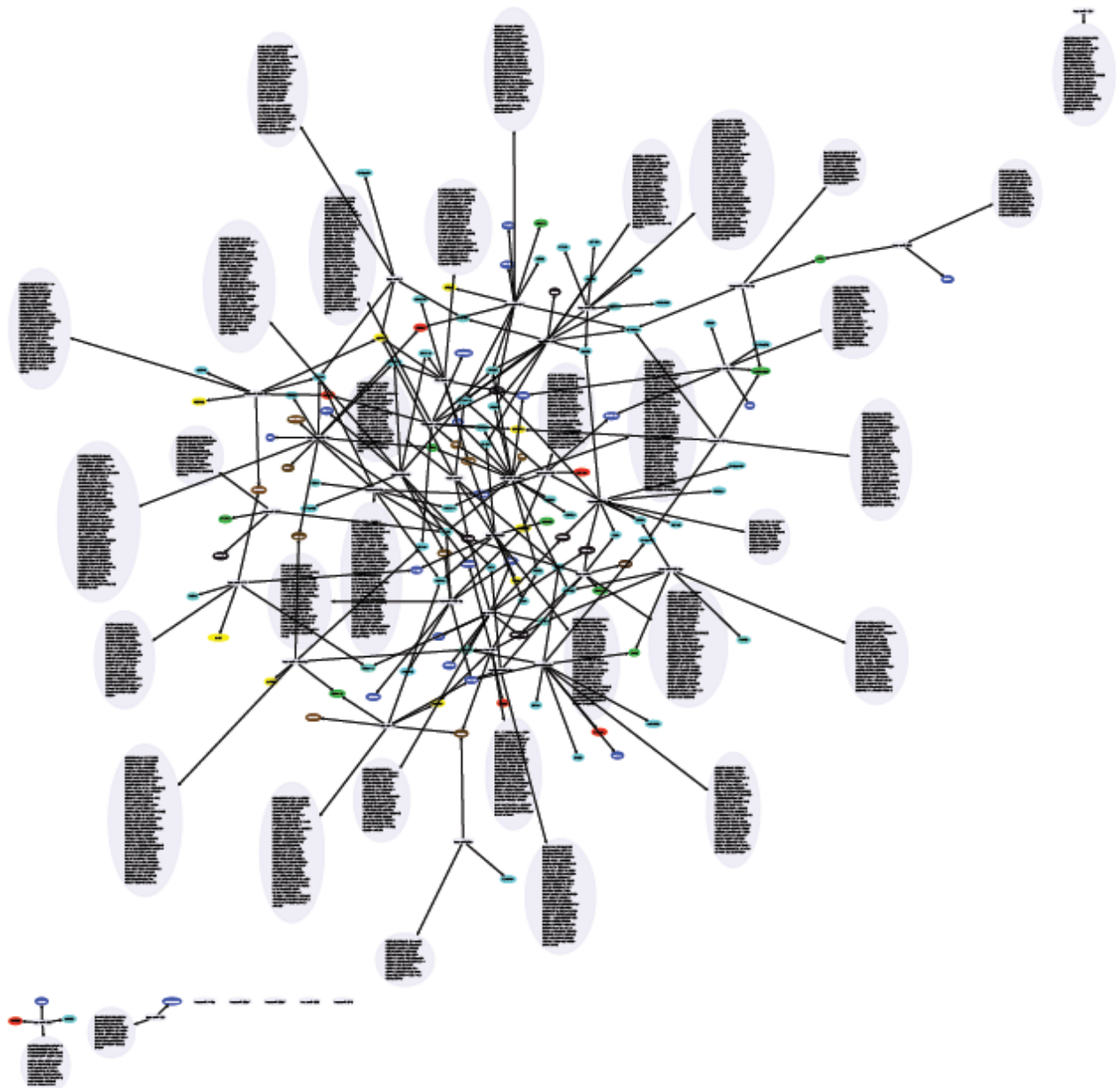miRNA other Module targets all mRNA modules including other mRNAs:

Figure 9: (g) miRNA targeted genes in rest of miRNAs

From Gene Ontology analysis, several biological processes captured in the miRNA targeted genes.

miRNA	GO ID	Biological Process Term
hsa-miR-18a		
hsa-miR-18b	GO:0055114	oxidation reduction,
hsa-miR-19a	GO:0002250	adaptive immune response,
	GO:0002252	immune effector process,
	GO:0002253	activation of immune response,
	GO:0002443	leukocyte mediated immunity,
	GO:0002449	lymphocyte mediated immunity,
	GO:0002455	humoral immune response mediated by circulating immunoglobulin,
	GO:0002460	adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains,
	GO:0002526	acute inflammatory response,
	GO:0002541	activation of plasma proteins involved in acute inflammatory response,
	GO:0002684	positive regulation of immune system process,
	GO:0006508	proteolysis,
	GO:0006952	defense response,
	GO:0006954	inflammatory response,
	GO:0006955	immune response,
	GO:0006956	complement activation,
	GO:0006958	complement activation, classical pathway,
	GO:0006959	humoral immune response,
	GO:0009611	response to wounding,
	GO:0016064	immunoglobulin mediated immune response,
	GO:0016485	protein processing,
	GO:0019724	B cell mediated immunity,
	GO:0045087	innate immune response,
	GO:0048584	positive regulation of response to stimulus,
	GO:0050778	positive regulation of immune response,
	GO:0051604	protein maturation,
	GO:0051605	protein maturation by peptide bond cleavage,
	GO:0006605	protein targeting,
	GO:0006612	protein targeting to membrane,
	GO:0006625	protein targeting to peroxisome,
	GO:0006886	intracellular protein transport,
	GO:0007031	peroxisome organization,
	GO:0008104	protein localization,
	GO:0015031	protein transport,

61

	GO:0016557	peroxisome membrane biogenesis,
	GO:0017038	protein import,
	GO:0034613	cellular protein localization,
	GO:0043574	peroxisomal transport,
	GO:0044091	membrane biogenesis,
	GO:0045046	protein import into peroxisome membrane,
	GO:0045184	establishment of protein localization,
	GO:0046907	intracellular transport,
	GO:0070727	cellular macromolecule localization
	GO:0006403	RNA localization,
	GO:0006405	RNA export from nucleus,
	GO:0006406	mRNA export from nucleus,
	GO:0006913	nucleocytoplasmic transport,
	GO:0006997	nucleus organization,
	GO:0006999	nuclear pore organization,
	GO:0007498	mesoderm development,
	GO:0008104	protein localization,
	GO:0015031	protein transport,
	GO:0015931	nucleobase, nucleoside, nucleotide and nucleic acid transport,
	GO:0034621	cellular macromolecular complex subunit organization,
	GO:0043933	macromolecular complex subunit organization,
	GO:0045184	establishment of protein localization,
	GO:0046907	intracellular transport,
	GO:0048339	paraxial mesoderm development,
	GO:0050657	nucleic acid transport,
	GO:0050658	RNA transport,
	GO:0051028	mRNA transport,
	GO:0051168	nuclear export,
	GO:0051169	nuclear transport,
	GO:0051236	establishment of RNA localization,
	GO:0055085	transmembrane transport,
hsa-miR-505		
hsa-miR-505*	GO:0000103	sulfate assimilation,
	GO:0001501	skeletal system development,
	GO:0006163	purine nucleotide metabolic process,
	GO:0006790	sulfur metabolic process,
	GO:0009165	nucleotide biosynthetic process,
	GO:0033865	nucleoside bisphosphate metabolic process,
	GO:0033866	nucleoside bisphosphate biosynthetic process,
	GO:0033875	ribonucleoside bisphosphate metabolic process,
	GO:0034030	ribonucleoside bisphosphate biosynthetic process,
	GO:0034032	purine nucleoside bisphosphate metabolic process,

	GO:0034033	purine nucleoside bisphosphate biosynthetic process,
	GO:0034035	purine ribonucleoside bisphosphate metabolic process,
	GO:0034036	purine ribonucleoside bisphosphate biosynthetic process,
	GO:0034404 '	nucleobase, nucleoside and nucleotide biosynthetic process,
	GO:0034654	nucleobase, nucleoside, nucleotide and nucleic acid biosynthetic process,
	GO:0044271	nitrogen compound biosynthetic process, 3'-phosphoadenosine 5'-phosphosulfate metabolic
	GO:0050427	process,
	GO:0050428	3'-phosphoadenosine 5'-phosphosulfate biosynthetic process,
hsa-let-7i hsa-miR-23a	GO:0002237	response to molecule of bacterial origin,
	GO:0009617	response to bacterium,
	GO:0009719	response to endogenous stimulus,
	GO:0009725	response to hormone stimulus,
	GO:0010033	response to organic substance,
	GO:0032496	response to lipopolysaccharide,
	GO:0032868	response to insulin stimulus,
	GO:0043434	response to peptide hormone stimulus,
	GO:0055085	transmembrane transport,
hsa-let-7e	GO:0002237	response to molecule of bacterial origin,
	GO:0009617	response to bacterium,
	GO:0009719	response to endogenous stimulus,
	GO:0009725	response to hormone stimulus,
	GO:0010033	response to organic substance,
	GO:0032496	response to lipopolysaccharide,
	GO:0032868	response to insulin stimulus,
	GO:0043434	response to peptide hormone stimulus,
	GO:0055085	transmembrane transport,
	GO:0001894	tissue homeostasis,
	GO:0007586	digestion,
	GO:0010669	epithelial structure maintenance,
	GO:0022600	digestive system process,
	GO:0030277	maintenance of gastrointestinal epithelium,
	GO:0042592	homeostatic process,
	GO:0048871	multicellular organismal homeostasis,
	GO:0060249	anatomical structure homeostasis,
hsa-miR-9*	GO:0009063	cellular amino acid catabolic process,
	GO:0009081	branched chain family amino acid metabolic process,

63

	GO:0009083	branched chain family amino acid catabolic process,
	GO:0009310	amine catabolic process,
	GO:0016054	organic acid catabolic process,
	GO:0046395	carboxylic acid catabolic process,
hsa-miR-26a	GO:0002474	antigen processing and presentation of peptide antigen via MHC classI,
	GO:0006955	immune response,
	GO:0019882	antigen processing and presentation,
	GO:0048002	antigen processing and presentation of peptide antigen,
hsa-miR-125a-5p		
hsa-miR-135b	GO:0001766	membrane raft polarization,
	GO:0001775	cell activation,
	GO:0002684	positive regulation of immune system process,
	GO:0002694	regulation of leukocyte activation,
	GO:0002696	positive regulation of leukocyte activation,
	GO:0006917	induction of apoptosis,
	GO:0007155	cell adhesion,
	GO:0007166	cell surface receptor linked signal transduction,
	GO:0010941	regulation of cell death,
	GO:0010942	positive regulation of cell death,
	GO:0012502	induction of programmed cell death,
	GO:0016044	membrane organization,
	GO:0016337	cell-cell adhesion,
	GO:0022610	biological adhesion,
	GO:0030101	natural killer cell activation,
	GO:0030885	regulation of myeloid dendritic cell activation,
	GO:0030887	positive regulation of myeloid dendritic cell activation,
	GO:0031579	membrane raft organization,
	GO:0031580	membrane raft distribution,
	GO:0042110	T cell activation,
	GO:0042981	regulation of apoptosis,
	GO:0043065	positive regulation of apoptosis,
	GO:0043067	regulation of programmed cell death,
	GO:0043068	positive regulation of programmed cell death,
	GO:0045321	leukocyte activation,
	GO:0045580	regulation of T cell differentiation,
	GO:0045619	regulation of lymphocyte differentiation,
	GO:0046649	lymphocyte activation,
	GO:0050863	regulation of T cell activation,
	GO:0050865	regulation of cell activation,
	GO:0050867	positive regulation of cell activation,
	GO:0051249	regulation of lymphocyte activation,
	GO:0051665	membrane raft localization,

	GO:0051668	localization within membrane,
hsa-miR-149, hsa-miR-423-5p		
hsa-miR-224 hsa-miR-629	GO:0003012	muscle system process,
	GO:0003056	regulation of vascular smooth muscle contraction,
	GO:0006936	muscle contraction,
	GO:0006937	regulation of muscle contraction,
	GO:0006939	smooth muscle contraction,
	GO:0006940	regulation of smooth muscle contraction,
	GO:0007166	cell surface receptor linked signal transduction,
	GO:0007186	G-protein coupled receptor protein signaling pathway,
	GO:0007213	muscarinic acetylcholine receptor signaling pathway,
	GO:0007586	digestion,
	GO:0008283	cell proliferation,
	GO:0019229	regulation of vasoconstriction,
	GO:0044057	regulation of system process,
	GO:0045933	positive regulation of muscle contraction,
	GO:0045987	positive regulation of smooth muscle contraction,
	GO:0051240	positive regulation of multicellular organismal process,
hsa-miR-339-3p	GO:0006091	generation of precursor metabolites and energy,
	GO:0022900	electron transport chain,
	GO:0055114	oxidation reduction,
hsa-miR-339-5p, hsa-miR-378*	GO:0001894	tissue homeostasis,
	GO:0001895	retina homeostasis,
	GO:0007267	cell-cell signaling,
	GO:0007268	synaptic transmission,
	GO:0007423	sensory organ development,
	GO:0007600	sensory perception,
	GO:0007601	visual perception,
	GO:0007605	sensory perception of sound,
	GO:0019226	transmission of nerve impulse,
	GO:0030030	cell projection organization,
	GO:0030182	neuron differentiation,
	GO:0042490	mechanoreceptor differentiation,
	GO:0042592	homeostatic process,
	GO:0043583	ear development,
	GO:0048666	neuron development,
	GO:0048839	inner ear development,
	GO:0048871	multicellular organismal homeostasis,
	GO:0050877	neurological system process,
	GO:0050890	cognition,
	GO:0050953	sensory perception of light stimulus,

65

	GO:0050954	sensory perception of mechanical stimulus,
	GO:0060113	inner ear receptor cell differentiation,
	GO:0060119	inner ear receptor cell development,
	GO:0060122	inner ear receptor stereocilium organization,
	GO:0060249	anatomical structure homeostasis
	GO:0001766	membrane raft polarization,
	GO:0001775	cell activation,
	GO:0002684	positive regulation of immune system process,
	GO:0002694	regulation of leukocyte activation,
	GO:0002696	positive regulation of leukocyte activation,
	GO:0006917	induction of apoptosis,
	GO:0007155	cell adhesion,
	GO:0007166	cell surface receptor linked signal transduction,
	GO:0010941	regulation of cell death,
	GO:0010942	positive regulation of cell death,
	GO:0012502	induction of programmed cell death,
	GO:0016044	membrane organization,
	GO:0016337	cell-cell adhesion,
	GO:0022610	biological adhesion,
	GO:0030101	natural killer cell activation,
	GO:0030885	regulation of myeloid dendritic cell activation,
	GO:0030887	positive regulation of myeloid dendritic cell activation,
	GO:0031579	membrane raft organization,
	GO:0031580	membrane raft distribution,
	GO:0042110	T cell activation,
	GO:0042981	regulation of apoptosis,
	GO:0043065	positive regulation of apoptosis,
	GO:0043067	regulation of programmed cell death,
	GO:0043068	positive regulation of programmed cell death,
	GO:0045321	leukocyte activation,
	GO:0045580	regulation of T cell differentiation,
	GO:0045619	regulation of lymphocyte differentiation,
	GO:0046649	lymphocyte activation,
	GO:0050863	regulation of T cell activation,
	GO:0050865	regulation of cell activation,
	GO:0050867	positive regulation of cell activation,
	GO:0051249	regulation of lymphocyte activation,
	GO:0051665	membrane raft localization,
	GO:0051668	localization within membrane,
	GO:0006694	steroid biosynthetic process,
	GO:0006695	cholesterol biosynthetic process,
	GO:0008202	steroid metabolic process,
	GO:0008203	cholesterol metabolic process,
	GO:0008610	lipid biosynthetic process,

hsa-miR-342-3p   66

	GO:0016125	sterol metabolic process,
	GO:0016126	sterol biosynthetic process,
hsa-miR-424		

Table 12: Gene Ontology Analysis for miRNA targeted genes

Using mir2disease database, miRNAs in Blue modules and other miRNAs related to the

disease are shown in the table below:

miRNA	Disease
hsa-miR-19a	hepatocellular carcinoma (HCC), anaplastic thyroid carcinoma (ATC), Cowden Syndrome, lung cancer, B-cell chronic lymphocytic leukemia, colorectal cancer, head and neck squamous cell carcinoma (HNSCC), malignant lymphoma, malignant melanoma, hepatocellular carcinoma (HCC), multiple myeloma (MM), medulloblastoma, glioma, colorectal cancer, Lung Cancer
hsa-miR-18a	anaplastic thyroid carcinoma (ATC), lung cancer, pancreatic ductal adenocarcinoma (PDAC), hepatocellular carcinoma (HCC), medulloblastoma, breast cancer, Hodgkin's lymphoma, colorectal cancer
hsa-miR-18b	cardiac hypertrophy, multiple sclerosis
hsa-let-7e	lung cancer, acute lymphoblastic leukemia (ALL), acute myeloid leukemia (AML), acute myeloid leukemia (AML), diffuse large B-cell lymphoma (DLBCL), head and neck squamous cell carcinoma (HNSCC), malignant melanoma, Oral Squamous Cell Carcinoma (OSCC), pituitary adenoma, psoriasis, retinoblastoma, lupus nephritis, non-alcoholic fatty liver disease (NAFLD), multiple myeloma (MM)
hsa-let-7i	Alzheimer's disease, breast cancer, head and neck squamous cell carcinoma (HNSCC), ovarian cancer (OC)

hsa-miR-23a	hepatocellular carcinoma (HCC), acute lymphoblastic leukemia (ALL), acute myeloid leukemia (AML), autism spectrum disorder (ASD), bladder cancer, cardiac hypertrophy, colorectal cancer, glioblastoma, heart failure, hepatocellular carcinoma (HCC), Oral Squamous Cell Carcinoma (OSCC), pancreatic cancer, prostate cancer, Acute Promyelocytic Leukemia (APL), lupus nephritis, ulcerative colitis (UC), cardiac hypertrophy
hsa-miR-26a	anaplastic thyroid carcinoma (ATC), acute lymphoblastic leukemia (ALL), acute myeloid leukemia (AML), acute myeloid leukemia (AML), cardiac hypertrophy, colorectal cancer, Duchenne muscular dystrophy (DMD), epithelial ovarian cancer (EOC), pituitary adenoma, primary biliary cirrhosis (PBC), prostate cancer, prostate cancer, serous ovarian cancer, Burkitt lymphoma, ulcerative colitis (UC), primary biliary cirrhosis (PBC), hepatocellular carcinoma (HCC), Oral Squamous Cell Carcinoma (OSCC), bladder cancer, glioma, kidney cancer, hepatocellular carcinoma (HCC), breast cancer, papillary thyroid carcinoma (PTC)
hsa-miR-125a-5p	lung cancer, multiple myeloma (MM)
hsa-miR-135b	colorectal cancer, malignant melanoma
hsa-miR-149	breast cancer, cardiac hypertrophy, cardiac hypertrophy, diffuse large B-cell lymphoma (DLBCL), follicular lymphoma (FL), glioblastoma multiforme (GBM), malignant melanoma, pituitary adenoma, prostate cancer
hsa-miR-224	hepatocellular carcinoma (HCC), epithelial ovarian cancer (EOC), hepatocellular carcinoma (HCC), hepatocellular carcinoma (HCC), lung cancer, malignant melanoma, Oral Squamous Cell Carcinoma (OSCC), pancreatic ductal adenocarcinoma (PDAC), prostate cancer, hepatocellular carcinoma (HCC), colorectal cancer
hsa-miR-339-5p	neurodegeneration
hsa-miR-342-3p	neurodegeneration, kidney cancer, prion disease
hsa-miR-378*	colorectal cancer
hsa-miR-423-5p	Gastric Cancer

hsa-miR-424	acute lymphoblastic leukemia (ALL), acute myeloid leukemia (AML), cardiac hypertrophy, head and neck squamous cell carcinoma (HNSCC), ovarian cancer (OC), pancreatic cancer, endometriosis kidney cancer, chronic lymphocytic leukemia (CLL), Intrahepatic cholangiocarcinoma (ICC)
hsa-miR-629	lupus nephritis, ulcerative colitis (UC)

Table 13: Involvement of miRNAs in disease

CHAPTER 4

DISCUSSION

miRNAs have been measured as one of the most important regulators; however, identifying their functions is a challenging task and to understand their biological process as regulators is even more difficult and crucial. In this study, we identified 49 miRNAs that are differentially expressed between 60 ER+ and 35 ER- breast cancer samples. However, a large proportion of miRNAs and mRNAs identified in the modules have been reported to have associations with ER+ and ER- subtypes of breast cancer.

The one cluster of 5 miRNAs (has-miR-18a,-18b,-19a,-505 and -505*) were up regulated in ER- samples but the other cluster of 5 miRNAs (has-miR-29b-2*,-29c,-29c*,-190b and-148b) was down regulated in ER- samples. The rest of 39 miRNAs, 12 miRNAs (hsa-miR-224, hsa-miR-146b-5p, hsa-miR-135b, hsa-miR-374a, hsa-miR-452, hsa-miR-223, hsa-miR-9, hsa-miR-9*, hsa-miR-378*, hsa-miR-23a, hsa-miR-148a, hsa-miR-424) was up regulated in ER- samples and 27 miRNAs (hsa-miR-149, hsa-miR-342-3p, hsa-miR-342-5p, hsa-miR-339-5p, hsa-miR-99b, hsa-let-7e*, hsa-miR-339-3p, hsa-miR-628-3p, hsa-miR-499-5p, hsa-miR-125a-5p, hsa-miR-625, hsa-miR-26a, hsa-miR-10b*, hsa-miR-629, hsa-miR-623, hsa-miR-181d, hsa-miR-423-5p, hsa-miR-326, hsa-miR-101*, hsa-miR-103, hsa-miR-432, hsa-let-7e, hsa-miR-26b*, hsa-miR-146a, hsa-let-7i, hsa-miR-30a*, hsa-miR-375) were down regulated in ER- samples.

From this study, we identified the association of several miRNAs with the biological process such as, cell differentiation, cell death, cell adhesion, cell proliferation, immune response, defense response, inflammatory response, signaling pathway, tissue

homeostasis and apoptosis. MicroRNA expression patterns are promising biomarkers for several tumor types including breast, ovarian, lung, pancreatic, kidney, prostate, bladder and colorectal cancer.

Our framework identified miRNA targeted genes, from that mutation in listed genes leads to breast cancer:

Gene Name	Mutation	Description
NUP133	G → V Glycine to Valine Position : 326 G → R Glycine to Arginine Position : 448	In a breast cancer sample, somatic mutation
CD2	C → Y Cysteine to Tyrosine Position : 217	In a breast cancer sample, somatic mutation

Table 14: (a) Involvement targeted genes in breast cancer

Literature citation: The consensus coding sequences of human breast and colorectal cancers. Science 314:268-274(2006)

Literatures suggest that following miRNAs are involved in Breast cancer:

miRNA	Literature citation
hsa-miR-18a	Differential expression profiles of microRNAs between breast cancer cells and mammary epithelial cells.
hsa-miR-26a	Widespread estrogen-dependent repression of micrornas involved in breast tumor cell growth.
hsa-miR-149	MicroRNA gene expression deregulation in human breast cancer.
hsa-let-7i	MicroRNA gene expression deregulation in human breast cancer.

Table 14: (b) Involvement miRNAs in breast cancer

LIMITATIONS OF PROPOSED FRAMEWORK

We used miRanda algorithm to predict the miRNA targets in combination with WGCNA methodology.

In this study, WGCNA methodology is not to able to produce productive result in Module identification as it conserved very little portion of the data. As we can see the result of hierarchical clustering of Module Detection for miRNA and mRNA; it doesn't preserved more region as expected.

Also in the statistical analysis, Fisher's exact test doesn't show more significance for miRNA modules to mRNA modules, which is quite dissatisfied. It gives only two miRNA-mRNA modules (Blue, Grey for miRNA and Black, Brown for mRNA modules), so we may miss important miRNA functional targets.

Computational methods have their own limitations in their accuracy. Our knowledge of miRNA regulatory modules remains incomplete.

In future, one can use this framework on multiple datasets of mRNA and miRNA expression data on two phenotypes but it may not work for more than two phenotypes. Therefore, the effectiveness of this approach has to be further investigated using other datasets. For that reason we converted our focus to the other miRNAs which are not related to any modules.

CHAPTER 5

CONCLUSION

In this study, we identified 49 miRNAs that are differentially expressed between 60 ER+ and 35 ER- breast cancer samples. However, a large proportion of miRNAs and mRNAs identified in the modules have been reported to have associations with ER+ and ER- subtypes of breast cancer. Here we identified one cluster of 5 miRNAs (has-miR-29b-2*,-29c,-29c*,-190b and-148b) was down regulated in ER- samples and up regulated ER+ samples. As we successfully identified miR 29 families that shows the validity of our module-based approach.

We found that targets were identified by more than one miRNA modules, which could be more fascinating to investigate their involvement in different biological process. From the functional enrichment analysis, we found that turquoise module of miRNAs, miR-29 family, involved in biological process of extracellular matrix and structure organization. Similarly, rest of the miRNAs found in biological processed such as cell migration, cell motility, cell and biological adhesion.

From this study, we identified the association of several miRNAs with the biological process. A miRNA, hsa-miR-19a, up regulated in ER- samples, is associated with immune response, defense response, inflammatory response and membrane biogenesis. A miRNA, hsa-miR-135b up regulated in ER- samples, are associated with regulation of cell death, cell adhesion, regulation of apoptosis. A miRNA, hsa-miR-224 up in ER- samples, are important in cell proliferation, regulation of muscle contraction

and G-protein signaling pathway. A miRNA, hsa-miR-378* up regulated in ER- samples, play role in cell differentiation, cell signaling and tissue homeostasis. A miRNA, hsa-miR-26a, up regulated in ER- samples, is related with immune response. miRNAs, hsa-let-7i and hsa-miR-23a, down and up regulated respectively in ER- samples, are associated with transporter activity. A miRNA, hsa-miR-9* up regulated in ER- samples, play role in catabolic process.

Using miR2disease database, we identified the relation of several miRNAs with the disease such as breast, ovarian, lung, pancreatic, kidney, prostate, bladder and colorectal cancer; Alzheimer's, Parkinson's and Huntington's disease.

Furthermore, many novel associations among miRNAs, mRNAs and biological processes have been identified in this study. Several miRNAs and mRNAs are highly related to cancers as reported by literature. Literature suggest that hsa-let-7i, hsa-miR-18a, -26a, -149 and mutation in NUP133 and CD2 genes, up and down regulated respectively in ER- samples, are related to breast cancer. We can say that CD2 targeted by hsa-miR-135b, ABCC5 targeted by hsa-miR-23a and CHRM3 targeted by has-miR-224 could be the novel biomarkers for Estrogen Receptor Negative subtype of breast cancer.

REFERENCES

1. Tingting Du and Philip D. Zamore : microPrimer: the biogenesis and function of microRNA. The company of Biologist 2005. doi: 10.1242/dev.02070

2. Bartel, D. P. (2004). MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell , 116* (2), 281 - 297.

3. Friedman, R. C., Farh, K. K.-H., Burge, C. B., & Bartel, D. P. (2009). Most mammalian mRNA are conserved targets of microRNAs. *Genome Res , 19* (1),92 - 105.

4. Yoon, S., & Micheli, G. D. (2006). Computational identification of microRNAs and their targets. *Birth Defects Research Part C: Embryo Today: Reviews , 78* (2),118 - 128.

5. Kim, S.-K., Nam, J.-W., Rhee, J.-K., Lee, W.-J., & Zhang, B.-T. (2006). miTarget: microRNA target gene prediction using a support vector machine. *BMC Bioinformatics , 7* (1), 411.

6. Lee, Y., Jeon, K., Lee, J. T., Kim, S., & Kim, V. N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. *EMBO , 21* (17), 4663 - 4670.

7. Nam, J.-w., Shin, K.-r., Han, J., Lee, Y., Kim, V. N., & Zhang, B.-t. (2005). Human microRNA prediction through a probabilistic co-learning model of sequence and structure. *Nucleic Acids Res* , 3570--3581.

8. Lindow, M., & Gorodkin, J. (2007). Principles and Limitations of Computational MicroRNA Gene and Target Finding. *DNA and Cell Biology , 26* (5), 339-351.

75

9. McDaneld, T. G. (2009). MicroRNA: mechanism of gene regulation and application to livestock. *Journal of animal science , 87*, E21-E28.

10. Griffiths-Jones S., Saini HK., van Dongen S., Enright AJ. (2008) miRBase: tools for microRNA genomics. *Nucleic Acids Res 36(Database issue): D154-8*

11. Griffiths-Jones S., Grocock RJ., van Dongen S., Bateman A., Enright AJ. (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res 34(Database issue): D140-4*

12. Enright AJ., John B., Gaul U., Tuschl T., Sander C., Marks DS. (2003) MicroRNA targets in Drosophila. *Genome Biol 5(1): R1*

13. Maziere, P., & Enright, A. J. (2007). Prediction of microRNA targets. *Drug Discovery Today , 12* (11), 452 - 458.

14. Hyeyoung Min and Sungroh Yoon, Got target?: computational methods for microRNA target prediction and their extension. EXPERIMENTAL and MOLECULAR MEDICINE, Vol. 42, No. 4, 233-244, April 2010

15. Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell , 75* (5), 843 - 854.

16. Enerly E, Steinfeld I, Kleivi K, Leivonen S-K, Aure MR, Russnes HG, Rønneberg JA, Johnsen H, Navon R, Rødland E, et al: miRNA-mRNA integrated analysis reveals roles for miRNAs in primary breast tumors. PLoS ONE 2011, (2):e16915.

17. Zhang B. Horvath S: A General Framework for Weighted Gene Co-expression Network Analysis. *Stat Appl Genet Mol Biol* 2005,4:Article 17.

18. Carey VJ, Gentry J, Whalen E, Gentleman R: Network structures and algorithms in Bioconductor. *Bioinformatics* 2005, 21:135-136.

19. Horvath S, Zhang B, Carlson M, Lu K, Zhu S, Felciano R, Laurance M, Zhao W, Shu Q, Lee Y, Scheck A, Liau L, Wu H, Geschwind D, Febbo P, Kornblum H, Cloughesy T, Nelson S, Mischel P: Analysis of Oncogenic Signaling Networks in Glioblastoma Identifies ASPM as a Novel Molecular Target. *Proc Natl Acad Sci USA* 2006, 103(46):17402-17407.

20. Horvath S, Dong J: Geometric interpretation of Gene Co-expression Network Analysis. *PLoS Computational Biology* 2008.

21. Langfelder P, Horvath S: Eigen-gene networks for studying the relationships between co-expression modules. *BMC Systems Biology* 2007, 1:54.

22. Carlson MR, Zhang B, Fang Z, Horvath S, Mishel PS, Nelson SF: Gene Connectivity, Function, and Sequence Conservation: Predictions from Modular Yeast Co-expression Networks. *BMC Genomics* 2006, 7(40):.

23. Yip A, Horvath S: Gene network interconnectedness and the generalized topological overlap measure. *BMC Bioinformatics* 2007, 8:22.

24. Langfelder P, Zhang B, Horvath S: Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* 2008, 24(5):719-720.

25. Dong J, Horvath S: Understanding network concepts in modules. *BMC Systems Biology* 2007, 1:24.

26. Dudoit S, Yang Y, Callow M, Speed T: Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Statistical Sinica* 2002.

27. Dennis G, Sherman B, Hosack D, Yang J, Gao W, Lane H, Lempicki R: DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 2003, 4(5):P3.

28. Gentleman R, Huber W, Carey V, Irizarry R, Dudoit S: Bioinformatics and Computational Biology Solutions Using R and Bioconductor. In *Book* Springer-Verlag New York; 2005.

29. Gentleman R, Carey V, Bates D, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, et al: Bioconductor: open software development for computational biology and bioinformatics. Genome Biology 2004, 5(10):R80.

30. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009;37(1):1-13.

31. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nature Protoc. 2009;4(1):44-57.,

32. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009;37(1):1-13.

33. Fisher, R. A. (1922). "On the interpretation of χ^2 from contingency tables, and the calculation of P". *Journal of the Royal Statistical Society* **85** (1): 87–94. doi:10.2307/2340521. JSTOR,2340521.

34. Fisher, R.A. (1954). *Statistical Methods for Research Workers.* Oliver and Boyd. ISBN,0-05-002170-2.

35. Fisher RA: on the 'Probable error' of a coefficient of correlation deduced from a small sample. *Metron* 1915, 1**:**1-32.

36. Roqueiro et al.: Identifying transcription factors and microRNAs as key regulators of pathways using Bayesian inference on known pathway structures. Proteome Science 2012 10 (Suppl 1):S15. doi:10.1186/1477-5956-10-S1-S15.

37. Ravasz E, Somera A, Mongru D, Oltvai Z, Barabási A: Hierarchical Organization of Modularity in Metabolic Networks. *Science* 2002, 297(5586):1551-1555.

38. Dudoit S, Fridlyand J: A prediction-based resampling method for estimating the number of clusters in a dataset. Genome Biol 2002, 3(7):RESEARCH0036

APPENDIX A: ADDITIONAL MATERIAL

1) **Supplementary Table 2**

    Differential Expression Analysis

    a. miRNA : pvalmir49.pdf

    b. mRNA : pvalmr12605.pdf

2) **Supplementary Table 6**

    miRNA – gene target prediction

    a. ER+ samples : mimTargetERpos.pdf

    b. ER- samples : mimTargetERneg.pdf

3) **Supplementary Table 7**

    Up/Down Regulation of mRNA expression in ER+ samples.

    a. For miRNA: shown in results

    b. For mRNA: updownregulation.pdf

4) **Supplementary Table 8**

Gene Ontology analysis and KEGG pathway information is shown for particular module.

    For Black mRNA module

ID	Gene Name	CHROMOSOME
PAPSS1	3'-phosphoadenosine 5'-phosphosulfate synthase 1	4,
GOTERM_BP_FAT	GO:0000103~sulfate assimilation, GO:0001501~skeletal system development, GO:0006163~purine nucleotide metabolic process, GO:0006790~sulfur metabolic process, GO:0009165~nucleotide biosynthetic process, GO:0033865~nucleoside bisphosphate metabolic process, GO:0033866~nucleoside bisphosphate biosynthetic process, GO:0033875~ribonucleoside bisphosphate metabolic process, GO:0034030~ribonucleoside bisphosphate biosynthetic process, GO:0034032~purine nucleoside bisphosphate metabolic process, GO:0034033~purine nucleoside bisphosphate biosynthetic process, GO:0034035~purine ribonucleoside bisphosphate metabolic process, GO:0034036~purine ribonucleoside bisphosphate biosynthetic process, GO:0034404~nucleobase, nucleoside and nucleotide biosynthetic process, GO:0034654~nucleobase, nucleoside, nucleotide and nucleic acid biosynthetic process, GO:0044271~nitrogen compound biosynthetic process, GO:0050427~3'-phosphoadenosine 5'-phosphosulfate metabolic process, GO:0050428~3'-phosphoadenosine 5'-phosphosulfate biosynthetic process,
GOTERM_CC_FAT	
GOTERM_MF_FAT	GO:0000166~nucleotide binding, GO:0001882~nucleoside binding, GO:0001883~purine nucleoside binding, GO:0004020~adenylylsulfate kinase activity, GO:0004779~sulfate adenylyltransferase activity, GO:0004781~sulfate adenylyltransferase (ATP) activity, GO:0005524~ATP binding,GO:0016779~nucleotidyltransferase activity, GO:0017076~purine nucleotide binding, GO:0030554~adenyl nucleotide binding, GO:0032553~ribonucleotide binding, GO:0032555~purine ribonucleotide binding, GO:0032559~adenyl ribonucleotide binding,

	GO:0070566~adenylyltransferase activity,	
KEGG_PATHWAY	hsa00230:Purine metabolism, hsa00450:Selenoamino acid metabolism, hsa00920:Sulfur metabolism,	
FBXO46	F-box protein 46	19,
GOTERM_BP_FAT	GO:0006508~proteolysis,GO:0009057~macromolecule catabolic process,GO:0019941~modification-dependent protein catabolic process,GO:0030163~protein catabolic process,GO:0043632~modification-dependent macromolecule catabolic process,GO:0044257~cellular protein catabolic process,GO:0044265~cellular macromolecule catabolic process,GO:0051603~proteolysis involved in cellular protein catabolic process,	
GOTERM_CC_FAT		
GOTERM_MF_FAT		
KEGG_PATHWAY		
ARHGAP29	Rho GTPase activating protein 29	1,
GOTERM_BP_FAT	GO:0007242~intracellular signaling cascade,GO:0007264~small GTPase mediated signal transduction,GO:0007265~Ras protein signal transduction,GO:0007266~Rho protein signal transduction,	
GOTERM_CC_FAT		
GOTERM_MF_FAT	GO:0005083~small GTPase regulator activity,GO:0005096~GTPase activator activity,GO:0005099~Ras GTPase activator activity,GO:0005100~Rho GTPase activator activity,GO:0008047~enzyme activator activity,GO:0008270~zinc ion binding,GO:0008289~lipid binding,GO:0019992~diacylglycerol binding,GO:0030695~GTPase regulator activity,GO:0043167~ion binding,GO:0043169~cation binding,GO:0046872~metal ion binding,	82

	GO:0046914~transition metal ion binding, GO:0060589~nucleoside-triphosphatase regulator activity,	
KEGG_PATHWAY		
BTNL9	butyrophilin-like 9	5,
GOTERM_BP_FAT		
GOTERM_CC_FAT	GO:0016021~integral to membrane,GO:0031224~intrinsic to membrane,	
GOTERM_MF_FAT		
KEGG_PATHWAY		
CR2	complement component (3d/Epstein Barr virus) receptor 2	1,
GOTERM_BP_FAT	GO:0002250~adaptive immune response,GO:0002252~immune effector process,GO:0002253~activation of immune response,GO:0002443~leukocyte mediated immunity,GO:0002449~lymphocyte mediated immunity,GO:0002455~humoral immune response mediated by circulating immunoglobulin,GO:0002460~adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains,GO:0002526~acute inflammatory response,GO:0002541~activation of plasma proteins involved in acute inflammatory response,GO:0002684~positive regulation of immune system process,GO:0006508~proteolysis,GO:0006952~defense response,GO:0006954~inflammatory response,GO:0006955~immune response,GO:0006956~complement activation,GO:0006958~complement activation, classical pathway,GO:0006959~humoral immune response,GO:0009611~response to wounding,	

83

	GO:0016064~immunoglobulin mediated immune response,GO:0016485~protein processing,GO:0019724~B cell mediated immunity,GO:0045087~innate immune response,GO:0048584~positive regulation of response to stimulus,GO:0050778~positive regulation of immune response,GO:0051604~protein maturation,GO:0051605~protein maturation by peptide bond cleavage,
GOTERM_CC_FAT	GO:0005886~plasma membrane,GO:0016021~integral to membrane,GO:0031224~intrinsic to membrane,
GOTERM_MF_FAT	GO:0001848~complement binding,GO:0004875~complement receptor activity,GO:0042802~identical protein binding,GO:0042803~protein homodimerization activity,GO:0046983~protein dimerization activity,
KEGG_PATHWAY	hsa04610:Complement and coagulation cascades,hsa04640:Hematopoietic cell lineage,hsa04662:B cell receptor signaling pathway,

DHRS1	dehydrogenase/reductase (SDR family) member 1	14,
GOTERM_BP_FAT	GO:0055114~oxidation reduction,	
GOTERM_CC_FAT	GO:0005739~mitochondrion,GO:0005740~mitochondrial envelope,GO:0005743~mitochondrial inner membrane,GO:0005783~endoplasmic reticulum,GO:0005794~Golgi apparatus,GO:0019866~organelle inner membrane,GO:0031090~organelle membrane,GO:0031966~mitochondrial membrane,GO:0031967~organelle envelope,GO:0031975~envelope,GO:0044429~mitochondrial part,	
GOTERM_MF_FAT		
KEGG_PATHWAY		
	84

NR1H4	nuclear receptor subfamily 1, group H, member 4	12,
GOTERM_BP_FAT	GO:0000122~negative regulation of transcription from RNA polymerase II promoter,GO:0002237~response to molecule of bacterial origin,GO:0006109~regulation of carbohydrate metabolic process,GO:0006350~transcription,GO:0006355~regulation of transcription, DNA-dependent,GO:0006357~regulation of transcription from RNA polymerase II promoter,GO:0008202~steroid metabolic process,GO:0008206~bile acid metabolic process,GO:0009617~response to bacterium,GO:0009743~response to carbohydrate stimulus,GO:0009746~response to hexose stimulus,GO:0009749~response to glucose stimulus,GO:0009890~negative regulation of biosynthetic process,GO:0009891~positive regulation of biosynthetic process,GO:0010033~response to organic substance,GO:0010557~positive regulation of macromolecule biosynthetic process,GO:0010558~negative regulation of macromolecule biosynthetic process,GO:0010604~positive regulation of macromolecule metabolic process,GO:0010605~negative regulation of macromolecule metabolic process,GO:0010628~positive regulation of gene expression,GO:0010629~negative regulation of gene expression,GO:0016481~negative regulation of transcription,GO:0031327~negative regulation of cellular biosynthetic process,GO:0031328~positive regulation of cellular biosynthetic process,GO:0032496~response to lipopolysaccharide,GO:0034284~response to monosaccharide stimulus,GO:0045449~regulation of transcription,GO:0045892~negative regulation of transcription, DNA-dependent,GO:0045893~positive regulation of transcription, DNA-dependent,GO:0045934~negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process,GO:0045935~positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process,GO:0045941~positive regulation of transcription,	85

	GO:0045944~positive regulation of transcription from RNA polymerase II promoter,GO:0048565~gut development,GO:0051172~negative regulation of nitrogen compound metabolic process,GO:0051173~positive regulation of nitrogen compound metabolic process,GO:0051252~regulation of RNA metabolic process,GO:0051253~negative regulation of RNA metabolic process,GO:0051254~positive regulation of RNA metabolic process,
GOTERM_CC_FAT	
GOTERM_MF_FAT	GO:0003677~DNA binding,GO:0003690~double-stranded DNA binding,GO:0003700~transcription factor activity,GO:0003707~steroid hormone receptor activity,GO:0003712~transcription cofactor activity,GO:0003713~transcription coactivator activity,GO:0003714~transcription corepressor activity,GO:0004879~ligand-dependent nuclear receptor activity,GO:0005496~steroid binding,GO:0008134~transcription factor binding,GO:0008270~zinc ion binding,GO:0008289~lipid binding,GO:0016563~transcription activator activity,GO:0016564~transcription repressor activity,GO:0030528~transcription regulator activity,GO:0031406~carboxylic acid binding,GO:0032052~bile acid binding,GO:0033293~monocarboxylic acid binding,GO:0042277~peptide binding,GO:0043167~ion binding,GO:0043169~cation binding,GO:0043565~sequence-specific DNA binding,GO:0043566~structure-specific DNA binding,GO:0046872~metal ion binding,GO:0046914~transition metal ion binding,GO:0047485~protein N-terminus binding,
KEGG_PATHWAY	
NUP133	nucleoporin 133kDa	1,
GOTERM_BP_FAT	GO:0006403~RNA localization,GO:0006405~RNA export from nucleus,GO:0006406~mRNA export from nucleus,GO:0006913~nucleocytoplasmic

	transport,GO:0006997~nucleus organization,GO:0006999~nuclear pore organization,GO:0007498~mesoderm development,GO:0008104~protein localization,GO:0015031~protein transport,GO:0015931~nucleobase, nucleoside, nucleotide and nucleic acid transport,GO:0034621~cellular macromolecular complex subunit organization,GO:0043933~macromolecular complex subunit organization,GO:0045184~establishment of protein localization,GO:0046907~intracellular transport,GO:0048339~paraxial mesoderm development,GO:0050657~nucleic acid transport,GO:0050658~RNA transport,GO:0051028~mRNA transport,GO:0051168~nuclear export,GO:0051169~nuclear transport,GO:0051236~establishment of RNA localization,GO:0055085~transmembrane transport,
GOTERM_CC_FAT	GO:0000775~chromosome, centromeric region,GO:0000776~kinetochore,GO:0000777~condensed chromosome kinetochore,GO:0000779~condensed chromosome, centromeric region,GO:0000793~condensed chromosome,GO:0000940~outer kinetochore of condensed chromosome,GO:0005635~nuclear envelope,GO:0005643~nuclear pore,GO:0005694~chromosome,GO:0012505~endomembrane system,GO:0016021~integral to membrane,GO:0031080~Nup107-160 complex,GO:0031224~intrinsic to membrane,GO:0031967~organelle envelope,GO:0031975~envelope,GO:0043228~non-membrane-bounded organelle,GO:0043232~intracellular non-membrane-bounded organelle,GO:0044427~chromosomal part,GO:0046930~pore complex,
GOTERM_MF_FAT	GO:0005487~nucleocytoplasmic transporter activity,
KEGG_PATHWAY	
OLFML3	olfactomedin-like 3

| OLFML3 | olfactomedin-like 3 | 1, | 87 |

GOTERM_BP_FAT	
GOTERM_CC_FAT	GO:0005576~extracellular region,
GOTERM_MF_FAT	
KEGG_PATHWAY	

PEX3	peroxisomal biogenesis factor 3	6,
GOTERM_BP_FAT	GO:0006605~protein targeting,GO:0006612~protein targeting to membrane,GO:0006625~protein targeting to peroxisome,GO:0006886~intracellular protein transport,GO:0007031~peroxisome organization,GO:0008104~protein localization,GO:0015031~protein transport,GO:0016557~peroxisome membrane biogenesis,GO:0017038~protein import,GO:0034613~cellular protein localization,GO:0043574~peroxisomal transport,GO:0044091~membrane biogenesis,GO:0045046~protein import into peroxisome membrane,GO:0045184~establishment of protein localization,GO:0046907~intracellular transport,GO:0070727~cellular macromolecule localization,	
GOTERM_CC_FAT	GO:0005777~peroxisome,GO:0005778~peroxisomal membrane,GO:0005779~integral to peroxisomal membrane,GO:0016021~integral to membrane,GO:0031090~organelle membrane,GO:0031224~intrinsic to membrane,GO:0031231~intrinsic to peroxisomal membrane,GO:0031300~intrinsic to organelle membrane,GO:0031301~integral to organelle membrane,GO:0031903~microbody membrane,GO:0042579~microbody,GO:0044438~microbody part,GO:0044439~peroxisomal part,	
GOTERM_MF_FAT	
KEGG_PATHWAY	88

RPL7L1	ribosomal protein L7-like 1; ribosomal protein L7 pseudogene 14; ribosomal protein L7 pseudogene 21; ribosomal protein L7 pseudogene 22; ribosomal protein L7 pseudogene 46	16,2,5,6,
GOTERM_BP_FAT	GO:0006412~translation,	
GOTERM_CC_FAT	GO:0005840~ribosome,GO:0015934~large ribosomal subunit,GO:0030529~ribonucleoprotein complex,GO:0033279~ribosomal subunit,GO:0043228~non-membrane-bounded organelle,GO:0043232~intracellular non-membrane-bounded organelle,	
GOTERM_MF_FAT	GO:0003735~structural constituent of ribosome,GO:0005198~structural molecule activity,GO:0030528~transcription regulator activity,	
KEGG_PATHWAY		
SPRED2	sprouty-related, EVH1 domain containing 2	2,
GOTERM_BP_FAT	GO:0000165~MAPKKK cascade,GO:0000188~inactivation of MAPK activity,GO:0006469~negative regulation of protein kinase activity,GO:0007242~intracellular signaling cascade,GO:0007243~protein kinase cascade,GO:0019220~regulation of phosphate metabolic process,GO:0033673~negative regulation of kinase activity,GO:0042325~regulation of phosphorylation,GO:0043086~negative regulation of catalytic activity,GO:0043405~regulation of MAP kinase activity,GO:0043407~negative regulation of MAP kinase activity,GO:0043549~regulation of kinase activity,GO:0044092~negative regulation of molecular function,GO:0045859~regulation of protein kinase activity,GO:0051174~regulation of phosphorus metabolic process,GO:0051338~regulation of transferase activity, GO:0051348~negative regulation of transferase activity, 89	

GOTERM_CC_FAT	GO:0005886~plasma membrane,GO:0009898~internal side of plasma membrane,GO:0031410~cytoplasmic vesicle,GO:0031982~vesicle,GO:0044459~plasma membrane part,	
GOTERM_MF_FAT	GO:0005173~stem cell factor receptor binding,	
KEGG_PATHWAY	hsa04630:Jak-STAT signaling pathway,	
TCP11L1	t-complex 11 (mouse)-like 1	11,
GOTERM_BP_FAT		
GOTERM_CC_FAT		
GOTERM_MF_FAT		
KEGG_PATHWAY		
WNT11	wingless-type MMTV integration site family, member 11	11,
GOTERM_BP_FAT	GO:0007166~cell surface receptor linked signal transduction,GO:0007223~Wnt receptor signaling pathway, calcium modulating pathway,GO:0016055~Wnt receptor signaling pathway,	
GOTERM_CC_FAT	GO:0005576~extracellular region,GO:0005578~proteinaceous extracellular matrix,GO:0031012~extracellular matrix,GO:0044421~extracellular region part,	
GOTERM_MF_FAT		
KEGG_PATHWAY	hsa04310:Wnt signaling pathway,hsa04340:Hedgehog signaling pathway,hsa04916:Melanogenesis,hsa05200:Pathways in cancer,hsa05217:Basal cell carcinoma,	

For Brown mRNA module

HIBCH	3-hydroxyisobutyryl-Coenzyme A hydrolase	2,
GOTERM_BP_FAT	GO:0009063~cellular amino acid catabolic process,GO:0009081~branched chain family amino acid metabolic process,GO:0009083~branched chain family amino acid catabolic process,GO:0009310~amine catabolic process,GO:0016054~organic acid catabolic process,GO:0046395~carboxylic acid catabolic process,	
GOTERM_CC_FAT	GO:0005739~mitochondrion,	
GOTERM_MF_FAT	GO:0003860~3-hydroxyisobutyryl-CoA hydrolase activity,GO:0016289~CoA hydrolase activity,GO:0016790~thiolester hydrolase activity,	
KEGG_PATHWAY	hsa00280:Valine, leucine and isoleucine degradation,hsa00410:beta-Alanine metabolism,hsa00640:Propanoate metabolism,	
C21orf99	chromosome 21 open reading frame 99	21,
GOTERM_BP_FAT		
GOTERM_CC_FAT		
GOTERM_MF_FAT		
KEGG_PATHWAY		
ABCC5	ATP-binding cassette, sub-family C (CFTR/MRP), member 5	3,
GOTERM_BP_FAT	GO:0002237~response to molecule of bacterial origin,GO:0009617~response to bacterium,GO:0009719~response to endogenous stimulus,GO:0009725~response to hormone stimulus,GO:0010033~response to organic	

91

	substance,GO:0032496~response to lipopolysaccharide,GO:0032868~response to insulin stimulus,GO:0043434~response to peptide hormone stimulus,GO:0055085~transmembrane transport,	
GOTERM_CC_FAT	GO:0000267~cell fraction,GO:0005624~membrane fraction,GO:0005626~insoluble fraction,GO:0005886~plasma membrane,GO:0005887~integral to plasma membrane,GO:0016021~integral to membrane,GO:0031224~intrinsic to membrane,GO:0031226~intrinsic to plasma membrane,GO:0044459~plasma membrane part,GO:0048471~perinuclear region of cytoplasm,	
GOTERM_MF_FAT	GO:0000166~nucleotide binding,GO:0001882~nucleoside binding,GO:0001883~purine nucleoside binding,GO:0005524~ATP binding,GO:0008509~anion transmembrane transporter activity,GO:0008514~organic anion transmembrane transporter activity,GO:0015399~primary active transmembrane transporter activity,GO:0015405~P-P-bond-hydrolysis-driven transmembrane transporter activity,GO:0016820~hydrolase activity, acting on acid anhydrides, catalyzing transmembrane movement of substances,GO:0016887~ATPase activity,GO:0017076~purine nucleotide binding,GO:0030554~adenyl nucleotide binding,GO:0032553~ribonucleotide binding,GO:0032555~purine ribonucleotide binding,GO:0032559~adenyl ribonucleotide binding,GO:0042623~ATPase activity, coupled,GO:0042626~ATPase activity, coupled to transmembrane movement of substances,GO:0043492~ATPase activity, coupled to movement of substances,	
KEGG_PATHWAY	hsa02010:ABC transporters,	
CD2	CD2 molecule	1,
GOTERM_BP_FAT	GO:0001766~membrane raft polarization,GO:0001775~cell activation,	92

	GO:0002684~positive regulation of immune system process,GO:0002694~regulation of leukocyte activation,GO:0002696~positive regulation of leukocyte activation,GO:0006917~induction of apoptosis,GO:0007155~cell adhesion,GO:0007166~cell surface receptor linked signal transduction,GO:0010941~regulation of cell death,GO:0010942~positive regulation of cell death,GO:0012502~induction of programmed cell death,GO:0016044~membrane organization,GO:0016337~cell-cell adhesion,GO:0022610~biological adhesion,GO:0030101~natural killer cell activation,GO:0030885~regulation of myeloid dendritic cell activation,GO:0030887~positive regulation of myeloid dendritic cell activation,GO:0031579~membrane raft organization,GO:0031580~membrane raft distribution,GO:0042110~T cell activation,GO:0042981~regulation of apoptosis,GO:0043065~positive regulation of apoptosis,GO:0043067~regulation of programmed cell death,GO:0043068~positive regulation of programmed cell death,GO:0045321~leukocyte activation,GO:0045580~regulation of T cell differentiation,GO:0045619~regulation of lymphocyte differentiation,GO:0046649~lymphocyte activation,GO:0050863~regulation of T cell activation,GO:0050865~regulation of cell activation,GO:0050867~positive regulation of cell activation,GO:0051249~regulation of lymphocyte activation,GO:0051665~membrane raft localization,GO:0051668~localization within membrane,
GOTERM_CC_FAT	GO:0005576~extracellular region,GO:0005886~plasma membrane,GO:0005887~integral to plasma membrane,GO:0009897~external side of plasma membrane,GO:0009898~internal side of plasma membrane,GO:0009986~cell surface,GO:0016021~integral to membrane,GO:0031224~intrinsic to membrane,GO:0031225~anchored to membrane,GO:0031226~intrinsic to plasma membrane,GO:0044459~plasma membrane 93

	part,GO:0046658~anchored to plasma membrane,	
GOTERM_MF_FAT	GO:0042802~identical protein binding,GO:0042803~protein homodimerization activity,GO:0043498~cell surface binding,GO:0043499~eukaryotic cell surface binding,GO:0046983~protein dimerization activity,	
KEGG_PATHWAY	hsa04514:Cell adhesion molecules (CAMs),hsa04640:Hematopoietic cell lineage,	
GPATCH8	G patch domain containing 8	17,
GOTERM_BP_FAT		
GOTERM_CC_FAT		
GOTERM_MF_FAT	GO:0008270~zinc ion binding,GO:0043167~ion binding,GO:0043169~cation binding,GO:0046872~metal ion binding,GO:0046914~transition metal ion binding,	
KEGG_PATHWAY		
CHRM3	cholinergic receptor, muscarinic 3	1,
GOTERM_BP_FAT	GO:0003012~muscle system process,GO:0003056~regulation of vascular smooth muscle contraction,GO:0006936~muscle contraction,GO:0006937~regulation of muscle contraction,GO:0006939~smooth muscle contraction,GO:0006940~regulation of smooth muscle contraction,GO:0007166~cell surface receptor linked signal transduction,GO:0007186~G-protein coupled receptor protein signaling pathway,GO:0007213~muscarinic acetylcholine receptor signaling pathway,GO:0007586~digestion,GO:0008283~cell proliferation,GO:0019229~regulation of vasoconstriction,GO:0044057~regulation of system process,GO:0045933~positive regulation of muscle contraction,GO:0045987~positive regulation of smooth muscle contraction,	

94

	GO:0051240~positive regulation of multicellular organismal process,
GOTERM_CC_FAT	GO:0005886~plasma membrane,GO:0005887~integral to plasma membrane,GO:0016021~integral to membrane,GO:0030054~cell junction,GO:0030424~axon,GO:0030425~dendrite,GO:0031224~intrinsic to membrane,GO:0031226~intrinsic to plasma membrane,GO:0032279~asymmetric synapse,GO:0033267~axon part,GO:0042995~cell projection,GO:0043005~neuron projection,GO:0043679~nerve terminal,GO:0044456~synapse part,GO:0044459~plasma membrane part,GO:0044463~cell projection part,GO:0045202~synapse,GO:0045211~postsynaptic membrane,
GOTERM_MF_FAT	GO:0004435~phosphoinositide phospholipase C activity,GO:0004620~phospholipase activity,GO:0004629~phospholipase C activity,GO:0004981~muscarinic acetylcholine receptor activity,GO:0008081~phosphoric diester hydrolase activity,GO:0008144~drug binding,GO:0008227~amine receptor activity,GO:0015464~acetylcholine receptor activity,GO:0016298~lipase activity,GO:0016907~G-protein coupled acetylcholine receptor activity,GO:0030594~neurotransmitter receptor activity,GO:0042165~neurotransmitter binding,GO:0042166~acetylcholine binding,GO:0043176~amine binding,
KEGG_PATHWAY	hsa04020:Calcium signaling pathway,hsa04080:Neuroactive ligand-receptor interaction,hsa04810:Regulation of actin cytoskeleton,
C21orf121	chromosome 21 open reading frame 121	21,
GOTERM_BP_FAT	
GOTERM_CC_FAT	95

GOTERM_MF_FAT		
KEGG_PATHWAY		
CYB561D1	cytochrome b-561 domain containing 1	1,
GOTERM_BP_FAT	GO:0006091~generation of precursor metabolites and energy,GO:0022900~electron transport chain,GO:0055114~oxidation reduction,	
GOTERM_CC_FAT	GO:0016021~integral to membrane,GO:0031224~intrinsic to membrane,	
GOTERM_MF_FAT	GO:0005506~iron ion binding,GO:0043167~ion binding,GO:0043169~cation binding,GO:0046872~metal ion binding,GO:0046914~transition metal ion binding,	
KEGG_PATHWAY		
CYC1	cytochrome c-1	8,
GOTERM_BP_FAT	GO:0006091~generation of precursor metabolites and energy,GO:0022900~electron transport chain,GO:0055114~oxidation reduction,	
GOTERM_CC_FAT	GO:0005739~mitochondrion,GO:0005740~mitochondrial envelope,GO:0005743~mitochondrial inner membrane,GO:0016021~integral to membrane,GO:0019866~organelle inner membrane,GO:0031090~organelle membrane,GO:0031224~intrinsic to membrane,GO:0031966~mitochondrial membrane,GO:0031967~organelle envelope,GO:0031975~envelope,GO:0044429~mitochondrial part,GO:0070469~respiratory chain,	
GOTERM_MF_FAT	GO:0005506~iron ion binding,GO:0009055~electron carrier activity,GO:0020037~heme binding,GO:0043167~ion binding,GO:0043169~cation binding,GO:0045155~electron transporter, transferring electrons from CoQH2-cytochrome c 96	

	reductase complex and cytochrome c oxidase complex activity,GO:0046872~metal ion binding,GO:0046906~tetrapyrrole binding,GO:0046914~transition metal ion binding,	
KEGG_PATHWAY	hsa00190:Oxidative phosphorylation,hsa04260:Cardiac muscle contraction,hsa05010:Alzheimer's disease,hsa05012:Parkinson's disease,hsa05016:Huntington's disease,	
DFNB31	deafness, autosomal recessive 31	9,
GOTERM_BP_FAT	GO:0001894~tissue homeostasis,GO:0001895~retina homeostasis,GO:0007267~cell-cell signaling,GO:0007268~synaptic transmission,GO:0007423~sensory organ development,GO:0007600~sensory perception,GO:0007601~visual perception,GO:0007605~sensory perception of sound,GO:0019226~transmission of nerve impulse,GO:0030030~cell projection organization,GO:0030182~neuron differentiation,GO:0042490~mechanoreceptor differentiation,GO:0042592~homeostatic process,GO:0043583~ear development,GO:0048666~neuron development,GO:0048839~inner ear development,GO:0048871~multicellular organismal homeostasis,GO:0050877~neurological system process,GO:0050890~cognition,GO:0050953~sensory perception of light stimulus,GO:0050954~sensory perception of mechanical stimulus,GO:0060113~inner ear receptor cell differentiation,GO:0060119~inner ear receptor cell development,GO:0060122~inner ear receptor stereocilium organization,GO:0060249~anatomical structure homeostasis,	
GOTERM_CC_FAT	GO:0001917~photoreceptor inner segment,GO:0002139~stereocilia coupling link,GO:0002141~stereocilia ankle link,GO:0002142~stereocilia ankle link complex,GO:0005856~cytoskeleton,GO:0005884~actin	97

	filament,GO:0005902~microvillus,GO:0005929~cilium,GO:0015629~actin cytoskeleton,GO:0030424~axon,GO:0030425~dendrite,GO:0030426~growth cone,GO:0030427~site of polarized growth,GO:0031513~nonmotile primary cilium,GO:0032391~photoreceptor connecting cilium,GO:0032420~stereocilium,GO:0032421~stereocilium bundle,GO:0042995~cell projection,GO:0043005~neuron projection,GO:0043025~cell soma,GO:0043228~non-membrane-bounded organelle,GO:0043232~intracellular non-membrane-bounded organelle,GO:0044430~cytoskeletal part,	
GOTERM_MF_FAT	GO:0019904~protein domain specific binding,	
KEGG_PATHWAY		
LSS	lanosterol synthase (2,3-oxidosqualene-lanosterol cyclase)	21,
GOTERM_BP_FAT	GO:0006694~steroid biosynthetic process,GO:0006695~cholesterol biosynthetic process,GO:0008202~steroid metabolic process,GO:0008203~cholesterol metabolic process,GO:0008610~lipid biosynthetic process,GO:0016125~sterol metabolic process,GO:0016126~sterol biosynthetic process,	
GOTERM_CC_FAT	GO:0000267~cell fraction,GO:0005624~membrane fraction,GO:0005626~insoluble fraction,GO:0005792~microsome,GO:0042598~vesicular fraction,	
GOTERM_MF_FAT	GO:0000250~lanosterol synthase activity,GO:0016866~intramolecular transferase activity,GO:0031559~oxidosqualene cyclase activity,	
KEGG_PATHWAY	hsa00100:Steroid biosynthesis,	
MR1	major histocompatibility complex, class I-related	1, 98

GOTERM_BP_FAT	GO:0002474~antigen processing and presentation of peptide antigen via MHC class I,GO:0006955~immune response,GO:0019882~antigen processing and presentation,GO:0048002~antigen processing and presentation of peptide antigen,
GOTERM_CC_FAT	GO:0005576~extracellular region,GO:0005783~endoplasmic reticulum,GO:0005886~plasma membrane,GO:0009897~external side of plasma membrane,GO:0009986~cell surface,GO:0016021~integral to membrane,GO:0031224~intrinsic to membrane,GO:0042611~MHC protein complex,GO:0042612~MHC class I protein complex,GO:0044459~plasma membrane part,
GOTERM_MF_FAT	GO:0032393~MHC class I receptor activity,
KEGG_PATHWAY	

MUC3A	mucin 3B, cell surface associated; similar to Mucin-3A precursor (Intestinal mucin-3A); mucin 3A, cell surface associated; similar to mucin 3	7,Un,
GOTERM_BP_FAT	GO:0001894~tissue homeostasis,GO:0007586~digestion,GO:0010669~epithelial structure maintenance,GO:0022600~digestive system process,GO:0030277~maintenance of gastrointestinal epithelium,GO:0042592~homeostatic process,GO:0048871~multicellular organismal homeostasis,GO:0060249~anatomical structure homeostasis,	
GOTERM_CC_FAT	GO:0005576~extracellular region,GO:0005886~plasma membrane,GO:0016021~integral to membrane,GO:0016324~apical plasma membrane,GO:0031224~intrinsic to membrane,GO:0044459~plasma membrane part,GO:0045177~apical part of cell,	
GOTERM_MF_FAT	GO:0005198~structural molecule activity,GO:0005201~extracellular matrix structural constituent,GO:0030197~extracellular matrix constituent,	99

	lubricant activity,	
KEGG_PATHWAY		
PHACTR2	phosphatase and actin regulator 2	6,
GOTERM_BP_FAT		
GOTERM_CC_FAT		
GOTERM_MF_FAT	GO:0003779~actin binding,GO:0004857~enzyme inhibitor activity,GO:0004864~phosphoprotein phosphatase inhibitor activity,GO:0008092~cytoskeletal protein binding,GO:0019208~phosphatase regulator activity,GO:0019212~phosphatase inhibitor activity,GO:0019888~protein phosphatase regulator activity,	
KEGG_PATHWAY		
PIGM	phosphatidylinositol glycan anchor biosynthesis, class M	1,
GOTERM_BP_FAT	GO:0006497~protein amino acid lipidation,GO:0006505~GPI anchor metabolic process,GO:0006506~GPI anchor biosynthetic process,GO:0006644~phospholipid metabolic process,GO:0006650~glycerophospholipid metabolic process,GO:0008610~lipid biosynthetic process,GO:0008654~phospholipid biosynthetic process,GO:0016254~preassembly of GPI anchor in ER membrane,GO:0019637~organophosphate metabolic process,GO:0030384~phosphoinositide metabolic process,GO:0042157~lipoprotein metabolic process,GO:0042158~lipoprotein biosynthetic process,GO:0045017~glycerolipid biosynthetic process,GO:0046474~glycerophospholipid biosynthetic process,GO:0046486~glycerolipid metabolic process,GO:0046489~phosphoinositide biosynthetic process,	
GOTERM_CC_FAT	GO:0005783~endoplasmic reticulum,GO:0005789~endoplasmic reticulum	100

	membrane,GO:0012505~endomembrane system,GO:0016021~integral to membrane,GO:0031090~organelle membrane,GO:0031224~intrinsic to membrane,GO:0042175~nuclear envelope-endoplasmic reticulum network,GO:0044432~endoplasmic reticulum part,	
GOTERM_MF_FAT	GO:0000030~mannosyltransferase activity,	
KEGG_PATHWAY	hsa00563:Glycosylphosphatidylinositol(GPI)-anchor biosynthesis,	
PRSS21	protease, serine, 21 (testisin)	16,
GOTERM_BP_FAT	GO:0006508~proteolysis,	
GOTERM_CC_FAT	GO:0000267~cell fraction,GO:0005624~membrane fraction,GO:0005626~insoluble fraction,GO:0005886~plasma membrane,GO:0031224~intrinsic to membrane,GO:0031225~anchored to membrane,	
GOTERM_MF_FAT	GO:0004175~endopeptidase activity,GO:0004252~serine-type endopeptidase activity,GO:0008233~peptidase activity,GO:0008236~serine-type peptidase activity,GO:0017171~serine hydrolase activity,GO:0070011~peptidase activity, acting on L-amino acid peptides,	
KEGG_PATHWAY		
RHPN1	rhophilin, Rho GTPase binding protein 1	8,
GOTERM_BP_FAT		
GOTERM_CC_FAT		
GOTERM_MF_FAT	GO:0005083~small GTPase regulator activity,GO:0030695~GTPase regulator activity,GO:0060589~nucleoside-triphosphatase regulator activity,	101

KEGG_PATHWAY	

5) Supplementary Table 9

How miRNAs in Blue and Grey modules related to the disease are shown in the table below using mir2disease databse:

miRNA	Disease	Reference
hsa-miR-19a	hepatocellular carcinoma (HCC)	Identification of metastasis-related microRNAs in hepatocellular carcinoma.
	anaplastic thyroid carcinoma (ATC)	Oncogenic role of miR-17-92 cluster in anaplastic thyroid cancer cells.
	Cowden Syndrome	Differential expression of PTEN-targeting microRNAs miR-19a and miR-21 in Cowden syndrome.
	lung cancer	A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation.
	B-cell chronic lymphocytic leukemia	MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias
	colorectal cancer	Identification by Real-time PCR of 13 mature microRNAs differentially expressed in colorectal cancer and non-tumoral tissues.
	head and neck squamous cell carcinoma (HNSCC)	High mobility group A2 is a target for miRNA-98 in head and neck squamous cell carcinoma.
	malignant lymphoma	A microRNA cluster as a target of genomic amplification in malignant lymphoma.

	malignant melanoma	MicroRNA let-7b targets important cell cycle molecules in malignant melanoma cells and interferes with anchorage-independent growth.
	hepatocellular carcinoma (HCC)	Elevated expression of the miR-17-92 polycistron and miR-21 in hepadnavirus-associated hepatocellular carcinoma contributes to the malignant phenotype.
	multiple myeloma (MM)	MicroRNAs regulate critical genes associated with multiple myeloma pathogenesis.
	medulloblastoma	The miR-17~92 cluster collaborates with the Sonic Hedgehog pathway in medulloblastoma.
	medulloblastoma	The miR-17/92 polycistron is up-regulated in sonic hedgehog-driven medulloblastomas and induced by N-myc in sonic hedgehog-treated cerebellar neural precursors.
	glioma	Identification and Functional Characterization of microRNAs Involved in the Malignant Progression of Gliomas
	colorectal cancer	Characterization of global microRNA expression reveals oncogenic potential of miR-145 in metastatic colorectal cancer
	Lung Cancer	MicroRNAs expressed during lung cancer development are expressed in human pseudoglandular lung embryogenesis.
hsa-miR-18a	anaplastic thyroid carcinoma (ATC)	Oncogenic role of miR-17-92 cluster in anaplastic thyroid cancer cells.

	lung cancer	A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation.
	pancreatic ductal adenocarcinoma (PDAC)	MicroRNA expression alterations are linked to tumorigenesis and non-neoplastic processes in pancreatic ductal adenocarcinoma.
	hepatocellular carcinoma (HCC)	Elevated expression of the miR-17-92 polycistron and miR-21 in hepadnavirus-associated hepatocellular carcinoma contributes to the malignant phenotype.
	medulloblastoma	The miR-17/92 polycistron is up-regulated in sonic hedgehog-driven medulloblastomas and induced by N-myc in sonic hedgehog-treated cerebellar neural precursors.
	breast cancer	Differential expression profiles of microRNAs between breast cancer cells and mammary epithelial cells.
	Hodgkin's lymphoma	Comparison of miRNA profiles of microdissected Hodgkin/Reed-Sternberg cells and Hodgkin cell lines versus CD77 B-cells reveals a distinct subset of differentially expressed miRNAs
	colorectal cancer	Characterization of global microRNA expression reveals oncogenic potential of miR-145 in metastatic colorectal cancer Over- and under-expressed microRNAs in human colorectal cancer.
hsa-miR-18b	cardiac hypertrophy	Expression of microRNAs is dynamically regulated during cardiomyocyte hypertrophy. 104

	multiple sclerosis	Differential micro RNA expression in PBMC from multiple sclerosis patients.
miRNA	Disease	Reference
hsa-let-7e	lung cancer	The tumor suppressor microRNA let-7 represses the HMGA2 oncogene.
	acute lymphoblastic leukemia (ALL)	MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia.
	acute myeloid leukemia (AML)	MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia.
	acute myeloid leukemia (AML)	Distinctive patterns of microRNA expression associated with karyotype in acute myeloid leukaemia.
	diffuse large B-cell lymphoma (DLBCL)	MicroRNA signatures characterize diffuse large B-cell lymphomas and follicular lymphomas.
	head and neck squamous cell carcinoma (HNSCC)	High mobility group A2 is a target for miRNA-98 in head and neck squamous cell carcinoma.
	malignant melanoma	MicroRNA let-7b targets important cell cycle molecules in malignant melanoma cells and interferes with anchorage-independent growth.
	Oral Squamous Cell Carcinoma (OSCC)	MicroRNA expression and identification of putative miRNA targets in ovarian cancer.
	pituitary adenoma	Identification of differentially expressed microRNAs by microarray: a possible role for microRNA genes in pituitary adenomas.

	psoriasis	MicroRNAs: novel regulators involved in the pathogenesis of Psoriasis?
	retinoblastoma	Identification of miRNAs associated with tumorigenesis of retinoblastoma by miRNA microarray analysis.
	lupus nephritis	Comprehensive analysis of microRNA expression patterns in renal biopsies of lupus nephritis patients.
	non-alcoholic fatty liver disease (NAFLD)	Effect of miRNA-10b in regulating cellular steatosis level by targeting PPAR-alpha expression, a novel mechanism for the pathogenesis of NAFLD
	multiple myeloma (MM)	Identification of microRNA expression patterns and definition of a microRNA/mRNA regulatory network in distinct molecular groups of multiple myeloma
hsa-let-7i	Alzheimer's disease	Loss of microRNA cluster miR-29a/b-1 in sporadic Alzheimer's disease correlates with increased BACE1/beta-secretase expression.
	breast cancer	MicroRNA gene expression deregulation in human breast cancer.
	head and neck squamous cell carcinoma (HNSCC)	High mobility group A2 is a target for miRNA-98 in head and neck squamous cell carcinoma.
	ovarian cancer (OC)	MicroRNA microarray identifies Let-7i as a novel biomarker and therapeutic target in human epithelial ovarian cancer.
hsa-miR-23a	hepatocellular carcinoma (HCC)	Upregulation of miR-23a approximately 27a approximately 24 decreases transforming growth factor-beta-induced tumor-suppressive activities in human hepatocellular carcinoma cells.

	acute lymphoblastic leukemia (ALL)	MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia.
	acute myeloid leukemia (AML)	MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia.
	autism spectrum disorder (ASD)	Heterogeneous dysregulation of microRNAs across the autism spectrum.
	bladder cancer	Micro-RNA profiling in kidney and bladder cancers.
	cardiac hypertrophy	MicroRNAs play an essential role in the development of cardiac hypertrophy.
	colorectal cancer	Differentially regulated micro-RNAs and actively translated messenger RNA transcripts by tumor suppressor p53 in colon cancer.
	glioblastoma	Extensive modulation of a set of microRNAs in primary glioblastoma.
	heart failure	A signature pattern of stress-responsive microRNAs that can evoke cardiac hypertrophy and heart failure.
	hepatocellular carcinoma (HCC)	Downregulation of miR-122 in the rodent and human hepatocellular carcinomas.
	Oral Squamous Cell Carcinoma (OSCC)	Exploration of tumor-suppressive microRNAs silenced by DNA hypermethylation in oral cancer.
	pancreatic cancer	MicroRNA expression patterns to differentiate pancreatic adenocarcinoma from normal pancreas and chronic pancreatitis.

107

	prostate cancer	MicroRNA expression profiling in prostate cancer.
	Acute Promyelocytic Leukemia (APL)	Transcriptional repression of microRNA genes by PML-RARA increases expression of key cancer proteins in acute promyelocytic leukemia.
	lupus nephritis	Comprehensive analysis of microRNA expression patterns in renal biopsies of lupus nephritis patients.
	ulcerative colitis (UC)	MicroRNAs are differentially expressed in ulcerative colitis and alter expression of macrophage inflammatory peptide-2 alpha.
	cardiac hypertrophy	miR-23a functions downstream of NFATc3 to regulate cardiac hypertrophy.
hsa-miR-26a	anaplastic thyroid carcinoma (ATC)	Specific microRNAs are downregulated in human thyroid anaplastic carcinomas.
	acute lymphoblastic leukemia (ALL)	MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia.
	acute myeloid leukemia (AML)	MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia.
	acute myeloid leukemia (AML)	Distinctive patterns of microRNA expression associated with karyotype in acute myeloid leukaemia.
	cardiac hypertrophy	MicroRNAs play an essential role in the development of cardiac hypertrophy.
	colorectal cancer	Differentially regulated micro-RNAs and actively translated messenger 108

		RNA transcripts by tumor suppressor p53 in colon cancer.
	Duchenne muscular dystrophy (DMD)	Distinctive patterns of microRNA expression in primary muscular disorders.
	epithelial ovarian cancer (EOC)	Genomic and epigenetic alterations deregulate microRNA expression in human epithelial ovarian cancer.
	pituitary adenoma	Identification of differentially expressed microRNAs by microarray: a possible role for microRNA genes in pituitary adenomas.
	primary biliary cirrhosis (PBC)	MicroRNA profile in peripheral blood T cells of patients with primary biliary cirrhosis.
	prostate cancer	MicroRNA expression profiling in prostate cancer.
	prostate cancer	Expression of microRNAs and protein-coding genes associated with perineural invasion in prostate cancer.
	serous ovarian cancer	MicroRNA expression profiles in serous ovarian carcinoma.
	Burkitt lymphoma	MYC stimulates EZH2 expression by repression of its negative regulator miR-26a.
	ulcerative colitis (UC)	MicroRNAs are differentially expressed in ulcerative colitis and alter expression of macrophage inflammatory peptide-2 alpha.
	primary biliary cirrhosis (PBC)	Primary biliary cirrhosis is associated with altered hepatic microRNA expression.
	hepatocellular carcinoma (HCC)	Therapeutic microRNA delivery suppresses tumorigenesis in a murine liver cancer model. 109

	Oral Squamous Cell Carcinoma (OSCC)	The expression profile of microRNAs in a model of 7,12-dimethyl-benz[a]anthrance-induced oral carcinogenesis in Syrian hamster.
	bladder cancer	Up-regulation of microRNA in bladder tumor tissue is not common.
	glioma	The PTEN-regulating microRNA miR-26a is amplified in high-grade glioma and facilitates gliomagenesis in vivo.
	kidney cancer	Differential expression profiling of microRNAs and their potential involvement in renal cell carcinoma pathogenesis.
	hepatocellular carcinoma (HCC)	MicroRNA expression, survival, and response to interferon in liver cancer
	breast cancer	Widespread estrogen-dependent repression of micrornas involved in breast tumor cell growth
	papillary thyroid carcinoma (PTC)	miRNA Expression in a Human Papillary Thyroid Carcinoma Cell Line Varies with Invasiveness
hsa-miR-125a-5p	lung cancer	Epidermal growth factor receptor-regulated miR-125a-5p - a metastatic inhibitor of lung cancer.
	multiple myeloma (MM)	Identification of microRNA expression patterns and definition of a microRNA/mRNA regulatory network in distinct molecular groups of multiple myeloma
hsa-miR-135b	colorectal cancer	Identification by Real-time PCR of 13 mature microRNAs differentially expressed in colorectal cancer and 110

		non-tumoral tissues.
	malignant melanoma	MicroRNA let-7b targets important cell cycle molecules in malignant melanoma cells and interferes with anchorage-independent growth.
hsa-miR-149	breast cancer	MicroRNA gene expression deregulation in human breast cancer.
	cardiac hypertrophy	MicroRNAs play an essential role in the development of cardiac hypertrophy.
	cardiac hypertrophy	MicroRNAs are aberrantly expressed in hypertrophic heart: do they play a role in cardiac hypertrophy?
	diffuse large B-cell lymphoma (DLBCL)	MicroRNA signatures characterize diffuse large B-cell lymphomas and follicular lymphomas.
	follicular lymphoma (FL)	MicroRNA signatures characterize diffuse large B-cell lymphomas and follicular lymphomas.
	glioblastoma multiforme (GBM)	miR-124 and miR-137 inhibit proliferation of glioblastoma multiforme cells and induce differentiation of brain tumor stem cells.
	malignant melanoma	MicroRNA let-7b targets important cell cycle molecules in malignant melanoma cells and interferes with anchorage-independent growth.
	pituitary adenoma	Identification of differentially expressed microRNAs by microarray: a possible role for microRNA genes in pituitary adenomas.
	prostate cancer	Diagnostic and prognostic implications of microRNA profiling in prostate carcinoma.

hsa-miR-224	hepatocellular carcinoma (HCC)	hepatocellular carcinoma (HCC)
	epithelial ovarian cancer (EOC)	MicroRNA signatures in human ovarian cancer.
	hepatocellular carcinoma (HCC)	Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues.
	hepatocellular carcinoma (HCC)	MicroRNA profiling in hepatocellular tumors is associated with clinical features and oncogene/tumor suppressor gene mutations.
	lung cancer	Unique microRNA molecular profiles in lung cancer diagnosis and prognosis.
	malignant melanoma	MicroRNA let-7b targets important cell cycle molecules in malignant melanoma cells and interferes with anchorage-independent growth.
	Oral Squamous Cell Carcinoma (OSCC)	Exploration of tumor-suppressive microRNAs silenced by DNA hypermethylation in oral cancer.
	pancreatic ductal adenocarcinoma (PDAC)	MicroRNA expression alterations are linked to tumorigenesis and non-neoplastic processes in pancreatic ductal adenocarcinoma.
	prostate cancer	Expression of microRNAs and protein-coding genes associated with perineural invasion in prostate cancer.
	hepatocellular carcinoma (HCC)	Bead-based microarray analysis of microRNA expression in hepatocellular carcinoma: miR-338 is downregulated.
	colorectal cancer	Characterization of global microRNA expression reveals oncogenic potential of miR-145 in metastatic colorectal cancer 112

hsa-miR-339-5p	neurodegeneration	A miRNA signature of prion induced neurodegeneration.
hsa-miR-342-3p	neurodegeneration	A miRNA signature of prion induced neurodegeneration.
	kidney cancer	Differential expression profiling of microRNAs and their potential involvement in renal cell carcinoma pathogenesis.
	prion disease	Upregulation of miRNA hsa-miR-342-3p in experimental and idiopathic prion disease.
hsa-miR-378*	colorectal cancer	Characterization of global microRNA expression reveals oncogenic potential of miR-145 in metastatic colorectal cancer
hsa-miR-423-5p	Gastric Cancer	A five-microRNA signature identified from genome-wide serum microRNA expression profiling serves as a fingerprint for gastric cancer diagnosis.
hsa-miR-424	acute lymphoblastic leukemia (ALL)	MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia.
	acute myeloid leukemia (AML)	MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia.

113

	cardiac hypertrophy	MicroRNAs are aberrantly expressed in hypertrophic heart: do they play a role in cardiac hypertrophy?
	head and neck squamous cell carcinoma (HNSCC)	High mobility group A2 is a target for miRNA-98 in head and neck squamous cell carcinoma.
	ovarian cancer (OC)	MicroRNA expression and identification of putative miRNA targets in ovarian cancer.
	pancreatic cancer	Expression profiling identifies microRNA signature in pancreatic cancer.
	endometriosis	MicroRNA-regulated pathways associated with endometriosis.
	kidney cancer	Differential expression profiling of microRNAs and their potential involvement in renal cell carcinoma pathogenesis.
	chronic lymphocytic leukemia (CLL)	miRNA deregulation by epigenetic silencing disrupts suppression of the oncogene PLAG1 in chronic lymphocytic leukemia.
	Intrahepatic cholangiocarcinoma (ICC)	MicroRNA profiling of human intrahepatic cholangiocarcinoma cell lines reveals biliary epithelial cell-specific microRNAs
hsa-miR-629	lupus nephritis	Comprehensive analysis of microRNA expression patterns in renal biopsies of lupus nephritis patients.
	ulcerative colitis (UC)	MicroRNAs are differentially expressed in ulcerative colitis and alter expression of macrophage inflammatory peptide-2 alpha.

114

6) Supplementary Figure 1

   a. miRNA Blue module and its targets : bluemiblackm.pdf and bluemiallm.pdf.

   b. miRNA Turquoise module and its targets : turmiallm.pdf

   c. other miRNAs and its targets : greymibrownm.pdf and greymiallm.pdf

VITA

NAME   Bhavisha Chapatwala

EDUCATION   B.E. Electronics, Sardar Vallabhbhai National Institute of Technology,

Surat-395007.

TEACHING   Department of Electronics, Sardar Vallabhbhai National Institute of

EXPERIENCE   Technology, Surat-395007. 2003

Department of Electronics and Instrument & Control, Sarvajanik College

of Engineering and Technology, Surat-395001. 2003 – 2009