Dynamic Remodeling of Context-Specific miRNAs Regulation Networks Facilitate *in silico* Cancer Drug Screening

Lida Zhu[1, §], Fengji Liang[2, §], Juan Liu[1,*], Simon Rayner[3], Yinghui Li[2], Shanguang Chen[2], Jianghui Xiong[2,1,*]

1 School of Computer Science, Wuhan University, Wuhan, P.R. China

2 Bioinformatics Group and Data Coordination Center, State Key Lab of Space Medicine Fundamentals and Application (SMFA), China Astronaut Research and Training Center (ACC), Beijing, P.R. China

3 Bioinformatics Group, State Key Laboratory of Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan, P.R. China

§ These authors contributed equally.

* Correspondence should be addressed to J.X. (laserxiong@gmail.com) or J.L. (liujuan@whu.edu.cn).

Abstract—Background: Much effort has been expended in exploring the connections between transcriptome, disease and drug, based on the premise that drug induced perturbations in the transcriptome will affect the phenotype and finally help to cure a disease. MicroRNAs (miRNAs) play a key role in the regulation of the transcriptome and have been identified as a key mediator in human disease and drug response. However, even if miRNA expression can be precisely detected, the information regarding miRNAs action on a particular part of the transcriptome is still lacking. Here, we introduced a novel concept, the Context-specific MiRNA activity (CoMi activity), to reflect a miRNA's regulation effect on a context specific gene set, by calculating the statistical difference between the distributions of its target gene expression and non-target gene expression. In this study we investigate whether CoMi activity could provide a novel perspective on miRNA mechanisms of action in disease and drug response, and facilitate *in silico* drug screening.

Results: Using breast cancer as an example, we examined the CoMi activity based on a Gene Ontology (GO) term as context. Then we constructed a differential CoMi activity network (cancer vs. normal), based on the comi activity represents as a link between miRNAs and GO terms. (e.g. hsa-miR-27a's regulation on GO term "apoptosis"). The topological analysis of the generated network demonstrated that the cancer specific CoMi network is a scale free network. The highly connected nodes highlighted a group of known onco-miRNAs (e.g. hsa-miR-183*) and tumor suppressor miRNAs (e.g. hsa-miR-34a), as well as some well-known cancer related GO biological processes (e.g. apoptosis). Interestingly, we found that chemotherapeutic drug treatment can counteract the dis-regulated CoMi activity in the cancer-specific network. For instance, 100% of down-regulated CoMi activities in a "core" breast cancer network contains apoptosis-related GO terms that could be counteracted by Paclitaxel treatment.

To perform *in silico* drug screening, the similarity of a query CoMi activity signature (e.g. differential CoMi activity in cancer vs. normal) to each of the reference CoMi activities (converted from reference mRNAs expression profiles in the Connectivity Map) was assessed. We found that the most negatively correlated compounds significantly overlapped with known cancer drugs.

Conclusions: By defining a Stability Index for *in silico* drug screening, we found CoMi activity signatures strikingly outperformed the traditional CMAP method or mRNA-based signatures. Thus, the dynamic remodeling of context-specific miRNAs regulation network could reveal the hidden miRNAs that act as key mediators of drug action and facilitate *in silico* cancer drug screening.

Keywords: Specific-context MiRNA activity; miRNA;in silico drug screening; network biology; network pharmacology

I. BACKGROUND

Gene-expression profiling has historically been applied in specific settings to connect human diseases with the drugs that treat them. For example, the Connectivity Map (CMAP) [1, 2] project set out to create a reference collection of gene-expression profiles from cultured human cells treated with bioactive small molecules. The similarity of a query signature (e.g. differential expressed gene set in cancer vs. normal) to each of the reference expression profiles in the data set was then assessed and the most negatively correlated drugs might represent potential new therapeutics. The CMAP can be considered as a conceptual framework of an *in silico* drug screening tool: given the query signature for a disease state (e.g. cancer vs. normal) and the CMAP reference data set, identifying the compounds that have the most negative correlation provides a candidate set of potential therapeutics. By applying the CMAP method, several novel agents have been discovered and characterized [3-5] .

Considering the important role microRNAs play in defining cancer related phenotypes, *in silico* drug screening based on miRNAs regulation features might be a useful approach [6-12]. Previously our group constructed a prognosis-guided synergistic gene-gene interaction network, and demonstrated its application as an efficient *in silico* tool for pre-clinical drug prioritization and rational design of combinatorial therapies [13]. An interesting discovery was that microRNAs play an important role in the network, yet how miRNAs regulation features connect a disease state with drugs is largely unknown.

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

Here we propose that miRNAs regulation on various parts of the transcriptome might provide a novel viewpoint to connect the disease state with the drugs that treat them, and facilitate *in silico* drug screening. We introduce a novel concept, the Context-specific MiRNA activity (CoMi activity), to reflect a miRNA's regulation effect on a context-specific gene set by calculating the statistical difference between the distributions of its target gene expression and non-target gene expression. The "specific context" can be defined from different perspectives, such as pathway information, protein-protein interaction (PPI) networks [14, 15] and Gene Ontology [16]. A set of genes with "specific context" is marked with a specific biological function background, and may share a common regulation mechanism.

In this paper, we first construct a differential CoMi activity network for breast cancer (cancer vs. normal), then analyze the topology structure and the highly connected miRNAs and GO terms. The perturbation pattern of therapeutic drugs on this network was then investigated. Finally we evaluate the performance of the CoMi-based *in silico* drug screening method by comparing with mRNAs-based methods.

II. Results and Discussion

A. CoMi activity

The definition and calculation of CoMi activity is similar to our paper "Xionghui Zhou et al. Context-specific miRNA regulation network predicts cancer prognosis" (Also submitted to this conference). Briefly, gene sets defined by Gene Ontology Terms (Gene Ontology Biological Process - GOBP, Gene Ontology cellular component - GOCC) were used to define "context". For each miRNA, we gained target information from online databases [17], to form a miRNA target gene set. To measure a miRNA's CoMi activity on a specific GO term, we separated its target genes and non-target genes from the GO term, and calculated the statistical variation difference between target genes and non-target genes. Through the combination of GO term gene sets and the miRNA target gene sets, we abstract a virtual probe –a "CoMi activity probe" – to test the miRNA regulation effect on the transcriptome (Fig 1.b).

B. In silico drug screening framework

Based on the CoMi activity concept, we developed a computational framework to construct the CoMi activity network, and search for specific cancer treatment drugs (Fig 1).

First, a disease-specific differential gene expression pattern was generated by a fold-change calculation or T-test for query (Fig1.a). At the same time, we characterized each drug perturbation profile in the similar way (drug perturbation vs. control), and constructed a library of drug induced differential gene expression patterns (fold-change here) (Fig1.c).

After building the GO term gene set database and the miRNA target gene set database, we paired each GO term and each miRNA target gene set to construct a CoMi activity probe by calculating the intersection of miRNA's target gene set and GO term gene set, and then collected all the significant probes to form a pool (Fig1.b).

Next, we performed a "CoMi activity calculation" to translate the gene expression pattern into the CoMi activity pattern (Fig1.d). In the process we mapped a specific CoMi activity probe to a gene expression fold-change pattern. When the statistical gene expression variation of miRNA target genes is significantly lower than the non-target genes, the CoMi activity is considered as up-regulated. Conversely, the CoMi activity is considered as down-regulated when non-target gene expression variation is significantly lower than target genes.

Using this "CoMi activity calculation", we translated a cancer specific gene expression fold-change pattern into a corresponding CoMi activity pattern. Thus the constructed CoMi activity network is representative of the specific cancer (Fig 1.e). In the same way, we transferred the drug library of gene expression fold-change patterns into a library of CoMi activity patterns (Fig 1.f).

From the cancer-specific CoMi activity pattern, we chose the most significant CoMi activities, and took the top k up-regulated and bottom k down-regulated CoMi activities as the query signature. We then performed a similarity search in the drug library of CoMi activity patterns that was created earlier based on Spearman correlation, and each instance in the library was given a Potential Efficacy Score (PES) according to the correlation (see **Method**).

Finally, we ranked the instances in the drug library according to their PESs (Fig 1.h).

C. Breast cancer CoMi activity network

By compiling differential CoMi activities for breast cancer (cancer vs. normal), a CoMi activity network was generated. Each CoMi activity is virtualized as an edge from a miRNA node to a node representing a specific GO term. Since a certain miRNA could have several CoMi activities, it means that this miRNA could affect several biological transcription activities according to GO terms. Conversely, a set of genes within the same GO term could be affected by several miRNAs. In this way, the CoMi activity pattern is virtualized as a network connected by GO terms and miRNAs.

We first investigated the topology of the CoMi activity network. Both the in-degree (GO term) and the out-degree (miRNA) of the network fit a power low distribution (data not shown), suggesting that the CoMi network is a typical scale free network.

We list several highly connected hubs in Table1. The highly connected GO term nodes in the network show close association with cancer, for example, "apoptosis", "regulation of apoptosis", "signal transduction" and "transport" (Fig 2).

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

At the same time, a few highly connected miRNA hubs in the network are also associated with breast cancer (hsa-miR-183*, hsa-miR-34a and hsa-miR-27a) (Fig 3). Interestingly, the miRNAs-GO term combination is highly relevant with the reported mechanism of miRNAs. For example, it was found that hsa-miR-183 could interact with β TrCP1 which is an important gene in signal transduction and the cell cycle [18] .It has been also found that hsa-miR-183 could regulate Ezrin expression in cancer, which is an important gene in the cytoskeleton [19]. Our network clearly shows that hsa-miR-183*, the complementary strand of hsa-miR-183 (which is often cleaved as a miRNA duplex that is loaded onto RISC), is associated with breast cancer via several GO terms include "signal transduction", "cell cycle" and "cytoskeleton". Another highly connected miRNA, hsa-miR-34a, is also highlighted. As an important tumor suppressor miRNA, hsa-miR-34a was found to associate with various types of cancer [20]. One of the reported target genes is E2F3, which is involved in "regulations of transcription, DNA-dependent" [21]. These results suggest that the CoMi activity network could not only reveal the key phenotype specific miRNAs and biological processes, but also reveal the miRNA mechanism of action.

D. CoMi activity network perturbed by therapeutic Drug

In the generated breast cancer specific CoMi activity network (Fig 2), we selected a "core" subnetwork that includes highly related GO terms including "anti-apoptosis", "regulation of apoptosis" and "positive regulation of cell proliferation". As shown in Figure 4, most of the CoMi activities in this subnetwork are down-regulated (14 green links vs. 1 red links). We next investigated the perturbation pattern of the first line chemotherapy drug Paclitaxel on this "core" subnetwork. As shown in Figure 4, 100% of the down-regulated CoMi activities in the "core" subnetwork was up-regulated by Paclitaxel treatment (all 14 green links change to red links). Especially, all the CoMi activity of has-miR-27a and hsa-miR-27b are perturbed by Paclitaxel which suggests a pivotal role for has-miR-27a/b. It is reported that hsa-miR-27a is a typical oncogenic miRNA that could regulate ZBTB10/RINZF, an Sp1 repressor, and RYBP/DEDAF, an apoptotic facilitator, in breast cancer cells [22]. It makes sense that hsa-miR-27a's target genes were up-regulated compared to the other genes in response to the breast cancer induced down-regulation of hsa-miR-27a. In treatment with Paclitaxel, hsa-miR-27a is induced to be up-regulated, which causes the target genes (involved in "anti-apoptosis" and "positive regulation of cell proliferation") to be down-regulated. Thus apoptosis is up-regulated and the cancer is inhibited by Paclitaxel.

We present another example, hsa-miR-125b, to illustrate the action of drug on CoMi activity network. Hsa-miR-125b was reported to be significantly deregulated in breast cancer tissues [12, 23]. As shown in Figure 5, hsa-miR-125b's activities on "apoptosis" and "metabolic process" were down-regulated in breast cancer. Interestingly, after Paclitaxel treatment, its activities on "apoptosis" and "metabolic process" were up-regulated. This is in line with the evidence that the hsa-miR-125b was over-expressed in Paclitaxel treatment on Taxol-resistant cells [24]. As shown in Figure 5, 80% of the down-regulated CoMi activities in the "apoptosis" subnetwork was up-regulated by Paclitaxel treatment (all 16 green links change to red links, 5 red links change to green links). This example highlights how we can use this technique as a bridge between cancer and drug with the remodeling of the CoMi activity network.

E. Performance of in silico drug screening based on CoMi network

We next designed a test bed to evaluate the performance of our method. First of all, we built a drug CoMi activity pattern library. The library contains 103 instances, 19 of these instances come from 4 known drugs that are specific to breast cancer treatment. Starting from a breast cancer gene expression fold-change pattern, we obtained a corresponding CoMi activity pattern. Then we chose the most significant CoMi activities (see **Method**), and selected the top k up-regulated and bottom k down-regulated CoMi activities in the pattern to create a query signature, and performed a similarity search in the drug CoMi activity pattern library. We then investigated whether the framework could recognize breast cancer specific drugs by awarding them a high rank.

The correlation (similarity) calculation is adapted from standard Connectivity Map query [2]. In CMAP, an up-tag (over-expressed gene set) and down-tag (under-expressed gene set) is generated by selecting a top/down k gene in the ranked gene expression fold-change. Then the Kolmogorov–Smirnov (KS) score is calculated [25] . In the above procedures, the parameter k is a critical factor affecting the performance of *in silico* drug screening. Previous studies demonstrated that different selection of up-tags/down-tags may generate diverse and even conflicting outcomes [26]. Thus we also defined a "Stability Index" (SI), to represent the ratio of good results in the entire result set according to different parameter options, to assess the performance of drug screening methods (see **Methods**).

We tested 3 different miRNAs target data sources: miRanda, ExprtargetDB and TargetScan&RNA22. ExprtargetDB achieved the highest SI. When the permutation p-value cutoff was set to 0.05, the SI reached 0.9406. When the permutation p-value cutoff was set to 0.1, the SI of another data source, TargetScan&RNA22, also reached 0.8689 (Fig 6). Thus our method is robust regardless of the miRNAs target gene data source.

F. Comparison with traditional CMAP drug screening

Our proposed method is based on the CoMi activity pattern: the query pattern and the reference drug library pattern are both CoMi activity patterns (we call this method the "CoMi-based method"). For comparison, we also conducted drug screening based on the gene expression

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

pattern: the query pattern and the reference drug library pattern (we call this method the "Gene-based method").

As described above, we use Spearman correlation analysis to conduct the similarity search. In the initial form of the CMAP method, the KS test was employed to conduct the similarity search (between the query cancer differential gene expression and the drug induced gene expression fold-change). Thus we also include the original CMAP method for comparison. (In the following discussion, we call this method the "CMAP method")

As shown in Figure 7, when the permutation p-value cutoff is set to 0.05, the CoMi-based method achieves an SI of 0.9406, while the Gene-based method only achieves 0.3810 and CMAP only achieves an SI of 0.0004. When the permutation p-value cutoff is set to 0.1, the CoMi-based method achieves an SI of 0.9525, for the Gene-based method SI achieves at 0.4925, and for CMAP SI achieves 0.0051. The low stability index of the original CMAP method suggests that CMAP is very sensitive to the selection of up-tags/down-tags [26].

We derived a drug-list result generated from the instance-list result using KS test, in the same way as CMAP online service [27]. In Table 2, a drug list of CoMi activity-based method and a drug list of CMAP method are shown respectively. It suggests that CoMi-based method could provide better performance than CMAP method in identifying breast cancer treatment drugs.

III. METHOD

A. Data source

Query data: The gene expression data of breast cancer vs. normal was downloaded from the Gene Expression Omnibus (GSE15852 [28]), with 43 breast carcinomas tissues and 43 patient-matched normal breast tissues.

Reference data: We selected 17 FDA approved oncology cancer drugs [29], including 4 breast cancer drugs: Paclitaxel [30-32], Tamoxifen [33], Vinblastine [34] and Mitoxantrone [35]. The drug induced gene expression data in the reference library was downloaded from Connectivity Map (build 02) [27].

MiRNAs target gene information was collected from 3 sources:

(1) ExprtargetDB, an integrated database combining miRanda, PicTar and TargetScan [17];

(2) miRanda (miRBase Targets Release Version v5) [10, 36, 37];

(3) Targetscan and RNA22 [38, 39]. In this paper, we use the union of these two sources, since our previous result demonstrates this appears to represent a more comprehensive set of miRNA target predictions [29].

B. CoMi activity calculation

In this paper, we calculated CoMi activity by T-test statistics, in order to evaluate the statistical variation difference between target genes expression fold-change and the non-target genes. The basic idea is to represent how significantly the miRNA can down-regulate its target genes by comparing the gene expression variation distribution of miRNA target genes with non target genes [25, 40].

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}}$$

We perform the T-test on gene expression fold-change value. It is important to notice that, the T-test to calculate the CoMi activity here is different from the calculation in our paper "Xionghui Zhou et al. Context-specific miRNA regulation network predicts cancer prognosis", which perform T-test on gene expression value. Based on the calculation from the gene expression fold-change value in our paper, the CoMi activity is denoted for differential miRNA activity changes in different conditions (e.g. cancer vs. normal, or drug perturbation vs. control).

C. Similarity calculation between drug induced CoMi activity pattern and cancer-specific CoMi activity pattern

We propose a similarity calculation method in the following steps based on Spearman correlation.

Step 1: After finishing the CoMi activity calculation for all the CoMi activity probes, we selected the most significant CoMi activities (p-value < 0.05) and constructed a cancer specific CoMi activity pattern.

Step 2: From the cancer specific CoMi activity pattern, we chose the top k up-regulated CoMi activities to generate the up-tag, and chose the bottom k down-regulated CoMi activities for the down-tag as well. By combining the up-tag and the down-tag, the query signature was generated.

Step 3: For each instance in the drug CoMi activity pattern library, the reference signature was chosen according to the query signature in Step 2,

Step 4: The Spearman correlation was performed between the query signature and each instance of reference signature in the drug CoMi activity pattern library. For each candidate instance, we could generate a Potential Efficacy Score (PES) based on the correlation between the query signature and each instance of the reference signature (the most negatively correlated instance is best).

For performance comparison, we also implemented the KS test for the similarity calculation method, which was performed in CMAP [2]. In brief, in the CMAP method, the KS test method was used to query the input feature. For each query, differentially expressed genes are partitioned as up- or down-regulated signatures. KS scores for both up and down-

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

regulated signature were calculated by applying the KS test respectively.

D. In silico drug screening performance evaluation method

To test whether the PES could significantly discriminate successful drugs from the candidate pool, we developed an evaluation method. For each option of k, all the n instances in the candidate instance library were ranked in an ascending order according to PES, and the Kolmogorov-Smirnov statistic was computed to test whether the rank positions of the t successful-drug-related instances are significantly enriched on the top of the ranked list [25]. We denote this KS test statistic as the "Drug Screening Performance index" (DSP index).

Furthermore, the rank position of successful-drug-related instances was permuted 1000 times, the "bootstrapped" KS score was calculated, and the fraction of runs for which the absolute value of the random KS score was larger than the absolute value of the real KS score is the p-value for the real KS score. We denote this p-value as the "Drug Screening Performance p-value" (DSP p-value).

E. Stability Index (SI)

Previous studies have demonstrated that the selection of parameter k (for generating up-tag/down-tag) is a critical factor affecting the performance of *in silico* drug screening [26]. Thus we defined a "Stability Index" (SI), the ratio of obtaining an acceptable/qualified estimate of the DSP index when k varies from 1 to N.

$$\text{Stability Index (SI)} = \frac{\text{Count of the acceptable DSP index}}{\text{N}}$$

In this paper, we defined an acceptable/qualified DSP index at different levels as following: Level 1 (DSP index > 0, and DSP p-value < 0.05); Level 2 (DSP index > 0, and DSP p-value < 0.1).

ACKNOWLEDGE

This work was supported by National Basic Research Program of China (973 Program, No2011CB711000), the National Natural Science Foundation of China to J.X. (30600759) and J.L. (60773010 and 60970063), the Ph.D. Programs Foundation of Ministry of Education of China (20090141110026) and the Fundamental Research Funds for the Central Universities (6081007) to J.L, grant from State Key Lab of Space Medicine Fundamentals and Application (SMFA) to J.X (SMFA09A07, SMFA10A03), and Advanced Space Medico-Engineering Research Project of China (01111014, 01111016). The study was carried out in SMFA.

REFERENCE

[1] J. Lamb, "The Connectivity Map: a new tool for biomedical research," *Nat Rev Cancer,* vol. 7, pp. 54-60, Jan 2007.

[2] J. Lamb, *et al.*, "The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease," *Science,* vol. 313, pp. 1929-35, Sep 29 2006.

[3] D. C. Hassane, *et al.*, "Discovery of agents that eradicate leukemia stem cells using an in silico screen of public gene expression data," *Blood,* vol. 111, pp. 5654-62, Jun 15 2008.

[4] G. Wei, *et al.*, "Gene expression-based chemical genomics identifies rapamycin as a modulator of MCL1 and glucocorticoid resistance," *Cancer Cell,* vol. 10, pp. 331-42, Oct 2006.

[5] D. R. Rhodes, *et al.*, "Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles," *Neoplasia,* vol. 9, pp. 166-80, Feb 2007.

[6] C. Blenkiron, *et al.*, "MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype," *Genome Biol,* vol. 8, p. R214, 2007.

[7] M. D. Mattie, *et al.*, "Optimized high-throughput microRNA expression profiling provides novel biomarker assessment of clinical prostate and breast cancer biopsies," *Mol Cancer,* vol. 5, p. 24, 2006.

[8] M. V. Iorio and C. M. Croce, "MicroRNAs in cancer: small molecules with a huge impact," *J Clin Oncol,* vol. 27, pp. 5848-56, Dec 1 2009.

[9] T. G. Vandenboom Ii, *et al.*, "MicroRNA and Cancer: Tiny Molecules with Major Implications," *Curr Genomics,* vol. 9, pp. 97-109, Apr 2008.

[10] S. Griffiths-Jones, *et al.*, "miRBase: tools for microRNA genomics," *Nucleic Acids Res,* vol. 36, pp. D154-8, Jan 2008.

[11] S. Griffiths-Jones, *et al.*, "miRBase: microRNA sequences, targets and gene nomenclature," *Nucleic Acids Res,* vol. 34, pp. D140-4, Jan 1 2006.

[12] M. Redova, *et al.*, "MicroRNAs and their target gene networks in renal cell carcinoma," *Biochem Biophys Res Commun,* vol. 405, pp. 153-6, Feb 11 2011.

[13] J. Xiong, *et al.*, "Pre-clinical drug prioritization via prognosis-guided genetic interaction networks," *PLoS One,* vol. 5, p. e13937, 2010.

[14] C. Besemann, *et al.*, "BISON: Bio-Interface for the Semi-global analysis Of Network patterns," *Source Code Biol Med,* vol. 1, p. 8, 2006.

[15] M. Krull, *et al.*, "TRANSPATH: an information resource for storing and visualizing signaling pathways and their pathological aberrations," *Nucleic Acids Res,* vol. 34, pp. D546-51, Jan 1 2006.

[16] "The Gene Ontology project in 2008," *Nucleic Acids Res,* vol. 36, pp. D440-4, Jan 2008.

[17] E. R. Gamazon, *et al.*, "Exprtarget: an integrative approach to predicting human microRNA targets," *PLoS One,* vol. 5, p. e13534, 2010.

[18] I. Elcheva, *et al.*, "CRD-BP protects the coding region of betaTrCP1 mRNA from miR-183-mediated degradation," *Mol Cell,* vol. 35, pp. 240-6, Jul 31 2009.

[19] G. Wang, *et al.*, "MicroRNA-183 regulates Ezrin expression in lung cancer cells," *FEBS Lett,* vol. 582, pp. 3663-8, Oct 29 2008.

[20] J. Cheng, *et al.*, "The impact of miR-34a on protein output in hepatocellular carcinoma HepG2 cells," *Proteomics,* vol. 10, pp. 1557-72, Apr 2010.

[21] W. C. Cho, "OncomiRs: the discovery and progress of microRNAs in cancers," *Mol Cancer,* vol. 6, p. 60, 2007.

[22] G. K. Scott, *et al.*, "Rapid alteration of microRNA levels by histone deacetylase inhibition," *Cancer Res,* vol. 66, pp. 1277-81, Feb 1 2006.

[23] M. V. Iorio, *et al.*, "MicroRNA gene expression deregulation in human breast cancer," *Cancer Res,* vol. 65, pp. 7065-70, Aug 15 2005.

[24] M. Zhou, *et al.*, "MicroRNA-125b confers the resistance of breast cancer cells to paclitaxel through suppression of pro-apoptotic Bcl-2 antagonist killer 1 (Bak1) expression," *J Biol Chem,* vol. 285, pp. 21496-507, Jul 9 2010.

[25] D. W. M. Hollander, *Nonparametric Statistical Methods* 2ed. New York: Wiley, 1999.

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

[26] Y. Li, *et al.*, "Gene expression module-based chemical function similarity search," *Nucleic Acids Res,* vol. 36, p. e137, Nov 2008.
[27] (2006). *Connectivity Map 02*. Available: http://www.broadinstitute.org/cmap/
[28] I. B. Pau Ni, *et al.*, "Gene expression patterns distinguish breast carcinomas from normal breast tissues: the Malaysian context," *Pathol Res Pract,* vol. 206, pp. 223-8, Apr 15 2010.
[29] J. M. Collins. *Developmental Therapeutics Program*. Available: http://dtp.nci.nih.gov/
[30] M. W. Saville, *et al.*, "Treatment of HIV-associated Kaposi's sarcoma with paclitaxel," *Lancet,* vol. 346, pp. 26-8, Jul 1 1995.
[31] R. Bharadwaj and H. Yu, "The spindle checkpoint, aneuploidy, and cancer," *Oncogene,* vol. 23, pp. 2016-27, Mar 15 2004.
[32] D. A. Brito, *et al.*, "Microtubules do not promote mitotic slippage when the spindle assembly checkpoint cannot be satisfied," *J Cell Biol,* vol. 182, pp. 623-9, Aug 25 2008.
[33] A. Hurtado, *et al.*, "Regulation of ERBB2 by oestrogen receptor-PAX2 determines response to tamoxifen," *Nature,* vol. 456, pp. 663-6, Dec 4 2008.
[34] D. Starling, "Two ultrastructurally distinct tubulin paracrystals induced in sea-urchin eggs by vinblastine sulphate," *J Cell Sci,* vol. 20, pp. 79-89, Jan 1976.
[35] J. Mazerski, *et al.*, "The geometry of intercalation complex of antitumor mitoxantrone and ametantrone with DNA: molecular dynamics simulations," *Acta Biochim Pol,* vol. 45, pp. 1-11, 1998.
[36] A. J. Enright, *et al.*, "MicroRNA targets in Drosophila," *Genome Biol,* vol. 5, p. R1, 2003.
[37] K. C. Miranda, *et al.*, "A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes," *Cell,* vol. 126, pp. 1203-17, Sep 22 2006.
[38] B. P. Lewis, *et al.*, "Prediction of mammalian microRNA targets," *Cell,* vol. 115, pp. 787-98, Dec 26 2003.
[39] B. P. Lewis, *et al.*, "Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets," *Cell,* vol. 120, pp. 15-20, Jan 14 2005.
[40] D. Augot, *et al.*, "Spatial and temporal patterning of bank vole demography and the epidemiology of the Puumala hantavirus in northeastern France," *Epidemiol Infect,* vol. 136, pp. 1638-43, Dec 2008.

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

FIGURES AND TABLES:

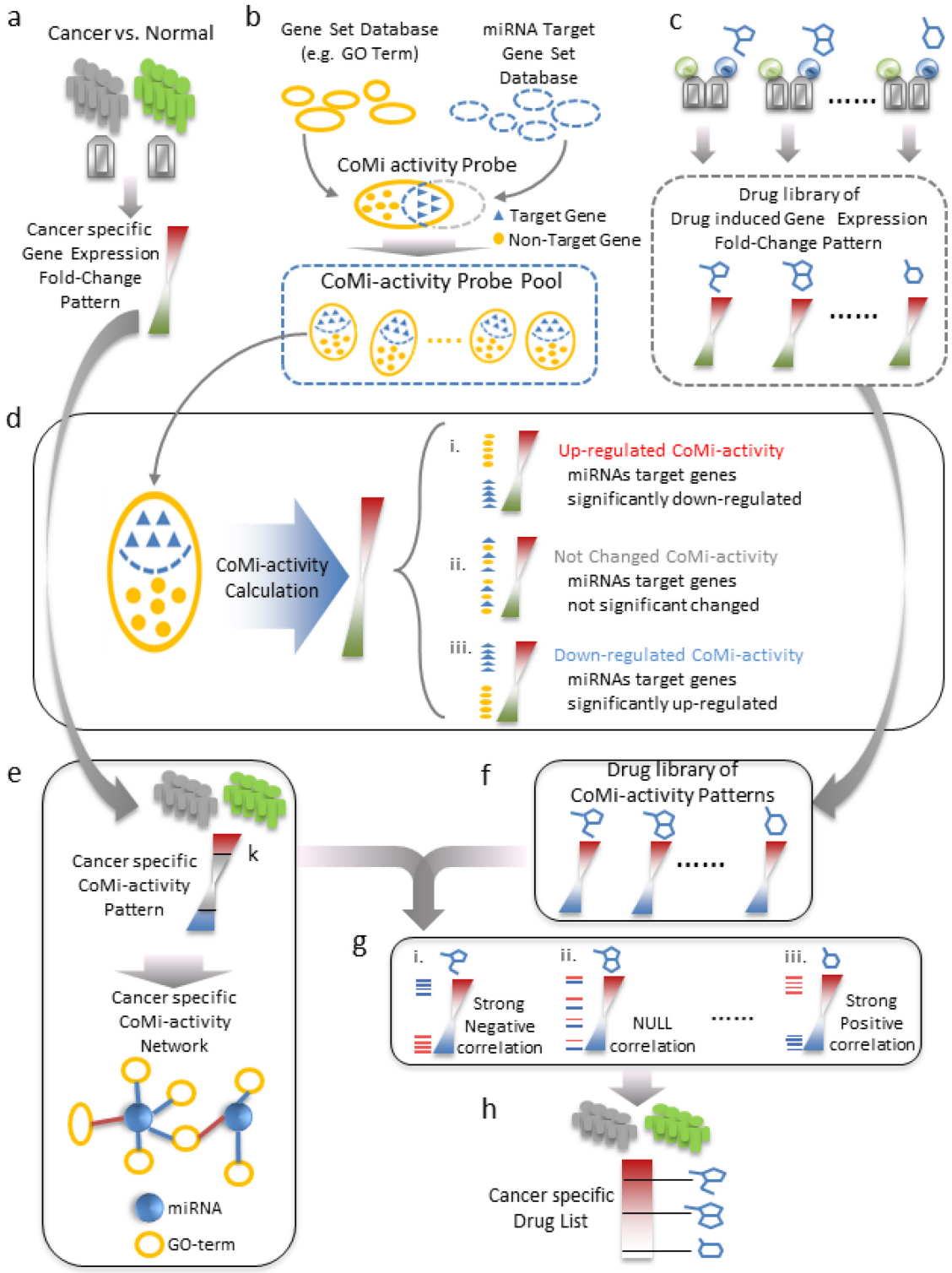

Figure 1. **Work flow of the framework**. (a) A disease specific gene expression pattern generated for the query data (cancer vs. normal); (b) Formation of CoMi activity probe pool. (c) Drug library of drug induced gene expression fold-change pattern. (d) CoMi activity calculation to transfer the gene expression pattern to CoMi activty pattern. (e) Cancer specific CoMi activity pattern. This can also be used to form a cancer specific CoMi-activity network. A parameter k was specified to generate the query signature. (f) Drug library of CoMi activity pattern. (g) Similarity search is performed based on Spearman correlation analysis, each instance in the library is assigned a Potential Efficacy Score (PES) . (h) Cancer specific ranked drug list generated according to PES.

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

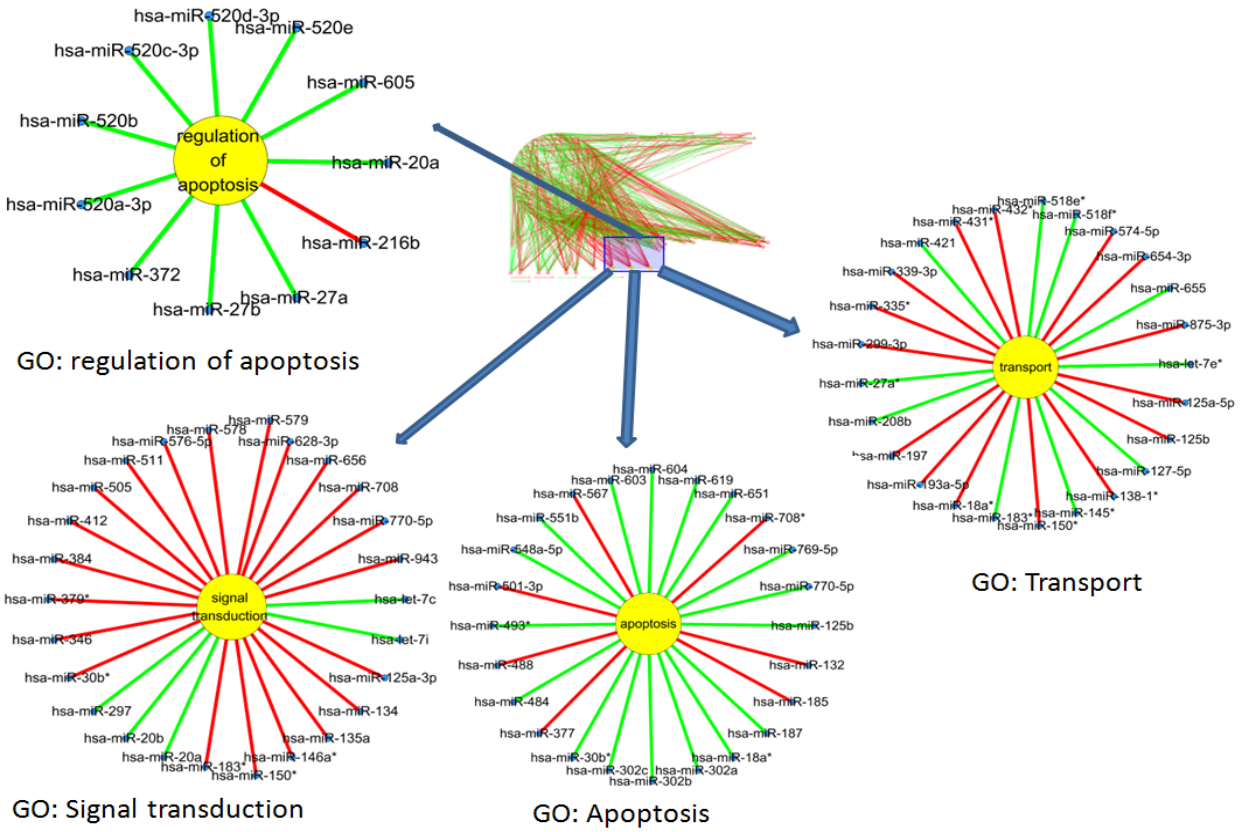

Figure 2. Breast cancer CoMi activity network sorted by the in-degree of the GO terms. Here shows four examples of GO term nodes

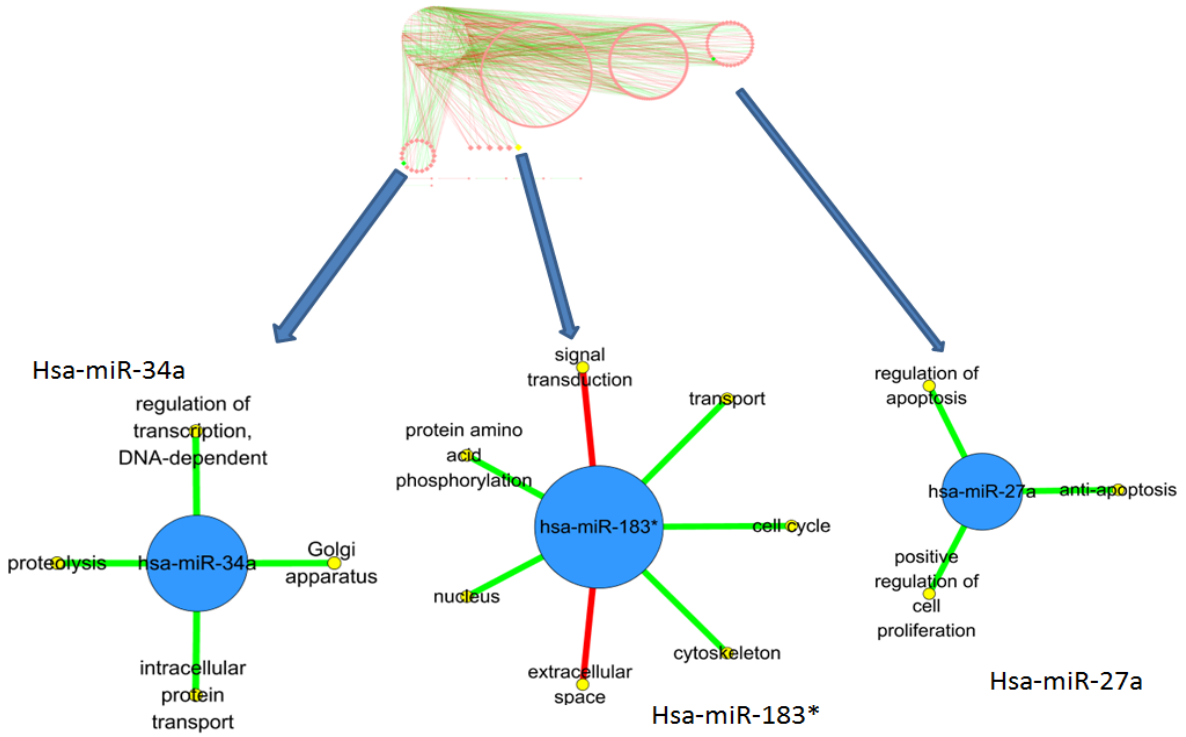

Figure 3. Breast cancer CoMi network sorted by miRNA's node degree. Here shows three examples of miRNA nodes.

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

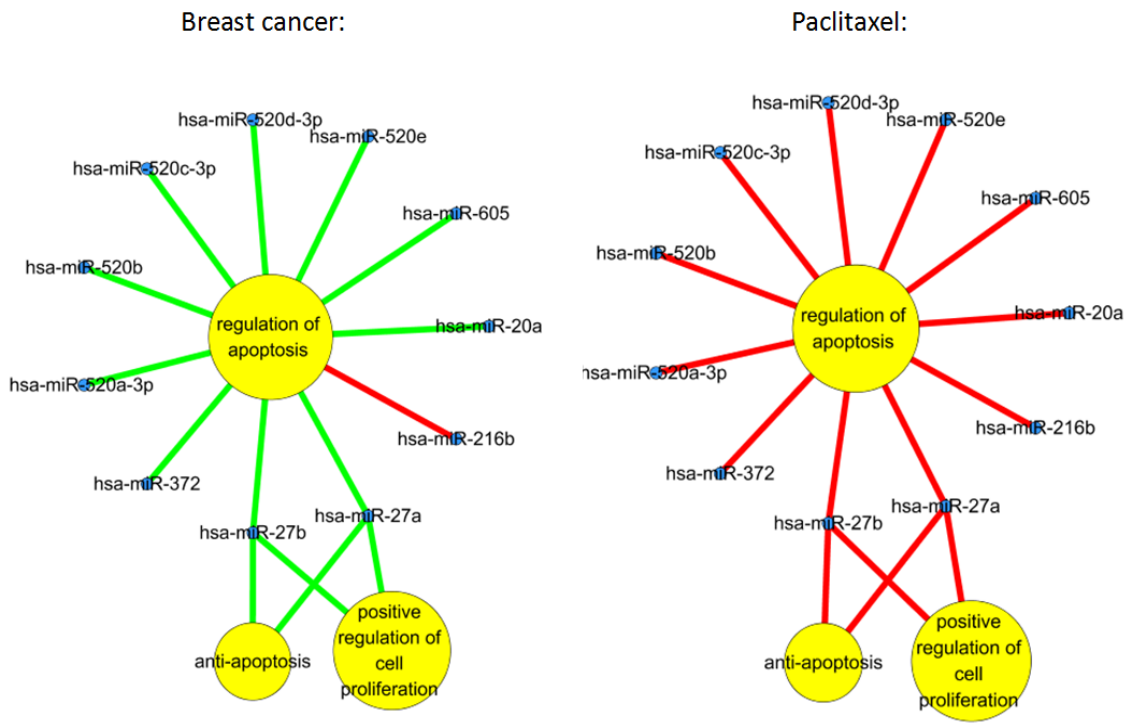

Figure 4. **Comparison between a "core" breast cancer CoMi activity subnetwork and corresponding Paclitaxel-induced subnetwork.** The subnetwork includes highly related GO terms including "anti-apoptosis", "regulation of apoptosis" and "positive regulation of cell proliferation"

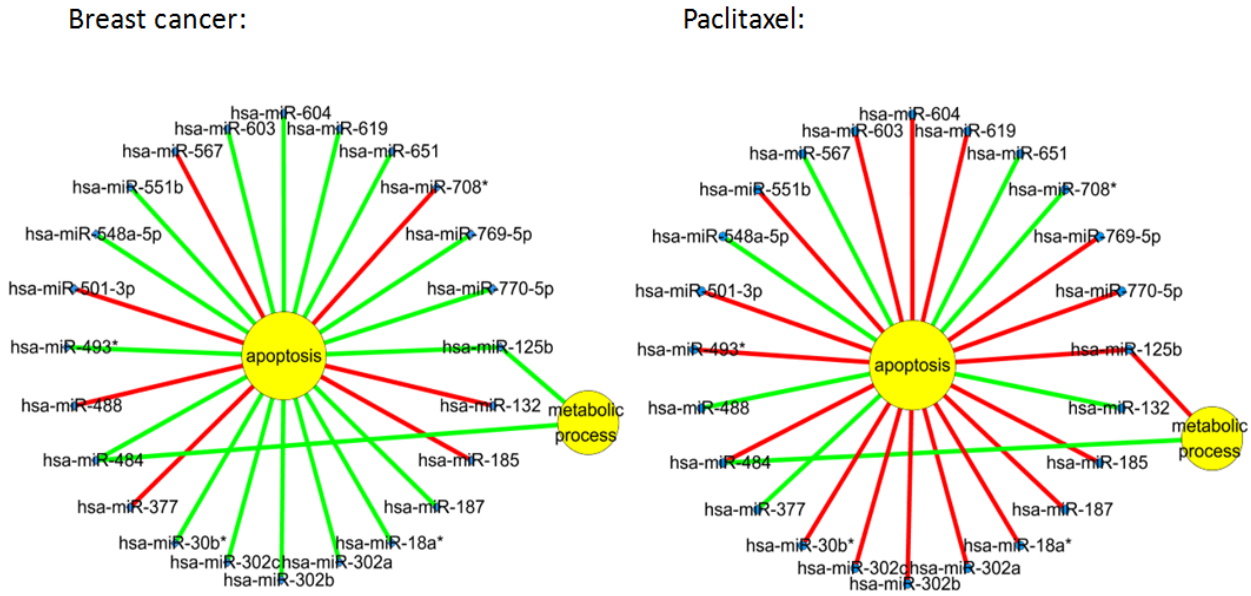

Figure 5. **Comparison between a "core" breast cancer CoMi activity subnetwork and corresponding Paclitaxel-induced subnetwork.** The subnetwork includes highly related GO terms including "apoptosis" and "metabolic process"

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

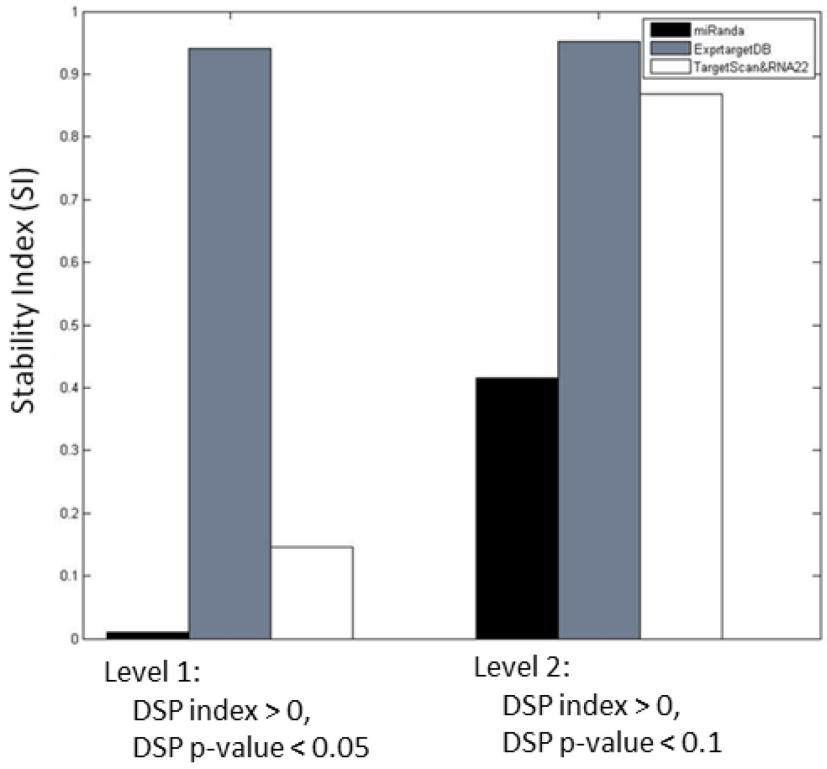

Figure 6. Performances comparison of different miRNA target databases (miRanda, ExprtargetDB and Targetscan&RNA22).

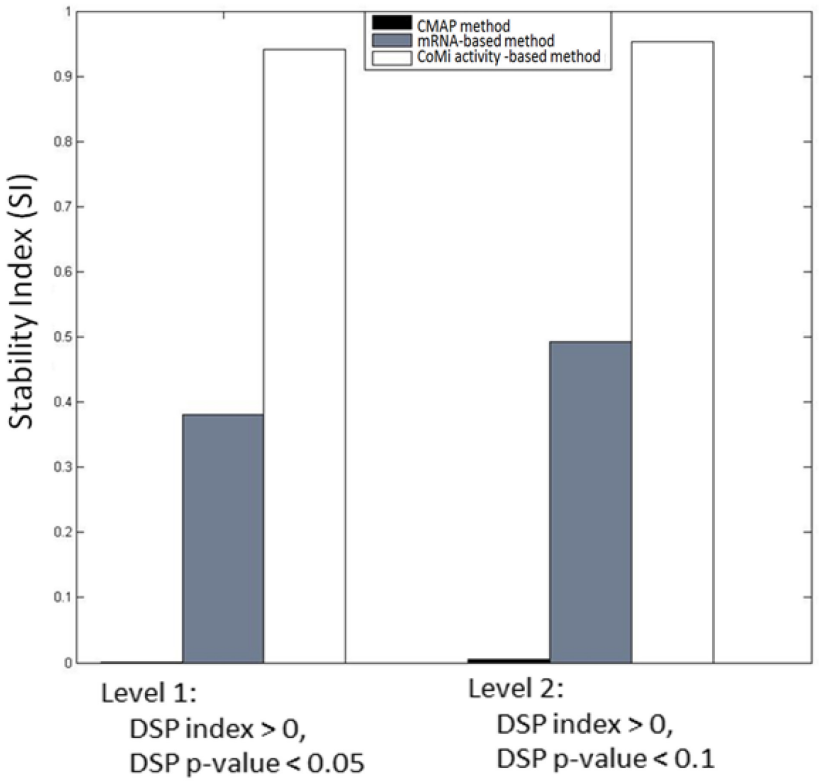

Figure 7. Performances comparison of (i) CMAP method, (ii) mRNA-based method and (iii) CoMi activity –based method.

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

Table 1 Degree of nodes in Breast cancer specific CoMi activity network

Degree	GO term node	Degree	miRNA node
26	signal transduction	7	hsa-mir-183*
25	Transport	6	hsa-miR-34a
24	Apoptosis	6	hsa-miR-103
21	protein amino acid phosphorylation	6	hsa-miR-339-3p
20	extracellular region	5	hsa-miR-892b
20	cell cycle	5	hsa-miR-216b
11	regulation of apoptosis	3	hsa-miR-27a

Table 2 Comparison between drug list of CoMi activity-based method and drug list of CMAP method

Drug list of CoMi activity-based method			**Drug list of CMAP method**		
Rank	Drug	KS Score	Rank	Drug	KS Score
1	mercaptopurine	0.9417	1	decitabine	0.6893
2	***mitoxantrone***	***0.6214***	2	lomustine	0.4587
3	***vinblastine***	***0.5825***	***3***	***tamoxifen***	***0.4397***
4	daunorubicin	0.5073	4	procarbazine	0.4369
5	doxorubicin	0.4563	5	chlorambucil	0.4223
6	lomustine	0.4029	***6***	***mitoxantrone***	***0.3883***
7	***tamoxifen***	***0.3329***	***7***	***paclitaxel***	***0.3576***
8	***paclitaxel***	***0.2427***	8	etoposide	0.2646
9	sirolimus	-0.2058	9	sirolimus	0.1534
10	azacitidine	-0.2524	10	daunorubicin	-0.3811
11	methotrexate	-0.2755	11	tetrandrine	-0.4393
12	etoposide	-0.3107	12	methotrexate	-0.4660
13	hycanthone	-0.3131	***13***	***vinblastine***	***-0.4919***
14	tetrandrine	-0.3204	14	hycanthone	-0.4951
15	chlorambucil	-0.3981	15	doxorubicin	-0.5146
16	procarbazine	-0.5696	16	azacitidine	-0.5728
17	decitabine	-0.8835	17	mercaptopurine	-0.9417

978-1-4577-1666-9/11/$26.00 ©2011 IEEE