

Research Article

Integrative Analysis with Monte Carlo Cross-Validation Reveals miRNAs Regulating Pathways Cross-Talk in Aggressive Breast Cancer

**Antonio Colaprico,^{1,2} Claudia Cava,³ Gloria Bertoli,³
Gianluca Bontempi,^{1,2} and Isabella Castiglioni³**

¹Interuniversity Institute of Bioinformatics in Brussels (IB)², 1050 Brussels, Belgium

²Machine Learning Group, ULB, 1050 Brussels, Belgium

³IBFM-CNR, Segrate, 20090 Milan, Italy

Correspondence should be addressed to Antonio Colaprico; acolapri@ulb.ac.be

Received 20 March 2015; Revised 31 May 2015; Accepted 8 June 2015

Academic Editor: Charles Plessy

Copyright © 2015 Antonio Colaprico et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this work an integrated approach was used to identify functional miRNAs regulating gene pathway cross-talk in breast cancer (BC). We first integrated gene expression profiles and biological pathway information to explore the underlying associations between genes differently expressed among normal and BC samples and pathways enriched from these genes. For each pair of pathways, a score was derived from the distribution of gene expression levels by quantifying their pathway cross-talk. Random forest classification allowed the identification of pairs of pathways with high cross-talk. We assessed miRNAs regulating the identified gene pathways by a mutual information analysis. A Fisher test was applied to demonstrate their significance in the regulated pathways. Our results suggest interesting networks of pathways that could be key regulatory of target genes in BC, including stem cell pluripotency, coagulation, and hypoxia pathways and miRNAs that control these networks could be potential biomarkers for diagnostic, prognostic, and therapeutic development in BC. This work shows that standard methods of predicting normal and tumor classes such as differentially expressed miRNAs or transcription factors could lose intrinsic features; instead our approach revealed the responsible molecules of the disease.

1. Introduction

In recent years, gene expression experiments identified an increasing number of disease biomarkers, for example, [1–4]. In Breast Cancer (BC) different gene signatures have been identified [5–10], some of them with similar role (e.g., BC grade classification), but their reproducibility and overlap are poor. Gene signatures of many gene-based classification methods are often generated by genes selected independently, even though their functional products may interact with each other, and the selected gene markers may contain redundant information. In this way, they may not synergistically improve the overall classification performance. In order to overcome these limitations, one of the key challenges of the postgenomic era is to understand the complex interaction

among genes, thus going a step forward the elucidation of essential principles of cellular systems and disease machinery [11, 12]. In addition, in order to obtain a correct interpretation of high-throughput genomic experiments, the identification of signaling and metabolic pathways involved in a given phenotype is a crucial step. Several studies [13, 14] have shown that pathway-based classifiers are more reproducible and often achieve comparable or better classification performance than classifier based on independent gene biomarkers [15].

Most methods, currently available, consider the pathways as independent mechanisms, and they do not treat the relation between pathways, which is referred to as cross-talk. A cross-talk among gene pathways can be mean of regulatory interaction among different pathways or can express the gene overlap among pathways.

In a normal condition many of the cellular signaling pathways are interconnected to maintain homeostasis [16]. However, the development of cancer might affect the pathway interaction and comprehensively alter the phenotype of a cell. Therefore, the interaction among pathways is a crucial step for understanding the regulatory mechanisms and the synergistic effects on certain biological processes.

Cancer develops through a complex multistep process, involving both genetics and epigenetic mechanisms. Epigenetic alterations include DNA methylation, histone modifications, and small noncoding RNA molecules, including microRNAs (miRNAs). They are involved in the modulation of gene expression on a large scale [17–20]. In recent years, miRNAs have been reported to control many biological processes, such as development, differentiation, growth, and even cancer development and progression [21, 22]. Therefore, to investigate and integrate cross-talk pathways that comprise both genes and miRNAs have become crucial. Recent studies have shown that miRNAs can mediate the cross-talk between pathways. Butz et al. [23] reported cross-talk between TGF- β signaling and the miRNA machinery to suggest potential novel therapeutic targets. Wang et al. [24] provided an overview description of the cross-talk between Notch signaling with many pathways and evidenced where miRNAs appear to play a major role. Zhao and Carrasco [25] discussed the appealing possibility of a functional link between miRNA30a/b/c/d/e-5p and Wnt/ β -catenin pathway for multiple myeloma therapy.

To our knowledge, there are few studies that examine the role of miRNAs for cross-talked pathway able to correctly discriminate normal versus BC samples. Recently, in myasthenia gravis Cao et al. [26] calculated the cross-talk between pathways, identified by pathway-enriched analysis based on a cumulative hypergeometric distribution. They obtained key genes regulated by miRNAs, and these miRNAs were found to mediate cross-talk.

Many pathways cross-talk was demonstrated in cancer, but few studies were focused on the regulatory role of miRNAs. For instance, Notch signaling pathway showed a cross-talk with multiple oncogenic signaling pathways, such as NF- κ B, Akt, Sonic hedgehog (Shh), mTOR, Ras, Wnt, estrogen receptor (ER), androgen receptor (AR), epidermal growth factor receptor (EGFR), and platelet-derived growth factor (PDGF), maybe playing critical roles in tumor aggressiveness [27]. Recently, Hiemer et al. [28] revealed novel cross-talk between the TGF- β pathway and TAZ/YAP in late-stage BC. Han et al. [29] presented an evidence of cross-talk between TGF- β signaling and Notch pathway through a basic helix-loop-helix transcription factor, HEYL.

Several studies suggested the potential role of pathway cross-talk in therapeutic approaches [30, 31]. Leehy et al. [30] showed that Aurora A kinase and progesterone receptor (PR) cross-talk may drive early BC progression in response to growth promoting signals. Aurora kinase inhibitors and antiprogestins, administered with conventional estrogen-blocking therapies, may increase survival outcomes by preventing progression to endocrine failure. Recent studies focused on understanding the molecular biology associated with acquired endocrine resistance, like cross-talk between

ER and peptide growth factor receptor pathways, such as EGFR and human epidermal growth factor receptor 2 (HER2) [31]. Future therapeutic approaches may combine endocrine therapy with inhibitors of growth factor receptors or downstream signaling pathways to treat or avoid critical resistance pathways that become active in ER+ tumors.

Although it is intuitive that different pathways could influence each other, the presence and amount of this scenario have not been completely studied and, most importantly, there is no currently available technique able to quantify the amount of such cross-talk for pairs of pathway [32].

Donato et al. [32] demonstrated, using Jaccard similarity index and Pearson correlation for pairs of pathways, that the cross-talk among two pathways can be explained by the presence of genes that are involved in more than one pathway. However, they focused on cross-talk due to the effect that pathways exercise on each other for the presence of overlapping genes. They did not explore cross-talk as regulatory interaction among different pathways. Lee et al. [33] proposed a pathway activity score, summarizing the activity level of each pathway from the gene-expression level of its condition-responsive genes (CORGs). CORGs were defined as the subset of genes in the pathway whose combined expression delivers optimal discriminative power for the disease phenotype. Yang et al. [34] proposed an average rank-based score capturing the relative expression levels of genes in a pathway and used Spearman test to evaluate similarity between pathways. Furthermore, they showed that both correlation of activity profiles between pathways and correlations of expression profiles for genes from the same pathways are reduced in tumors with respect to normal samples or cell lines. Guo et al. [35] used the mean or median expression value of the member genes (that belong to the same pathway) as the activity level of a given pathway. Su et al. [15] proposed a probabilistic pathway activity inference method that uses the log-likelihood ratio between different phenotypes based on the expression level of each member gene. Cava et al. [36] presented some distance measures using pathway information among BC patients showing that, although not improving the classification of BC, their possible use in the pathway cross-talk expression could be beneficial for more reproducible and biologically valid gene signatures in BC. McCormack et al. [37] proposed CrossTalkZ, a statistical method to identify the significance of cross-talk enrichment between pairs of gene or protein groups in large biological networks.

In this work our goal was to develop an approach that could be able to quantify the cross-talk between pathways and to identify key miRNAs regulating pathway cross-talk. Figure 1 shows our proposed methodology. We integrated the information of differentially expressed genes (DEGs) in a Differential Expression Analysis (DEA) (1st step) between BC and normal samples (NS) with the pathway information (2nd step) in a Pathway Enrichment Analysis (PEA). We thus found a subset of pathways enriched from DEGs and, starting from a matrix containing mean of genes for each pathway (3rd step), we created a matrix score for each pair of pathways (4th step). These procedures were tested through a Monte

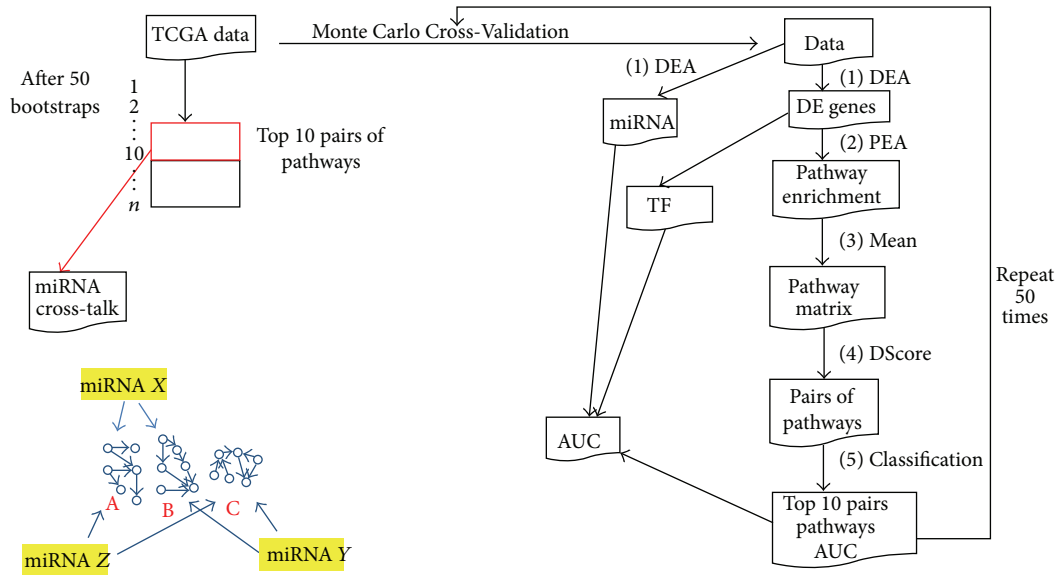


FIGURE 1: Workflow of the proposed methodology.

Carlo Cross-Validation (50 bootstraps). A training dataset was used for each bootstrap in order to find a matrix score for the top 10 pairs of pathways with the best AUC value between NS versus BC samples (Random Forest classification). Finally, we considered the top 10 pairs of pathways ranked for all 50 bootstraps (5th step) as those most frequently found. We then focused on miRNAs targeting the higher number of genes within each pathway pairs (6th step). We found an interesting subset of miRNAs that can be involved in the cross-talk regulation among different pathways in aggressive BC. We assessed pairs of pathways that could be key regulatory of BC and miRNAs that control these pairs of pathways as potential biomarkers for the diagnosis of BC.

The paper is organized as follows. Section 2 covers material and methods by detailing the proposed approach and defining the adopted evaluation procedure. Sections 3 and 4 report and discuss the results obtained by applying the proposed approach on two different datasets. Section 5 draws our conclusions.

2. Materials and Methods

2.1. Datasets. We applied our approach on two BC datasets obtained from the Cancer Genome Atlas (TCGA) and from the Gene Expression Omnibus (GEO): GSE39004. TCGA is a comprehensive and coordinated effort to accelerate our understanding of the molecular basis of cancer by providing a huge amount of genomic data obtained with the ultimate genome analysis technologies. The data are organized in four levels: (1) raw, (2) processed, (3) interpreted, and (4) region of interest.

In our approach, we adopted the level 3. We used the package TCGAbiolinks [38] to create two Expression Matrices, after downloading files of miRNAs and mRNAs from TCGA. The BC dataset contains: (1) the expression levels of 1046 miRNAs from 720 tumor samples and 87 NS,

obtained with Illumina Genome Analyzer miRNA Sequencing, which reports the calculated expression for each miRNA sequence isoform observed and (2) the expression level of 15243 genes (after quantile analysis) and 1052 tumor samples and 113 NS obtained with IlluminaHiSeq RNASeqV2, which reports the normalized results for the expression of a gene.

In our analysis, we used 142 BC and 113 NS from BC RNAseqV2, in order to obtain primary BC patients with stage greater than 3 and matched samples between mRNA and miRNA and 142 BC and 87 NS from BC illuminahiSeq mirnaseq.

In order to avoid cohort-specific biases, we used also a BC dataset from Gene Expression Omnibus (GEO): GSE 39004. It contains 108 BC samples: 61 samples of macrodissected tumor tissue and 47 adjacent noncancerous tissues. The dataset comes from the Affymetrix Gene Chip Human Gene 1.0 ST Arrays platform.

2.2. The Proposed Approach. Our goal was to develop an approach to quantify the cross-talk between pathways and to identify key miRNAs regulating pathway cross-talk. Figure 1 shows the proposed methodology. In Monte Carlo Cross-Validation approach all steps were repeated for 50 bootstraps. For each bootstrap in our approach we obtained (1) DEA, (2) PEA, (3) Mean, (4) DScore, and (5) Classification. After 50 bootstraps we focused on top 10 pairs of pathways with the best AUC and their miRNAs. Our approach was compared using the genomic biomarkers selected by the Monte Carlo Cross-Validation: transcription factors (TF) and miRNAs from DEA.

Step 1: Differential Expression Analysis. We applied the DEA on those mRNA transcripts and miRNA, which had mean across the 142 samples, higher than the $0.25 * \text{quantile mean across all samples}$. To determine whether a gene or a miRNA is expressed in a differential way, we applied a test

of hypothesis and the fold-change between the two starting conditions, in aggressive BC and normal conditions.

In particular, we used the edgeR package from Bioconductor that uses the quantile-adjusted conditional maximum likelihood (qCML) method for experiments with single factor to determine genes differentially expressed [39]. Compared against several other estimators, qCML is the most reliable in terms of bias on a wide range of conditions and specifically performs best in the situation of many small samples with a common dispersion. The p values, generated from the analysis sorted in ascending order, are corrected using the Benjamini-Hochberg procedure for multiple testing correction [40]. We considered significant DEGs or miRNA differentially expressed between BC and NS if log fold change (FC) > 1 and FDR < 0.01.

Step 2: Pathway Enrichment Analysis. In order to identify a group of pathways significantly enriched by DEGs in BC with respect to NS, we used a Pathway Enrichment Analysis from DEGs (PEA-DEGs). In particular, we focused on 589 biological pathways derived from the IPA (Ingenuity Pathways Analysis) tool (<http://www.ingenuity.com/>).

The enrichment was evaluated using Fisher's exact test. The aim was to place DEGs within a regulatory context (IPA pathways) and identify the pathways responsible for coordinating their activity, thus highlighting the regulatory apparatus driving phenotypic differentiation. A Fisher's test was applied between DEGs and genes of IPA pathways and we thus obtained pathway enriched with p value < 0.01. p values were adjusted using the Benjamini-Hochberg procedure for multiple testing correction [40]. IPA pathways were filtered considering only the genes obtained from the quantile function.

Steps 3-4: Discriminating Score for Pathway Cross-Talk. We computed a discriminating score (DS) by comparing the gene expression levels of each pair of IPA pathways enriched from DEGs, in each sample (e.g., we applied a DS(x, y) in each sample for the pair of pathways x and y). DS was defined as

$$DS = \frac{|(M_x - M_y)|}{S_x + S_y}, \quad (1)$$

where M_x and S_x represent mean and standard deviation of expression levels of genes in a pathway x and M_y and S_y in a pathway y . DS score indicates the relationships between pairs of pathway, with a larger value indicating relatively higher difference of activity between pathways. The considered DSs were already used in previous studies by Golub et al. [41] and Orsetti et al. [42] for the comparison of expression levels between the subgroup of samples presenting amplification and the subgroup of samples without amplification. We used the score for the first time at our knowledge for pathway cross-talk analysis.

We compared DS with the method proposed by Cava et al. [36], which uses the Euclidean distance as metric to quantify pathway cross-talk.

Step 5: Selection of the Best Pairs of Pathways. In order to evaluate the performance of the proposed methodology, we developed a Random Forest (RF) classification model using the R-package [43]. The model was used to classify the considered BC versus NS. AUC was estimated by cross-validation method (k -fold cross-validation, $k = 10$). We adopted the following parameters: mtry (number of variables randomly sampled as candidates at each split) = \sqrt{p} , p being the number of variables in the matrix of data; ntree (number of trees grown) = 500. Classification was applied on pairs of pathways using DS for each sample.

We implemented a Monte Carlo Cross-Validation method. It randomly selected some fraction of TCGA data (60% of original dataset) to form the training set and then assigned the rest of the points to the testing set (40% of original dataset). This process was then repeated multiple times (50 bootstraps), generating (at random) new training and test partitions each time. For each bootstrap we analyzed DEGs, pathways PEA-DEGs, and a matrix score for pairs of pathways. Figure 1 summarizes the procedure. Each bootstrap gives pairs of pathway significantly enriched from DEGs and from a matrix score (DS); RF classifier establishes an AUC value. We thus considered the top 10 pairs of pathway for each bootstrap that obtained the best classification performance in the training dataset. For each bootstrap a testing dataset was then used to validate the top 10 pairs of pathways. At the end of all 50 bootstraps (runs) we selected a list of the top 10 pairs of pathways ordering according a decreasing frequency that each pair of pathway was selected in the 50 runs.

We ordered each pathway with respect to their AUC and we got the first 10 pathways with the best 10 AUC. Specifically, we considered for further analysis top 10 pairs with better ranked AUC value for all 50 bootstraps. Figure 1 shows how the selection of the top 10 pairs of pathways is performed for each bootstrap and the final result after 50 bootstraps in the training dataset.

Step 6: miRNA Regulating Pathway Cross-Talk. Consider the following.

(1) miRNA Regulon Estimation. Network inference, which is the reconstruction of biological networks from high-throughput data, can provide valuable information about the regulation of gene expression in cells. Several methods have been proposed in literature [44], such as TIGRESS and Lasso based on linear regression, Aracne and CLR based on mutual information, correlation, and Bayesian networks. We adopted mutual information as it is particularly effective for large datasets [20, 45, 46]. The mutual information provides an index of dependence between miRNAs and genes.

For each gene belonging to the top 10 pairs of pathways obtained for all 50 bootstraps, we analyzed their miRNAs. We calculated mutual information between dataset of miRNAs and genes, thus creating an index of dependence between them. Mutual information was calculated using entropy estimates from K -nearest neighbor distances [47] with the R-package: Parmigene [48]. Mutual information was applied in previous studies [20, 45, 46].

(2) *miRNA Master Regulator Analysis*. Master Regulator Analysis (MRA) [49] is an algorithm used to identify transcription factors whose targets (e.g., as represented in an ARACNe-generated interactome) are enriched for a particular gene signature (e.g., a list of differentially expressed genes). The enrichment is evaluated using a statistical test such as Fisher exact test or Gene Set Enrichment Analysis (GSEA) [50]. The objective is to place the signature genes within a regulatory context and to identify the master regulators responsible for coordinating their activity, thus highlighting the regulatory apparatus driving phenotypic differentiation. We identified miRNA master regulators (MRs) of pair of pathways. A miRNA was defined as MR, when its targets are enriched for a particular gene signature, such as genes involved in one of the two coupled pathways at the same time. Namely, we modified MRA instead of TFs, we used miRNAs, and instead of DEG we used genes annotated in a pathway. We used a Fisher's exact test to identify miRNAs significantly enriched by their target genes from the top 10 pathways as obtained for all 50 bootstraps. We found miRNAs that have a significant target genes in the pair of pathways with both p values < 0.01 . p values were adjusted using the Benjamini-Hochberg procedure for multiple testing correction [40]. Thus, we then focused on miRNAs found differentially expressed in Section 2.2.

Step 7: Comparison with Other Approaches. The input of the classifier for our approach was applied on

- (i) DS obtained from each couple of pathways for each sample.

Our approach was compared using the genomic biomarkers selected by the Monte Carlo Cross-Validation. The inputs of the classifier for the comparison were

- (i) the expression levels of TFs in DEGs as obtained from DEA in Section 2.2.
The resource of transcription factors in human is obtained from TRANSFAC database [51],
- (ii) the expression levels of miRNAs as obtained from DEA in Section 2.2.

3. Results

3.1. Steps 1-2. From quantile analysis, we obtained 15243 genes and 764 miRNAs. From the DEA between aggressive BC and NS for all samples BC TCGA we obtained 3225 DEGs and 254 differentially expressed miRNAs, and a list of 48 significantly pathways enriched from DEGs in total represented a list of 2214 unique genes. Table 1 shows pathways significantly enriched by DEGs with their FDR score, number of genes for each pathway, and number of common genes between DEGs and genes in the pathway.

3.2. Steps 3-4. Figure 2 shows AUC classification in the two different approaches. DS obtained a better performance than Euclidean distance. Median values are around 0.65 with euclidean distance and 0.7 with DS.

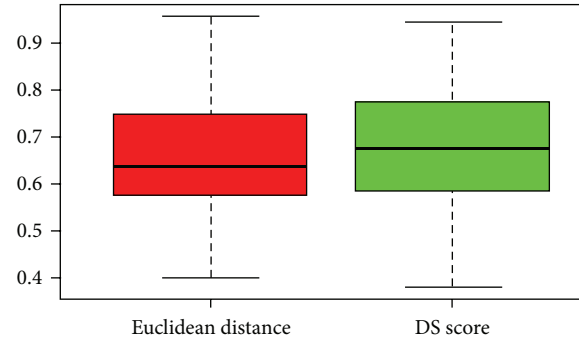


FIGURE 2: Boxplot of AUC with two different approaches: red (euclidean distance) and DS (green) in the TCGA dataset.

3.3. Step 5. Figure 3 shows the results for each bootstrap. We obtained a heatmap where blue square indicates pairs of pathways in the top 10 positions for classification in the training dataset for that bootstrap. We found (1) the pair of pathways Acute Phase Response Signaling and HIF1 Signaling in 26 bootstraps; (2) the pair of pathways HIF1 Signaling and Fatty Acid oxidation in 22 bootstraps; (3) the Ethanol Degradation IV and Estrogen Receptor Signaling in 21 bootstraps; and (4) the HIF1 Signaling and Tryptophan Degradation X (Mammalian, via Tryptamine) in 21 bootstraps.

We ordered each pathway with respect to their AUC and we found 32 paired pathways with AUROC > 0.90 between BC and NS. We focused on pathways with the best 10 AUC. Table 2 presents the top 10 pairs of pathways that have the best performance for classification of aggressive BC and NS for all 50 bootstraps. It showed AUC value for TCGA and GSE39004 dataset. The pair of pathways, Acute Phase Response Signaling and HIF1 Signaling, with its DScore that was in the top 10 in 26 bootstraps, obtained an AUC value 0.92 in TCGA dataset and 0.889 in GSE39004 dataset, showing a good balance of performance between the two datasets. The pair of pathways Bladder Cancer Signaling and Fatty Acid oxidation showed also a good performance with 0.932 in TCGA dataset and 0.84 in GSE39004 dataset; similar performance was found in the pair of pathways Bladder Cancer Signaling and Tryptophan Degradation X (Mammalian, via Tryptamine) with 0.91 in TCGA dataset and 0.802 in GSE39004 dataset. In TCGA dataset we obtained a mean AUC value 0.92 while in GSE39004 0.70. Figure 4 shows overall AUC performances when we give only top 10 pairs of pathways for each bootstrap after 50 bootstraps. Results for AUC values are good; 25th percentile does not fall down 0.95.

3.4. Step 6: miRNA Regulating Pathway Cross-Talk. We identified 12 miRNAs that may have an important role in the regulation of the four pairs of pathways.

Table 3 reports for each pair of pathways their potential miRNAs together with the log fold change (FC) obtained from the DEA, the miRNA expression levels in BC and NS, and the delta value, an index to quantify the difference

TABLE 1: Pathways enriched from differentially expressed genes.

Pathway	FDR	Number of genes in pathway	Number of common genes
Agranulocyte Adhesion and Diapedesis	$FDR = 6.38e - 10$	(ng = 173)	(nc = 62)
Atherosclerosis Signaling	$FDR = 6.25e - 08$	(ng = 119)	(nc = 44)
Granulocyte Adhesion and Diapedesis	$FDR = 4.26e - 07$	(ng = 163)	(nc = 53)
LXR/RXR Activation	$FDR = 6.58e - 07$	(ng = 121)	(nc = 42)
Inhibition of Matrix Metalloproteases	$FDR = 9.55e - 06$	(ng = 38)	(nc = 19)
Hepatic Fibrosis/Hepatic Stellate Cell Activation	$FDR = 1.20e - 05$	(ng = 137)	(nc = 44)
Acute Phase Response Signaling	$FDR = 2.38e - 05$	(ng = 167)	(nc = 49)
ILK Signaling	$FDR = 4.95e - 05$	(ng = 181)	(nc = 52)
Ethanol Degradation II	$FDR = 7.41e - 05$	(ng = 30)	(nc = 15)
Intrinsic Prothrombin Activation Pathway	$FDR = 1.26e - 04$	(ng = 28)	(nc = 14)
Coagulation System	$FDR = 1.26e - 04$	(ng = 35)	(nc = 16)
Extrinsic Prothrombin Activation Pathway	$FDR = 1.26e - 04$	(ng = 16)	(nc = 10)
Noradrenaline and Adrenaline Degradation	$FDR = 1.36e - 04$	(ng = 32)	(nc = 15)
Bladder Cancer Signaling	$FDR = 1.36e - 04$	(ng = 86)	(nc = 29)
Axonal Guidance Signaling	$FDR = 1.44e - 04$	(ng = 421)	(nc = 98)
Ethanol Degradation IV	$FDR = 1.96e - 04$	(ng = 17)	(nc = 10)
Colorectal Cancer Metastasis Signaling	$FDR = 2.46e - 04$	(ng = 232)	(nc = 60)
Wnt/-catenin Signaling	$FDR = 2.89e - 04$	(ng = 167)	(nc = 46)
Cardiac- adrenergic Signaling	$FDR = 4.24e - 04$	(ng = 132)	(nc = 38)
Protein Kinase A Signaling	$FDR = 6.29e - 04$	(ng = 365)	(nc = 84)
LPS/IL-1 Mediated Inhibition of RXR Function	$FDR = 7.24e - 04$	(ng = 210)	(nc = 52)
EIF2 Signaling	$FDR = 1.20e - 03$	(ng = 171)	(nc = 11)
Epithelial Adherens Junction Signaling	$FDR = 1.47e - 03$	(ng = 144)	(nc = 38)
HIF1.Signaling	$FDR = 1.67e - 03$	(ng = 100)	(nc = 29)
Oxidative Ethanol Degradation III	$FDR = 1.67e - 03$	(ng = 15)	(nc = 8)
Retinoate Biosynthesis I	$FDR = 1.67e - 03$	(ng = 29)	(nc = 12)
Factors Promoting Cardiogenesis in Vertebrates	$FDR = 1.74e - 03$	(ng = 87)	(nc = 26)
GADD45 Signaling	$FDR = 2.15e - 03$	(ng = 19)	(nc = 9)
Mitotic Roles of Polo-Like Kinase	$FDR = 2.15e - 03$	(ng = 62)	(nc = 20)
TR/RXR Activation	$FDR = 2.17e - 03$	(ng = 85)	(nc = 25)
Aryl Hydrocarbon Receptor Signaling	$FDR = 2.61e - 03$	(ng = 133)	(nc = 35)
Role of Osteoblasts, Osteoclasts ...	$FDR = 3.18e - 03$	(ng = 214)	(nc = 50)
Estrogen-mediated S-phase Entry	$FDR = 3.30e - 03$	(ng = 24)	(nc = 10)
Calcium Signaling	$FDR = 3.64e - 03$	(ng = 168)	(nc = 41)
Estrogen Receptor Signaling	$FDR = 3.73e - 03$	(ng = 126)	(nc = 8)
Triacylglycerol Biosynthesis	$FDR = 4.27e - 03$	(ng = 33)	(nc = 12)
Mitochondrial Dysfunction	$FDR = 4.27e - 03$	(ng = 139)	(nc = 10)
Thyroid Cancer Signaling	$FDR = 4.27e - 03$	(ng = 40)	(nc = 14)
Human Embryonic Stem Cell Pluripotency	$FDR = 4.61e - 03$	(ng = 132)	(nc = 34)
FXR/RXR Activation	$FDR = 5.24e - 03$	(ng = 82)	(nc = 22)
Glycogen Degradation II	$FDR = 8.40e - 03$	(ng = 9)	(nc = 5)
Putrescine Degradation III	$FDR = 8.49e - 03$	(ng = 16)	(nc = 7)
Tryptophan Degradation X	$FDR = 8.49e - 03$	(ng = 16)	(nc = 7)
Dopamine Degradation	$FDR = 9.05e - 03$	(ng = 21)	(nc = 8)
Glioma Invasiveness Signaling	$FDR = 9.89e - 03$	(ng = 57)	(nc = 17)
Actin Cytoskeleton Signaling	$FDR = 9.89e - 03$	(ng = 210)	(nc = 47)
NAD biosynthesis II (from tryptophan)	$FDR = 9.89e - 03$	(ng = 13)	(nc = 6)
Fatty Acid-oxidation	$FDR = 9.89e - 03$	(ng = 13)	(nc = 6)

BRCA heatmap top 10 pairs of pathways with 50 bootstraps with training 60% sample

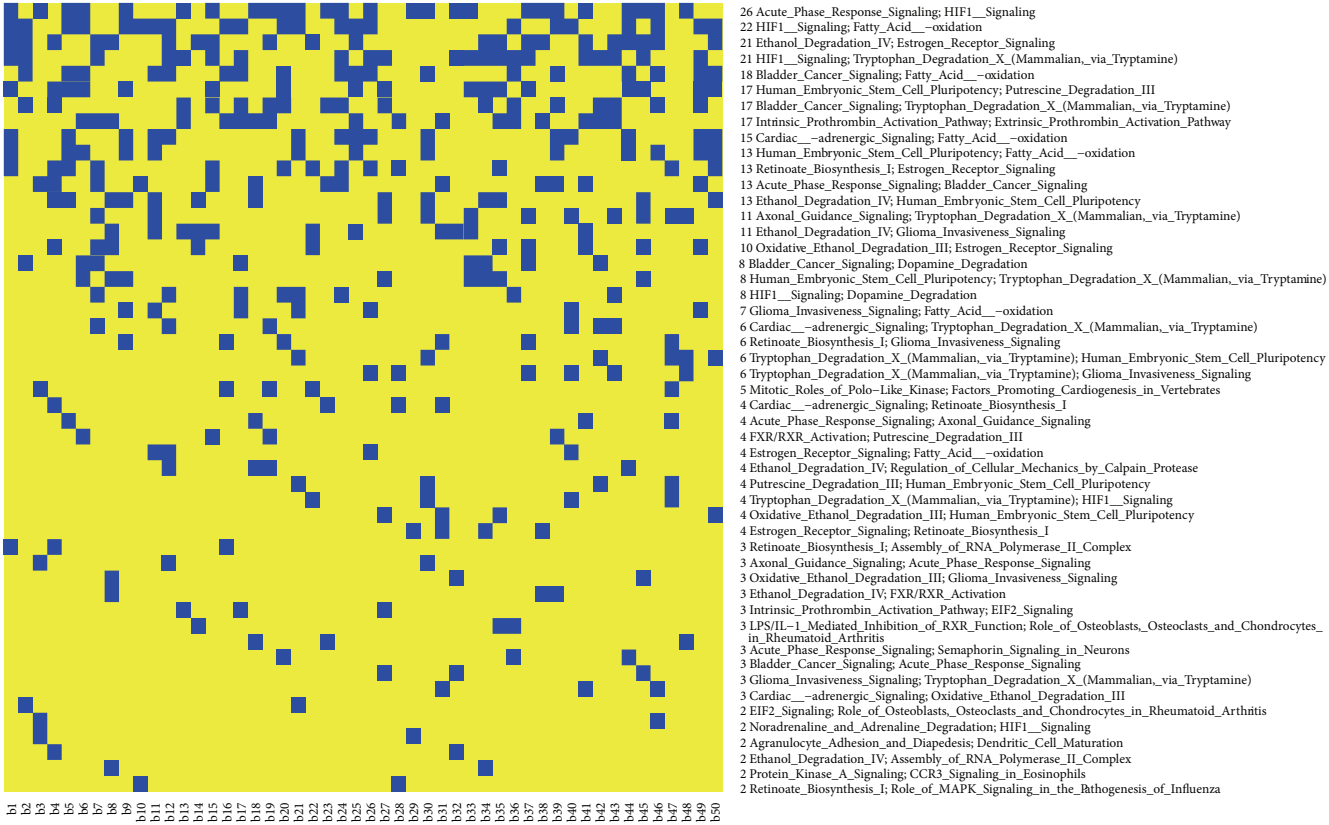


FIGURE 3: Heatmap with top 10 pairs of pathways (blue square) that obtained the best AUC values in the training dataset.

TABLE 2: Top 10 pairs of pathways with AUC value for TCGA testing and GEO dataset.

Pairs of pathways	TCGA AUC	GSE39004 AUC
(1a) Acute Phase Response Signaling		
(1b) HIF1 Signaling	0.92	0.889
(2a) HIF1 Signaling		
(2b) Fatty Acid oxidation	0.955	0.654
(3a) Ethanol Degradation IV		
(3b) Estrogen Receptor Signaling	0.925	0.656
(4a) HIF1 Signaling		
(4b) Tryptophan Degradation X (Mammalian, via Tryptamine)	0.922	0.629
(5a) Bladder Cancer Signaling		
(5b) Fatty Acid oxidation	0.932	0.84
(6a) Human Embryonic Stem Cell Pluripotency		
(6b) Putrescine Degradation III	0.931	0.613
(7a) Bladder Cancer Signaling		
(7b) Tryptophan Degradation X (Mammalian, via Tryptamine)	0.91	0.802
(8a) Intrinsic Prothrombin Activation Pathway		
(8b) Extrinsic Prothrombin Activation Pathway	0.911	0.581
(9a) Cardiac adrenergic Signaling		
(9b) Fatty Acid oxidation	0.916	0.732
(10a) Human Embryonic Stem Cell Pluripotency		
(10b) Fatty Acid oxidation	0.926	0.615

TABLE 3: Differentially expressed miRNA-regulating pathway cross-talk. ex.BC indicates miRNA expression levels in BC and ex.NS indicates miRNA expression levels in NS.

Pairs Pathways	miRNA	logFC	ex.BC	ex.NS	Delta
(1a) Human Embryonic Stem Cell Pluripotency					
(1b) Putrescine Degradation III	<i>hsa-let-7c</i>	-1.507	12801.77	27353.99	21938.10
(2a) Intrinsic Prothrombin Activation Pathway					
(2b) Extrinsic Prothrombin Activation Pathway	<i>hsa-mir-210</i>	3.273	2702.47	317.80	6433.72
(3a) Intrinsic Prothrombin Activation Pathway					
(3b) Extrinsic Prothrombin Activation Pathway	<i>hsa-mir-9-1</i>	1.925	4679.87	1339.12	7806.52
(4a) Intrinsic Prothrombin Activation Pathway					
(4b) Extrinsic Prothrombin Activation Pathway	<i>hsa-mir-483</i>	-1.786	22.80	67.85	28.86
(5a) Intrinsic Prothrombin Activation Pathway					
(5b) Extrinsic Prothrombin Activation Pathway	<i>hsa-mir-592</i>	4.156	16.77	0.7	80.44
(6a) Ethanol Degradation IV					
(6b) Estrogen Receptor Signaling	<i>hsa-mir-103-2</i>	1.134	48.07	19.65	32.24
(7a) Intrinsic Prothrombin Activation Pathway					
(7b) Extrinsic Prothrombin Activation Pathway	<i>hsa-mir-887</i>	1.094	50.56	24.18	66.81
(8a) Acute Phase Response Signaling;					
(8b) HIF1 Signaling	<i>hsa-mir-181b-2</i>	1.432	23.47	8.02	22.14
(9a) Intrinsic Prothrombin Activation Pathway					
(9b) Extrinsic Prothrombin Activation Pathway	<i>hsa-mir-301b</i>	2.608	4.75	0.78	10.36
(10a) Intrinsic Prothrombin Activation Pathway					
(10b) Extrinsic Prothrombin Activation Pathway	<i>hsa-mir-147b</i>	2.312	3.99	0.701	7.61
(11a) Intrinsic Prothrombin Activation Pathway					
(11b) Extrinsic Prothrombin Activation Pathway	<i>hsa-mir-665</i>	-1.447	2.28	5.56	3.88
(12a) Intrinsic Prothrombin Activation Pathway					
(12b) Extrinsic Prothrombin Activation Pathway	<i>hsa-mir-939</i>	1.110	5.93	2.43	4.74

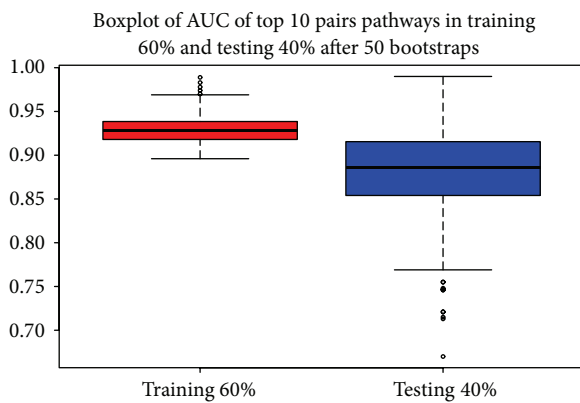


FIGURE 4: Boxplot: overall AUC performance on both TCGA training (red) and testing dataset (blue).

of expression between BC and NS for each miRNA (e.g., $(\text{ex.BC} - \text{ex.NS}) * \log \text{FC}$; difference of expression between BC and NS multiplied log FC). *hsa-let-7c*, *hsa-mir-210*, and *hsa-mir-9-1*, according to the Delta index, showed a higher difference between NS and BC.

Table 4 reports for each pair of pathways their potential regulatory miRNA, the number of genes for each pathway,

and the number of the potential target genes. Table 5 reports for pathways their potential regulatory miRNAs and the more frequent genes target of their miRNAs.

hsa-mir-181b-2 seemed to regulate the pair of pathways that obtained the best AUC values in both TCGA dataset and independent GSE39004 dataset. Figure 5 shows the network context of the two pathways Acute Phase Response Signaling and HIF1 Signaling and the possible role of *hsa-mir-181b-2* using IPA software. Figure 6 shows *hsa-let-7c* and target genes in Human Embryonic Stem Cell Pluripotency, Estrogen-Dependent Breast Cancer Signaling, and Molecular Mechanism of Cancer, according to IPA software.

3.5. Step 7: Comparison with Other Approaches. Figures 7 and 8 show DEGs and differentially expressed miRNAs as obtained from Monte Carlo Cross-Validation for each bootstrap. We obtained a heatmap, where blue square indicates DEGs (Figure 7) and miRNAs (Figure 8) in the top 10 positions for classification for that bootstrap. Figure 9 shows AUC value for top 10 genes that obtained the best classification for all 50 bootstraps in GSE39004. Figure 10 shows a boxplot with AUC value comparison for each approach. All methods have a good performance. GSE39004 has a poorer AUC, consistently with the limited number of samples available for the k -fold validation.

TABLE 4: Differentially expressed miRNA-regulating pathway cross-talk with mutual information: p.a ng indicates the number of genes for pathway a; mirna.p.a indicates the number of miRNA targets for pathway a; p.b ng indicates the number of genes for pathway b; mirna.p.b indicates the number of miRNA targets for pathway b.

Pairs Pathways	miRNA	p.a ng	mirna.p.a	p.b ng	mirna.p.b
(1a) Human Embryonic Stem Cell Pluripotency					
(1b) Putrescine Degradation III	<i>hsa-let-7c</i>	104	40	16	8
(2a) Intrinsic Prothrombin Activation Pathway					
(2b) Extrinsic Prothrombin Activation Pathway	<i>hsa-mir-9-1</i>	19	9	12	4
(3a) Intrinsic Prothrombin Activation Pathway					
(3b) Extrinsic Prothrombin Activation Pathway	<i>hsa-mir-210</i>	19	9	12	8
(4a) Intrinsic Prothrombin Activation Pathway					
(4b) Extrinsic Prothrombin Activation Pathway	<i>hsa-mir-592</i>	19	5	12	4
(5a) Intrinsic Prothrombin Activation Pathway					
(5b) Extrinsic Prothrombin Activation Pathway	<i>hsa-mir-887</i>	19	7	12	6
(6a) Ethanol Degradation IV					
(6b) Estrogen Receptor Signaling	<i>hsa-mir-103-2</i>	17	4	112	15
(7a) Intrinsic Prothrombin Activation Pathway					
(7b) Extrinsic Prothrombin Activation Pathway	<i>hsa-mir-483</i>	19	6	12	5
(8a) Acute Phase Response Signaling					
(8b) HIF1 Signaling	<i>hsa-mir-181b-2</i>	146	26	94	19
(9a) Intrinsic Prothrombin Activation Pathway					
(9b) Extrinsic Prothrombin Activation Pathway	<i>hsa-mir-301b</i>	19	7	12	5
(10a) Intrinsic Prothrombin Activation Pathway					
(10b) Extrinsic Prothrombin Activation Pathway	<i>hsa-mir-147b</i>	19	5	12	5
(11a) Intrinsic Prothrombin Activation Pathway					
(11b) Extrinsic Prothrombin Activation Pathway	<i>hsa-mir-939</i>	19	8	12	5
(12a) Intrinsic Prothrombin Activation Pathway					
(12b) Extrinsic Prothrombin Activation Pathway	<i>hsa-mir-665</i>	19	7	12	4

4. Discussion

Given the substantial difference in the activities of many pairs of pathways between BC and NS, we examined the effectiveness to classify BC and NS based on their pairwise activity profiles. The final purpose of our work was to find, for the best pairs of pathways able to discriminate BC versus NS, their miRNA regulators. Different pathways often act in a coordinated manner to participate in many biological process. We calculated DScore to understand the interrelationships among pathways PEA-DEGs between BC and NS.

We applied a RF classifier, using the DS obtained by DEGs of paired pathways significantly associated, and through Monte Carlo Cross-Validation we identified the top 10 pairs of pathways that obtained the best classification for all the considered bootstraps (50).

We found 32 paired pathways with AUROC > 0.90 between BC and NS, but we focused on pathways with the best 10 AUC. From these pairs, by a mutual information approach, we found 4 pairs of pathways potentially targeted by 12 miRNAs.

DS obtained a slightly more improvement than the Euclidean distance, as measure to quantify the cross-talk.

Although the performances of our approach with pairs of pathways are similar to these of standard investigations (DEGs and miRNAs) based on traditional gene and miRNA expression analysis, some novel aspects are needed to be investigated:

- (1) Our method to select the epigenetic signatures (based on the pairs of pathway combinations) has been implemented for the first time, to our knowledge in BC diagnosis with a robust Monte Carlo Cross-Validation approach.
- (2) Our proposed methodology could be very useful for understanding the interactions between miRNAs and pathways cross-talk. Further studies should be conducted to these purposes.
- (3) Standard approach with miRNA expression analysis using the same dataset does not found the same our miRNAs give a missing information. For instance, *hsa-let-7c* and *hsa-mir-210* that in our approach could be important miRNA regulating pathway cross-talk in the miRNA analysis are found as differentially expressed only in 2 and 1 bootstraps with the best AUC, respectively (Figure 8).

TABLE 5: Principal target genes of miRNA-regulating pathway cross-talk in NS versus BC.

Pairs Pathways	miRNAs	Genes (a)	Genes (b)
(1a) Human Embryonic Stem Cell Pluripotency (1b) Putrescine Degradation III	<i>hsa-let-7c</i>	ACVR1 BMP4	ALDH1A1 ALDH1A21
		BMP5 BMP6	ALDH2 ALDH4A1
		FGF2 FZD7	ALDH9A1 IL41I
		KLK3 LEFTY2	MAOA SMOX
		MRAS NTRK2	
		NTRK3 PDGFRA	
		PDGFRB PIK3CB	
		PIK3CD PIK3R6	
		S1PR1 SPHK1	
		TCF4 TCF7L1	
		TGFB1 TGFB2	
		WNT11 WNT6	
(2a) Ethanol Degradation IV (2b) Estrogen Receptor Signaling	<i>hsa-mir-103-2</i>	ACSS2 ALDH1A1	EP300 GTF2B
		ALDH3A1 ALDH9A1	GTF2H3 MED15
			MEDI7 MED6
			NCOR1 NCOR2
			PCK1 POLR2E
			TAF3 TAF7
(3a) Acute Phase Response Signaling (3b) HIF1 Signaling	<i>hsa-mir-181b-2</i>	ELK1 HNRNPK	ARNT EGLN2
		IKBKB KRAS	EP300 EPO
		MAP2K7 MAPK1	KRAS MAPK1
		MAPK3 MRAS	MAPK3 MMP23B
		NRAS OSM	MRAS NRAS
		PIK3CB PIK3R2	PIK3CB PIK3R2
		PTPN11 SOCS4	SLC2A3 VHL

- (4) A reduced number of miRNAs could be suitable to be translated in a clinical environment, acting on important network of pathways.
- (5) This kind of approach is even more interesting from a biological point of view, as the study of the pairs of pathways involved in BC could help in defining the molecular mechanisms leading to the onset and progression of the pathology, a feature that a single TF or a miRNA does not allow understanding.

4.1. miRNAs Regulating Pathway Cross-Talk in BC with Stage Greater Than 3

hsa-mir-181b-2: Acute Phase Response Signaling and HIF1 Signaling. The pair of pathways Acute Phase Response Signaling and HIF1 Signaling that obtained a good performance in both TCGA dataset and GSE39004 (AUC value 0.92 and 0.88, resp.) seems to be regulated by *hsa-mir-181b-2*. Acute Phase Response Signaling pathway contains 146 genes, hypoxia-inducible factor 1 (HIF-1) signaling contains 94 genes and 25 common genes among them.

We found that *hsa-miR-181b* upregulation is able to control the expression of 26 over 146 genes belonging to Acute Phase Response Signaling and 19 over 94 genes belonging

to HIF1 pathway. Altered expression of *hsa-miR-181b* has been found in various malignancies, especially associated with poorer clinical prognosis [52–54]. Several publications linked the inflammatory process with hypoxia pathway [55, 56], but none has demonstrated the link between the two pathways and *hsa-miR-181* expression (Figure 5). In our analysis *hsa-miR-181* has been proposed as a novel marker for inflammatory response, as its upregulation is strongly correlated with the expression of interleukin (IL)-1 β , IL-6, and tumor necrosis factor alpha [57]. Recently, it seems that the production of IL-1b induces the expression and secretion of the stem cell factor, a growth factor involved in the control of the proliferation of epithelial BC cells [58]. It is thus possible that the *hsa-miR-181* expression, modulating IL-1b levels, could influence the proliferation of BC, by increasing the production of the stem cell factor.

The acute phase response is a rapid inflammatory response that provides protection against infections [59]. The presence of this pathway in our analysis complies with the generally accepted observation that inflammation is often observed in tumors and appears to play a dominant function in the pathogenesis of various cancer types [59]. Oxygen homeostasis underlies many developmental and physiological processes. Important consequences of rapid tumor growth include poor vascularization and insufficient oxygen delivery

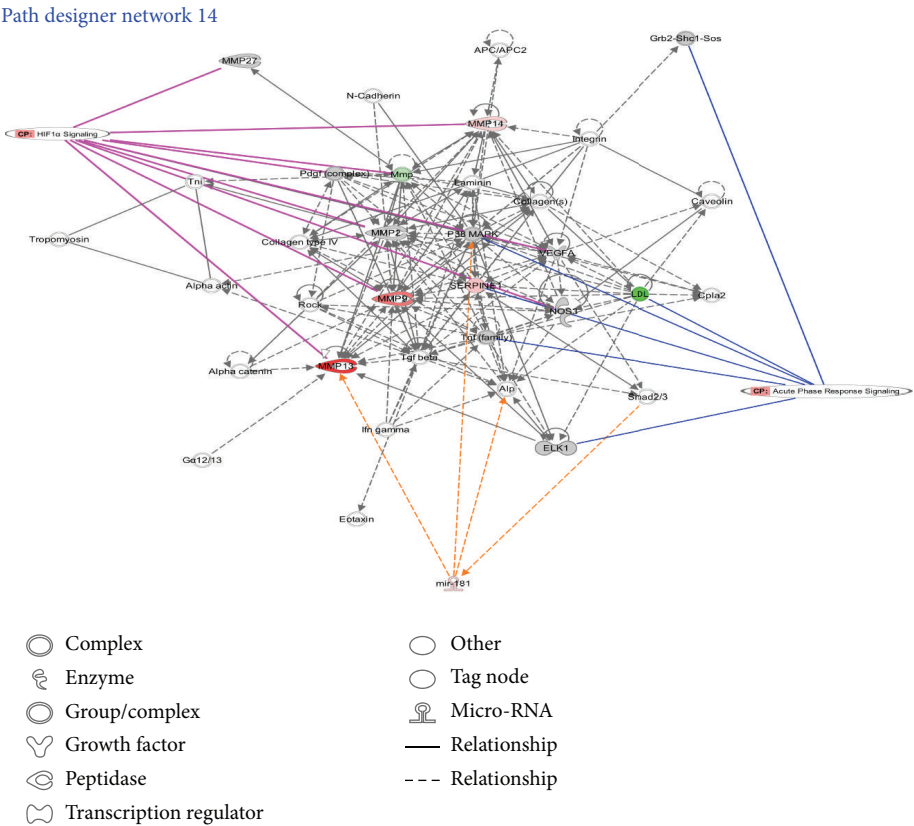


FIGURE 5: Examples of IPA software outcome, the pair of pathways Acute Phase Response Signaling and HIF1 Signaling, are regulated by *hsa-mir-181*.

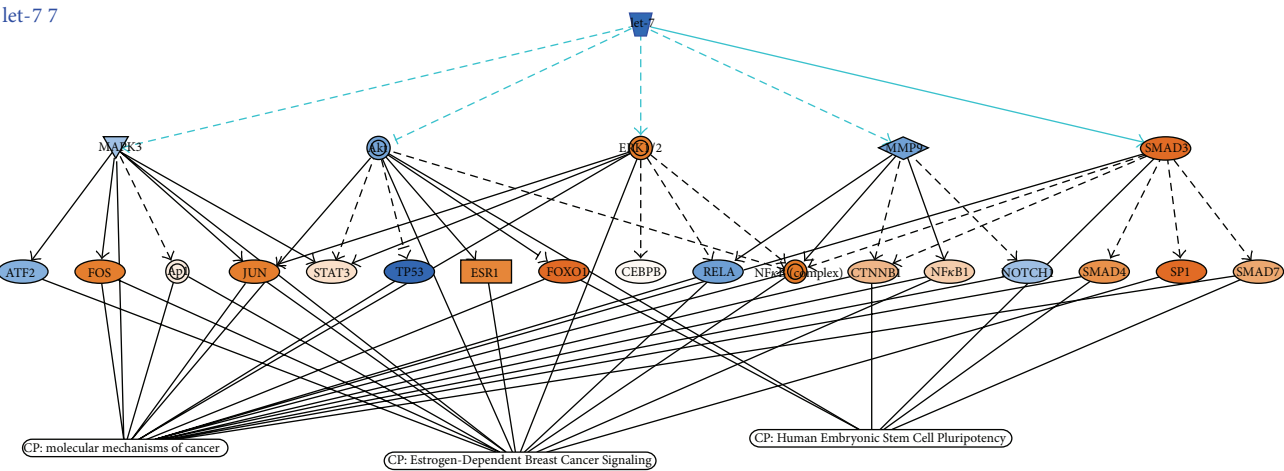


FIGURE 6: Examples of IPA software outcome: *hsa-let-7* and their target genes in Human Embryonic Stem Cell Pluripotency Pathway.

involving formation of hypoxic (poorly oxygenated) areas [60, 61]. Adaptation to hypoxia is facilitated by the activation of transcriptional machinery, in which hypoxia inducible factor (HIF) plays a principal role in coordinating angiogenesis. Several studies found cross-talk between HIF-1 signaling and inflammatory pathways suggesting that the development of inflammation in response to hypoxia is clinically relevant

[62]: HIF-1 plays a crucial role in hypoxic T-cell and neutrophil survival, an important determinant of tissue inflammation [63]; Zampell et al. [64] found that HIF-1 α inhibition by small molecule inhibitors (YC-1 and 2-methoxyestradiol) results in delayed lymphatic repair, decreased local vascular endothelial growth factor-C (VEGF-C) expression, reduced numbers of VEGF-C+ cells, and reductions in inflammatory

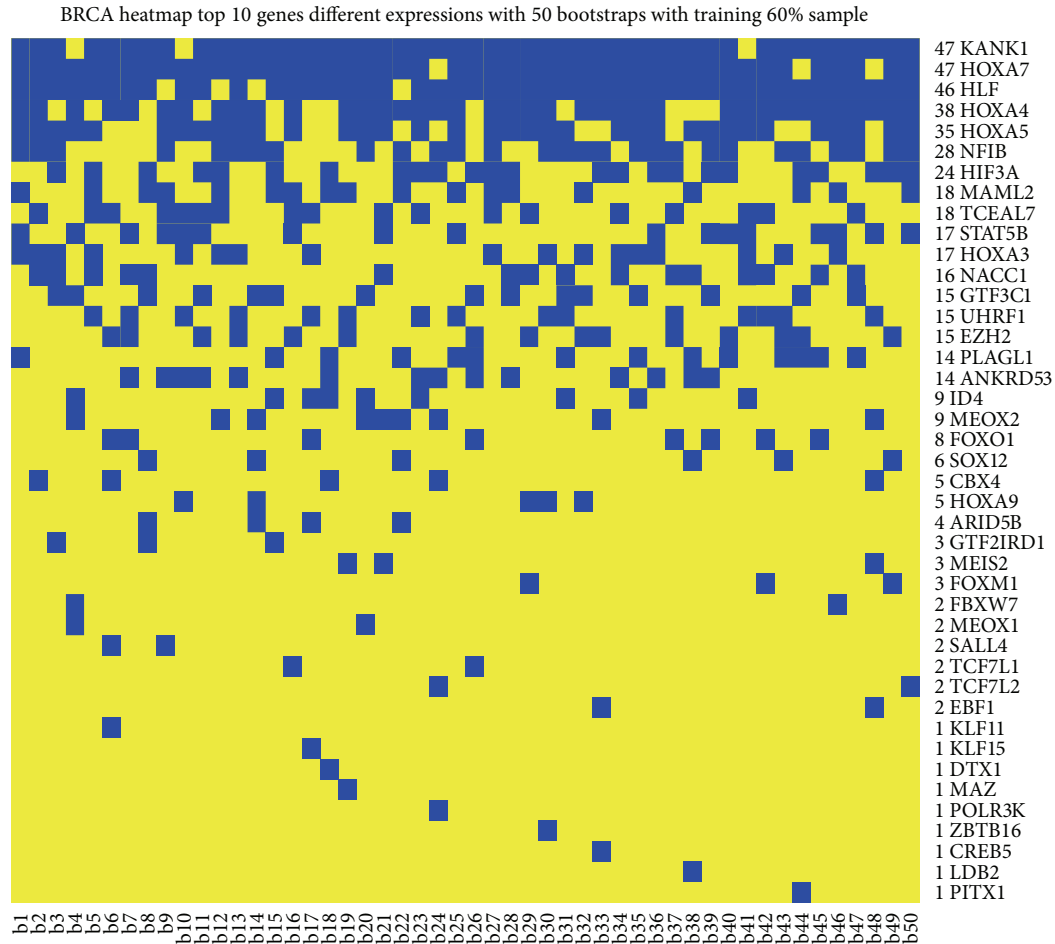


FIGURE 7: Heatmap with top 10 genes (blue square) that obtained the best AUC values in the training dataset.

lymphangiogenesis. Cramer et al. [65] have shown that HIF-1 overexpression is essential for myeloid cell-mediated inflammation. A recent review [62] discussed the regulation of immune responses by hypoxia-induced signaling and the emergent molecular aspects between hypoxia and inflammation in certain cancers.

As shown in Table 5, among genes regulated by *hsa-miR-181*, in both pathways members of the RAS family are present. RAS family members are already described to be involved in BC development [66]. It is relevant that HIF1 activity could influence the activation of RAS/MAPK/ERK1 [67].

hsa-mir-103-2: Ethanol Degradation IV and Estrogen Receptor Signaling. *hsa-mir-103-2* corresponds to *hsa-mir-103/107* human homologous and has been already found upregulated in human BC cells, associated with metastatic process [68], and in serum of BC patients by next-generation-sequencing technique [69]. In BC, *hsa-mir-103/107* has an important role of epithelial mesenchymal transition (EMT) regulator [70]. EMT is a key process for metastatic spreading of the BC cells. Our analysis reveals that this miRNA is able to modulate the expression of 4 over 17 genes involved in ethanol degradation and 15 over 112 involved in ER signaling pathway. Looking to the list of target genes of *hsa-mir-103*, we found that a lot

of these genes belong to the family of ALDH genes, and, although having a role in ethanol detoxification, are also considered biomarkers of cancer stem cells (CSC) [71]. It is not thus surprising that these genes and their upstream miRNA, *hsa-mir-103*, are associated with higher grade BC samples. In fact, this tumor grade BC contains probably a smaller population of highly invasive and aggressive CSCs [72].

This hypothesis is further supported by the fact that among the top ten cross-talk pathways depicted in Table 2, Human Embryonic Stem Cell Pluripotency Pathway is among the more discriminating metabolic processes between BC and NS. *hsa-mir-103* expression has been already associated with ER status of BC samples, being more abundant in ER+ BC [73].

As shown in Table 5, among the genes controlled by *hsa-mir-103* in estrogen receptor signaling pathway, the family of MED complex is emerging (i.e., MED6, MED15, and MED16). This family contains several members of coactivators of transcription of RNA polymerase II-controlled genes. These proteins have a known regulatory role in metastatic process, as demonstrated by the silencing of MED15 that decreases the metastatic potential of a highly aggressive BC cell line by reducing TGF β /Smad signaling [74].

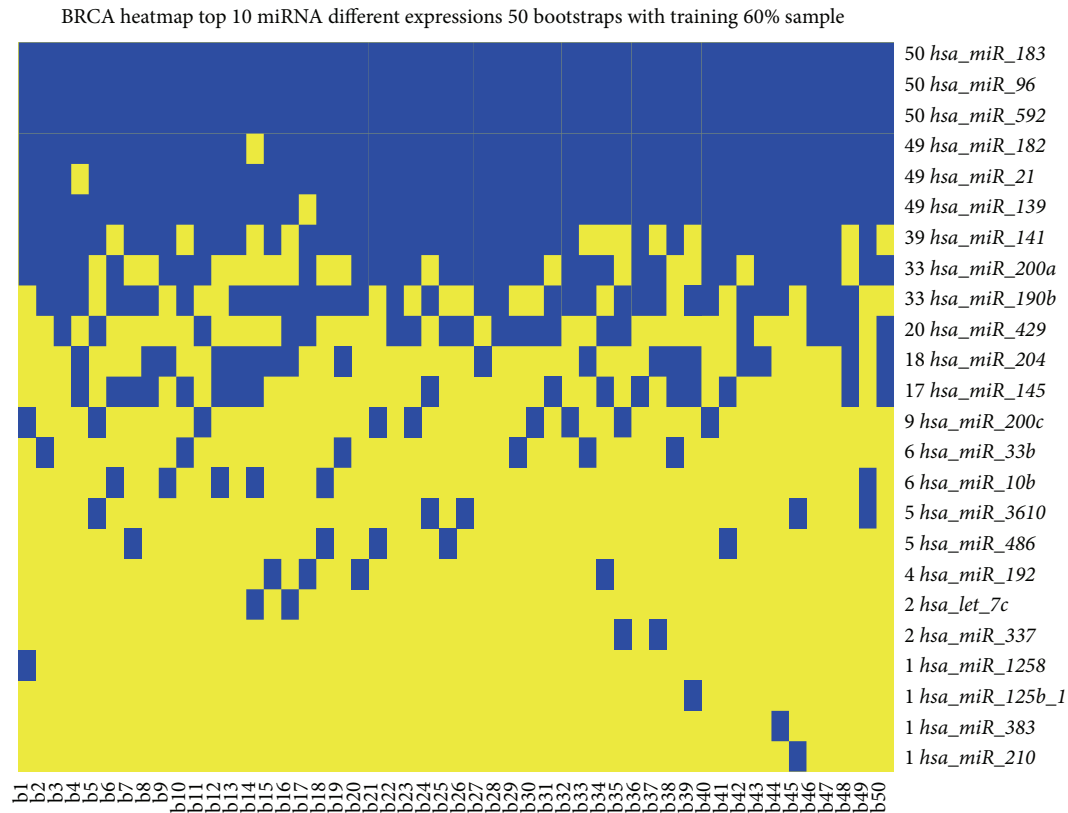


FIGURE 8: Heatmap with top 10 miRNAs (blue square) that obtained the best AUC values in the training dataset.

Intrinsic Prothrombin Activation Pathway and Extrinsic Prothrombin Activation Pathway. We found that two cross-talked pathways, extrinsic and intrinsic prothrombin pathways, are regulated by a group of 9 miRNAs (*hsa-miR-147b*, *hsa-miR-210*, *hsa-miR-301b*, *hsa-miR-483*, *hsa-miR-592*, *hsa-miR-665*, *hsa-miR-887*, *hsa-miR-9-1*, and *hsa-miR-939*). The frequency of miRNA regulation over the genes belonging to the two pathways goes from 33% (3/13) up to 66% (8/12). Analyzing the functional role of the pathways, in almost every patient with a progressive metastasized tumor, a constitutive activation of the coagulation cascade can be found [75]. The coagulation pathway is essential for the establishment of metastasis also in experimental model of cancer. Tumor cells express factors that trigger coagulation in different ways [76]. The extrinsic pathway is triggered by tissue factor (TSF) expressed on the surface of the tumor cells, on microparticles released by the tumor cells, or on the tumor stroma, leading to fibrin formation. TSF binds and activates factor VIIa, initiating the coagulation cascade and leading to thrombin activation. Some of the TSFs involved in the control of the intrinsic pathway are able also to activate the extrinsic pathway. For example, the FXIIa, a factor involved in the intrinsic pathway, indirectly facilitates the extrinsic pathway by converting FVII to FVIIa. The activation of FVII also occurs through the action of thrombin or FXa, mostly generated through the intrinsic pathway. The ability of FXa to activate FVII creates a link between the intrinsic and extrinsic pathways. In reality, the main function of the extrinsic pathway is to magnify

the activity of the intrinsic pathway. Metastatic cancer cells have been found to express exceptionally high levels of TSF (up to 100-fold higher than nonmetastatic cells) [77] and it has been suggested that cancer stem cells (CSCs) may express higher levels of TSF [78, 79]. Enhanced cancer cell TSF expression can also lead to increased tumor growth. This may be due to signaling via proteases from tumor environment, such as FVIIa, FXa, or thrombin [80] or to signaling through the TSF cytoplasmic domain [81], which contribute to prooncogenic signals affecting mainly the CSC behavior [82, 83]. The expression of these procoagulant molecules guides development of inappropriate coagulation in cancer, leading to the onset of coagulopathies related with BC [84]. Several are the miRNAs able to control genes belonging to both intrinsic and extrinsic coagulation pathways. Among them, *hsa-miR-210* emerges as one of the miRNAs, whose expression is more altered in higher grade BC [85]. Also the altered expression of other miRNAs, such as *hsa-miR-301b* or *hsa-miR-9-1*, has been found in our miRNA analysis and has been already reported in BC, being associated with proliferation and invasion control [86] or stem cell phenotype [87], respectively.

hsa-let-7c: Human Embryonic Stem Cell Pluripotency and Putrescine DegradationIII. The last couple of pathways able to discriminate among BC and NS is the couple formed by Human Embryonic Stem Cell Pluripotency genes and those involved in the process of putrescine degradation. Putrescine

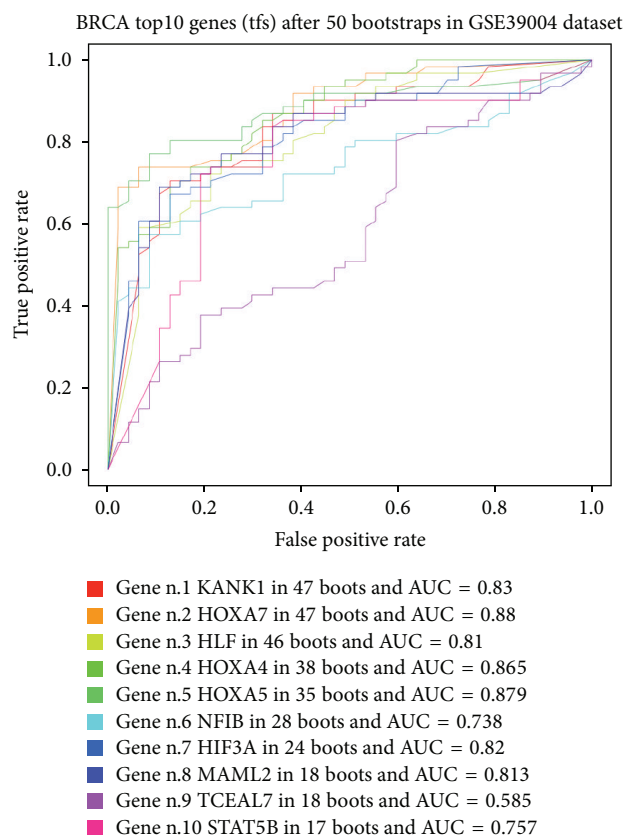


FIGURE 9: AUC curve for each top 10 genes for all 50 bootstraps.

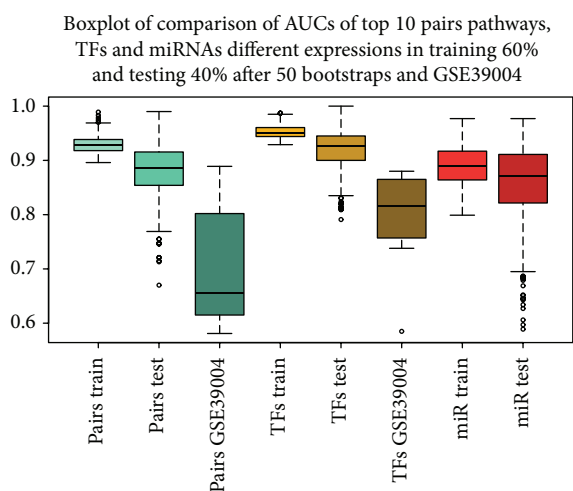


FIGURE 10: Comparison Boxplot among our approach (pairs train, pairs test, and GSE39004), TF (TFs train, TFs test, GSE39004), and miRNA (miRNA train and miRNA test).

is a known metabolite that plays an important role in cancer and cancer stem cells [88]. Putrescine belongs to the class of polyamine that has been shown to affect numerous processes in normal and cancer cells, such as proliferation, apoptosis, cell-cell interactions, and angiogenesis [89]. Total polyamine levels are higher in highly proliferative cells, like cancer cells,

and lower in cells with low proliferation rates [89]. It is not surprisingly to find that putrescine degradation pathway is one of the pathways important to distinguish BC from NS. *hsa-let-7c*, the miRNA regulating this metabolite, has a known role in cancer stem cell phenotype control. In particular, it has been reported that altered levels of *hsa-let-7c* are associated with higher grade BC [9, 90], as *hsa-let-7c* is responsible for tumor proliferation control [91]. Emerging evidences have linked the cellular levels of polyamines, such as putrescine, to the regulation of the level of expression of let-7 family members. In particular, polyamines are emerging as potential oncometabolites that influence specific aspects of tumorigenesis by regulating pluripotency associated factors, such as LIN28 [92]. Our speculation is that, in more aggressive BC samples, an alteration in polyamine metabolism alters the level of expression of *hsa-let-7c*, which in turn regulates the pluripotency capacity of the cancer stem cell population. In this way we can explain the link among the two BC-discriminating pathways, the putrescine degradation pathway and the stem cell pluripotency pathway, and the altered level of expression of *hsa-let-7c*.

As shown in Table 5, among the genes regulated by *hsa-let-7c* in Human Embryonic Stem Cell Pluripotency, some are already known stem cell transcription factors, necessary to induce or maintain pluripotency potential (i.e., WNT11 or WNT6, BMP5, and BMP6) [93, 94]. Among the genes controlled by *hsa-let-7c* in putrescine degradation pathway, several of them are related to stem cell phenotype (i.e., the family of aldehyde dehydrogenase, such as ALDH1A1, ALDH1A21, ALDH2, ALDH4A1, ALDH9A1) [95].

5. Conclusions

In conclusion, the approach used in our work allowed identifying (1) four pairs of pathways able to accurately classify aggressive BC versus NS with an important role in the regulations of several mechanisms in BC; (2) an epigenetic signature of 12 miRNAs able to regulate those pairs of pathways in BC, and (3) a discriminating score able to quantify the cross-talk among pathways, with a potential diagnostic and therapeutic roles in BC. These interesting suggestions, obtained by in silico analysis, should be further validated in laboratories by using, for example, BC cell lines. Due to the reduced number of miRNAs, our epigenetic signature could be suitable to be translated in a clinical environment.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Authors' Contribution

Antonio Colaprico and Claudia Cava contributed equally to this work.

Acknowledgments

This work was supported by a research project funded by the BridgeIRIS project funded by INNOVIRIS, Region de Bruxelles Capitale, Brussels, Belgium, and by INTEROMICS flagship project (<http://www.interomics.eu/it/home>), National Research Council CUP Grant B91J12000190001.

References

- [1] L. Gu, S. C. Frommel, C. C. Oakes et al., “BAZ2A (TIP5) is involved in epigenetic alterations in prostate cancer and its overexpression predicts disease recurrence,” *Nature Genetics*, vol. 47, no. 1, pp. 22–30, 2014.
- [2] M. A. Knowles and C. D. Hurst, “Molecular biology of bladder cancer: new insights into pathogenesis and clinical diversity,” *Nature Reviews Cancer*, vol. 15, no. 1, pp. 25–41, 2014.
- [3] S. T. Bailey, T. Westerling, and M. Brown, “Loss of estrogen-regulated microRNA expression increases her2 signaling and is prognostic of poor outcome in luminal breast cancer,” *Cancer Research*, vol. 75, no. 2, pp. 436–445, 2015.
- [4] C. Cava, I. Zoppis, M. Gariboldi, I. Castiglioni, G. Mauri, and M. Antoniotti, “Combined analysis of chromosomal instabilities and gene expression for colon cancer progression inference,” *Journal of Clinical Bioinformatics*, vol. 4, article 2, 2014.
- [5] T. M. Severson, J. Peeters, I. Majewski et al., “BRCA1-like signature in triple negative breast cancer: molecular and clinical characterization reveals subgroups with therapeutic potential,” *Molecular Oncology*, 2015.
- [6] S. G. Zhao, M. Shilkut, C. Speers et al., “Development and validation of a novel platform-independent metastasis signature in human breast cancer,” *PLOS ONE*, vol. 10, no. 5, Article ID e0126631, 2015.
- [7] C. Sotiriou, P. Wirapati, S. Loi et al., “Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis,” *Journal of the National Cancer Institute*, vol. 98, no. 4, pp. 262–272, 2006.
- [8] A. V. Ivshina, J. George, O. Senko et al., “Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer,” *Cancer Research*, vol. 66, no. 21, pp. 10292–10301, 2006.
- [9] C. Cava, G. Bertoli, M. Ripamonti et al., “Integration of mRNA expression profile, copy number alterations, and microRNA expression levels in breast cancer to improve grade definition,” *PLOS ONE*, vol. 9, no. 5, Article ID e97681, 2014.
- [10] P. C. Miller, J. Clarke, T. Koru-Sengul, J. Brinkman, and D. El-Ashry, “A novel mapk-microrna signature is predictive of hormone-therapy resistance and poor outcome in er-positive breast cancer,” *Clinical Cancer Research*, vol. 21, no. 2, pp. 373–385, 2015.
- [11] D. Jin and H. Lee, “A computational approach to identifying gene-microrna modules in cancer,” *PLOS Computational Biology*, vol. 11, no. 1, Article ID e1004042, 2015.
- [12] M. Du, T. Yuan, K. F. Schilter et al., “Prostate cancer risk locus at 8q24 as a regulatory hub by physical interactions with multiple genomic loci across the genome,” *Human Molecular Genetics*, vol. 24, no. 1, pp. 154–166, 2015.
- [13] F. Rapaport, A. Zinovyev, M. Dutreix, E. Barillot, and J.-P. Vert, “Classification of microarray data using gene networks,” *BMC Bioinformatics*, vol. 8, no. 1, article 35, 2007.
- [14] J. Tomfohr, J. Lu, and T. B. Kepler, “Pathway level analysis of gene expression using singular value decomposition,” *BMC Bioinformatics*, vol. 6, no. 1, article 225, 2005.
- [15] J. Su, B.-J. Yoon, and E. R. Dougherty, “Accurate and reliable cancer classification based on probabilistic inference of pathway activity,” *PLoS ONE*, vol. 4, no. 12, Article ID e8161, 2009.
- [16] R. Bernards, “A missing link in genotype-directed cancer therapy,” *Cell*, vol. 151, no. 3, pp. 465–468, 2012.
- [17] F. Wang, W. C. Cho, L. W. Chan et al., “Gene network exploration of crosstalk between apoptosis and autophagy in chronic myelogenous leukemia,” *BioMed Research International*, vol. 2015, Article ID 459840, 9 pages, 2015.
- [18] S. L. Anwar and U. Lehmann, “DNA methylation, microRNAs, and their crosstalk as potential biomarkers in hepatocellular carcinoma,” *World Journal of Gastroenterology*, vol. 20, no. 24, pp. 7894–7913, 2014.
- [19] M. Leotta, L. Biamonte, L. Raimondi et al., “A p53-dependent tumor suppressor network is induced by selective mir-125a-5p inhibition in multiple myeloma cells,” *Journal of Cellular Physiology*, vol. 229, no. 12, pp. 2106–2116, 2014.
- [20] G. Maria Ventola, A. Colaprico, F. D. Angelo et al., “An approach to identify mirna associated with cancer altered pathways,” in *New Trends in Image Analysis and Processing—ICIAP 2013*, pp. 399–408, Springer, 2013.
- [21] X. Jiang, L. Du, L. Wang et al., “Serum microRNA expression signatures identified from genome-wide microRNA profiling serve as novel noninvasive biomarkers for diagnosis and recurrence of bladder cancer,” *International Journal of Cancer*, vol. 136, no. 4, pp. 854–862, 2015.
- [22] E. van Schooneveld, H. Wildiers, I. Vergote, P. B. Vermeulen, L. Y. Dirix, and S. J. van Laere, “Dysregulation of microRNAs in breast cancer and their potential role as prognostic and predictive biomarkers in patient management,” *Breast Cancer Research*, vol. 17, no. 1, p. 21, 2015.
- [23] H. Butz, K. Rácz, L. Hunyady, and A. Patócs, “Crosstalk between TGF- β signaling and the microRNA machinery,” *Trends in Pharmacological Sciences*, vol. 33, no. 7, pp. 382–393, 2012.
- [24] Z. Wang, Y. Li, D. Kong, A. Ahmad, S. Banerjee, and F. H. Sarkar, “Cross-talk between miRNA and Notch signaling pathways in tumor development and progression,” *Cancer Letters*, vol. 292, no. 2, pp. 141–148, 2010.
- [25] J. Zhao and R. D. Carrasco, “Crosstalk between microRNA30a/b/c/d/e-5p and the canonical wnt pathway: implications for multiple myeloma therapy,” *Cancer Research*, vol. 74, no. 19, pp. 5351–5358, 2014.
- [26] Y. Cao, J. Wang, H. Zhang et al., “Detecting key genes regulated by mirnas in dysfunctional crosstalk pathway of myasthenia gravis,” *BioMed Research International*, vol. 2015, Article ID 724715, 10 pages, 2015.
- [27] Z. Wang, Y. Li, D. Kong, A. Ahmad, S. Banerjee, and F. H. Sarkar, “Cross-talk between miRNA and notch signaling pathways in tumor development and progression,” *Cancer Letters*, vol. 292, no. 2, pp. 141–148, 2010.
- [28] S. E. Hiemer, A. D. Szymaniak, and X. Varelas, “The transcriptional regulators TAZ and YAP direct transforming growth factor β -induced tumorigenic phenotypes in breast cancer cells,” *The Journal of Biological Chemistry*, vol. 289, no. 19, pp. 13461–13474, 2014.
- [29] L. Han, A. Adam Diehl, N. K. Nguyen et al., “The notch pathway inhibits tgfb signaling in breast cancer through heylmediated crosstalk,” *Cancer Research*, vol. 74, no. 22, pp. 6509–6518, 2014.
- [30] K. A. Leehy, T. M. Regan Anderson, A. R. Daniel, A. B. D’Assoro, and C. A. Lange, “Aurora a kinase and progesterone receptor cross talk in breast cancer,” *Cancer Research*, vol. 74, no. 19, supplement, p. 2107, 2014.

- [31] S. R. D. Johnston, "New strategies in estrogen receptor-positive breast cancer," *Clinical Cancer Research*, vol. 33, no. 7, pp. 382–393, 2012.
- [32] M. Donato, Z. Xu, A. Tomoiaga et al., "Analysis and correction of crosstalk effects in pathway analysis," *Genome Research*, vol. 23, no. 11, pp. 1885–1893, 2013.
- [33] E. Lee, H.-Y. Chuang, J.-W. Kim, T. Ideker, and D. Lee, "Inferring pathway activity toward precise disease classification," *PLoS Computational Biology*, vol. 4, no. 11, Article ID e1000217, 2008.
- [34] H. Yang, C. Cheng, and W. Zhang, "Average rank-based score to measure deregulation of molecular pathway gene sets," *PLoS ONE*, vol. 6, no. 11, Article ID e27579, 2011.
- [35] Z. Guo, T. Zhang, X. Li et al., "Towards precise classification of cancers based on robust gene functional expression profiles," *BMC Bioinformatics*, vol. 6, article no. 58, 2005.
- [36] C. Cava, G. Bertoli, and I. Castiglioni, "Pathway-based expression profile for breast cancer diagnoses," in *Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '14)*, pp. 1151–1154, IEEE, Chicago, Ill, USA, August 2014.
- [37] T. McCormack, O. Frings, A. Alexeyenko, and E. L. L. Sonnhammer, "Statistical assessment of crosstalk enrichment between gene groups in biological networks," *PLoS ONE*, vol. 8, no. 1, Article ID e54945, 2013.
- [38] A. Colaprico, T. Silva, C. Olsen et al., "TCGAbiolinks: an R package to download and analyze data from TCGA," *Nature Methods and Bioconductor*. Submitted.
- [39] M. D. Robinson, D. J. McCarthy, and G. K. Smyth, "edgeR: a bioconductor package for differential expression analysis of digital gene expression data," *Bioinformatics*, vol. 26, no. 1, pp. 139–140, 2010.
- [40] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the Royal Statistical Society, Series B: Methodological*, vol. 57, no. 1, pp. 289–300, 1995.
- [41] T. R. Golub, D. K. Slonim, P. Tamayo et al., "Molecular classification of cancer: class discovery and class prediction by gene expression monitoring," *Science*, vol. 286, no. 5439, pp. 531–527, 1999.
- [42] B. Orsetti, M. Nugoli, N. Cervera et al., "Genetic profiling of chromosome 1 in breast cancer: mapping of regions of gains and losses and identification of candidate genes on 1q," *British Journal of Cancer*, vol. 95, no. 10, pp. 1439–1447, 2006.
- [43] A. Liaw and M. Wiener, "Classification and regression by randomforest," *R News*, vol. 2, no. 3, pp. 18–22, 2002.
- [44] R. de Smet and K. Marchal, "Advantages and limitations of current network inference methods," *Nature Reviews Microbiology*, vol. 8, no. 10, pp. 717–729, 2010.
- [45] R. De Smet and K. Marchal, "Advantages and limitations of current network inference methods," *Nature Reviews Microbiology*, vol. 8, no. 10, pp. 717–729, 2010.
- [46] D. Marbach, J. C. Costello, R. Küffner et al., "Wisdom of crowds for robust gene network inference," *Nature Methods*, vol. 9, no. 8, pp. 796–804, 2012.
- [47] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 69, no. 6, Article ID 066138, 2004.
- [48] G. Sales and C. Romualdi, "Parmigene—a parallel R package for mutual information estimation and gene network reconstruction," *Bioinformatics*, vol. 27, no. 13, pp. 1876–1877, 2011.
- [49] C. Lefebvre, P. Rajbhandari, M. J. Alvarez et al., "A human B-cell interactome identifies MYB and FOXM1 as master regulators of proliferation in germinal centers," *Molecular Systems Biology*, vol. 6, no. 1, article 377, 2010.
- [50] A. Subramanian, P. Tamayo, V. K. Mootha et al., "Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 43, pp. 15545–15550, 2005.
- [51] E. Wingender, P. Dietze, H. Karas, and R. Knüppel, "TRANSFAC: a database on transcription factors and their DNA binding sites," *Nucleic Acids Research*, vol. 24, no. 1, pp. 238–241, 1996.
- [52] J. Jiang, X. Zheng, X. Xu et al., "Prognostic significance of miR-181b and miR-21 in gastric cancer patients treated with S-1/Oxaliplatin or Doxifluridine/Oxaliplatin," *PLoS ONE*, vol. 6, no. 8, Article ID e23271, 2011.
- [53] R. Visone, A. Veronese, L. Z. Rassenti et al., "miR-181b is a biomarker of disease progression in chronic lymphocytic leukemia," *Blood*, vol. 118, no. 11, pp. 3072–3079, 2011.
- [54] F. Zhi, X. Chen, S. Wang et al., "The use of hsa-miR-21, hsa-miR-181b and hsa-miR-106a as prognostic indicators of astrocytoma," *European Journal of Cancer*, vol. 46, no. 9, pp. 1640–1649, 2010.
- [55] C. Zhang, S. Cao, B. P. Toole, and Y. Xu, "Cancer may be a pathway to cell survival under persistent hypoxia and elevated ROS: a model for solid-cancer initiation and early development," *International Journal of Cancer*, vol. 136, no. 9, pp. 2001–2011, 2015.
- [56] J. L. Heinecke, L. A. Ridnour, R. Y. S. Cheng et al., "Tumor microenvironment-based feed-forward regulation of NOS2 in breast cancer progression," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 17, pp. 6323–6328, 2014.
- [57] W. Xie, Z. Li, M. Li, N. Xu, and Y. Zhang, "mir-181a and inflammation: mirna homeostasis response to inflammatory stimuli in vivo," *Biochemical and Biophysical Research Communications*, vol. 430, no. 2, pp. 647–652, 2013.
- [58] R. W. Wyszynski, B. F. Gibbs, L. Varani, D. Iannotta, and V. V. Sumbayev, "Interleukin-1 beta induces the expression and production of stem cell factor by epithelial cells: crucial involvement of the PI-3K/mTOR pathway and HIF-1 transcription complex," *Cellular & Molecular Immunology*, 2014.
- [59] K. Davalieva, S. Kiprijanovska, S. Komina, G. Petrushevska, N. C. Zografiska, and M. Polenakovic, "Proteomics analysis of urine reveals acute phase response proteins as candidate diagnostic biomarkers for prostate cancer," *Proteome Science*, vol. 13, no. 1, article 2, 2015.
- [60] J. Obacz, S. Pastorekova, B. Vojtesek, and R. Hrstka, "Crosstalk between HIF and p53 as mediators of molecular responses to physiological and genotoxic stresses," *Molecular Cancer*, vol. 12, no. 1, article 93, 2013.
- [61] D. Liao and R. S. Johnson, "Hypoxia: a key regulator of angiogenesis in cancer," *Cancer and Metastasis Reviews*, vol. 26, no. 2, pp. 281–290, 2007.
- [62] R. S. Schwartz, H. K. Eltzschig, and P. Carmeliet, "Hypoxia and inflammation," *The New England Journal of Medicine*, vol. 364, no. 7, pp. 656–665, 2011.
- [63] D. Lukashev, C. Caldwell, A. Ohta, P. Chen, and M. Sitkovsky, "Differential regulation of two alternatively spliced isoforms of hypoxia-inducible factor-1 α in activated T lymphocytes," *The Journal of Biological Chemistry*, vol. 276, no. 52, pp. 48754–48763, 2001.

- [64] J. C. Zampell, A. Yan, T. Avraham, S. Daluvoy, E. S. Weitman, and B. J. Mehrara, "HIF-1 α coordinates lymphangiogenesis during wound healing and in response to inflammation," *The FASEB Journal*, vol. 26, no. 3, pp. 1027–1039, 2012.
- [65] T. Cramer, Y. Yamanishi, B. E. Clausen et al., "HIF-1 α is essential for myeloid cell-mediated inflammation," *Cell*, vol. 112, no. 5, pp. 645–657, 2003.
- [66] H.-L. Jiang, H.-F. Sun, S.-P. Gao et al., "Loss of RAB1B promotes triple-negative breast cancer metastasis by activating TGF- β /SMAD signaling," *Oncotarget*, vol. 5, 2015.
- [67] X. Jia, Q. Hong, L. Lei et al., "Basal and therapy-driven hypoxia-inducible factor-1 α confers resistance to endocrine therapy in estrogen receptor-positive breast," *Oncotarget*, vol. 6, no. 11, pp. 8648–8662, 2015.
- [68] G. Martello, A. Rosato, F. Ferrari et al., "A microRNA targeting dicer for metastasis control," *Cell*, vol. 141, no. 7, pp. 1195–1207, 2010.
- [69] Q. Wu, C. Wang, Z. Lu, L. Guo, and Q. Ge, "Analysis of serum genome-wide microRNAs for breast cancer detection," *Clinica Chimica Acta*, vol. 413, no. 13–14, pp. 1058–1065, 2012.
- [70] S. Chan and L. Wang, "Regulation of cancer metastasis by microRNAs," *Journal of Biomedical Science*, vol. 22, article 9, 2015.
- [71] P. Marcato, C. A. Dean, C. A. Giacomantonio, and P. W. K. Lee, "Aldehyde dehydrogenase its role as a cancer stem cell marker comes down to the specific isoform," *Cell Cycle*, vol. 10, no. 9, pp. 1378–1384, 2011.
- [72] M. L. Oon, A. A. Thike, S. Y. Tan, and P. H. Tan, "Cancer stem cell and epithelial–mesenchymal transition markers predict worse outcome in metaplastic carcinoma of the breast," *Breast Cancer Research and Treatment*, vol. 150, no. 1, pp. 31–41, 2015.
- [73] D. R. Cochrane, D. M. Cittelly, E. N. Howe et al., "MicroRNAs link estrogen receptor alpha status and dicer levels in breast cancer," *Hormones and Cancer*, vol. 1, no. 6, pp. 306–319, 2010.
- [74] M. Zhao, X. Yang, Y. Fu et al., "Mediator MED15 modulates transforming growth factor beta (TGF β)/Smad signaling and breast cancer cell metastasis," *Journal of Molecular Cell Biology*, vol. 5, no. 1, pp. 57–60, 2013.
- [75] A. Falanga, M. N. Levine, R. Consonni et al., "The effect of very-low-dose warfarin on markers of hypercoagulation in metastatic breast cancer: results from a randomized trial," *Thrombosis and Haemostasis*, vol. 79, no. 1, pp. 23–27, 1998.
- [76] A. Falanga and M. Marchetti, "Venous thromboembolism in the hematologic malignancies," *Journal of Clinical Oncology*, vol. 27, no. 29, pp. 4848–4857, 2009.
- [77] B. M. Mueller, R. A. Reisfeld, T. S. Edgington, and W. Ruf, "Expression of tissue factor by melanoma cells promotes efficient hematogenous metastasis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, no. 24, pp. 11832–11836, 1992.
- [78] C. Milsom, J. Yu, L. May et al., "The role of tumor-and host-related tissue factor pools in oncogene-driven tumor progression," *Thrombosis Research*, vol. 120, pp. S82–S91, 2007.
- [79] Y. W. van den Berg, S. Osanto, P. H. Reitsma, and H. H. Versteeg, "The relationship between tissue factor and cancer progression: insights from bench and bedside," *Blood*, vol. 119, no. 4, pp. 924–932, 2012.
- [80] J. L. Yu, L. May, V. Lhotak et al., "Oncogenic events regulate tissue factor expression in colorectal cancer cells: implications for tumor progression and angiogenesis," *Blood*, vol. 105, no. 4, pp. 1734–1741, 2005.
- [81] W. Ruf, J. Disse, T. C. Carneiro-Lobo, N. Yokota, and F. Schaffner, "Tissue factor and cell signalling in cancer progression and thrombosis," *Journal of Thrombosis and Haemostasis*, vol. 9, no. s1, pp. 306–315, 2011.
- [82] D. Garnier, N. Magnus, E. D'Asti et al., "Genetic pathways linking hemostasis and cancer," *Thrombosis Research*, vol. 129, supplement 1, pp. S22–S29, 2012.
- [83] F. Schaffner, N. Yokota, and W. Ruf, "Tissue factor proangiogenic signaling in cancer progression," *Thrombosis Research*, vol. 129, supplement 1, pp. S127–S131, 2012.
- [84] C. Robier, M. Neubauer, C. Beham-Schmid, and H. Sill, "Thrombotic microangiopathy and disseminated intravascular coagulation associated with carcinocythemia in a patient with breast cancer," *Journal of Clinical Oncology*, vol. 29, no. 34, pp. e825–e826, 2011.
- [85] F. Rothé, M. Ignatiadis, C. Chaboteaux et al., "Global microRNA expression profiling identifies MiR-210 associated with tumor proliferation, invasion and poor clinical outcome in breast cancer," *PLoS ONE*, vol. 6, no. 6, Article ID e20980, 2011.
- [86] W. Shi, K. Gerster, N. M. Alajez et al., "MicroRNA-301 mediates proliferation and invasion in human breast cancer," *Cancer Research*, vol. 71, no. 8, pp. 2926–2937, 2011.
- [87] C. Delaloy, L. Liu, J.-A. Lee et al., "MicroRNA-9 coordinates proliferation and migration of human embryonic stem cell-derived neural progenitors," *Cell Stem Cell*, vol. 6, no. 4, pp. 323–335, 2010.
- [88] M. Lessard, C. Zhao, S. M. Singh, and R. Poulin, "Hormonal and feedback regulation of putrescine and spermidine transport in human breast cancer cells," *The Journal of Biological Chemistry*, vol. 270, no. 4, pp. 1685–1694, 1995.
- [89] K. Igarashi and K. Kashiwagi, "Modulation of cellular function by polyamines," *The International Journal of Biochemistry and Cell Biology*, vol. 42, no. 1, pp. 39–51, 2010.
- [90] X.-X. Li, S.-Y. Gao, P.-Y. Wang et al., "Reduced expression levels of let-7c in human breast cancer patients," *Oncology Letters*, vol. 9, no. 3, pp. 1207–1212, 2015.
- [91] S. Björner, P. A. Fitzpatrick, Y. Li et al., "Epithelial and stromal microRNA signatures of columnar cell hyperplasia linking let-7c to precancerous and cancerous breast cancer cell proliferation," *PLoS ONE*, vol. 9, no. 8, Article ID e105099, 2014.
- [92] E. A. Paz, B. Lafleur, and E. W. Gerner, "Polyamines are oncometabolites that regulate the LIN28/let-7 pathway in colorectal cancer cells," *Molecular Carcinogenesis*, vol. 53, no. 1, pp. E96–E106, 2014.
- [93] Y. Yang, K. Adachi, M. A. Sheridan et al., "Heightened potency of human pluripotent stem cell lines created by transient bmp4 exposure," *Proceedings of the National Academy of Sciences*, vol. 112, no. 18, pp. E2337–E2346, 2015.
- [94] K. Hawkins, "Cell signalling pathways underlying induced pluripotent stem cell reprogramming," *World Journal of Stem Cells*, vol. 6, no. 5, article 620, 2014.
- [95] D. Zhao, Y. Mo, M. Li et al., "NOTCH-induced aldehyde dehydrogenase 1A1 deacetylation promotes breast cancer stem cells," *The Journal of Clinical Investigation*, vol. 124, no. 12, pp. 5453–5465, 2014.