Cancer Genetics 239 (2019) 62–74

Contents lists available at ScienceDirect

Cancer Genetics

journal homepage: www.elsevier.com/locate/cancergen

Original Article

Transcriptomics-based screening of molecular signatures associated with patients overall survival and their key regulators in subtypes of breast cancer

Elaheh Eskandari[a,b], Jamshid Motalebzadeh[a,b,*]

[a] Department of Clinical Genetics, Institute of Medical Biotechnology, National Institute of Genetic Engineering and Biotechnology (NIGEB), Tehran, Iran
[b] Medical Informatics and Genetics Center (MedInfoGene), Tehran, Iran

ARTICLE INFO

Article history:
Received 8 April 2019
Revised 18 August 2019
Accepted 15 September 2019

Keywords:
Breast cancer subtypes
Molecular signature
Prognosis
Gene regulatory network
Transcription factors
Micro-RNAs

ABSTRACT

Molecular subtypes of breast cancer are associated with differences in prognosis and strategies of molecular targeted therapies. Gene regulatory mechanisms as one of the reasons might modulate these differences. In the present study, we proposed a comprehensive data analysis and systems biology approach to explore molecular signatures which reduce the chance of patients overall survival and the possible mechanisms of their regulation by transcription factors (TFs) and microRNAs (miRNAs) in the main subtypes of breast tumor consist of Basal like, Her2 enriched, Luminal A and Luminal B breast cancer. In this regards, we used available microarray datasets to assess common differentially expressed genes (DEGs) in breast cancer subtypes. Using Kaplan-Meier curve analysis we identified common DEGs which are associated with decreasing in the overall survival of breast cancer patients. Furthermore, gene regulatory networks (GRNs) were depicted based on TFs and miRNAs with interest target genes. Then GRNs were analyzed and using five algorithms (Control centrality, Betweenness, Degree, Classification, and MCDS) the key regulators were identified for each subtype. In this study, we highlighted mechanisms underlying the regulation of breast cancer molecular signatures by TFs and miRNAs which their alteration reduce the chance of survival rate in each subtype of breast cancer. Our current study in a holistic insight revealed the importance of some genes and their regulators as potential prognostic markers and/or therapeutic targets in breast cancer patients.

Introduction

Breast cancer as one of the leading cause of cancer morbidity and mortality among women is a heterogeneous disease with many distinct biological differences comprised different clinical outcomes, different responses to the therapies, tumor relapse and patients' survival rate [1,2]. Due to the importance of prognosis and targeted therapies, molecular subtypes of breast cancer is known as one of the most important classifications of this malignancy [3,4]. According to the molecular subtypes, breast cancer is divided into four main groups comprised of Luminal A (LuA), Luminal B (LuB), Her2 enriched (Her2) and Basal like (BL) subtypes. Each subtype has distinct characteristics which are briefly described in the following part.

Both LuA and LuB breast cancer represent positive expression of ER and PR receptors in the cells. LuB tumor compared to the LuA shows positive expression of Her2 receptor, high expression in Ki-67 gene, and high grade of tumor cells [5]. These breast cancer subtypes are well responded to hormone therapy. Her2 breast cancer is identified according to the negative expression of ER and PR but positive expression of Her2 receptor. This type of breast cancer has high grade and grows faster than luminal subtypes and may have a worse prognosis [6]. Molecular targeted therapy using herceptin is an appropriate treatment for Her2 breast cancer. BL subtype of breast cancer is known based on lack or low expression of ER, PR and Her2 receptors in the cancer cells [7]. This subtype of breast cancer has high grade with a larger size of tumor [8]. There is no suitable targeted therapy for BL, thus identification of molecular targeted therapies for this subtype of breast cancer is known as a crucial strategy for the patients treatment.

The mechanisms of gene regulation in cancer have been a controversial topic during the last decades. The clear role of regulatory elements in cancer is largely unknown, especially, it is not completely understood how micro-RNAs (miRNAs) and transcription factors (TFs) regulate and influence cancer related processes [9]. TFs and miRNAs are known as two crucial gene regulatory fac-

---

[*] Corresponding author at: Shahrak-e Pajoohesh, km 15, Tehran - Karaj Highway P.O. Box: 14965/161, Tehran, Iran.
E-mail address: motalebzade@gmail.com (J. Motalebzadeh).

https://doi.org/10.1016/j.cancergen.2019.09.004

Table 1
Microarray datasets used in this study to achieve DEGs, DETFs, and DEmiRs in each subtype of breast cancer.

GEO Accession	Author	Platform	NO. Used Samples
GSE10886	Parker JS [13]	Agilent Human 1A Oligo UNC custom Microarrays	Normal: 7 BL: 11 Her2: 12 LuA: 40 LuB: 13
GSE26304	Muggerud AA [14]	Agilent-012391 Whole Human Genome Oligo Microarray G4112A (Probe Name version)	Normal: 6 BL: 20 Her2: 16 LuA: 24 LuB: 30
GSE59246	Lesurf R [15]	Agilent-028004 SurePrint G3 Human GE 8 × 60 K Microarray (Feature Number version)	Normal: 3 BL: 16 Her2: 17 LuA: 32 LuB: 21
GSE58606	Matamala N [16]	miRCURY LNA microRNA Array 7th generation - human, mouse & rat [product# 208502] (miRBase 18.0)	Normal: 11 BL: 31 Her2: 27 LuA: 31 LuB: 33
GSE97811	Hironaka-Mitsuhashi A [17]	3D-Gene Human miRNA V21_1.0.0	Normal: 6 BL: 7 Her2: 4 LuA: 20 LuB: 12

tors in the cells. Many studies have highlighted the significant role of TFs and miRNAs in the initiation and progression of cancer. In addition, several TFs and miRNAs were known to be differentially expressed in different types of cancer compared to the normal cells. Effective treatment of cancer may require a systemic understanding of genomic regulators and identifying the gene regulatory networks (GRN) involved in TFs and miRNAs in cancer helps us to find an effective strategy for cancer molecular targeted therapy [10]. However, systematically identification of the molecular interactions by experimental methods is time consuming, difficult and expensive. Therefore, it is valuable to develop computational approaches capable to identify such GRNs.

Alteration in the level of the gene regulatory elements in the cell may result in changes in the level of the genes transcription and induces some abnormalities. Many biological researches require GRNs to provide clear insight and understanding of the cellular process in living cells due to the interactions between genes and their products which play an important role in many molecular processes [11]. It is known that alteration in gene regulation induces changes in normal cell function that lead to the progression of pathogenesis in cancer. GRN can act as a blueprint for the researchers to observe the relationships among genes, so understanding the GRN by TFs and miRNAs will shed light on the mechanisms of cancer development and progression [12].

There are some studies in which the GRN in breast cancer was discussed. However, to the best of our knowledge, there is no comprehensive study that highlights the GRNs for the molecular signatures as prognostic markers which are associated with overall survival for each molecular subtype of breast cancer using a robust computational approach. Our approach allows a deeper level of understanding of gene regulatory mechanism of molecular signatures associated with overall survival by TFs and miRNAs in breast cancer. Finally, we highlighted the key regulators in each subtype of breast cancer. In our current study, we aimed to explore a GRN for each subtype of breast cancer which would be useful for clinical and pathological research in breast cancer.

In this study, we used five microarray datasets in patients with breast cancer to assess differentially expressed genes (DEGs), differentially expressed transcription factors (DETFs) and differentially expressed miRNAs (DEmiRs) which are involved with each subtype of breast cancer. Afterward, we selected common DEGs with potential as prognostic markers for each subtype. Then we depicted GRNs for the candidate prognostic markers using acquired DETFs and DEmiRs. Thereafter, we revealed the architecture and features of the GRNs based on DETFs, DEmiRs and their target genes (the candidate prognostic markers) to find the key regulators. The workflow of our study is depicted in Fig. 1.

Materials and methods

Data collection and microarray data analysis to identify common DEGs

We searched for the datasets involved with gene expression data for each subtype of breast cancer in GEO database (https://www.ncbi.nlm.nih.gov/gds/) and ArrayExpress (https://www.ebi.ac.uk/arrayexpress/). In this regards, three independent breast cancer tissue microarray datasets for mRNA expression with patients pathological information were downloaded. All datasets and samples comprised breast cancer cell lines, patients with neoadjuvant therapy, patients without necessary clinical data and hereditary breast cancer patients were excluded from the study. In this way, we selected three reliable microarray datasets with sufficient number of sample size. All selected datasets were based on Agilent platform included GSE10886, GSE26304, and GSE59246 which were shown in Table 1. Tumor samples were classified into four groups consist of BL, Her2, LuA and LuB subtypes of breast cancer. Totally 252 breast tumor samples containing each subtype of breast cancer were analyzed for screening DEGs. Details of all datasets and the number of patients involved in each subtype of breast cancer were mentioned in Table 1. To identify DEGs involved with breast cancer subtypes, we compared the gene expression of each subtype of breast tumor samples with breast normal tissues. In this regards, each dataset was analyzed separately using limma package in R to assess DEGs. Genes with $|FC| > 1.5$ and p-value < 0.05 were selected as meaningful DEGs in each dataset. We selected all common DEGs in each subtype and afterward, we analyzed each common DEG for its association with patients overall survival rate.

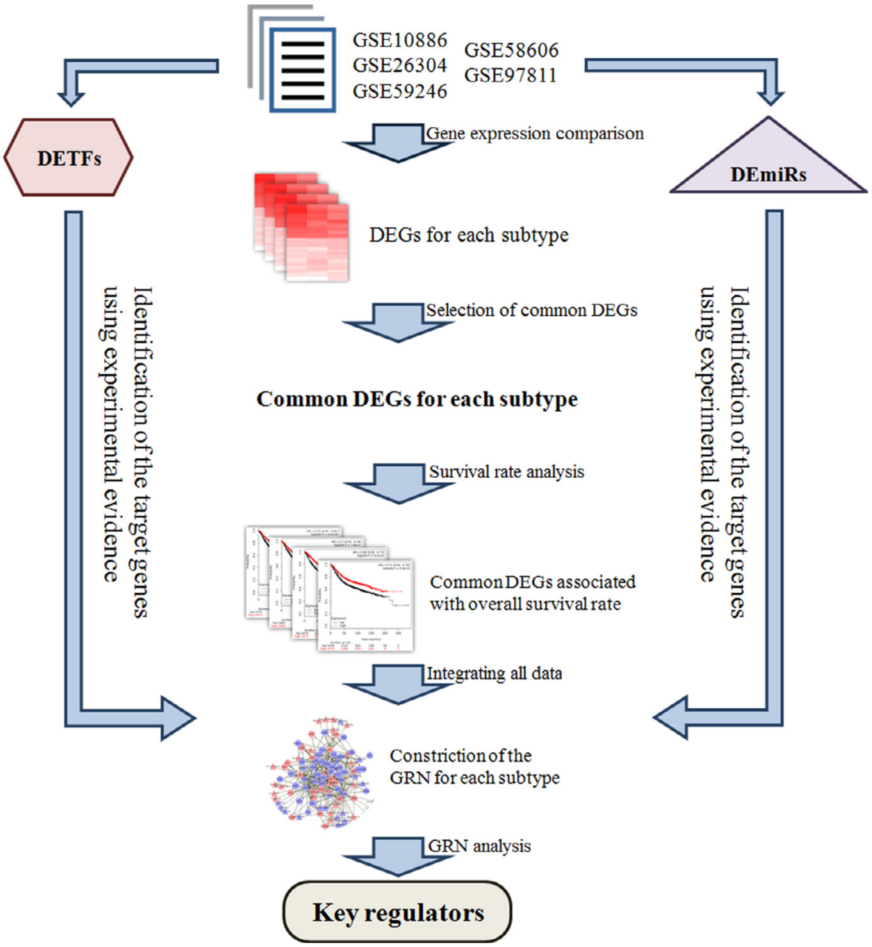

Fig. 1. The workflow and analysis steps.

Patients overall survival rate analysis to identify molecular signatures

Using Kaplan-Meier curve we estimated each common DEG with potential as a prognostic marker for breast cancer subtypes. We used KM-plotter web-based tool [18] which had high throughput data from 5143 breast cancer patients to acquire the association of each DEG with breast cancer patients survival rate. All patients were classified into four main breast cancer subtypes and then, we performed overall survival rate analysis using Kaplan-Meier curve for each common DEG. In this way, all common DEGs were tested for their association with patients survival rate in each breast cancer subtype. We selected all common up-regulated DEGs with $HR > 1$ and $p\text{-value} < 0.05$, and also all down-regulated DEGs with $HR < 1$ and $p\text{-value} < 0.05$ for each subtype of breast cancer. In this regards, we selected all common DEGs (Up/Down-regulated genes) which changes in their expression are associated with a lower chance of the patients to be survived in each breast cancer subtype. In other words, we selected common DEGs which were negatively associated with overall survival rate for each subtype of breast cancer. We used these candidate prognostic markers for the construction of the GRN for each breast cancer subtype.

Identification of DEmiRs and their target genes

Two miRNA expression datasets consist of GSE97811 and GSE58606 were downloaded from GEO database and analyzed using limma package in R. Both datasets are shown in Table 1. Similar to the identification of DEGs, DEmiRs expression comparison was performed between each breast cancer subtype with breast normal tissues. miRNAs with $|FC| > 1.5$ and $p\text{-value} < 0.05$ were selected as candidate DEmiRs. To identify DEmiRs target genes, miRNet web base tool was utilized [19]. We used all experimental evidence for identification of miRNAs target genes. In this way, we identified RNA-RNA interactions between DEmiRs and interest target genes, which were identified as prognostic markers and DETFs, based on experimental evidence for each breast cancer subtype.

Identification of DETFs and their target genes

TFs with $|FC| > 1.5$ and $p\text{-value} < 0.05$ were selected as DETFs in each subtype of breast cancer. To identify TFs target genes we utilized ChEA and TRANSFAC databases [20] which are based on experimental evidence for identification of TFs target genes. These databases utilized ChIP-X methods to identify the interactions between TFs and the target genes. A p-value less than 0.05 was considered as significant. Therefore, we identified Protein-DNA interactions between the candidate prognostic markers and DETFs based on experimental evidence for each breast cancer subtype. To identify the interaction between DETFs and DEmiRs we used Transmir database [21] which is designed based on experimental evidence.

Gene regulatory network construction and the network analyzing

GRNs based on DETFs, DEmiRs and their target DEGs for each subtype were constructed and visualized in Cytoscape3.4.0 software. Constructed networks were analyzed using Centiscape [22], CytoCtrlAnalyser [23], and MCDS [24] plugins. We used nodes degree, betweenness, classification, control centrality, and MCDS (Minimum Connected Dominating Set) algorithms to find each

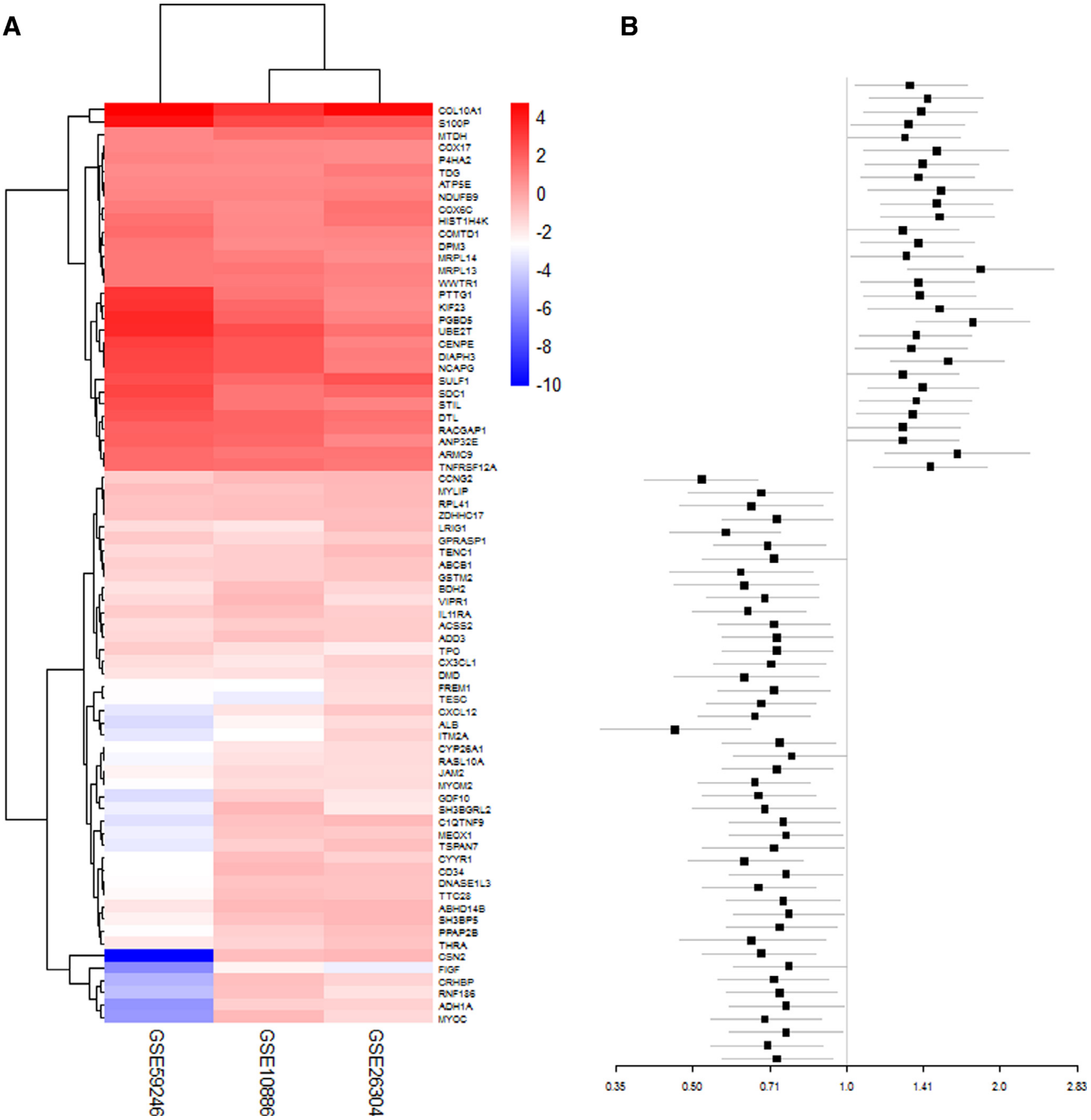

Fig. 2. Basal like potential prognostic markers. **(A)** Heatmap shows the patterns of gene expression for common DEGs in three used datasets. In this figure up-regulated genes are shown at the top and down-regulated genes are shown at the beneath them. **(B)** We portrayed a forest plot to show the overall burden of HR for common DEGs from 879 BL patients in KM-plotter.

GRN key regulators. Node degree is determined based on the number of other nodes that are directly connected to a node [22]. The betweenness of a node measures the number of regulators that control a node over the interactions in the whole GRN [25]. The node classification was used to identify the controllability of a node in the GRN. The classification for a node is defined as "indispensable", "neutral" or "dispensable" which are correlated with increasing, no effect or decreasing the number of the minimum number of driver nodes needed to control the network when a specific node is absent [26]. The control centrality quantifies a node ability in controlling the whole GRN which shows the power of a node in controlling the whole network [27]. MCDS classifies a node as "dominator", "connector" or "none". A dominator node is a node that provides full control over the GRN. A connector is a node which connects the dominators in the GRN [24]. We also used GOlorize plugin of Cytoscape [28] and PANTHER database [29] to attain the interest genes ontology (GO) involved in the constructed GRNs with considering *p*-value level less than 0.05.

Protein-Protein interaction network

We utilized STRING database (https://string-db.org/) to assess the possible protein-protein interactions (PPIs) among the elements of each GRN. In this database, we used experimentally determined interaction with score more than 0.4 and then the results were shown in Cytoscape software.

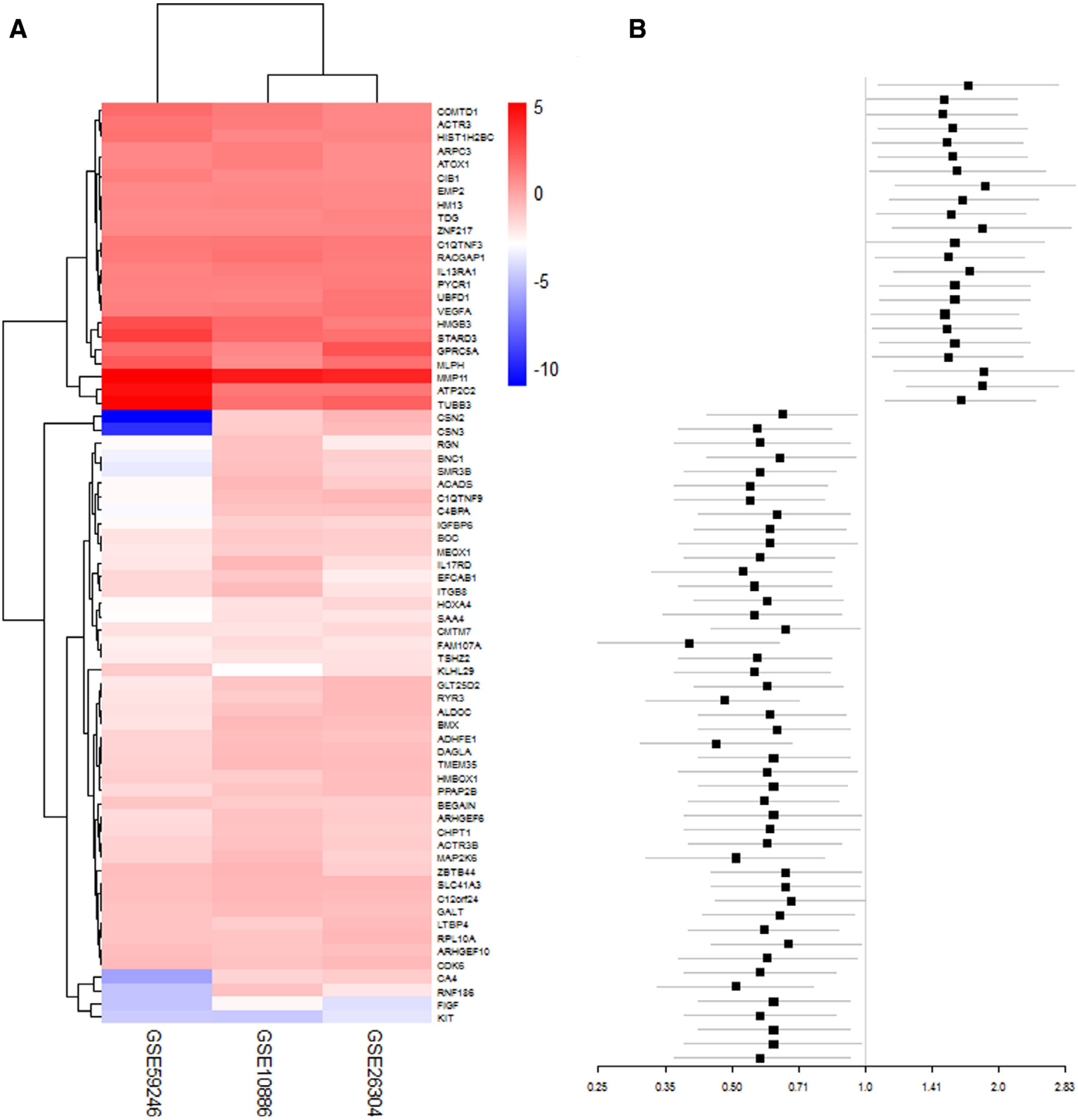

Fig. 3. Her2 enriched potential prognostic markers. (**A**) Heatmap shows the patterns of gene expression for common DEGs in three used datasets. (**B**) Forest plot indicates the overall burden of HR for these common DEGs from 335 Her2 patients in KM-plotter.

Results

Common DEGs involved in each subtype of breast cancer

To reach a set of common DEGs involved in each subtype of breast cancer we performed expression data analysis of 252 tumors integrating mRNA expression profiles of three available datasets. According to our criteria, DEGs were divided into Up- or Down-regulated genes. In this regards, we found 541, 858, 710 and 948 common DEGs in BL, Her2, LuA and LuB breast cancer respectively. Meanwhile, in BL a list of 253 DEGs were up-regulated and 288 DEGs were down-regulated. Also, in Her2 breast cancer tumors we found 402 common up-regulated and 456 common down-regulated DEGs compared to the normal tissues. Subsequently, in LuA and LuB breast cancer patients we found 265 and 363 common DEGs which were up-regulated and a total of 445 and 585 common down-regulated DEGs respectively.

Molecular signatures as potential prognostic markers for subtypes of breast cancer

Using patients survival analysis from 5143 breast cancer patients we highlighted all common DEGs which were negatively associated with overall survival of breast cancer subtypes. Alteration in the expression of these genes exerts a reduction in overall survival of the patients. Therefore, the expression levels of these genes are considered as potential prognostic markers for each subtype of breast cancer.

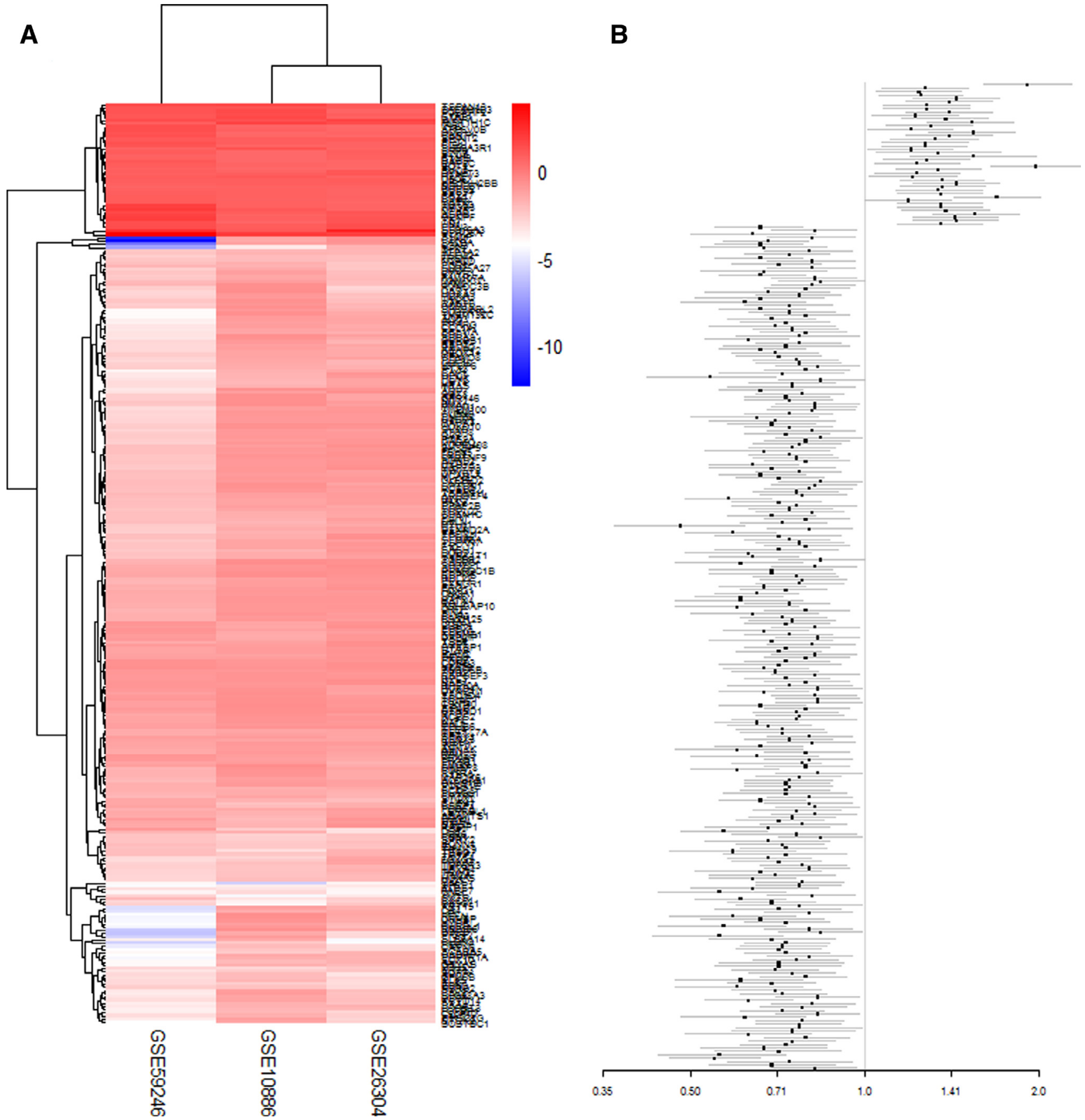

Fig. 4. Luminal A potential prognostic markers. (**A**) Heatmap shows the patterns of gene expression for common DEGs in three used datasets. (**B**) Forest plot shows the overall burden of HR for common DEGs from 2504 LuA patients in KM-plotter.

Potential prognostic markers for Basal like subtype

Analysis of the association of common DEGs with BL breast cancer patients survival rate showed a list of 75 DEGs considered as potential prognostic markers for this type of breast cancer. Among these genes, 30 DEGs were up-regulated with HR > 1 and 45 DEGs were down-regulated with HR < 1 in BL breast cancer patients. In Fig. 2, FC in three datasets and HR for these genes are indicated.

Potential prognostic markers for Her2 enriched subtype

Common DEGs of Her2 subtype showed a list of 69 DEGs with prognostic potential for this type of breast cancer. Among these genes, 23 DEGs were up-regulated with HR > 1 and 46 DEGs were down-regulated with HR < 1 in Her2 breast cancer patients. In Fig. 3, FC and HR for these genes are indicated.

Potential prognostic markers for Luminal A subtype

The association of common DEGs with survival rate of LuA breast cancer showed 291 DEGs with prognostic importance in this subtype of breast cancer. In this regards, we found 42 up-regulated DEGs with HR > 1 and 249 down-regulated DEGs with HR < 1 in LuA subtype. In Fig. 4, FC and HR for the mentioned genes are illustrated.

Potential prognostic markers for Luminal B subtype

LuB patients survival rate analysis showed a list of 293 DEGs as potential prognostic markers which comprised 92 up-regulated

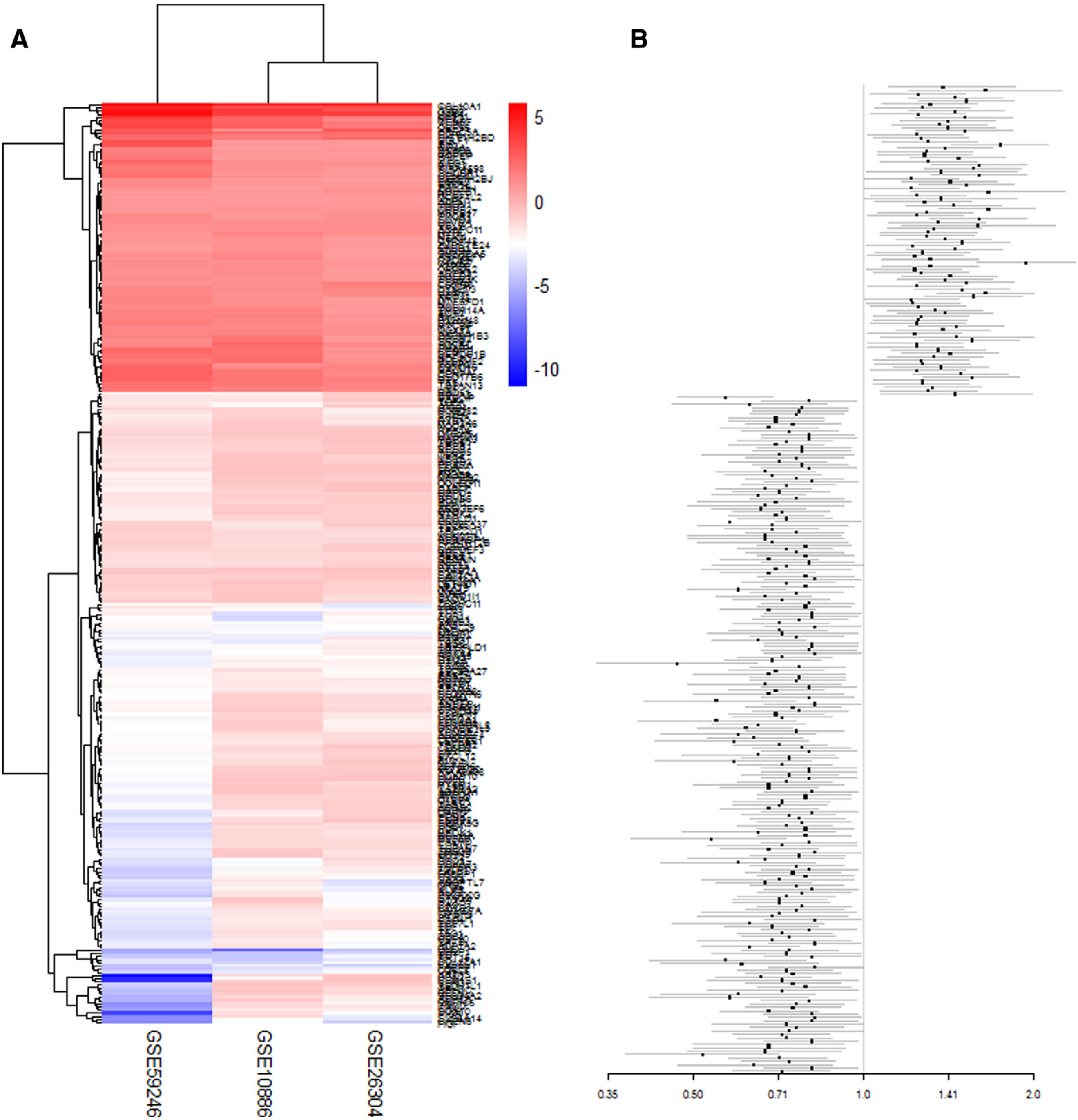

Fig. 5. Luminal B potential prognostic markers. **(A)** Heatmap shows the patterns of gene expression for common DEGs in three used datasets. **(B)** Forest plot shows the overall burden of HR for these common DEGs from 1425 LuB patients in KM-plotter.

DEGs with HR > 1 and 201 down-regulated DEGs with HR < 1 in this subtype of breast cancer patients. FC and HR for these genes are shown in Fig. 5.

Comparison of the prognostic markers among the subtypes of breast cancer

In Fig. 6A venn diagram was depicted to show comparison of the prognostic markers in each subtype of breast cancer. In this regards, we found 5 common prognostic markers among the subtypes of breast cancer. These common prognostic markers are MEOX1, FIGF/VEGF-D, PPAP2B/LPP3, RNF186, and CSN2. The prognostic value of MEOX1 [30], FIGF [31] and PPAP2B [32] have been reported in breast cancer. RNF186 has a role in apoptosis and inflammation of cells. Besides, a higher expression level of RNF186 induces ER stress and impairs insulin signaling pathway [33]. CSN2 gene up-regulation leads to chromosome instability in cells and can lead to VEGF production [34].

Network characteristics

Based on DETFs, DEmiRs and their targets as prognostic markers for each subtype of breast cancer we constructed four GRNs. In our study, BL subtype had 16 DETFs and 173 DEmiRs as regulators of the GRN which were complicated with this subtype. We found 13 DETFs and 130 DEmiRs for Her2 patients. We also found 40 DETFs

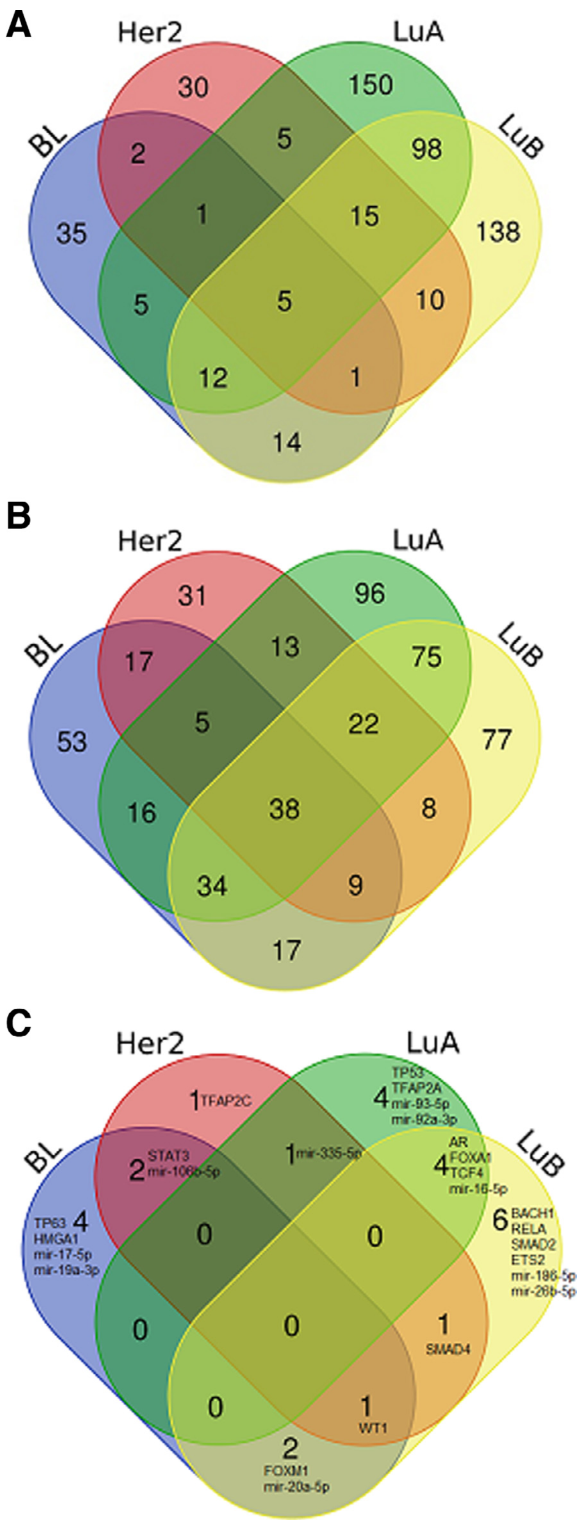

Fig. 6. Venn diagrams indicate for comparison of the prognostic markers, regulators and the key regulators among the subtypes of breast cancer. **(A)** Comparison of the candidate prognostic markers in GRNs among BL, Her2, LuA and LuB subtypes. As shown there are 5 common DEGs and 35, 30, 150 and 138 specific DEGs for BL, Her2, LuA, and LuB respectively. **(B)** Comparison of the regulators (DETFs and DEmiRs) in GRNs among BL, Her2, LuA, and LuB subtypes. As shown there are 38 common regulators and 35, 31, 96 and 77 specific regulators for BL, Her2, LuA, and LuB respectively. **(C)** Comparison of the key regulators (DETFs and DEmiRs) in GRNs among BL, Her2, LuA, and LuB subtypes.

and 259 DEmiRs in patients with LuA breast cancer. In addition, LuB subtype comprised of a list of 30 DETFs and 250 DEmiRs. In this regards the depicted GRNs for BL, Her2, LuA and LuB respectively comprised 630, 504, 3498 and 2922 interactions.

Network GO analysis

Analysis of the gene ontology (GO) showed similar results for LuA and LuB, however, BL and Her2 subtypes showed different results. In this regards, GO analysis of the candidate prognostic markers for LuA and LuB subtypes showed multicellular organism development, developmental process and system development as the most significant biological process. GO analysis for prognostic markers of BL subtype showed regulation of gene expression, regulation of macromolecule biosynthetic process, regulation of cellular process. Besides, Her2 subtype GO analysis showed regulation of chemotaxis, regulation of cell migration, regulation of cellular component movement as the most significant biological process for the candidate prognostic markers.

Regulators of the candidate prognostic markers in each subtype of breast cancer

Comparison of the regulators among the subtypes of breast cancer showed a list of 38 common regulators for these subtypes. We found all intersected regulators among the subtypes which are shown in Fig. 6B. Next, we analyzed each subtype to find the key regulators. In this way, we found 9, 6, 9, and 14 key regulators consist of DETFs and DEmiRs for BL, Her2, LuA, and LuB subtypes respectively. In Fig. 7A–D we depicted all interactions which are complicated with the key regulators in each GRN. The overlapped key regulators among BL, Her2, LuA and LuB subtypes are shown in Fig. 6C.

Basal like key regulators

Using Centiscape, CytoCtrlAnalyser and MCDS plugins we found 5 out of 16 DETFs as key regulators in BL subtype comprised HMGA1, STAT3, WT1, TP63, and FOXM1. Among these DETFs, HMGA1, WT1, and FOXM1 were up-regulated and STAT3 and TP63 were down-regulated. Regarding to DEmiRs we found mir-19a-3p, mir-106b-5p, mir-20a-5p, and mir-17–5p as key DEmiRs which were up-regulated in the depicted GRN (Fig. 7A). Results of the GRN analysis for these TFs and DEmiRs are shown in Table 2.

Her2 enriched key regulators

We found 4 DETFs comprised SMAD4, WT1, STAT3, and TFAP2C as key regulators for Her2 enriched depicted GRN. SMAD4, STAT3, and TFAP2C were down-regulated and WT1 was up-regulated in Her2 breast cancer subtype. Regarding DEmiRs we found mir-335-5p and mir-106b-5p as key DEmiRs in the aforementioned GRN which were down-regulated and up-regulated respectively in Her2 breast cancer (Fig. 7B). The key regulators criteria were shown in Table 3.

Luminal A key regulators

Among 40 DETFs in LuA we found 5 DETFs comprised TFAP2A, TCF4, FOXA1, TP53, and AR as key regulators in this subtype. TCF4 and TP53 were down-regulated and TFAP2A, FOXA1and AR were up-regulated in LuA breast cancer subtype. Regarding to DEmiRs we found mir-335-5p, mir-16-5p, mir-92a-3p and mir-93-5p as key DEmiRs. Among these miRNAs, mir-16-5p and mir-93-5p were up-regulated and mir-335-5p and mir-92a-3p were down-regulated in breast cancer LuA subtype (Fig. 7C). The criteria for these TFs and DEmiRs are shown in Table 4.

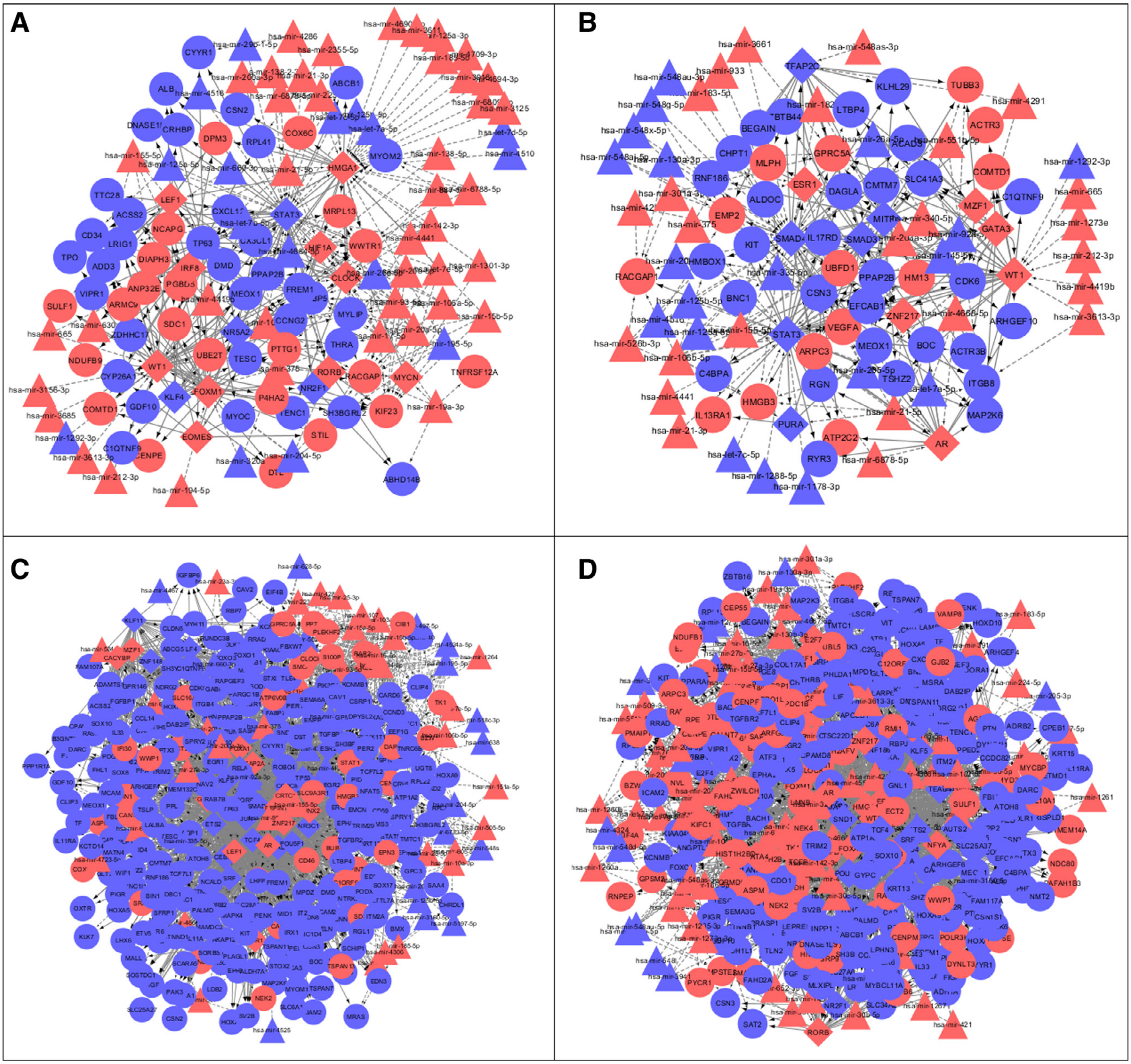

Fig. 7. GRNs involved with the key regulators and their target genes in each subtype. **(A)** All interactions which are complicated with the key regulators in BL subtype. BL key regulators are complicated with 58% of all interactions in the whole GRN. **(B)** All interactions which are complicated with the key regulators in Her2 subtype. Her2 key regulators are complicated with 59% of all interactions in the whole GRN. **(C)** All interactions which are complicated with the key regulators in LuA subtype. LuA key regulators are complicated with 70% of all interactions in the whole GRN. **(D)** All interactions which are complicated with the key regulators in LuB subtype. LuB key regulators are complicated with 77% of all interactions in the whole GRN. (Circular nodes: DEGs; Diamond nods: DETFs; Triangle nodes: DEmiRs; Solid line: Protein-DNA interaction; Dash line: RNA-RNA interaction; Red color: Up-regulated genes; Blue color: Down-regulated genes). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Basal like subtype GRN analysis results based on seven parameters.

Regulators name	Expression pattern	Control centrality	Betweenness	Out degree	In degree	Classification	MCDS
HMGA1	Up	9	1970.25	17	28	Indispensable	Dominator
STAT3	Down	8	2548.87	26	25	Indispensable	Dominator
WT1	Up	8	1043.95	25	10	Indispensable	Connector
TP63	Down	8	1618.41	25	3	Indispensable	Connector
FOXM1	Up	7	957.3	27	5	Indispensable	Dominator
mir-19a-3p	Up	10	0	5	0	Neutral	None
mir-106b-5p	Up	9	0	9	0	Neutral	None
mir-20a-5p	Up	9	0	9	0	Neutral	None
mir-17-5p	Up	9	0	9	0	Dispensable	None

Table 3
Her2 enriched subtype GRN analysis results based on seven parameters.

Regulators name	Expression pattern	Control centrality	Betweenness	Out degree	In degree	Classification	MCDS
SMAD4	Down	7	1279.70	19	24	Indispensable	Dominator
WT1	Up	6	714.39	23	8	Indispensable	Dominator
STAT3	Down	5	1221.58	27	17	Indispensable	Dominator
TFAP2C	Down	4	352.44	14	4	Indispensable	Connector
mir–335–5p	Down	5	0	13	0	Neutral	None
mir-106b-5p	Up	8	0	9	0	Dispensable	None

Table 4
Luminal A subtype GRN analysis results based on seven parameters.

Regulators name	Expression pattern	Control centrality	Betweenness	Out degree	In degree	Classification	MCDS
TFAP2A	Up	25	4371.67	83	19	Indispensable	Connector
TCF4	Down	25	3645.63	145	14	Indispensable	Connector
FOXA1	Up	25	3063.55	82	15	Indispensable	Connector
TP53	Down	25	8159.45	92	34	Indispensable	Dominator
AR	Up	25	6135.09	146	20	Indispensable	Dominator
mir–335–5p	Down	26	0	74	0	Neutral	None
mir–16–5p	Up	26	0	36	0	Neutral	None
mir–92a-3p	Down	26	0	22	0	Neutral	None
mir–93–5p	Up	26	0	35	0	Neutral	None

Table 5
Luminal B subtype GRN analysis results based on seven parameters.

Regulators name	Expression pattern	Control centrality	Betweenness	Out degree	In degree	Classification	MCDS
WT1	Up	14	3500.35	88	15	Indispensable	Connector
RELA	Down	14	2805.65	48	12	Indispensable	Connector
FOXM1	Up	14	2138.64	76	10	Indispensable	Connector
ETS2	Down	14	2937.69	36	10	Indispensable	Connector
FOXA1	Up	14	3577.05	54	11	Indispensable	Connector
AR	Up	14	5482.99	105	17	Indispensable	Connector
SMAD2	Down	14	4274.7	53	23	Indispensable	Dominator
TCF4	Down	14	3785.55	105	10	Indispensable	Dominator
SMAD4	Down	14	5690.66	102	28	Indispensable	Dominator
BACH1	Down	14	5271.40	46	26	Indispensable	Dominator
mir-26b-5p	Up	15	0	44	0	Neutral	None
mir–16–5p	Up	15	0	29	0	Neutral	Dominator
mir-20a-5p	Up	15	0	27	0	Dispensable	None
mir–186–5p	Up	15	0	20	0	Neutral	None

Luminal B key regulators

In addition we found 10 DETFs consist of WT1, RELA, FOXM1, ETS2, FOXA1, AR, SMAD2, TCF4, SMAD4, and BACH1 as key regulators for LuB subtype. Among these TFs SMAD4, TCF4, BACH1, RELA, SMAD2, and ETS2 were down-regulated and AR, WT1, FOXM1 and FOXA1 were up-regulated in LuB breast cancer. Regarding to DEmiRs we found mir-26b-5p, mir-16-5p, mir-20a-5p and mir-186-5p as key DEmiRs which were up-regulated in the aforementioned GRN (Fig. 7D). The criteria for these TFs and DEmiRs are shown in Table 5.

Protein–protein interactions

Protein-Protein Interactions among the elements of each GRN showed some protein complexes particularly in LuA and LuB subtypes. PPI networks for LuA and LuB were shown in Fig. 8. As shown in this figure, CTNNB1 was known as a hub node for both PPI networks. According to the used algorithm, which was based on experimentally determined interaction with score more than 0.4, we found no remarkable PPI network for Her2 and BL subtypes.

Discussion

Main subtypes of breast cancer are currently defined by a combination of morphologic, genomic, and proteomic characteristics. Breast cancer is highly heterogeneous disease which has different risk factors, clinical presentation, histopathologic characteristics, clinical outcome and response to systemic treatments [35].

There are many layers of gene regulation in the cells. Transcription factors (TFs) and microRNAs (miRNAs) are considered as two elements of the best-studied genes regulatory mechanisms [36]. The mechanisms of gene regulation during development and progression of breast cancer is known as a concern among oncologists because gene regulation process is likely comprised of complex genetic interactions. Due to this complexity, computational studies using bioinformatics and systems biology approaches have been done to construct GRNs for various kind of cancer such as breast cancer [37]. GRNs indicate the regulatory relationships among the genes and therefore are necessary to understand gene regulatory mechanisms.

This study aimed to deepen our understanding of gene regulatory mechanisms of breast cancer molecular signatures which their alteration reduce the chance of survival rate for patients in each subtype of breast cancer. In this regards, three mRNA microarray datasets were analyzed to reach a list of common DEGs which were complicated with each subtype of breast cancer. We then tested each common DEG for its association with patients survival rate. Using Kaplan–Meier curve analysis the association of common DEGs with patients survival rate for each subtype of breast cancer was assessed. Therefore, we highlighted a list of prognostic markers for each subtype of breast cancer which up- or down-regulation in these genes reduce the overall sur-

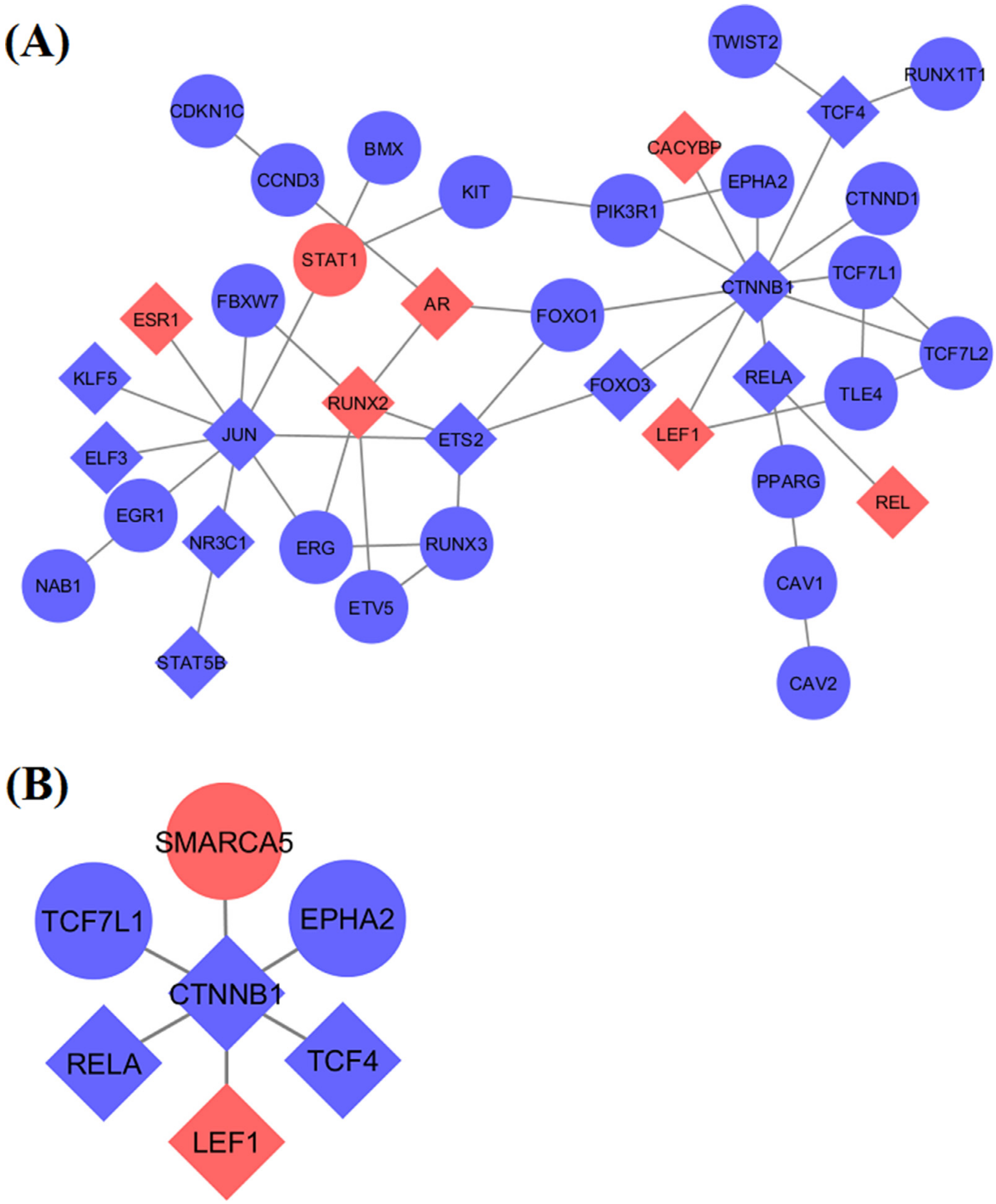

Fig. 8. PPIs between the elements of LuA and LuB GRNs. (A) PPIs for LuA subtype which shows CTNNB1 and JUN as two down-regulated hubs in this network. (B) PPIs for LuB subtype which shows CTNNB1 as a hub. (Circular nodes: DEGs; Diamond nods: DETFs; Red color: Up-regulated genes; Blue color: Down-regulated genes). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

vival rate of the patients. Furthermore, we depicted a GRN for each subtype and analyzed the GRN to find the key regulators. Thus, we elucidated the possible regulatory mechanisms of the candidate prognostic marker in each subtype of breast cancer. GO analysis for candidate prognostic markers showed regulation of gene expression, regulation of macromolecule biosynthetic process, regulation of cellular process were known as the main biological process for BL subtype. The most significant biological process for candidate prognostic marker of Her2 subtype were regulation of chemotaxis, regulation of cell migration, regulation of cellular component movement. We also found multicellular organism development, developmental process and system development as the most significant biological process for both luminal subtypes prognostic markers that indicated similar biological process which are complicated with LuA and LuB subtypes of breast cancer.

Next, we decided to consider only experimentally validated DETFs-DEmiRs-DEGs of interest interactions to construct the GRN. Using bioinformatics and systems biology approaches, and based on experimental evidence, we constructed GRN for each subtype based on DETFs and DEmiRs with their target genes to elucidate the regulatory mechanism of candidate prognostic markers by TFs and miRNAs. Then, using robust approaches and algorithms we analyzed each GRN to find the key regulators. There are many studies on the alteration of these key regulators in breast cancer. In the following part, we will discuss our finding with some of the other studies.

In our current study, we found 9 key regulators for Basal like subtype comprised HMGA1, STAT3, WT1, TP63, FOXM1, mir-19a-3p, mir-106b-5p, mir-20a-5p and mir-17-5p. HMGA1, TP63, mir-19a-3p, and mir-17-5p which were identified as specific key regulators (but not specific regulators) for BL subtype of breast cancer. According to our analysis and other literature HMGA1 and Tp63 play a pivotal role in biological process of Basal like subtype of breast cancer. Studies have shown HMGA1 promotes metastasis in BL breast cancer subtype, the presence of HMGA1 protein has been correlated with a higher cancer grade in mammary epithelial cancer also suggested HMGA1 may be a key player in sustaining breast cancer [38]. It has been also shown that TP63 highly expressed in BL subtype of breast cancer cell line. TP63 gene knockout results in cells proliferation more slowly and adhere less tightly. It was shown TP63 suppress the expression of luminal subtype related genes and maintain the basal phenotype in breast cancer [39].

In our study, we found 6 key regulators comprised SMAD4, WT1, STAT3, TFAP2C, mir-335-5p and mir-106b-5p for Her2 subtype. Among these regulators, TFAP2C was identified as specific key regulator (not a specific regulator) for Her2 subtype of breast cancer. Studies showed TFAP2C transcription factor plays a critical role in gene regulation in multiple breast cancer subtypes. Previous studies demonstrate an important role of TFAP2C in regulation of HER family tyrosine kinase receptors and results in the cell growth, tumorgenesis of HER2 enriched breast cancer subtype [40].

Luminal A network showed 9 DETFs and DEmiRs as key regulators comprised TFAP2A, TCF4, FOXA1, TP53, AR, mir-335-5p, mir-16-5p, mir-92a-3p, and mir-93-5p. TFAP2A, TP53, mir-92a-3p, and mir-93-5p which were identified as specific key regulators (not specific regulators) for LuA subtype in breast cancer. Studies showed FOXA1 and TFAP2A role as regulators of the luminal phenotype associated genes in breast cancer [41]. Studies have also shown the clinical outcome of TP53 overexpression and its usefulness as a prognostic factor in each subtype of breast cancer, especially in luminal A breast cancer [42].

Analysis of the depicted GRN for LuB subtype showed 14 key regulators consist of WT1, RELA, FOXM1, ETS2, FOXA1, AR, SMAD2, TCF4, SMAD4, BACH1, mir-26b-5p, mir-16-5p, mir-20a-5p, and mir-186-5p. Our study demonstrated RELA, ETS2, SMAD2, BACH1, mir-26b-5p, and mir-186-5p were known as specific key regulators (not specific regulators) for LuB subtype. RELA which known as NF-kappa-B p65 subunit may be a useful marker for identifying metastasis-initiating tumor cells [43]. NF-kappaB pathway is required for the formation of luminal breast cancer [44].

In this study, we found some overlapped results between both luminal subtypes of breast cancer. DEGs, network analysis, GO analysis showed LuA and LuB subtypes share some similarities in gene expression and molecular functions in their pathogenesis, however, they are not identical. This finding supports previous studies [45]. There are some similarities between LuA and LuB subtypes. Some evidence suggested that LuB evolves from LuA subtype of breast cancer. However, other evidence suggested that each subtype of breast cancer develop independently of each other [46,47]. In this study, we found 130 overlapped prognostic markers and 169 overlapped regulators (DETFs and DEmiRs) between LuA and LuB (Fig. 6).

As shown in Fig. 6C we found STAT3 and mir-106b-5p as two overlapped key regulators for both BL and Her2 subtypes. Studies showed that STAT3 transcription level was lower in invasive breast cancer compared to the normal tissue [48]. STAT3 has a key role in metastasis of breast cancer and is associated with triple-negative breast cancer [49]. We also found WT1 as a common key regulator for BL, Her2 and LuB subtypes. WT1 plays an oncogenic role in various types of cancer and has prognostic value in breast cancer [50]. WT1 gene expression silencing in breast cancer cell lines results in the inhibition of growth and proliferation of the cancer cells [51,52]. In addition, we found AR, FOXA1, TCF4, and mir-16-5p as overlapped key regulators for both luminal subtypes.

AR (Androgen receptor) contains unique functional domains with relevance to its altered role in human breast cancer. Studies indicate that AR is an emerging hormonal target in breast cancer with potential clinical benefit in both ER positive and negative. The role of AR as an important biomarker in breast cancer was reviewed [53].

We also demonstrated PPIs for elements of LuA and LuB networks which provide insights on the mechanism of their interactions. PPIs form the backbone of signaling pathways and networks in various biological processes. Analysis of PPIs may suggest a new method of targeted therapies [54]. Both LuA and LuB subtypes PPI networks showed CTNNB1 plays as a hub. CTNNB1 (Catenin beta 1) is known as a key downstream element of Wnt signaling pathway [55].

In summary, we introduced a comprehensive transcriptome data analysis and systems biology approaches through the analysis of the expression profile of breast cancer patients to assess a list of DEGs which are associated with lower chance of patients survival rate in each subtype. Then we proposed the possible regulatory mechanisms of these DEGs by TFs and miRNAs for each breast cancer subtype. The approach used in our current study allows us to integrate various data into a GRN for each subtype to find their key regulators. The validation of these findings is valuable in clinical and pathological research in breast cancer. The results of this study suggested some key elements that could be used as potential prognostic markers and/or therapeutic targets in breast cancer patients.

Supplementary material

All data will be made available on request.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Acknowledgment

This research did not receive any external funding.

References

[1] Prat A, Fan C, Fernández A, Hoadley KA, Martinello R, Vidal M, Viladot M, Pineda E, Arance A, Muñoz M. Response and survival of breast cancer intrinsic subtypes following multi-agent neoadjuvant chemotherapy. BMC Med 2015;13:303.

[2] Ambs S. Prognostic significance of subtype classification for short-and long-term survival in breast cancer: survival time holds the key. PLoS Med. 2010;7:e1000281.

[3] Yersal O, Barutca S. Biological subtypes of breast cancer: prognostic and therapeutic implications. World J Clin Oncol 2014;5:412.

[4] Andre F, Pusztai L. Molecular classification of breast cancer: implications for selection of adjuvant chemotherapy. Nat Rev Clin Oncol 2006;3:621.

[5] Inic Z, Zegarac M, Inic M, Markovic I, Kozomara Z, Djurisic I, Inic I, Pupic G, Jancic S. Difference between luminal A and luminal B subtypes according toKi-67, tumor size, and progesterone receptor negativity providing prognostic information. Clin Med Insights 2014;8 CMO. S18006.

[6] Zhang M-Z, Guan Y-X, Zhong J-X, Chen X-Z. Preparation and identification of HER2 radioactive ligands and imaging study of breast cancer-bearing nude mice. Transl Oncol 2017;10:518–26.

[7] Hon JDC, Singh B, Sahin A, Du G, Wang J, Wang VY, Deng F-M, Zhang DY, Monaco ME, Lee P. Breast cancer molecular subtypes: from TNBC to QNBC. Am J Cancer Res 2016;6:1864.

[8] Dai X, Li T, Bai Z, Yang Y, Liu X, Zhan J, Shi B. Breast cancer intrinsic subtype classification, clinical use and future trends. Am J Cancer Res 2015;5:2929.

[9] Ahmadi M, Jafari R, Marashi S, Farazmand A. Indirect role of microRNAs and transcription factors in the regulation of important cancer genes: a network biology approach. Cell Mol Biol 2015;61:100–7.

[10] Afshar AS, Xu J, Goutsias J. Integrative identification of deregulated miRNA/TF-mediated gene regulatory loops and networks in prostate cancer. PLoS ONE 2014;9:e100806.

[11] de Sousa Abreu R, Penalva LO, Marcotte EM, Vogel C. Global signatures of protein and mRNA expression levels. Mol Biosyst 2009;5:1512–26.

[12] Chai LE, Loh SK, Low ST, Mohamad MS, Deris S, Zakaria Z. A review on the computational approaches for gene regulatory network construction. Comput Biol Med 2014;48:55–65.

[13] Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 2009;27:1160.

[14] Muggerud AA, Hallett M, Johnsen H, Kleivi K, Zhou W, Tahmasebpoor S, Amini R-M, Botling J, Børresen-Dale A-L, Sørlie T. Molecular diversity in ductal carcinoma in situ (DCIS) and early invasive breast cancer. Mol Oncol 2010;4:357–68.

[15] Lesurf R, Aure MR, Mørk HH, Vitelli V, Lundgren S, Børresen-Dale A-L, Kristensen V, Wärnberg F, Hallett M, Sørlie T. Molecular features of subtype-specific progression from ductal carcinoma in situ to invasive breast cancer. Cell Rep 2016;16:1166–79.

[16] Matamala N, Vargas MT, González-Cámpora R, Miñambres R, Arias JI, Menéndez P, Andrés-León E, Gómez-López G, Yanowsky K, Calvete-Candenas J. Tumor microRNA expression profiling identifies circulating microRNAs for early breast cancer detection. Clin Chem 2015;61:1098–106.

[17] Hironaka-Mitsuhashi A, Matsuzaki J, Takahashi R-U, Yoshida M, Nezu Y, Yamamoto Y, Shiino S, Kinoshita T, Ushijima T, Hiraoka N. A tissue microRNA signature that predicts the prognosis of breast cancer in young women. PLoS ONE 2017;12:e0187638.

[18] Lánczky A, Nagy Á, Bottai G, Munkácsy G, Szabó A, Santarpia L, Győrffy B. miRpower: a web-tool to validate survival-associated miRNAs utilizing expression data from 2178 breast cancer patients. Breast Cancer Res Treat 2016;160:439–46.

[19] Fan Y, Siklenka K, Arora SK, Ribeiro P, Kimmins S, Xia J. miRNet-dissecting miRNA-target interactions and functional associations through network-based visual analysis. Nucleic Acids Res 2016;44:W135–WW41.

[20] Lachmann A, Xu H, Krishnan J, Berger SI, Mazloom AR, Ma'ayan A. ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. Bioinformatics 2010;26:2438–44.

[21] Wang J, Lu M, Qiu C, Cui Q. TransmiR: a transcription factor–microRNA regulation database. Nucleic Acids Res 2009;38:D119–DD22.

[22] Scardoni G, Tosadori G, Faizan M, Spoto F, Fabbri F, Laudanna C. Biological network analysis with CentiScaPe: centralities and experimental dataset integration. F1000Res 2014;3.

[23] Wu L, Min L, Wang J, Wu F-X. CytoCtrlAnalyser: a Cytoscape app for biomolecular network controllability analysis. Bioinformatics 2017;1:3.

[24] Nazarieh M, Wiese A, Will T, Hamed M, Helms V. Identification of key player genes in gene regulatory networks. BMC Syst Biol 2016;10:88.

[25] Barabasi A-L, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet 2004;5:101.

[26] Vinayagam A, Gibson TE, Lee H-J, Yilmazel B, Roesel C, Hu Y, Kwon Y, Sharma A, Liu Y-Y, Perrimon N. Controllability analysis of the directed human protein interaction network identifies disease genes and drug targets. Proc Natl Acad Sci 2016;113:4976–81.

[27] Liu Y-Y, Slotine J-J, Barabási A-L. Control centrality and hierarchical structure in complex networks. PLoS ONE 2012;7:e44459.

[28] Garcia O, Saveanu C, Cline M, Fromont-Racine M, Jacquier A, Schwikowski B, Aittokallio T. GOlorize: a Cytoscape plug-in for network visualization with Gene Ontology-based layout and coloring. Bioinformatics 2006;23:394–6.

[29] Mi H, Poudel S, Muruganujan A, Casagrande JT, Thomas PD. PANTHER version 10: expanded protein families and functions, and analysis tools. Nucleic Acids Res 2015;44:D336–DD42.

[30] Sun L, Burnett J, Gasparyan M, Xu F, Jiang H, Lin C-C, Myers I, Korkaya H, Liu Y, Connarn J. Novel cancer stem cell targets during epithelial to mesenchymal transition in PTEN-deficient trastuzumab-resistant breast cancer. Oncotarget 2016;7:51408.

[31] Nakamura Y, Yasuoka H, Tsujimoto M, Yang Q, Imabun S, Nakahara M, Nakao K, Nakamura M, Mori I, Kakudo K. Prognostic significance of vascular endothelial growth factor D in breast carcinoma with long-term follow-up. Clin Cancer Res 2003;9:716–21.

[32] Lei FJ, Cheng BH, Liao PY, Wang HC, Chang WC, Lai HC, Yang JC, Wu YC, Chu LC, Ma WL. Survival benefit of sphingosin-1-phosphate and receptors expressions in breast cancer patients. Cancer Med 2018;7:3743–54.

[33] Tong X, Zhang Q, Wang L, Ji Y, Zhang L, Xie L, Chen W, Zhang H. RNF186 impairs insulin sensitivity by inducing ER stress in mouse primary hepatocytes. Cell. Signal 2018;52:155–62.

[34] Lee M-H, Zhao R, Phan L, Yeung S-CJ. Roles of COP9 signalosome in cancer. Cell Cycle 2011;10:3057–66.

[35] Dai X, Xiang L, Li T, Bai Z. Cancer hallmarks, biomarkers and breast cancer molecular subtypes. J Cancer 2016;7:1281.

[36] Chen K, Rajewsky N. The evolution of gene regulation by transcription factors and microRNAs. Nat Rev Genet 2007;8:93.

[37] Ahmad FK, Deris S, Othman N. The inference of breast cancer metastasis through gene regulatory networks. J Biomed Inform 2012;45:350–62.

[38] Pegoraro S, Ros G, Piazza S, Sommaggio R, Ciani Y, Rosato A, Sgarra R, Del Sal G, Manfioletti G. HMGA1 promotes metastatic processes in basal-like breast cancer regulating EMT and stemness. Oncotarget 2013;4:1293.

[39] Orzol P, Nekulova M, Holcakova J, Muller P, Votesek B, Coates PJ. ΔNp63 regulates cell proliferation, differentiation, adhesion, and migration in the BL2 subtype of basal-like breast cancer. Tumor Biol 2016;37:10133–40.

[40] Park J, Wu T, Cyr A, Woodfield G, De Andrade J, Spanheimer P, Li T, Sugg S, Lal G, Domann F. The role of Tcfap2c in tumorigenesis and cancer growth in an activated Neu model of mammary carcinogenesis. Oncogene 2015;34:6105.

[41] Bogachek MV, Chen Y, Kulak MV, Woodfield GW, Cyr AR, Park JM, Spanheimer PM, Li Y, Li T, Weigel RJ. Sumoylation pathway is required to maintain the basal breast cancer subtype. Cancer Cell 2014;25:748–61.

[42] Lee SK, Bae SY, Lee JH, Lee H-C, Yi H, Kil WH, Lee JE, Kim SW, Nam SJ. Distinguishing low-risk luminal a breast cancer subtypes with Ki-67 and p53 is more predictive of long-term survival. PLoS ONE 2015;10:e0124658.

[43] Helbig G, Christopherson KW, Bhat-Nakshatri P, Kumar S, Kishimoto H, Miller KD, Broxmeyer HE, Nakshatri H. NF-κ B promotes breast cancer cell migration and metastasis by inducing the expression of the chemokine receptor CXCR4. J Biol Chem 2003;278:21631–8.

[44] Pratt M, Tibbo E, Robertson S, Jansson D, Hurst K, Perez-Iratxeta C, Lau R, Niu M. The canonical NF-κB pathway is required for formation of luminal mammary neoplasias and is activated in the mammary progenitor population. Oncogene 2009;28:2710.

[45] Ades F, Zardavas D, Bozovic-Spasojevic I, Pugliano L, Fumagalli D, De Azambuja E, Viale G, Sotiriou C, Piccart M. Luminal B breast cancer: molecular characterization, clinical management, and future perspectives. J Clin Oncol 2014;32:2794–803.

[46] Creighton CJ. The molecular profile of luminal B breast cancer. Biologics 2012;6:289.

[47] Allred DC, Brown P, Medina D. The origins of estrogen receptor alpha-positive and estrogen receptor alpha-negative human breast cancer. Breast Cancer Res 2004;6:240.

[48] Wu H-T, Liu J, Li G-W, Shen J-X, Huang Y-T. The transcriptional STAT3 is a potential target, whereas transcriptional STAT5A/5B/6 are new biomarkers for prognosis in human breast carcinoma. Oncotarget 2017;8:36279.

[49] McDaniel JM, Varley KE, Gertz J, Savic DS, Roberts BS, Bailey SK, Shevde LA, Ramaker RC, Lasseigne BN, Kirby MK. Genomic regulation of invasion by STAT3 in triple negative breast cancer. Oncotarget 2017;8:8226.

[50] Qi X-W, Zhang F, Wu H, Liu J-l, Zong B-G, Xu C, Jiang J. Wilms' tumor 1 (WT1) expression and prognosis in solid cancer patients: a systematic review and meta-analysis. Sci Rep 2015;5:8924.

[51] Chesor M, Roytrakul S, Graidist P, Kanokwiroon K. Proteomics analysis of siRNA-mediated silencing of Wilms' tumor 1 in the MDA-MB-468 breast cancer cell line. Oncol Rep 2014;31:1754–60.

[52] Chavaboon Dechsukum M. Growth inhibition of breast cancer cell line MCF-7 by siRNA silencing of Wilm tumor 1 gene. J Med Assoc Thai 2007;90:2416–21.

[53] Iacopetta D, Rechoum Y, Fuqua SA. The role of androgen receptor in breast cancer. Drug Discov Today 2012;9:e19–27.

[54] Li Z, Ivanov AA, Su R, Gonzalez-Pecchi V, Qi Q, Liu S, Webber P, McMillan E, Rusnak L, Pham C. The OncoPPi network of cancer-focused protein–protein interactions to inform biological insights and therapeutic strategies. Nat Commun 2017;8:14356.

[55] Liu Y, Patel L, Mills GB, Lu KH, Sood AK, Ding L, Kucherlapati R, Mardis ER, Levine DA, Shmulevich I. Clinical significance of CTNNB1 mutation and Wnt pathway activation in endometrioid endometrial carcinoma. JNCI 2014;106.