

RESEARCH ARTICLE

Integrated bioinformatics and statistical approaches to explore molecular biomarkers for breast cancer diagnosis, prognosis and therapies

Md. Shahin Alam^{1*}, Adiba Sultana^{1,2}, Md. Selim Reza¹, Md Amanullah^{1,3}, Syed Rashel Kabir⁴, Md. Nurul Haque Mollah^{1*}

1 Bioinformatics Lab (Dry), Department of Statistics, University of Rajshahi, Rajshahi, Bangladesh, **2** Center for Systems Biology, Soochow University, Suzhou, China, **3** Department of Respiratory Medicine, Sir Run Run Shaw Hospital and Institute of Translational Medicine, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China, **4** Department of Biochemistry and Molecular Biology, Rajshahi University, Rajshahi, Bangladesh

* mollah.stat.bio@ru.ac.bd (MNHM); shahin4824@gmail.com (MSA)



OPEN ACCESS

Citation: Alam M.S, Sultana A, Reza M.S, Amanullah M, Kabir SR, Mollah M.NH (2022) Integrated bioinformatics and statistical approaches to explore molecular biomarkers for breast cancer diagnosis, prognosis and therapies. PLoS ONE 17(5): e0268967. <https://doi.org/10.1371/journal.pone.0268967>

Editor: Muhammad Tarek Abdel Ghafar, Tanta University Faculty of Medicine, EGYPT

Received: November 2, 2021

Accepted: May 11, 2022

Published: May 26, 2022

Copyright: © 2022 Alam et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: In this study, we analyzed publicly available third-party data. We downloaded three microarray gene expression profiling datasets from the National Center of Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database with accession numbers GSE53566 [20], GSE119552 [21] and GSE152322 [22] using the weblink <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53566>, <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE119552> and <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE152322>

Abstract

Integrated bioinformatics and statistical approaches are now playing the vital role in identifying potential molecular biomarkers more accurately in presence of huge number of alternatives for disease diagnosis, prognosis and therapies by reducing time and cost compared to the wet-lab based experimental procedures. Breast cancer (BC) is one of the leading causes of cancer related deaths for women worldwide. Several dry-lab and wet-lab based studies have identified different sets of molecular biomarkers for BC. But they did not compare their results to each other so much either computationally or experimentally. In this study, an attempt was made to propose a set of molecular biomarkers that might be more effective for BC diagnosis, prognosis and therapies, by using the integrated bioinformatics and statistical approaches. At first, we identified 190 differentially expressed genes (DEGs) between BC and control samples by using the statistical LIMMA approach. Then we identified 13 DEGs (*AKR1C1*, *IRF9*, *OAS1*, *OAS3*, *SLCO2A1*, *NT5E*, *NQO1*, *ANGPT1*, *FN1*, *ATF6B*, *HPGD*, *BCL11A*, and *TP53INP1*) as the key genes (KGs) by protein-protein interaction (PPI) network analysis. Then we investigated the pathogenetic processes of DEGs highlighting KGs by GO terms and KEGG pathway enrichment analysis. Moreover, we disclosed the transcriptional and post-transcriptional regulatory factors of KGs by their interaction network analysis with the transcription factors (TFs) and micro-RNAs. Both supervised and unsupervised learning's including multivariate survival analysis results confirmed the strong prognostic power of the proposed KGs. Finally, we suggested KGs-guided computationally more effective seven candidate drugs (NVP-BHG712, Nilotinib, GSK2126458, YM201636, TG-02, CX-5461, AP-24534) compared to other published drugs by cross-validation with the state-of-the-art alternatives top-ranked independent receptor proteins. Thus, our findings might be played a vital role in breast cancer diagnosis, prognosis and therapies.

ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE152322, respectively. The necessary meta-drug agents were selected by using the online database GSCALite (<http://bioinfo.life.hust.edu.cn/web/GSCALite/>) and reviewing the published articles (see Table S2 and S3 in [S1 File](#)).

Funding: This work was supported by Rajshahi University Research Project (A-289/5/52/RU/Science-24/2021-2022). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Breast cancer (BC) is one of the most common types of invasive cancers among women according to the World Health Organization (WHO), which affected around 2.3 million women in 2020. It is also the cause of large number of cancer-related deaths among women worldwide [1]. Symptoms of BC include a change in breast shape, dimpling of the skin, nipple discharge, or a red scaly patch of skin, and a lump in the breast [2]. Based on the existing treatment facilities, the average 5-year survival rate with BC is 86%, but BC with distant metastasis, the average 5-year survival rate drops down to 28% [3]. Thus, the performance of existing therapeutic treatments on BC is not yet reach to the satisfactory level. Therefore, in-depth molecular research is essential to explore BC causing more effective biomarkers and candidate drugs.

However, new drug discovery is a tremendous challenging, time consuming and expensive task. The main challenges are to explore drug target proteins (receptors) responsible for diseases and drug agents (small molecules) that can reduce the diseases by the interaction with the target proteins. Genomic biomarkers induced proteins are considered as the key receptors. Transcriptomics analysis is a widely used popular approach to explore genomic biomarkers [4–8]. The repurposing of existing drugs for other diseases could reduce the time and cost compared to de novo drug development. By this time, several authors suggested several sets of genomic biomarkers to explore molecular mechanisms and pathogenetic processes of BC [9–17]. Some of them also suggested candidate drugs for the treatment against BC [18, 19]. However, their published data did not display any common set of receptors and/or drugs, and so far, none of them yet investigated the resistance of their suggested drugs against the independent receptors proposed by others. Obviously, a question may be raised, how a drug can be effective globally for all peoples around the world. Therefore, in this study, our main objectives are (i) computational identification of genomic biomarkers (drug targets) for BC highlighting their functions, pathways and regulatory factors, (ii) exploring genomic biomarker guided candidate drugs for the treatment against BC, and (iii) In-silico validation on the resistance performance of the proposed candidate drugs against the state-of-the-art alternatives top-ranked independent receptors associated with BC published by others.

Materials and methods

To reach the goal of this study, we considered both raw-data (gene expression profiles) and meta-data associated with BC. Integrated bioinformatics and statistical approaches were used to analyze the datasets to explore KGs highlighting their functions, pathways, regulatory factors, prognosis power and repurposable drugs. The pipeline of this study is given in [Fig 1](#).

Data sources and descriptions

Collection of gene expression profiles for exploring KGs. The microarray gene expressions profile dataset with accession number GSE53566 [20] was downloaded from the National Center of Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database. The dataset was generated based on two different BC cell lines (BT-20 and MDA-MB-231), either overexpression (BT-20) or knock-down (MDA-MB-231). Untreated cell lines served as controls. The whole genome expression profiles were consisted of 8 treated (case) and 8 control samples with 41078 probes. To investigate the prognostic performance of KGs unbiasedly, we collected two independent microarray gene expression datasets with accession numbers GSE119552 [21] and GSE152322 [22], respectively. More information about these datasets were given in [S1 Table in S1 File](#).

Collection of meta-drug agents for exploring candidate drugs. We collected meta-drug agents (small molecules) from the online database GSCALite [23] by the significant correlation

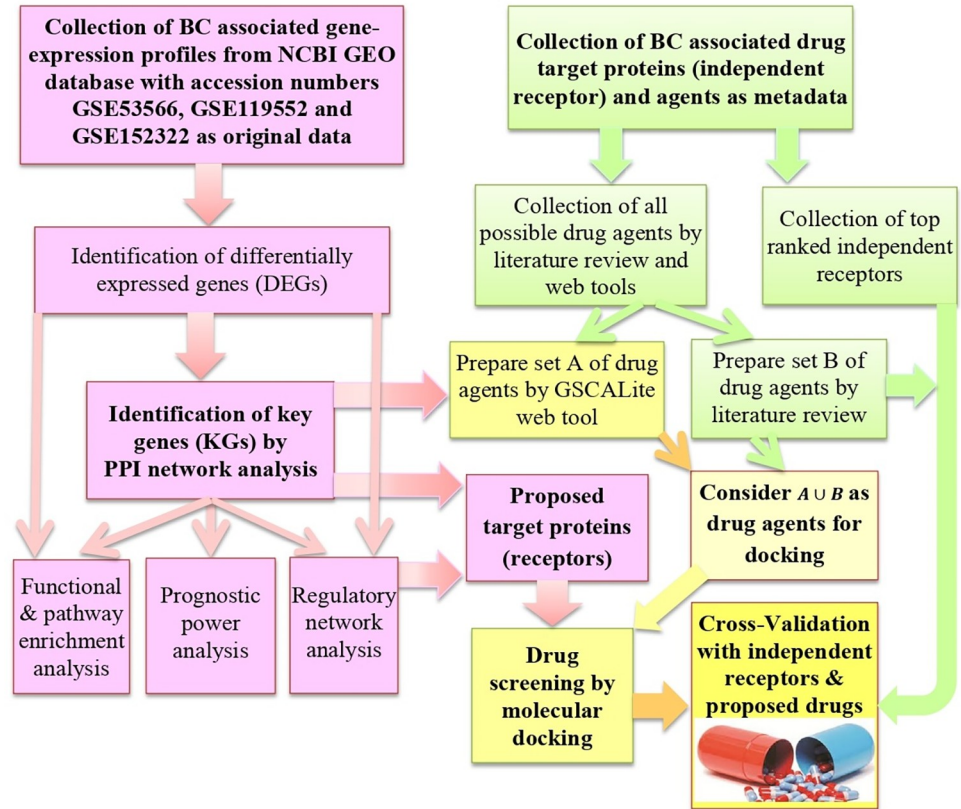


Fig 1. The pipeline of this study.

<https://doi.org/10.1371/journal.pone.0268967.g001>

with our proposed target proteins (set-A) and published articles (set-B) to explore candidate drugs (set-C) by molecular docking with our proposed target proteins (genomic biomarkers). Both A and B sets of meta-drug agents were given in the supplementary file (S2 Table in S1 File).

Collection of independent meta-receptors for cross-validation with the proposed drugs. To select the top-ranked hub-genes (meta-receptors) associated with BC, we reviewed 78 published articles and selected the top-ranked 13 target proteins as the meta-receptors (S3 Table in S1 File).

Identification of DEGs

To identify DEGs between BC and normal conditions, we considered the linear models for microarray (LIMMA) data analysis suggested by Smyth (2004) [24], which can be written as

$$y_g = X\alpha_g + \epsilon_g \tag{1}$$

Where $y_g = (y_{g1}, y_{g2}, \dots, y_{gn})'$ is the vector of expressions (responses) for gth gene with $n = n_1 + n_2$ samples ($g = 1, 2, \dots, m$), X is an $n \times 2$ design matrix, $\alpha_g = (\alpha_{g1}, \alpha_{g2})'$ is 2×1 vector ($2 < n$) of effects for two different groups of n samples and the error vector $\epsilon_g \sim N(0, W_g \sigma_g^2)$. Here W_g is a positive definite weight matrix. We want to test the null hypothesis (H_0): $\alpha_{g1} = \alpha_{g2} = \gamma_g = (\alpha_{g1} - \alpha_{g2}) = 0$ (that is, gth gene is equally expressed gene (EEG) in both case and control groups) against the alternative hypothesis (H_1): $\alpha_{g1} \neq \alpha_{g2} = \gamma_g \neq 0$ (that is, gth gene is differentially expressed gene (DEG) between case and control groups). To test H_0 against H_1 ,

the moderated t-statistic was formulated by hybridizing the classical and Bayesian approaches in which the posterior variance is substituted into to the classical t-statistic in place of the classical sample variance. The moderated t-statistic was defined as,

$$\tilde{t}_g = \frac{\hat{\gamma}_g - \gamma_g}{\tilde{s}_g \sqrt{\delta_g}} \tag{2}$$

which follows t-distribution with d_g+d_0 degrees of freedom under H_0 .

Adjusted *P*-value based on the moderated *t*-statistic and the average of log2 fold-change (aLog₂FC) values of treatment group with respect to the control group were used to select DEGs or EEGs as follows

$$DEG_g = \begin{cases} \text{DEG (Upregulated)}, & \text{if adj.p.value} < 0.01 \text{ and aLog}_2\text{FC}_g > +1.0 \\ \text{EEG}, & \text{if adj.p.value} < 0.01 \text{ and } -1.0 < \text{aLog}_2\text{FC}_g < 1.0 \\ \text{DEG (Downregulated)}, & \text{if adj.p.value} < 0.01 \text{ and aLog}_2\text{FC}_g < -1.0 \end{cases} \tag{3}$$

where

$$\text{aLog}_2\text{FC}_g = \begin{cases} \frac{1}{n_1} \sum_i^{n_1} \log_2(y_{gi}^T) - \frac{1}{n_2} \sum_j^{n_2} \log_2(y_{gj}^C), & \text{if } n_1 \neq n_2 \\ \frac{1}{n} \sum_i^n \log_2\left(\frac{y_{gi}^T}{y_{gi}^C}\right), & \text{if } n_1 = n_2 = n. \end{cases}$$

Here y_{gi}^T and y_{gj}^C are the expressions for the *g*th gene with the *i*th treatment and *j*th control samples, respectively. We implemented the limma R-package [25] for calculating the *P*-values and aLog₂FC values to select the DEGs, significantly.

Construction of PPI network of DEGs

Protein-protein interaction (PPI) network was constructed to identify key-genes (KGs). The online STRING-v11 database [26] was used to construct the PPI network of DEGs. The STRING database provides critical assessment and integration of protein interactions, including direct (physical) and indirect (functional) associations. To construct PPI network, the distance ‘D’ between pair of proteins (u,v) is calculated as

$$D(u, v) = \frac{2|N_u \cap N_v|}{|N_u| + |N_v|} \tag{4}$$

Where N_u is the neighbor set of *u* and N_v is the neighbor set of *v*. Cytoscape plug-in cyto-Hubba is used to rank the nodes of PPI network for identifying KGs in the network [27, 28]. In the present study five topological methods including Degree [29], BottleNeck [30], Betweenness [31], Stress [32], and Clustering Coefficient was utilized to identify KGs.

GO terms and KEGG pathway enrichment analysis of DEGs highlighting KGs

The GO (Gene Ontology) functions and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway enrichment analysis were performed to understand the pathogenetic processes and pathways of DEGs highlighting KGs. The GO terms have three categories: Biological Process (BP), Cellular Component (CC), and Molecular Function (MF). To explore the significantly

Table 1. Contingency table.

Annotated Gene-sets	DEGs (proposed)	EEGs (proposed)	Marginal total (Annotated)
i^{th} GO term/KEGG pathway (S_i)	k_i	$M_i - k_i$	M_i
Complement of S_i (S_i^c)	$n - k_i$	$N - M_i - n + k_i$	$N - M_i$
Marginal total	n	$N - n$	N (Grand total)

<https://doi.org/10.1371/journal.pone.0268967.t001>

enriched GO terms and KEGG pathways by DEGs including KGs, let S_i is the annotated gene-set corresponding to i^{th} type of biological functions or pathways given in the database and M_i is the number of genes in S_i ($i = 1, 2, \dots, r$); N is the total number of annotated genes those construct the entire combine set $S = \bigcup_{i=1}^r S_i = S_i \cup S_i^c$ such that $N \leq \sum_{i=1}^r M_i$; where S_i^c is the complement set of S_i . Again let n is the total number of DEGs of interest and k_i is the number of DEGs belonging to the annotated gene-set S_i . This problem is summarized by the following contingency table (Table 1).

To find the significantly enriched GO terms and KEGG pathways by our proposed DEGs, the P -value was calculated by the Fisher exact test statistic based on hypergeometric distribution. We used DAVID online tool (version 6.8) to perform Fisher exact test [33].

Regulatory network analysis of KGs

To identify key transcription factors (TFs) as the transcriptional regulators of KGs, the TFs-KGs interaction network was constructed using the publicly available database JASPAR [34]. The interaction network was generated using NetworkAnalyst [35]. To identify key microRNAs (miRNAs) as the post-transcriptional regulators of KGs, the KGs-miRNAs interaction network was constructed by using the publicly available online tool miRNet 2.0 [36]. The top degree miRNAs were selected from the network and considered them as key miRNAs.

Prognostic power analysis of KGs

To investigate the prognostic power of KGs, we performed cluster analysis, survival analysis and developed two prediction models using random forest (RF) and support vector machine (SVM) classifiers. The survival curve and ROC curve were used to assess the prognosis performance. The online SurvExpress computational tool [37] was used to produce survival curve. The R-packages 'gplots' and 'ROCR' were used to produce heatmap and ROC curve, respectively.

Molecular docking simulation for exploring candidate drugs

To propose *in-silico* validated effective drugs for the treatment against BC, we employed molecular docking simulation between the target receptor proteins and drug agents. We considered our proposed KGs based hub-proteins and associated TFs proteins as the drug target receptor proteins and meta-drug agents collected from online databases and published articles for docking analysis. The molecular docking simulation requires 3-Dimensional (3D) structures of both receptor proteins and candidate drugs. We downloaded 3D structure of all targeted receptor proteins from Protein Data Bank (PDB) [38] and SWISS-MODEL [39]. The 3D structures of drug agents were downloaded from PubChem and DrugBank database [40, 41]. The 3D structure of the target proteins was visualized using Discovery Studio Visualizer 2020 and the water molecules, co-crystal ligands which were bound to the protein were removed. Further, the protein was prepared using USCF Chimera and Autodock vina¹³⁶⁵¹ in PyRx open

source software by adding charges and minimizing the energy of the protein and subsequently converting it to pdbqt format [42–44]. The exhaustiveness parameter was set to 8. The Protein-Ligand Interaction Profiler (PLIP) web service [45] and PyMol was used to analyze the docked complexes for surface complexes, types and distances of non-covalent bonds. Let A_{ij} denotes the binding affinity between i^{th} target protein ($i = 1, 2, \dots, m$) and j^{th} drug agent ($j = 1, 2, \dots, n$). Then target proteins are ordered according to the descending order of row means $\sum_{j=1}^n A_{ij}/m, j = 1, 2, \dots, m$, and drug agents are ordered according to the descending order of column means $\sum_{i=1}^m A_{ij}/n, j = 1, 2, \dots, n$, to select the top ranking few drug agents as the candidate drugs. Then we validated the proposed candidate drugs by molecular docking simulation with the top ordered independent receptors associated with BC published by others.

Results

Identification of DEGs

We identified 190 DEGs, including 138 downregulated and 52 upregulated genes (S1 Table in S2 File) in BC tissue, using $\text{adj.}P.\text{Val} < 0.01$ and $\log\text{FC} > 1$ as the threshold for upregulated DEGs, and $\text{adj.}P.\text{Val} < 0.01$ and $\log\text{FC} < -1$ for downregulated DEGs. The upregulated and downregulated DEGs were displayed on the right and left sides respectively in the volcano plot by the green color in Fig 2A. A heatmap was constructed to show the clustering performance of case and control samples by the up and down regulated DEGs in Fig 2B. We observed that both DEGs and samples separated each other between their contrast groups accurately.

Identification of key genes (KGs) from DEGs

To identify KGs, the PPI network of DEGs was constructed which includes 180 nodes and 218 edges, with an average node degree 2.42 and $P\text{-value} < 1.0e-16$. In the PPI network, pink color

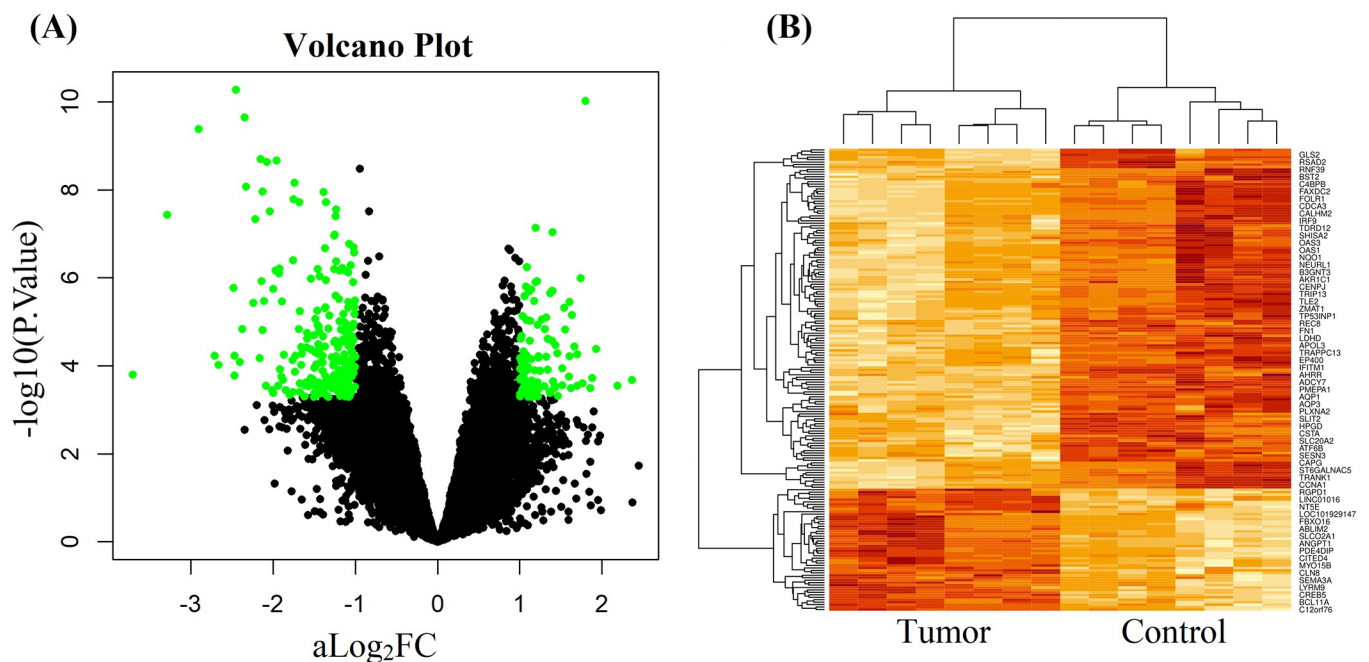


Fig 2. (A) Volcano plot of $-\log_{10}(P\text{-value})$ against $\log_2\text{FC}$ values to display significantly upregulated and downregulated DEGs. (B) Heatmap of significantly upregulated and downregulated DEGs to observe the clustering performance of tumor and control groups by hierarchical clustering approach.

<https://doi.org/10.1371/journal.pone.0268967.g002>

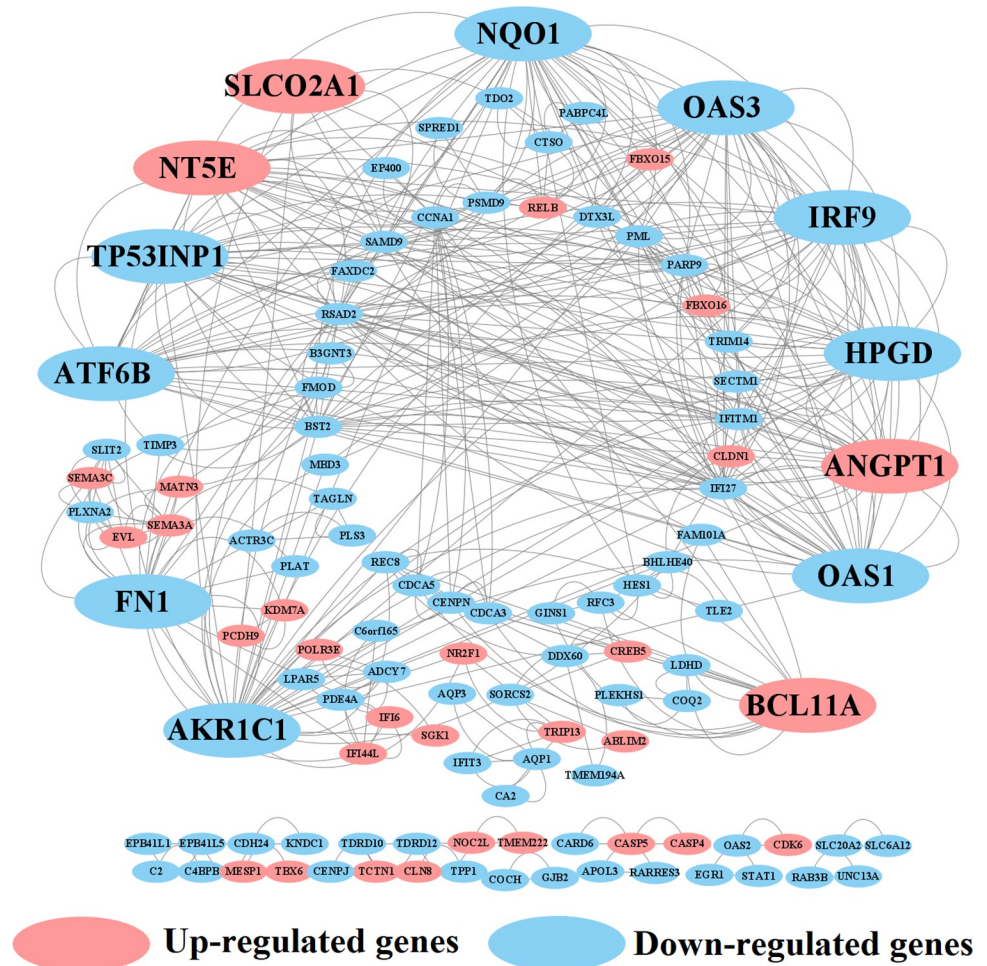


Fig 3. Protein-protein interaction (PPI) network of DEGs to select the key genes (KGs). Blue color indicates downregulated and pink color indicates upregulated DEGs, big size and octagon shape indicate the KGs.

<https://doi.org/10.1371/journal.pone.0268967.g003>

indicates up regulated and blue color indicates down regulated DEGs, big size and octagon shape indicate KGs (see Fig 3). We used five topological measures (Degree, BottleNeck, Betweenness, Stress and Clustering Coefficient) to select top-ranked 13 KGs that are *AKR1C1*, *IRF9*, *OAS1*, *OAS3*, *SLCO2A1*, *NT5E*, *NQO1*, *ANGPT1*, *FN1*, *ATF6B*, *HPGD*, *BCL11A*, and *TP53INP1*, where 4 KGs (*SLCO2A1*, *NT5E*, *BCL11A* and *ANGPT1*) were upregulated and the rest 9 KGs were downregulated (Table 2). Further information of 13 kg is included in (S2 Table in S2 File).

GO terms and KEGG pathway enrichment analysis of DEGs highlighting KGs

The GO functional enrichment analysis of DEGs showed that 46 GO-BP terms, 11 GO-CC terms and 14 GO-MF terms are enriched by the downregulated genes, where KGs were involved with 25 BPs, 8 CCs and 10 MFs. On the other hand, 15 BPs, 3 CCs and 4 MFs are enriched by the upregulated genes, where KGs were directly involved with 5 BPs (S3 Table in S2 File). Among the enriched GO functions including downregulated KGs, 5 GO-BP terms (GO:0060337~type I interferon signaling pathway, GO:0051607~defense response to virus,

Table 2. Selection of KGs by taking the union of five-sets of top-ranked 8 genes produced by five topological measures with the PPI network.

Degree (D)	BottleNeck (E)	Betweenness (F)	Stress (G)	Clustering Coefficient (H)	Key genes (<i>DUJEUJGUH</i>)
<i>STAT1</i>	<i>FN1</i>	<i>EGR1</i>	<i>EGR1</i>	<i>SLCO2A1</i>	<i>OAS1, FN1, SLCO2A1, HPGD, IRF9, NQO1, AKR1C1, OAS3, NT5E, TP53INP1, ANGPT1, ATF6B, BCL11A</i>
<i>EGR1</i>	<i>EGR1</i>	<i>FN1</i>	<i>FN1</i>	<i>HPGD</i>	
<i>OAS1</i>	<i>STAT1</i>	<i>STAT1</i>	<i>STAT1</i>	<i>TIMP3</i>	
<i>OAS2</i>	<i>REC8</i>	<i>REC8</i>	<i>REC8</i>	<i>NQO1</i>	
<i>IRF9</i>	<i>TRIP13</i>	<i>TRIP13</i>	<i>TRIP13</i>	<i>AKR1C1</i>	
<i>OAS3</i>	<i>NT5E</i>	<i>NT5E</i>	<i>NR2F1</i>	<i>CENPN</i>	
<i>RSAD2</i>	<i>NR2F1</i>	<i>NR2F1</i>	<i>ATF6B</i>	<i>TP53INP1</i>	
<i>IFIT3</i>	<i>RELB</i>	<i>RELB</i>	<i>BCL11A</i>	<i>ANGPT1</i>	

<https://doi.org/10.1371/journal.pone.0268967.t002>

GO:0060333~interferon-gamma-mediated signaling pathway and GO:0055114~oxidation-reduction process), 2 GO-CC terms (GO:0005737~cytoplasm and GO:0005578~proteinaceous extracellular matrix), and 1 GO-MF terms (GO:0016491~oxidoreductase activity) were reported by other researchers that association with BC (see [Table 3](#) and discussion section for more details). The upregulated KGs involving 2 GO-BP terms (GO:0045944~positive regulation of transcription from RNA polymerase II promoter and GO:0000122~negative regulation of transcription from RNA polymerase II promoter), 1 GO-CC terms (GO:0072559~NLRP3 inflammasome complex) and 1 GO-MF terms (GO:0097153~cysteine-type endopeptidase activity involved in apoptotic process) were also reported by other researchers that association with BC (see [Table 3](#) and discussion section for more details as before). The KEGG pathway enrichment analysis of DEGs showed that 8 and 2 pathways are enriched by the downregulated and upregulated KGs, respectively. Among them, downregulated KGs involving hsa05168: Herpes simplex infection pathway and upregulated KGs involving hsa04151:PI3K-Akt signaling pathway were also reported by other researchers as the pathways of BC development ([Table 3](#)).

Regulatory network analysis of KGs

We constructed KGs versus transcription factors (KGs-TFs) interaction network to identify top ranking TFs as the key transcriptional regulators of KGs. We selected top 4 key TFs (FOXC1, FOXL1, JUN, and GATA2) as the vital transcriptional regulators of KGs with degree > 4, where large blue ellipses indicate top degree key TFs and pink octagons indicate KGs in [Fig 4A](#). To identify top ranking micro-RNA (miRNA) as the key post-transcriptional regulators of KGs, we constructed KGs-miRNAs interaction network. We selected top 4 key miRNAs (hsa-miR-27a-5p, hsa-miR-124-3p, hsa-miR-1-3p, and hsa-miR-210-3p) as the vital regulators of KGs with degree > 7, where large blue ellipses indicate top degree key miRNAs and pink octagons indicate KGs in [Fig 4B](#).

Prognostic power analysis

We considered both supervised and unsupervised learning's including multivariate survival analysis to investigate the prognostic power of 13 KGs ([Fig 5](#)). The [Fig 5A](#) shows that KGs are able to classify case and control samples accurately by the unsupervised hierarchical clustering (HC). The multivariate survival curves based on the expressions of 13 KGs, separated the low (control) and high (BC) risk groups significantly (see [Fig 5B](#)). In the case of supervised learning, we trained two popular classifiers (RF and SVM) by taking the expression profiles of 13 KGs from all samples (8 BC and 8 controls) of the study dataset with the NCBI accession

Table 3. Significantly enriched GO functions and KEGG pathways by the DEGs involving KGs that were also supported by the literature review about their association with BC and other cancers.

GO Terms/Functions	DEGs (Counts)	P-Value	Associated KGs
Downregulated DEGs			
GO Terms of Biological Processes (BPs)			
GO:0060337~type I interferon signaling pathway [46]	12	8.41E-13	<i>OAS1, OAS3, IRF9</i>
GO:0051607~defense response to virus [47]	12	2.72E-08	<i>OAS1, OAS3, IRF9</i>
GO:0060333~interferon-gamma-mediated signaling pathway [47]	6	1.53E-04	<i>OAS1, OAS3, IRF9, ATF6B</i>
GO:0045071~negative regulation of viral genome replication [48]	5	1.86E-04	<i>OAS1, OAS3</i>
GO:0055114~oxidation-reduction process [49]	9	0.06191	<i>NQO1, HPGD, AKR1C1</i>
GO Terms of Cellular Components (CCs)			
GO:0005737~cytoplasm [50]	57	4.91E-05	<i>OAS1, OAS3, IRF9, HPGD, TP53INP1, NQO1</i>
GO:0005578~proteinaceous extracellular matrix [8]	7	0.009721	<i>FN1</i>
GO Terms of Molecular Function (MF)			
GO:0016491~oxidoreductase activity [51]	5	0.032918	<i>HPGD, AKR1C1</i>
Upregulated DEGs			
GO Terms of BP			
GO:0045944~positive regulation of transcription from RNA polymerase II promoter [52]	7	0.014821	<i>BCL11A, SLCO2A1</i>
GO:0000122~negative regulation of transcription from RNA polymerase II promoter [53]	6	0.015973	<i>BCL11A, NT5E</i>
GO Terms of CC			
GO:0072559~NLRP3 inflammasome complex [54]	2	0.017428	<i>FN1</i>
GO Terms of MF			
GO:0097153~cysteine-type endopeptidase activity involved in apoptotic process [55]	2	0.02663	<i>FN1</i>
KEGG pathways			
Downregulated DEGs			
hsa05168:Herpes simplex infection [56]	6	0.012895	<i>OAS1, OAS3, ATF6B, IRF9</i>
Upregulated DEGs			
hsa04151:PI3K-Akt signaling pathway [57]	4	0.030138	<i>ANGPT1, SLCO2A1, BCL11A</i>

<https://doi.org/10.1371/journal.pone.0268967.t003>

number GSE53566. Then we investigated their prediction performance using both training and independent test datasets. We investigated the training performance by taking all samples (8 BC and 8 controls). To investigate their test performance unbiasedly, we considered the expression profiles of 13 KGs from two independent GEO datasets with the NCBI accession numbers GSE119552 and GSE152322, respectively. The dataset GSE119552 consisted of 12 tumors and 4 control samples, the other test dataset GSE152322 consisted of 11 tumors and 12 control samples. We classified all samples from each dataset by the prediction models. Fig 5C showed the training performance (green color) and independent test performance (blue and red color). We observed that both training and independent test performance are good and reasonable (AUC>0.90) for each of training and independent test datasets.

Exploring candidate drugs by molecular docking analysis

To explore candidate drugs for breast cancer (BC), we considered 13 KGs based proteins (*AKR1C1, IRF9, OAS1, OAS3, SLCO2A1, NT5E, NQO1, ANGPT1, FN1, ATF6B, HPGD, BCL11A, and TP53INP1*) and its regulatory key 4 TFs proteins (*FOXC1, GATA2, FOXL1 and JUN*) as the $m = 17$ drug target receptors. The 3-Dimension (3D) structure of *NT5E, HPGD, NQO1, OAS1, ANGPT1, IRF9, BCL11A, SLCO2A1, FN1, OAS3, AKR1C1, JUN* and *GATA2* were downloaded from Protein Data Bank (PDB) with the PDB codes 6S7F, 2GDZ, 5FU, 4RWP, 4JYO, 5OEN, 6KI6, 3MRR, 2HAZ, 4S3N, 3C3U, 1A02 and 5O9B and rest of them

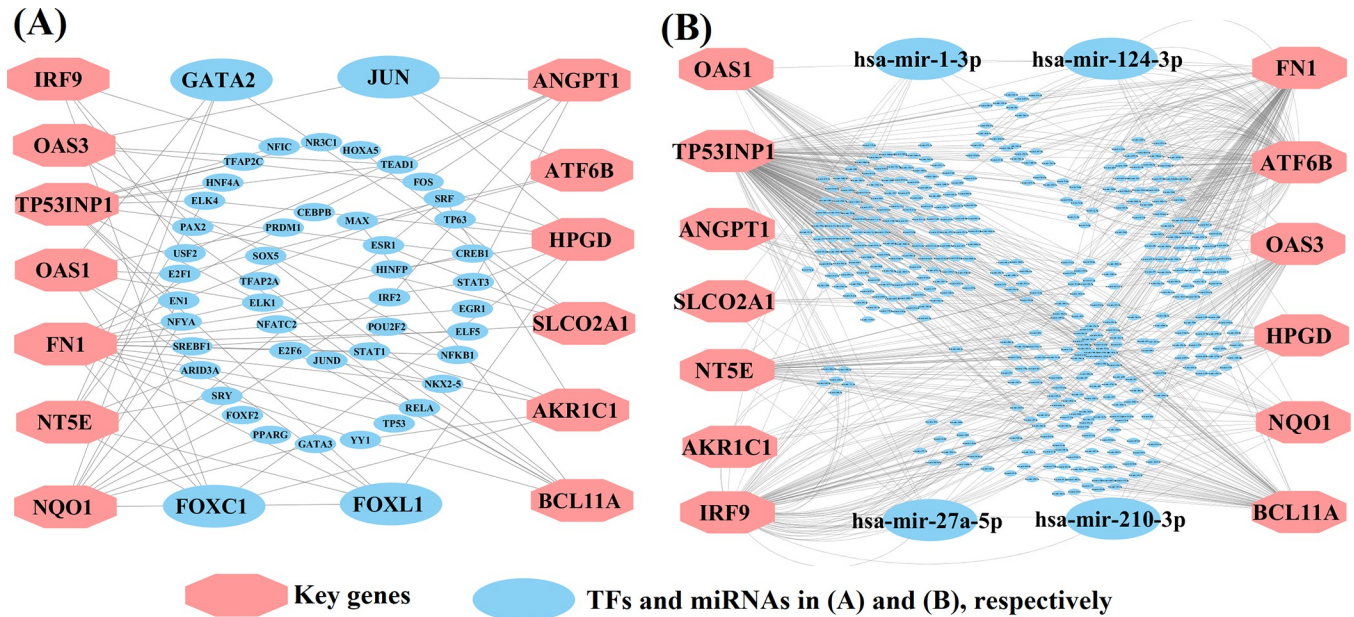


Fig 4. KGs regulatory network analysis results (A) KGs-TFs interaction network to identify key transcriptional regulators of KGs, (B) KGs-miRNAs interaction network to identify key post-transcriptional regulators of KGs. Here pink color octagon indicates the KGs in both A and B, blue color bigger size ellipse indicates key TFs in A and key miRNAs in B.

<https://doi.org/10.1371/journal.pone.0268967.g004>

such as ATF6B, TP53INP1, FOXC1, and FOXL1 targets were downloaded from SWISS-MODEL using UniProt with IDs Q99941, Q96A56, Q12948 and Q12952 respectively. Then we considered 82 meta-drug molecules from the GSCALite database and 47 meta-drugs from the published articles as drug agents (see S1 File (Tables 2 and 3)). The 3D structures of drug agents were downloaded from the PubChem database. Then we performed molecular docking analysis between our proposed receptors and meta-drug agents. The binding affinity score matrix between the ordered receptors and ordered drug-agents were displayed in Fig 6A. We observed that top order four lead compounds/drugs (NVP-BHG712, Nilotinib, GSK2126458, and YM201636) produce highly significant binding affinity scores with all $m = 17$ target proteins, and their average binding affinity scores across all receptors were -8.65, -8.55, -8.50, and -8.45 (kcal/mol), respectively. The next two top ordered drugs (TG-02 and CX-5461) produced highly significant binding affinity scores with 16 target proteins, and their average binding affinity scores across all $m = 17$ targets were -8.40 and -8.22, respectively. The 7th top ordered drug AP-24534 produced significant binding affinity scores with 14 target proteins and the average binding affinity score was -8.0. The other drugs (lead compounds) produced significant binding affinity scores with less than 12 target proteins out of 17 and their average binding affinity scores were negatively smaller than -7.5. Therefore, we considered top ordered seven drugs (NVP-BHG712, Nilotinib, GSK2126458, YM201636, TG-02, CX-5461 and AP-24534) as the candidate drugs in our study and highlighted them in Fig 6B. We also examined their complete interaction profile including hydrogen bonds, hydrophobic, halogen/ salt Bridge and electrostatic interactions in Fig 7.

Performance investigation of proposed drugs by cross-validation

To investigate the resistance performance of the proposed drugs against the state-of-the-art alternative receptors for BC compared to the transcriptome-guided 47 published drugs, we performed molecular docking analysis of our proposed drugs including all published drugs

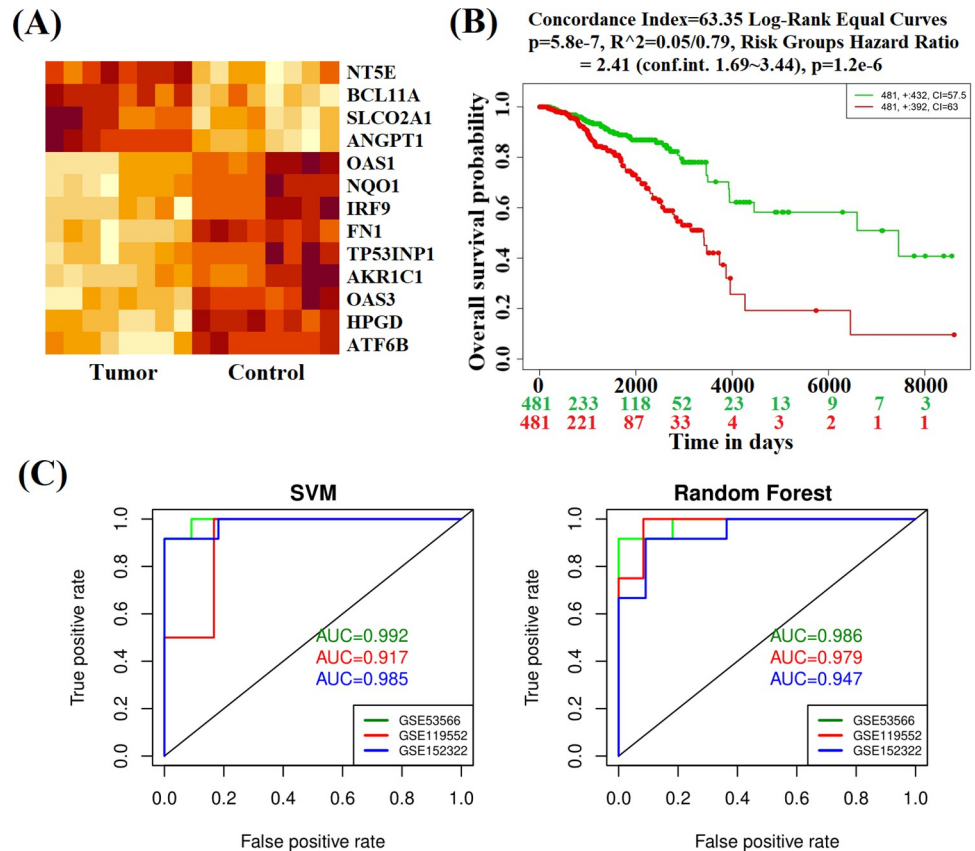


Fig 5. The prognostic powers of KGs were displayed by (A) Heatmap of hierarchical clustering (B) Multivariate survival curves with KGs and (C) ROC curves of prediction models with KGs.

<https://doi.org/10.1371/journal.pone.0268967.g005>

with the top ranked independent receptors (KIF11, RRM2, BUB1, CDC20, FOS, FN1, BUB1B, CCNB2, CCNA2, CDK1, TOP2A, CCNB1, and EGFR) published by others for BC in different 78 articles (see [S1 File \(Table 3\)](#)). The 3D structure of KIF11, RRM2, BUB1, CDC20, FOS, FN1, BUB1B, CCNA2, CDK1, TOP2A, CCNB1, and EGFR were downloaded from PDB database with the PDB codes 1Q0B, 3BS9, 4a1g, 1DUJ, 1FXL, 2HAZ, 2WVI, 1VIN, 6GU6, 1ZXM, 2B9R, and 3G5Z respectively and for another one CCNB2 downloaded from SWISS-MODEL using UniProt with ID O95067. The [Fig 6C](#) showed the resistance performance of our proposed drugs in a comparison of the publicly available drugs against the top ranked 13 independent receptors. We observed that our proposed drugs showed better performance compare to the published drugs in terms of negatively highest binding affinities with the independent receptors. Therefore, we can strongly recommend that the proposed drugs might be more effective candidate than the published drugs for the treatment against BC.

Discussion

In this study, we identified key genomic biomarkers highlighting their pathogenetic processes for breast cancer (BC) diagnosis, prognosis and therapies. At first, we identified 190 DEGs (138 downregulated and 52 upregulated) from the publicly available microarray gene-expression profiles. Then we detected 13 DEGs (*AKR1C1*, *IRF9*, *OAS1*, *OAS3*, *SLCO2A1*, *NT5E*, *NQO1*, *ANGPT1*, *FN1*, *ATF6B*, *HPGD*, *BCL11A*, and *TP53INP1*) as the KGs that drive the progression of BC. Some literatures also suggested that these KGs are BC causing genes [9, 58–78]

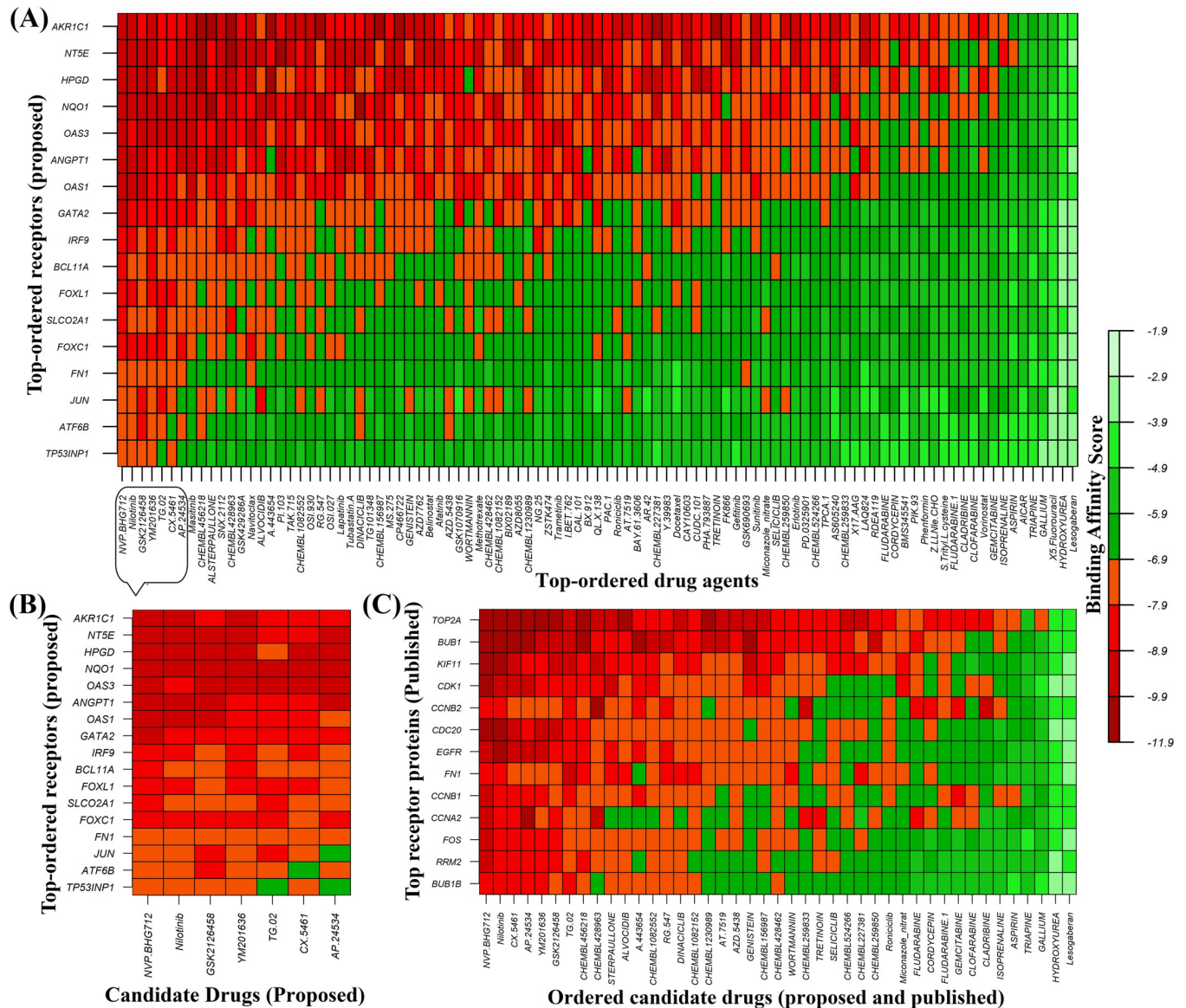


Fig 6. Molecular docking simulation results for exploring candidate drugs against BC. (A) Image of binding affinity scores of proposed ordered receptor proteins with the top 97 ordered meta-drug agents, (B) Image of binding affinity scores of proposed ordered receptor proteins with the proposed ordered candidate drugs only (C) Image of binding affinity scores of ordered proposed and already published candidate drugs against the top-ranked independent receptors published by others.

<https://doi.org/10.1371/journal.pone.0268967.g006>

(see Fig 8A). For example, the expression of two genes (AKR1C1 and AKR1C2) in carcinoma cells and stromal fibroblasts and their positive correlation are favorable tumor characteristics in primary BC patients [58]. Also these two genes appear to be an interesting target for new hormone-based therapy strategies in primary BC. The *IRF9* gene with overexpression has the potential to be a surrogate marker of response and may be associated with drug resistance for BC [59]. The expression of the *OAS1* gene that was inversely associated with multiple MSGs in the BC cell line [60]. The *OAS3* gene plays a prognostic role in BC patients with potential mechanical value [61]. The *SLCO2A1* is ubiquitously expressed and marked as a prostaglandin transporter due to its high affinity [62]. The Gene *NT5E* is regulated epigenetically in BC, the epigenetic status of this gene influencing metastasis and clinical outcome, and suggests that

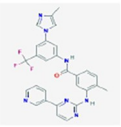
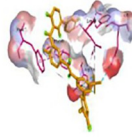
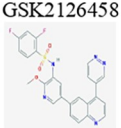
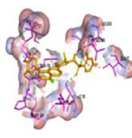
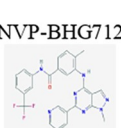
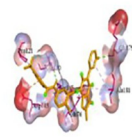
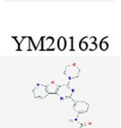

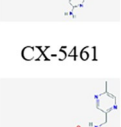
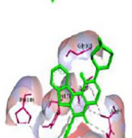
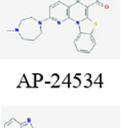
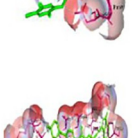
Potential Targets	Structure of Lead Compounds	Binding Affinity score (kcal/mol)	The 3D View and Interactions of Complex	Interacting Amino Acids			
				Hydrogen Bond	Hydrophobic Interactions	Halogen/ Salt Bridge	Electrostatic
NT5E	 Nilotinib	-9.9		Phe183	Leu184, Phe417, Phe500	Glu180, Asp506	-
HPGD	 GSK2126458	-9.9		Ser138, Pro144, Gln147, Gln148, Val186, Thr188	Ile17, Leu139, Val145, Ile190, Ile210, Met213,	Val186, Thr188	-
OAS3	 NVP-BHG712	-9.9		Gln130, Ser145, Glu179	Pro121, Ala181	Gln130	Glu76, Asp143
FOXL1	 YM201636	-9.7		Gly74, Tyr122, Glu113, Gly74	Val70, Tyr122		
FOXC1	 CX-5461	-9.6		Pro95, Asp96, Gly103	Ala94, Ile99, Ala94, Ile99, Pro146	-	-
OAS1	 AP-24534	-9.2		Lys65, Ala130, Glu119, Glu119, Glu233, Ser132	Ala130, Pro123, Pro128, Leu149, Pro123, Arg129, Pro123	-	Asp76, Asp147

Fig 7. The 3D views of the selected strong binding interactions between drug targets and agents were displayed. The key interacting amino acids and their binding types with potential targets were also shown.

<https://doi.org/10.1371/journal.pone.0268967.g007>

NT5E CpG island methylation is a promising BC epigenetic biomarker [63]. The Pharmacological prohibition of *NQO1* and *GCLC* is a new therapeutic strategy for overcoming tamoxifen-resistance and also shows that the prediction of *NQO1* as a biomarker has the significant prognostic value of tumor recurrence in BC patients [64]. β -lapachone (bL) may be a therapeutic targeting for BC stem-cells with appropriate *NQO1* expression [65]. Germline Genetic Variants in *ANGPT1*, *ANGPT2*, *TEK*, *MMP9*, *VEGFA* and *FGF2* are involved with Pathologic complete reaction to Bevacizumab in BC Patients [66]. Several studies have suggested that the gene *FN1* is highly associated with BC [9, 67–72]. They suggested that genetic variants of the *ATF6B* gene were associated with modified relationships between reproductive factors and BC [73]. A new BC risk variant rs8752 in *HPGD* in Chinese women's through a systematic case-control study of microRNA binding site SNPs [74, 75]. The key gene *BCL11A* plays a crucial role in BC tumorigenicity and stemness maintenance through activating Wnt/ β -catenin

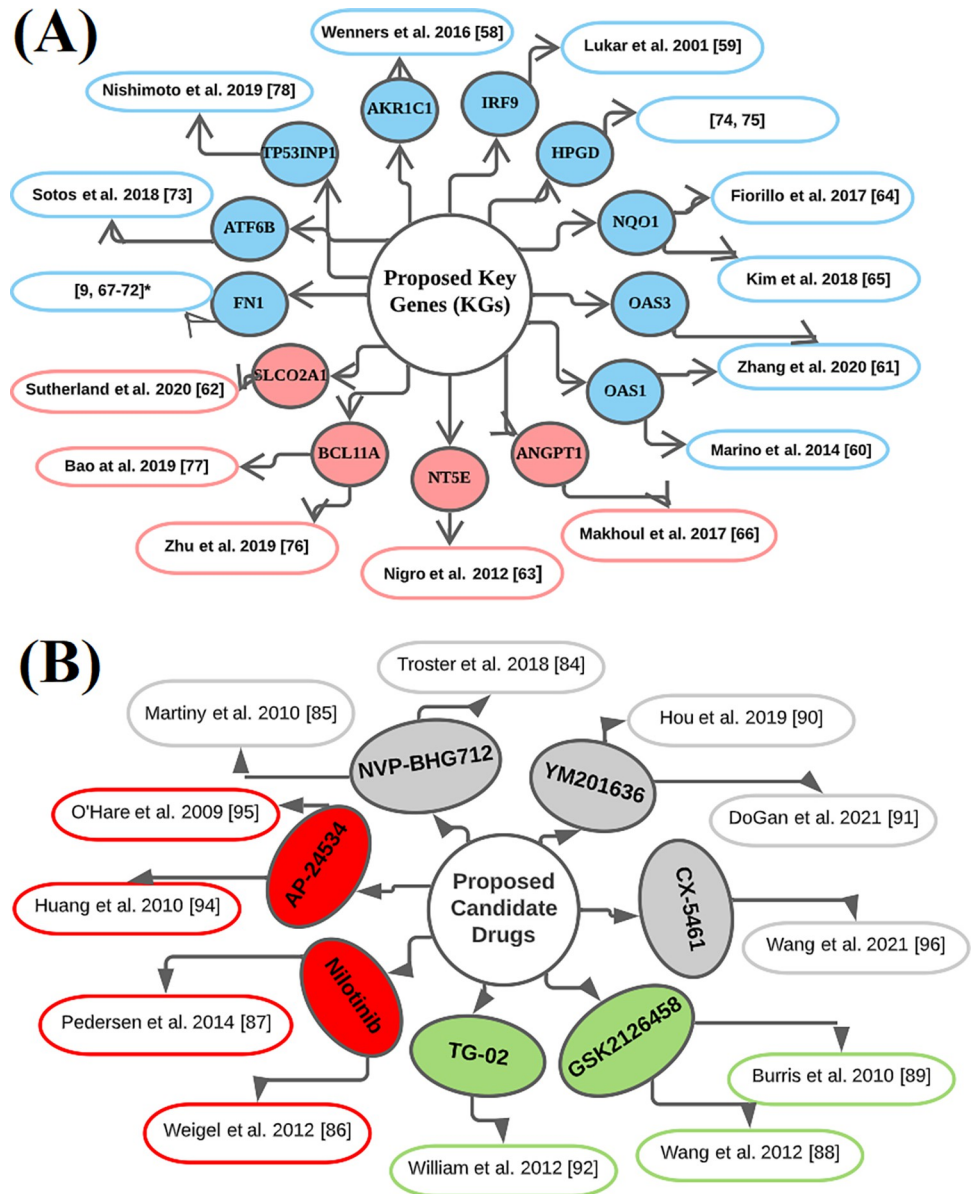


Fig 8. Validation of the proposed KGs (receptors) and candidate drugs in favor of BC by the literature review (A) Validation of the proposed KGs: circles with blue color indicate downregulated KGs and pink color indicates upregulated KGs, and each connected network with a circle indicates the reference in which the KG is associated with BC. (B) Validation of the proposed candidate drugs: circles with red color indicate FDA approved and investigational drugs, green color indicate investigational drugs and ash color indicate unapproved drugs, and each connected network with a circle indicates the references in which our suggested drugs might be effective against BC treatment.

<https://doi.org/10.1371/journal.pone.0268967.g008>

signaling pathway, and may become a potential target for the treatment of BC [76]. Moreover, another study identified 4 hub genes (*BCL11A*, *FOXCI*, *RGMA*, and *FAM171A1*) that showed a highly positive correlation with the triple-negative BC subtype [77]. Low expression of *TP53INP1* is an independent factor of poor prognosis in BC patients, especially ER α -positive patients and may become a potential therapeutic target in ER α -positive BC patients [78].

The GO functional and KEGG pathway enrichment analyses of DEGs significantly revealed some GO terms of BPs, MFs and CCs, and KEGG pathways by involving KGs that are highly

linked with BC patients (see [Table 3](#)). Our literature review also supported their link with BC. As for examples with the enriched BPs, the cell growth inhibition of MCF-7 (the BC cell line) significantly increases after treatment of BC with XN by influencing the *type I interferon signaling pathway* [46] that is associated with three KGs (*OAS1*, *OAS3* and *IRF9*). Two GO terms *defense response to virus* (associated with 3 KGs: *OAS1*, *OAS3* and *IRF9*) and *interferon-gamma-mediated signaling pathway* (associated with 4 KGs: *OAS1*, *OAS3*, *IRF9* and *ATF6B*) were reported as two important BPs for BC progression [47]. Two KGs (*OAS1*, *OAS3*) involving *negative regulation of viral genome replication* process influence the BC enriched dysregulated subnetworks and play a potential role in cardiotoxicity [48]. The *oxidation-reduction process* associated 3 KGs (*NQO1*, *HPGD* and *AKR1C1*) is functionally enriched with multiple cancer type-specific metastasis over-expression signatures [49]. Biological process analysis has shown that *positive regulation of transcription from RNA polymerase II promoter* (associated with *BCL11A* and *SLCO2A1*) and *negative regulation of transcription from RNA polymerase II promoter* (associated *BCL11A* and *NT5E*) are functionally enriched for DEGs of BC [52, 53]. Among the enriched CCs, The *cytoplasm* (associated with 6 KGs: *OAS1*, *OAS3*, *IRF9*, *HPGD*, *TP53INP1* and *NQO1*) has been found to be associated with most proteins that are highly expressed for cancer [50]. Upregulated DEGs for BC are functionally enriched in the *proteinaceous extracellular matrix* (associated KGs: *FNI*) pathway [79]. It was suggested that activation of the *NLRP3 inflammasome complex* (associated KGs: *FNI*) would be an innovative therapeutic pathway to control tumor growth [54]. Inhibition of NADH: ubiquinone *oxidoreductase activity* (associated KGs: *HPGD* and *AKR1C1*) blocks multiple signal transduction pathways in MCF-7 human BC cells through rotenoids drug [51]. The 320 differential expressed microRNAs targeted genes for BC were functionally enriched in negative regulation of *cysteine-type endopeptidase activity involved in apoptotic process* (Associated KGs: *FNI*) [55]. After treatment of a metastatic BC patient with *herpes simplex infection pathway* (associated KGs: *OAS1*, *OAS3*, *ATF6B*, *IRF9*), there has created a case of Sweet Syndrome [56]. Among the enriched KEGG pathways, Paclitaxel inhibits the proliferation and invasion of the MCF-7 cell in *PI3K-AKT signaling pathway* (Associated KGs: *ANGPT1*, *SLCO2A1* and *BCL11A*) to prevent BC [57].

The KGs-TFs interaction network analysis indicated that 4 TFs proteins (*FOXC1*, *FOXL1*, *GATA2*, and *JUN*) are the key transcriptional regulatory factors of Kgs (see [Fig 4A](#)). Among them *FOXC1* (a regulator of *NT5E*, *IRF9*, *AKR1C1*, *HPGD*, *FNI* and *OAS1*) is connected with lymphatic vessel formation, arterial cell specification, and cardiovascular development [80]. The expression of TF-protein *FOXL1* (a regulator of *NT5E*, *TP53INP1*, *HPGD*, *FNI*, *OAS1* and *NQO1*) is connected with numerous cancer [81]. The TF-protein *GATA2* (a regulator of *NT5E*, *SLCO2A1*, *OAS1* and *NQO1*) is connected with Hematopoietic and immune defects [82]. The TF-protein *JUN* (a regulator of *ANGPT1*, *HPGD*, *ATF6B* and *OAS3*) is associated with bladder cancer disease [83]. We also constructed the proteins-disease interaction network to detect other diseases that are also connected with the proposed KGs. Total 8 KGs out of 11 were associated with others 156 diseases that can be considered as the non-causal risk factors of BC. Especially, two diseases "Autosomal recessive predisposition" and "Schizophrenia" were mostly related with our target proteins.

To investigate the prognostic power of KGs, we performed multivariate survival analysis and developed two prediction models through two classifiers (SVM and RF) in [Fig 5](#). Our developed two prediction models showed good performance with both training and test datasets generated from the main data collected from NCBI with accession number GSE53566. The AUC values were 0.992 and 0.986 for SVM and RF based models for the training dataset, respectively. To investigate their performance unbiasedly, we also considered two independent test datasets from other NCBI sources with accession numbers GSE119552 and GSE152322,

respectively. We observed that both predictors show good performance with both independent test datasets. The values of AUC were 0.917 and 0.979 for independent test dataset-1 and 0.985 and 0.947 for the independent test dataset-2 based on SVM and RF models, respectively. These results indicate the good prediction performance for the identified KGs, so we suggested the prognostic model for the two classifiers (SVM and RF).

To explore our proposed KGs-guided new and repurposable candidate drugs for the treatment against BC, we considered the proposed KGs based 13 key proteins (AKR1C1, IRF9, OAS1, OAS3, SLCO2A1, NT5E, NQO1, ANGPT1, FN1, ATF6B, HPGD, BCL11A, and TP53INP1) and their regulatory 4 TFs proteins (FOXC1, FOXL1, GATA2, and JUN) as the drug target receptors and performed their docking simulation with 129 drug molecules collected from the GSCALite database and published articles (see Fig 6A). Then we selected top-ranked 7 drugs (NVP-BHG712, Nilotinib, GSK2126458, YM201636, TG-02, CX-5461, and AP-24534) as the most probable repurposable candidate drugs for BC patients based on their strong binding affinity scores (less than -7.0 kcal/mol) with all the target proteins (see Fig 6A and 6B). Then we investigated the resistance performance of both the proposed and already published candidate drugs against the state-of-the-art alternatives top-ranked 13 independent receptors suggested by others for BC and observed that our proposed candidate drugs are more effective compared to the already published drugs against the independent receptors also (see Fig 6C). We also validated our proposed drugs in favor of BC by the literature review (see Fig 8B).

Among the identified candidate drugs NVP-BHG712 had the knack to inhibit EphB4 kinase activity and EphA2 with an IC_{50} of 3 nM in HEK293 T cells. Besides, NVP-BHG712 had a good binding score for other Eph targets as well, with IC_{50} s ranging from 0.3 nM to 303 nM for EphA3 and EphA1 respectively. Overall the isomers had a low binding score with IC_{50} ranging from 163 to 1660 nM for EphA2 and EphB4, respectively, which revealed that small changes could be made a significant effect on Eph target binding [84]. NVP-BHG712 had the ability to inhibit VEGFR2 as well, but the compound has a 200 times higher binding score for EphB4 [85]. Preclinical studies showed that nilotinib had a growth inhibitory effect on LTED (long-term estrogen deprived) MCF-7 BC cells via ER [86]. Also the nilotinib and sorafenib were considered as potential new treatment options for tamoxifen-resistant BC [87]. GSK2126458 had been considered as potential therapies for BC and were highly selective and effective small compounds inhibitors that receptor both multiple class I PI3K isoforms and mTOR kinase activity [88, 89]. YM201636 was exposed with validate through vivo analysis that YM201636 have an inhibitory effect on tumour cell growth without any side effects for both liver cancer and non-small cell lung cancer [90, 91]. TG-02 (Zotiraciclib) is used as enzyme inhibitor of CDKs to treat cancer disease and also approved as an orphan drug by FDA to treat glioma disease [92, 93]. AP-24534 (Ponatinib) was proposed as an inhibitor of multi-target drugs to treat chronic myeloid leukemia disease and approved by the FDA in December 2012 as a candidate drug [94, 95]. CX-5461 is an inhibited drug for colorectal cancer (CRC) development in *Znf545Δ/ΔApcMin/+* mice [96]. Among the proposed seven candidate drugs, Nilotinib and AP-25534 are approved by the FDA in 2007 and 2012 respectively, TG-02 and GSK2126458 are investigational drugs and three other drugs (NVP-BHG712, YM201636 and CX-5461) are not yet approved. The unapproved drugs should be further assessed in molecular level by the wet-lab experiments in prior to clinical investigation in the treatment of BC.

Conclusion

The main purpose of this study was to identify potential KGs highlighting their function, pathways, and regulatory factors for breast cancer (BC) diagnosis, prognosis and therapies by using

the integrated bioinformatics and statistical approaches. We identified BC causing 13 DEGs (*AKR1C1*, *IRF9*, *OAS1*, *OAS3*, *SLCO2A1*, *NT5E*, *NQO1*, *ANGPT1*, *FN1*, *ATF6B*, *HPGD*, *BCL11A*, and *TP53INP1*) as the KGs by using the five topological measures in the PPI networking results. Their association with BC was also reported by several other studies directly or indirectly that we mentioned in the discussion section. We detected four TFs proteins (*FOXC1*, *FOXL1*, *JUN*, and *GATA2*) and four microRNAs (*hsa-miR-27a-5p*, *hsa-miR-124-3p*, *hsa-miR-1-3p*, and *hsa-miR-210-3p*) as the key transcriptional and post-transcriptional regulators of KGs. These regulatory factors play the vital role for the regulation of KGs. The GO terms (BPs, MFs and CCs) and KEGG pathway enrichment analysis revealed some vital GO terms from each of BPs, MFs and CCs that are significantly enriched by DEGs including KGs. The enriched GO terms and KEGG pathways were considered as the key pathogenetic processes of BC progression. These findings were also supported by the literature review directly or indirectly. We investigated the prognostic performance of KGs by using multivariate survival analysis including unsupervised hierarchical clustering and supervised classification. In each case, we observed the strong prognostic performance of the proposed KGs. Then we considered the proposed 13 key proteins and their regulatory 4 TFs-proteins as the drug target receptors to explore effective drugs for BC by molecular docking simulation with the 129 meta-drug agents. We detected 7 small molecules (*NVP-BHG712*, *Nilotinib*, *GSK2126458*, *YM201636*, *TG-02*, *CX-5461*, and *AP-24534*) as the top ranked candidate drugs for the treatment against BC. Then we investigated the resistance performance of both the proposed and already published candidate drugs against the state-of-the-art alternatives already published top-ranked 13 independent receptors for BC and observed that our proposed candidate drugs are computationally more effective against the independent receptors also. Therefore, the proposed candidate drugs might be played the vital role for the treatment against BC.

Supporting information

S1 File. Supplementary information on datasets.

(DOCX)

S2 File. Supplementary results.

(PDF)

Acknowledgments

We are grateful to the editor and reviewers for their valuable comments that help us to improve the quality of the manuscript. We are also grateful and thankful to the authors whose articles help us to write this paper.

Author Contributions

Conceptualization: Md. Shahin Alam, Md. Nurul Haque Mollah.

Data curation: Md. Shahin Alam, Adiba Sultana.

Formal analysis: Md. Shahin Alam, Adiba Sultana, Md. Selim Reza.

Investigation: Md. Shahin Alam, Adiba Sultana.

Methodology: Md. Shahin Alam, Adiba Sultana, Md. Selim Reza, Md. Nurul Haque Mollah.

Project administration: Md. Nurul Haque Mollah.

Supervision: Md. Nurul Haque Mollah.

Validation: Md. Shahin Alam, Adiba Sultana.

Visualization: Md. Shahin Alam, Adiba Sultana.

Writing – original draft: Md. Shahin Alam, Adiba Sultana, Md. Selim Reza.

Writing – review & editing: Md. Shahin Alam, Md Amanullah, Syed Rashel Kabir, Md. Nurul Haque Mollah.

References

1. Lei S, Zheng R, Zhang S, Wang S, Chen R, Sun K, et al. Global patterns of breast cancer incidence and mortality: A population-based cancer registry data analysis from 2000 to 2020. *Cancer Commun (Lond)*. 2021; 41(11):1183–94. <https://doi.org/10.1002/cac2.12207> PMID: 34399040.
2. Koo MM, von Wagner C, Abel GA, McPhail S, Rubin GP, Lyrtzopoulos G. Typical and atypical presenting symptoms of breast cancer and their associations with diagnostic intervals: Evidence from a national audit of cancer diagnosis. *Cancer Epidemiol*. 2017; 48:140–6. <https://doi.org/10.1016/j.canep.2017.04.010> PMID: 28549339.
3. Cancer.Net. Breast Cancer: Statistics 2021. Available from: <https://www.cancer.net/cancer-types/breast-cancer/statistics>.
4. Dong H, Zhang S, Wei Y, Liu C, Wang N, Zhang P, et al. Bioinformatic analysis of differential expression and core GENes in breast cancer. *Int J Clin Exp Pathol*. 2018; 11(3):1146–56 PMID: 31938209.
5. Mosharaf MP, Reza MS, Kibria MK, Ahmed FF, Kabir MH, Hasan S, et al. Computational identification of host genomic biomarkers highlighting their functions, pathways and regulators that influence SARS-CoV-2 infections and drug repurposing. *Sci Rep*. 2022; 12(1):4279. <https://doi.org/10.1038/s41598-022-08073-8> PMID: 35277538.
6. Islam T, Rahman R, Gov E, Turanli B, Gulfidan G, Haque A, et al. Drug Targeting and Biomarkers in Head and Neck Cancers: Insights from Systems Biology Analyses. *OMICS*. 2018; 22(6):422–36. <https://doi.org/10.1089/omi.2018.0048> PMID: 29927717.
7. Reza MS, Harun-Or-Roshid M, Islam MA, Hossen MA, Hossain MT, Feng S, et al. Bioinformatics Screening of Potential Biomarkers from mRNA Expression Profiles to Discover Drug Targets and Agents for Cervical Cancer. *Int J Mol Sci*. 2022; 23(7). <https://doi.org/10.3390/ijms23073968> PMID: 35409328
8. Ahmed FF, Reza MS, Sarker MS, Islam MS, Mosharaf MP, Hasan S, et al. Identification of host transcriptome-guided repurposable drugs for SARS-CoV-1 infections and their validation with SARS-CoV-2 infections by using the integrated bioinformatics approaches. *PLoS One*. 2022; 17(4):e0266124. <https://doi.org/10.1371/journal.pone.0266124> PMID: 35390032.
9. Amjad E, Asnaashari S, Sokouti B, Dastmalchi S. Systems biology comprehensive analysis on breast cancer for identification of key gene modules and genes associated with TNM-based clinical stages. *Sci Rep*. 2020; 10(1):10816. <https://doi.org/10.1038/s41598-020-67643-w> PMID: 32616754.
10. Li MX, Jin LT, Wang TJ, Feng YJ, Pan CP, Zhao DM, et al. Identification of potential core genes in triple negative breast cancer using bioinformatics analysis. *Onco Targets Ther*. 2018; 11:4105–12. <https://doi.org/10.2147/OTT.S166567> PMID: 30140156.
11. Jin H, Huang X, Shao K, Li G, Wang J, Yang H, et al. Integrated bioinformatics analysis to identify 15 hub genes in breast cancer. *Oncol Lett*. 2019; 18(2):1023–34. <https://doi.org/10.3892/ol.2019.10411> PMID: 31423162.
12. Liu S, Liu X, Wu J, Zhou W, Ni M, Meng Z, et al. Identification of candidate biomarkers correlated with the pathogenesis and prognosis of breast cancer via integrated bioinformatics analysis. *Medicine (Baltimore)*. 2020; 99(49):e23153. <https://doi.org/10.1097/MD.00000000000023153> PMID: 33285689.
13. Wei LM, Li XY, Wang ZM, Wang YK, Yao G, Fan JH, et al. Identification of hub genes in triple-negative breast cancer by integrated bioinformatics analysis. *Gland Surg*. 2021; 10(2):799–806. <https://doi.org/10.21037/gs-21-17> PMID: 33708561.
14. Hong Z, Wang Q, Hong C, Liu M, Qiu P, Lin R, et al. Identification of Seven Cell Cycle-Related Genes with Unfavorable Prognosis and Construction of their TF-miRNA-mRNA regulatory network in Breast Cancer. *J Cancer*. 2021; 12(3):740–53. <https://doi.org/10.7150/jca.48245> PMID: 33403032.
15. Yuan Q, Zheng L, Liao Y, Wu G. Overexpression of CCNE1 confers a poorer prognosis in triple-negative breast cancer identified by bioinformatic analysis. *World J Surg Oncol*. 2021; 19(1):86. <https://doi.org/10.1186/s12957-021-02200-x> PMID: 33757543.

16. Li Y, Zhou X, Liu J, Yin Y, Yuan X, Yang R, et al. Differentially expressed genes and key molecules of BRCA1/2-mutant breast cancer: evidence from bioinformatics analyses. *PeerJ*. 2020; 8:e8403. <https://doi.org/10.7717/peerj.8403> PMID: 31998560.
17. Alam MS, Rahaman MM, Sultana A, Wang G, Mollah MNH. Statistics and network-based approaches to identify molecular mechanisms that drive the progression of breast cancer. *Comput Biol Med*. 2022; 145:105508. <https://doi.org/10.1016/j.compbimed.2022.105508> PMID: 35447458.
18. Hao M, Liu W, Ding C, Peng X, Zhang Y, Chen H, et al. Identification of hub genes and small molecule therapeutic drugs related to breast cancer with comprehensive bioinformatics analysis. *PeerJ*. 2020; 8:e9946. <https://doi.org/10.7717/peerj.9946> PMID: 33083112.
19. Peng Z, Xu B, Jin F. Circular RNA hsa_circ_0000376 Participates in Tumorigenesis of Breast Cancer by Targeting miR-1285-3p. *Technol Cancer Res Treat*. 2020; 19:1533033820928471. <https://doi.org/10.1177/1533033820928471> PMID: 32462972.
20. Martin NH, Silke S, Rui NP, Karl S, Carolin S, René D, et al. CapG a putative oncogen affects gene expression in breast cancer cells 2017. Available from: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53566>.
21. Lecomte S, Demay F, Pham T, Moulis S, Efstathiou T, Chalmel F, et al. Effect of estradiol, zearalenone and apigenin on ER-positive breast cancer cells MCF-7 2020. Available from: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE119552>.
22. Cataldo A, Romero-Cordoba S, Plantamura I, Cosentino G, Hidalgo-Miranda A, Tagliabue E, et al. MiR-302b as a Combinatorial Therapeutic Approach to Improve Cisplatin Chemotherapy Efficacy in Human Triple-Negative Breast Cancer. *Cancers (Basel)*. 2020; 12(8). <https://doi.org/10.3390/cancers12082261> PMID: 32806777
23. Liu CJ, Hu FF, Xia MX, Han L, Zhang Q, Guo AY. GSCALite: a web server for gene set cancer analysis. *Bioinformatics*. 2018; 34(21):3771–2. <https://doi.org/10.1093/bioinformatics/bty411> PMID: 29790900.
24. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*. 2004; 3:Article3. <https://doi.org/10.2202/1544-6115.1027> PMID: 16646809.
25. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015; 43(7):e47. <https://doi.org/10.1093/nar/gkv007> PMID: 25605792.
26. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019; 47(D1):D607–D13. <https://doi.org/10.1093/nar/gky1131> PMID: 30476243.
27. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003; 13(11):2498–504. <https://doi.org/10.1101/gr.1239303> PMID: 14597658.
28. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol*. 2014; 8 Suppl 4:S11. <https://doi.org/10.1186/1752-0509-8-S4-S11> PMID: 25521941.
29. Jeong H, Mason SP, Barabasi AL, Oltvai ZN. Lethality and centrality in protein networks. *Nature*. 2001; 411(6833):41–2. <https://doi.org/10.1038/35075138> PMID: 11333967.
30. Przulj N, Wigle DA, Jurisica I. Functional topology in a network of protein interactions. *Bioinformatics*. 2004; 20(3):340–8. <https://doi.org/10.1093/bioinformatics/btg415> PMID: 14960460.
31. Freeman LC. A Set of Measures of Centrality Based on Betweenness. *JSTOR*. 1977; 40: 35–41.
32. Shimbel A. Structural parameters of communication networks. *Bulletin of Mathematical Biology*. 1953; 15:501–7. <https://doi.org/10.1007/BF02476438>.
33. Dennis G Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol*. 2003; 4(5):P3 PMID: 12734009.
34. Khan A, Fornes O, Stigliani A, Gheorghe M, Castro-Mondragon JA, van der Lee R, et al. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res*. 2018; 46(D1):D1284. <https://doi.org/10.1093/nar/gkx1188> PMID: 29161433.
35. Zhou G, Soufan O, Ewald J, Hancock REW, Basu N, Xia J. NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic Acids Res*. 2019; 47(W1):W234–W41. <https://doi.org/10.1093/nar/gkz240> PMID: 30931480.
36. Chang L, Zhou G, Soufan O, Xia J. miRNet 2.0: network-based visual analytics for miRNA functional analysis and systems biology. *Nucleic Acids Res*. 2020; 48(W1):W244–W51. <https://doi.org/10.1093/nar/gkaa467> PMID: 32484539.

37. Aguirre-Gamboa R, Gomez-Rueda H, Martinez-Ledesma E, Martinez-Torteya A, Chacolla-Huaringa R, Rodriguez-Barrientos A, et al. SurvExpress: an online biomarker validation tool and database for cancer gene expression data using survival analysis. *PLoS One*. 2013; 8(9):e74250. <https://doi.org/10.1371/journal.pone.0074250> PMID: 24066126.
38. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, et al. The Protein Data Bank. *Acta Crystallogr D Biol Crystallogr*. 2002; 58(Pt 6 No 1):899–907. <https://doi.org/10.1107/s0907444902003451> PMID: 12037327.
39. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. 2018; 46(W1):W296–W303. <https://doi.org/10.1093/nar/gky427> PMID: 29788355.
40. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, et al. PubChem 2019 update: improved access to chemical data. *Nucleic Acids Res*. 2019; 47(D1):D1102–D9. <https://doi.org/10.1093/nar/gky1033> PMID: 30371825.
41. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res*. 2018; 46(D1):D1074–D82. <https://doi.org/10.1093/nar/gkx1037> PMID: 29126136.
42. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem*. 2004; 25(13):1605–12. <https://doi.org/10.1002/jcc.20084> PMID: 15264254.
43. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*. 2010; 31(2):455–61. <https://doi.org/10.1002/jcc.21334> PMID: 19499576.
44. Dallakyan S, Olson AJ. Small-molecule library screening by docking with PyRx. *Methods Mol Biol*. 2015; 1263:243–50. https://doi.org/10.1007/978-1-4939-2269-7_19 PMID: 25618350.
45. Salentin S, Schreiber S, Haupt VJ, Adasme MF, Schroeder M. PLIP: fully automated protein-ligand interaction profiler. *Nucleic Acids Res*. 2015; 43(W1):W443–7. <https://doi.org/10.1093/nar/gkv315> PMID: 25873628
46. Roehrer S, Stork V, Ludwig C, Minceva M, Behr J. Analyzing bioactive effects of the minor hop compound xanthohumol C on human breast cancer cells using quantitative proteomics. *PLoS One*. 2019; 14(3):e0213469. <https://doi.org/10.1371/journal.pone.0213469> PMID: 30875365.
47. Xu Y, Zhang Z, Zhang L, Zhang C. Novel module and hub genes of distinctive breast cancer associated fibroblasts identified by weighted gene co-expression network analysis. *Breast Cancer*. 2020; 27(5):1017–28. <https://doi.org/10.1007/s12282-020-01101-3> PMID: 32383139.
48. Zhang Y, Tang X, Pang Y, Huang L, Wang D, Yuan C, et al. The Potential Mechanism of Bufadienolide-Like Chemicals on Breast Cancer via Bioinformatics Analysis. *Cancers (Basel)*. 2019; 11(1). <https://doi.org/10.3390/cancers11010091> PMID: 30646630.
49. Chen F, Zhang Y, Varambally S, Creighton CJ. Molecular Correlates of Metastasis by Systematic Pan-Cancer Analysis Across The Cancer Genome Atlas. *Mol Cancer Res*. 2019; 17(2):476–87. <https://doi.org/10.1158/1541-7786.MCR-18-0601> PMID: 30401717.
50. Yang C, Guo WB, Zhang WS, Bian J, Yang JK, Zhou QZ, et al. Comprehensive proteomics analysis of exosomes derived from human seminal plasma. *Andrology*. 2017; 5(5):1007–15. <https://doi.org/10.1111/andr.12412> PMID: 28914500.
51. Rowlands JC, Casida JE. NADH: ubiquinone oxidoreductase inhibitors block induction of ornithine decarboxylase activity in MCF-7 human breast cancer cells. *Pharmacol Toxicol*. 1998; 83(5):214–9. <https://doi.org/10.1111/j.1600-0773.1998.tb01471.x> PMID: 9834970.
52. Zhong G, Lou W, Shen Q, Yu K, Zheng Y. Identification of key genes as potential biomarkers for triple-negative breast cancer using integrating genomics analysis. *Mol Med Rep*. 2020; 21(2):557–66. <https://doi.org/10.3892/mmr.2019.10867> PMID: 31974598.
53. Fang L, Wang Y, Gao Y, Chen X. Overexpression of CXXC5 is a strong poor prognostic factor in ER+ breast cancer. *Oncol Lett*. 2018; 16(1):395–401. <https://doi.org/10.3892/ol.2018.8647> PMID: 29928427.
54. Tezcan G, Garanina EE, Alsaadi M, Gilazieva ZE, Martinova EV, Markelova MI, et al. Therapeutic Potential of Pharmacological Targeting NLRP3 Inflammasome Complex in Cancer. *Front Immunol*. 2020; 11:607881. <https://doi.org/10.3389/fimmu.2020.607881> PMID: 33613529.
55. Xiong H, Chen Z, Chen W, Li Q, Lin B, Jia Y. FKBP-related ncRNA-mRNA axis in breast cancer. *Genomics*. 2020; 112(6):4595–607. <https://doi.org/10.1016/j.ygeno.2020.08.017> PMID: 32814092.
56. Coskun U, Gunel N, Senol E, Ilter N, Dursun A, Tuzun D. A case of Sweet's syndrome developed after the treatment of herpes simplex infection in a metastatic breast cancer patient. *J Cutan Pathol*. 2002; 29(5):301–4. <https://doi.org/10.1034/j.1600-0560.2002.290508.x> PMID: 12100632.

57. Li G, Xu D, Sun J, Zhao S, Zheng D. Paclitaxel inhibits proliferation and invasion and promotes apoptosis of breast cancer cells by blocking activation of the PI3K/AKT signaling pathway. *Adv Clin Exp Med*. 2020; 29(11):1337–45. <https://doi.org/10.17219/acem/127681> PMID: 33269821.
58. Wenners A, Hartmann F, Jochens A, Roemer AM, Alkatout I, Klapper W, et al. Stromal markers AKR1C1 and AKR1C2 are prognostic factors in primary human breast cancer. *Int J Clin Oncol*. 2016; 21(3):548–56. <https://doi.org/10.1007/s10147-015-0924-2> PMID: 26573806.
59. Luker KE, Pica CM, Schreiber RD, Piwnica-Worms D. Overexpression of IRF9 confers resistance to antimicrotubule agents in breast cancer cells. *Cancer Res*. 2001; 61(17):6540–7 PMID: 11522652.
60. Marino N, Collins JW, Shen C, Caplen NJ, Merchant AS, Gokmen-Polar Y, et al. Identification and validation of genes with expression patterns inverse to multiple metastasis suppressor genes in breast cancer cell lines. *Clin Exp Metastasis*. 2014; 31(7):771–86. <https://doi.org/10.1007/s10585-014-9667-0> PMID: 25086928.
61. Zhang Y, Yu C. Prognostic characterization of OAS1/OAS2/OAS3/OASL in breast cancer. *BMC Cancer*. 2020; 20(1):575. <https://doi.org/10.1186/s12885-020-07034-6> PMID: 32560641.
62. Sutherland R, Meeson A, Lowes S. Solute transporters and malignancy: establishing the role of uptake transporters in breast cancer and breast cancer metastasis. *Cancer Metastasis Rev*. 2020; 39(3):919–32. <https://doi.org/10.1007/s10555-020-09879-6> PMID: 32388639.
63. Lo Nigro C, Monteverde M, Lee S, Lattanzio L, Vivenza D, Comino A, et al. NT5E CpG island methylation is a favourable breast cancer biomarker. *Br J Cancer*. 2012; 107(1):75–83. <https://doi.org/10.1038/bjc.2012.212> PMID: 22653144.
64. Fiorillo M, Sotgia F, Sisci D, Cappello AR, Lisanti MP. Mitochondrial "power" drives tamoxifen resistance: NQO1 and GCLC are new therapeutic targets in breast cancer. *Oncotarget*. 2017; 8(12):20309–27. <https://doi.org/10.18632/oncotarget.15852> PMID: 28411284.
65. Kim DW, Cho JY. NQO1 is Required for beta-Lapachone-Mediated Downregulation of Breast-Cancer Stem-Cell Activity. *Int J Mol Sci*. 2018; 19(12). <https://doi.org/10.3390/ijms19123813> PMID: 30513573
66. Makhoul I, Todorova VK, Siegel ER, Erickson SW, Dhakal I, Raj VR, et al. Germline Genetic Variants in TEK, ANGPT1, ANGPT2, MMP9, FGF2 and VEGFA Are Associated with Pathologic Complete Response to Bevacizumab in Breast Cancer Patients. *PLoS One*. 2017; 12(1):e0168550. <https://doi.org/10.1371/journal.pone.0168550> PMID: 28045923.
67. Yang X, Hu Q, Hu LX, Lin XR, Liu JQ, Lin X, et al. miR-200b regulates epithelial-mesenchymal transition of chemo-resistant breast cancer cells by targeting FN1. *Discov Med*. 2017; 24(131):75–85 PMID: 28972876.
68. Yuan CL, Jiang XM, Yi Y, E JF, Zhang ND, Luo X, et al. Identification of differentially expressed lncRNAs and mRNAs in luminal-B breast cancer by RNA-sequencing. *BMC Cancer*. 2019; 19(1):1171. <https://doi.org/10.1186/s12885-019-6395-5> PMID: 31795964.
69. Yan LR, Wang A, Lv Z, Yuan Y, Xu Q. Mitochondria-related core genes and TF-miRNA-hub mrDEGs network in breast cancer. *Biosci Rep*. 2021; 41(1). <https://doi.org/10.1042/BSR20203481> PMID: 33439992.
70. Wang Y, Xu H, Zhu B, Qiu Z, Lin Z. Systematic identification of the key candidate genes in breast cancer stroma. *Cell Mol Biol Lett*. 2018; 23:44. <https://doi.org/10.1186/s11658-018-0110-4> PMID: 30237810.
71. Liu X, Ma Y, Yang W, Wu X, Jiang L, Chen X. Identification of therapeutic targets for breast cancer using biological informatics methods. *Mol Med Rep*. 2015; 12(2):1789–95. <https://doi.org/10.3892/mmr.2015.3565> PMID: 25824986.
72. Zhang M, Gao CE, Li WH, Yang Y, Chang L, Dong J, et al. Microarray based analysis of gene regulation by mesenchymal stem cells in breast cancer. *Oncol Lett*. 2017; 13(4):2770–6. <https://doi.org/10.3892/ol.2017.5776> PMID: 28454465.
73. Dierssen-Sotos T, Palazuelos-Calderon C, Jimenez-Moleon JJ, Aragonés N, Altzibar JM, Castano-Vinyals G, et al. Reproductive risk factors in breast cancer and genetic hormonal pathways: a gene-environment interaction in the MCC-Spain project. *BMC Cancer*. 2018; 18(1):280. <https://doi.org/10.1186/s12885-018-4182-3> PMID: 29530003.
74. He N, Zheng H, Li P, Zhao Y, Zhang W, Song F, et al. miR-485-5p binding site SNP rs8752 in HPGD gene is associated with breast cancer risk. *PLoS One*. 2014; 9(7):e102093. <https://doi.org/10.1371/journal.pone.0102093> PMID: 25003827.
75. Bhar A, Haubrock M, Mukhopadhyay A, Maulik U, Bandyopadhyay S, Wingender E. Coexpression and coregulation analysis of time-series gene expression data in estrogen-induced breast cancer cell. *Algorithms Mol Biol*. 2013; 8(1):9. <https://doi.org/10.1186/1748-7188-8-9> PMID: 23521829.
76. Zhu L, Pan R, Zhou D, Ye G, Tan W. BCL11A enhances stemness and promotes progression by activating Wnt/beta-catenin signaling in breast cancer. *Cancer Manag Res*. 2019; 11:2997–3007. <https://doi.org/10.2147/CMAR.S199368> PMID: 31114347.

77. Bao C, Lu Y, Chen J, Chen D, Lou W, Ding B, et al. Exploring specific prognostic biomarkers in triple-negative breast cancer. *Cell Death Dis.* 2019; 10(11):807. <https://doi.org/10.1038/s41419-019-2043-x> PMID: 31649243.
78. Nishimoto M, Nishikawa S, Kondo N, Wanifuchi-Endo Y, Hato Y, Hisada T, et al. Prognostic impact of TP53INP1 gene expression in estrogen receptor alpha-positive breast cancer patients. *Jpn J Clin Oncol.* 2019; 49(6):567–75. <https://doi.org/10.1093/jjco/hyz029> PMID: 30855679.
79. Wang Y, Zhang Y, Huang Q, Li C. Integrated bioinformatics analysis reveals key candidate genes and pathways in breast cancer. *Mol Med Rep.* 2018; 17(6):8091–100. <https://doi.org/10.3892/mmr.2018.8895> PMID: 29693125.
80. Kume T. The cooperative roles of Foxc1 and Foxc2 in cardiovascular development. *Adv Exp Med Biol.* 2009; 665:63–77. https://doi.org/10.1007/978-1-4419-1599-3_5 PMID: 20429416.
81. Chen X, Deng M, Ma L, Zhou J, Xiao Y, Zhou X, et al. Inhibitory effects of forkhead box L1 gene on osteosarcoma growth through the induction of cell cycle arrest and apoptosis. *Oncol Rep.* 2015; 34(1):265–71. <https://doi.org/10.3892/or.2015.3969> PMID: 26062977.
82. Collin M, Dickinson R, Bigley V. Haematopoietic and immune defects associated with GATA2 mutation. *Br J Haematol.* 2015; 169(2):173–87. <https://doi.org/10.1111/bjh.13317> PMID: 25707267.
83. Zhang L, Feng C, Zhou Y, Zhou Q. Dysregulated genes targeted by microRNAs and metabolic pathways in bladder cancer revealed by bioinformatics methods. *Oncol Lett.* 2018; 15(6):9617–24. <https://doi.org/10.3892/ol.2018.8602> PMID: 29928337.
84. Troster A, Heinzlmeir S, Berger BT, Gande SL, Saxena K, Sreeramulu S, et al. NVP-BHG712: Effects of Regioisomers on the Affinity and Selectivity toward the EPHrin Family. *ChemMedChem.* 2018; 13(16):1629–33. <https://doi.org/10.1002/cmdc.201800398> PMID: 29928781.
85. Martiny-Baron G, Holzer P, Billy E, Schnell C, Brueggen J, Ferretti M, et al. The small molecule specific EphB4 kinase inhibitor NVP-BHG712 inhibits VEGF driven angiogenesis. *Angiogenesis.* 2010; 13(3):259–67. <https://doi.org/10.1007/s10456-010-9183-z> PMID: 20803239.
86. Weigel MT, Ghazoui Z, Dunbier A, Pancholi S, Dowsett M, Martin LA. Preclinical and clinical studies of estrogen deprivation support the PDGF/Abl pathway as a novel therapeutic target for overcoming endocrine resistance in breast cancer. *Breast Cancer Res.* 2012; 14(3):R78. <https://doi.org/10.1186/bcr3191> PMID: 22608253.
87. Pedersen AM, Thrane S, Lykkesfeldt AE, Yde CW. Sorafenib and nilotinib resensitize tamoxifen resistant breast cancer cells to tamoxifen treatment via estrogen receptor alpha. *Int J Oncol.* 2014; 45(5):2167–75. <https://doi.org/10.3892/ijo.2014.2619> PMID: 25175082.
88. Wang M, Gao M, Miller KD, Sledge GW, Zheng QH. [11C]GSK2126458 and [18F]GSK2126458, the first radiosynthesis of new potential PET agents for imaging of PI3K and mTOR in cancers. *Bioorg Med Chem Lett.* 2012; 22(4):1569–74. <https://doi.org/10.1016/j.bmcl.2011.12.136> PMID: 22297110.
89. Burris JR H., Sharma S., Herbst R. S., Tabernero J., Infante A J. R. Silva, Demanse D., Hackl W., Base-lga J. First-in-human phase I study of the oral PI3K inhibitor BEZ235 in patients (pts) with advanced solid tumors. *Journal of Clinical Oncology.* 2010; 28:3005–. https://doi.org/10.1200/jco.2010.28.15_suppl.3005
90. Hou JZ, Xi ZQ, Niu J, Li W, Wang X, Liang C, et al. Inhibition of PIKfyve using YM201636 suppresses the growth of liver cancer via the induction of autophagy. *Oncol Rep.* 2019; 41(3):1971–9. <https://doi.org/10.3892/or.2018.6928> PMID: 30569119.
91. DoGan E, DU Z, Yildirim Z, Ozdl B, Aktu GH, Bozok Cetlnta SV. The effects of PIKfyve inhibitor YM201636 on claudins and malignancy potential of nonsmall cell cancer cells. *Turk J Biol.* 2021; 45(1):26–34. <https://doi.org/10.3906/biy-2010-32> PMID: 33597819.
92. William AD, Lee AC, Goh KC, Blanchard S, Poulsen A, Teo EL, et al. Discovery of kinase spectrum selective macrocycle (16E)-14-methyl-20-oxa-5,7,14,26-tetraazatetracyclo[19.3.1.1(2,6).1(8,12)]heptaco sa-1(25),2(26),3,5,8(27),9,11,16,21,23-decaene (SB1317/TG02), a potent inhibitor of cyclin dependent kinases (CDKs), Janus kinase 2 (JAK2), and fms-like tyrosine kinase-3 (FLT3) for the treatment of cancer. *J Med Chem.* 2012; 55(1):169–96. <https://doi.org/10.1021/jm201112g> PMID: 22148278.
93. Administration USFaD. FDA grants orphan drug designation to zotiraciclib for the treatment of glioma. Available from: <https://ccr.cancer.gov/news/article/fda-grants-orphan-drug-designation-to-zotiraciclib-for-the-treatment-of-glioma>.
94. Huang WS, Metcalf CA, Sundaramoorthi R, Wang Y, Zou D, Thomas RM, et al. Discovery of 3-[2-(imidazo[1,2-b]pyridazin-3-yl)ethynyl]-4-methyl-N-{4-[(4-methylpiperazin-1-yl)methyl]-3-(trifluoromethyl)phenyl}benzamide (AP24534), a potent, orally active pan-inhibitor of breakpoint cluster region-abelson (BCR-ABL) kinase including the T315I gatekeeper mutant. *J Med Chem.* 2010; 53(12):4701–19. <https://doi.org/10.1021/jm100395q> PMID: 20513156.

95. O'Hare T, Shakespeare WC, Zhu X, Eide CA, Rivera VM, Wang F, et al. AP24534, a pan-BCR-ABL inhibitor for chronic myeloid leukemia, potently inhibits the T315I mutant and overcomes mutation-based resistance. *Cancer Cell*. 2009; 16(5):401–12. <https://doi.org/10.1016/j.ccr.2009.09.028> PMID: [19878872](https://pubmed.ncbi.nlm.nih.gov/19878872/).
96. Wang S, Wong CC, Zhang Y, Huang J, Li C, Zhai J, et al. ZNF545 loss promotes ribosome biogenesis and protein translation to initiate colorectal tumorigenesis in mice. *Oncogene*. 2021. <https://doi.org/10.1038/s41388-021-01938-8> PMID: [34615997](https://pubmed.ncbi.nlm.nih.gov/34615997/).