Heliyon 9 (2023) e16151

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Bioinformatics and network biology approach to identifying type 2 diabetes genes and pathways that influence the progression of breast cancer

Md Sumon Sarkar [a], Md Misor Mia [a], Md Al Amin [b], Md Sojib Hossain [c], Md Zahidul Islam [d,*]

[a] Department of Pharmacy, Islamic University, Kushtia-7003, Bangladesh
[b] Department of Computer Science & Engineering, Prime University, Dhaka-1216, Bangladesh
[c] Department of Mathematics, Govt. Bangla College, Dhaka-1216, Bangladesh
[d] Department of Information & Communication Technology, Islamic University, Kushtia-7003, Bangladesh

ARTICLE INFO

Dataset link: https://www.ncbi.nlm.nih.gov/gds

Keywords:
Type-2 diabetes
Breast cancer
Bioinformatics
Computational biology
Differentially expressed genes
Gene set enrichment analysis
Gene ontology
Protein–protein interaction (PPI)
Hub gene
Drug molecule
Disease association

ABSTRACT

Breast cancer is the second most prevalent malignancy affecting women. Postmenopausal women breast tumor is one of the top causes of death in women, accounting for 23% of cancer cases. Type 2 diabetes, a worldwide pandemic, has been connected to a heightened risk of several malignancies, although its association with breast cancer is still uncertain. In comparison to non-diabetic women, women with T2DM had a 23% elevated likelihood of developing breast cancer. It is difficult to determine causative or genetic susceptibility that connect T2DM and breast cancer. We created a large-scale network-based quantitative approach employing unbiased methods to discover abnormally amplified genes in both T2DM and breast cancer, to solve these issues. We performed transcriptome analysis to uncover identical genetic biomarkers and pathways to clarify the connection between T2DM and breast cancer patients. In this study, two RNA-seq datasets (GSE103001 and GSE86468) from the Gene Expression Omnibus (GEO) are used to identify mutually differentially expressed genes (DEGs) for breast cancer and T2DM, as well as common pathways and prospective medicines. Firstly, 45 shared genes (30 upregulated and 15 downregulated) between T2D and breast cancer were detected. We employed gene ontology and pathway enrichment to characterize prevalent DEGs' molecular processes and signal transduction pathways and observed that T2DM has certain connections to the progression of breast cancer. Using several computational and statistical approaches, we created a protein-protein interactions (PPI) network and revealed hub genes. These hub genes can be potential biomarkers, which may also lead to new therapeutic strategies for investigated diseases. We conducted TF-gene interactions, gene-microRNA interactions, protein-drug interactions, and gene-disease associations to find potential connections between T2DM and breast cancer pathologies. We assume that the potential drugs that emerged from this study could be useful therapeutic values. Researchers, doctors, biotechnologists, and many others may benefit from this research.

* Corresponding author.
E-mail address: zahidimage@gmail.com (M.Z. Islam).

https://doi.org/10.1016/j.heliyon.2023.e16151
Received 30 November 2022; Received in revised form 28 April 2023; Accepted 7 May 2023

Available online 12 May 2023

1. Introduction

The second commonest malignancy affecting women is breast cancer [1]. Postmenopausal women's breast tumor is among the leading causes of mortality in females, contributing to 23% of all cancer cases [2]. In 2020, approximately 2.3 million new cases were detected and about 8 million alive women were identified with breast tumors in the previous five years [3]. The incidence of breast tumors in men ranges from 0.5% to 1% [4]. The chance of developing breast cancer is affected by several variables (such as heredity, lifestyle, hormonal factors, benign breast disease, and the environment) [5]. Obesity is associated with an elevated risk of post-menopausal women's breast cancer [6].

Type 2 diabetes mellitus (T2DM) is a persistent hyperglycemic metabolic illness defined by insulin deficiency and multiple organ system dysfunctions that causes problems in carbohydrate, protein, and lipid metabolism [7]. Diabetes mellitus, a worldwide pandemic, is the ninth leading cause of death globally. Diabetes mellitus affects approximately 1 in every 11 persons worldwide, with 90% of those cases having T2DM [8]. By 2030, approximately 450 million people globally and 7.7% of adults (age 20 to 79) are predicted to have diabetes [9]. Insulin insensitivity is a hallmark of the most common form of diabetes. Type 2DM develops from insulin resistance, decreasing insulin levels, and eventually the destruction of pancreatic beta cells [10]. T2DM lifestyle variables include lack of education, tobacco smoking, obesity, an inheritance of diabetes, hypertension, and physical inactivity [11].

Type 2 DM correlates with a lower survival rate for various cancers. According to several epidemiological studies, diabetic people have a greater likelihood of liver malignancies, biliary tract, pancreas, colon, stomach, bladder, kidney, breast, and uterus. Diabetes, on the other hand, has been linked to a lower incidence of prostate cancer [12]. Individuals with T2DM have a greater possibility of death from a variety of malignancies, featuring a 30%-40% rise in pancreatic malignancy, a 2.5-fold rise in liver cancer, a 30% rise in uterine cancer, a 15%-30% rise in breast tumors, and a 20%-50% rise in intestinal malignancy [13]. Hyperinsulinemia, hyperglycemia, plasma testosterone levels, obesity, insulin restriction, raised insulin-like growth factor-1 (IGF-1) expression, hyperlipidemia, inflammatory cytokines, higher concentrations of leptin, and reduced amounts of adiponectin have been involved in increased cancer risk [14][13].

According to results from meta-analysis, diabetic women have a 23% elevated risk of breast tumors relative to non-diabetic women [15]. Findings from several studies; diabetes linked to breast tumors via some mechanisms:

- The insulin-like growth factor pathway is activated. IGF can bind to both the insulin and IGF receptors, initiating a chain of events that eventually leads to the development of breast tumors [15][16][17].
- Hyperinsulinemia is also linked to high estrogen levels and decreased sex hormone-binding globulin (SHBG) densities, causing a greater estradiol bioavailability. Furthermore, diabetes-induced hyperglycemia promotes tumor cell growth and proliferation [15][18].
- Being overweight and having T2DM have been linked to higher leptin quantities and lower adiponectin values, which may cause a threat of breast cancer, particularly in the more aggressive types [15][19].
- Increased insulin can bind to its receptors, which activates phosphatidylinositol 3-kinase, which in turn promotes the AKT pathway, which plays a vital role in carcinogenesis [15][20].

Numerous studies have examined the relationship between T2DM and various molecular subtypes of breast cancer and the findings suggest that there may be some differences in the association between T2DM and the risk of developing certain subtypes of breast cancer. For example, some studies have reported that T2DM is associated with an increased risk of developing ER+ breast cancer, which is the most common subtype of breast cancer. Other studies have suggested a stronger association with HER2+ or triple-negative breast cancer [21]. Differences in hormone receptor status, insulin resistance, inflammation, and other things, may impact the link between T2DM and breast cancer subtypes. To understand the underlying mechanisms and to clarify any potential variations in the correlation between T2DM and various breast cancer subtypes, additional study is necessary.

In conclusion, strong evidence suggests histological and therapeutically substantial correlations between T2DM and breast cancer, but this relation is not yet investigated [22]. The scientific relations and molecular mechanisms that underpin this association are yet unknown because the etiology of T2DM and breast cancer are highly complicated, and their threat variables intersect slightly. T2D and breast cancer connections are difficult to find, yet they are of substantial attention in clinical endocrinology. T2D and breast cancer have complicated clinical presentations, making them difficult to examine using traditional assumptions-based endocrinology inquiry, despite their therapeutic relevance. Furthermore, bioinformatics research on the relationship between T2D and breast tumors is still lacking. The goal of this study was to find such linkages between breast cancer and T2DM because knowing the nature of these links could reveal crucial information about the disease mechanisms. Therefore, we employed a bioinformatics system approach to evaluating genomic information from infected tissue investigations for insights into the nature of the T2D-breast cancer link. Differentially Expressed (DE) genes were identified for additional system biological analysis to understand molecular mechanisms of disease conditions. DE genes were statistically significant differences in expression levels between groups. There were several statistical approaches for identifying significant DE genes in the literature. In the assessment of RNA-seq data set analysis, the DESeq2 algorithm is the most common. For identifying DE genes in T2D and breast cancer datasets, we used RNA-seq datasets. We used a systems biology approach in this study to represent commonly dysfunctional proteins and pathways linked to T2D and breast cancer. This research aimed to identify inappropriately expressed proteins or genes in T2D individuals leading to the development of breast cancer in females. This investigation can also help to understand the genetic underpinnings of female breast cancer. This research may help to create generic drugs by determining the functions of these proteins. For female breast cancer patients with

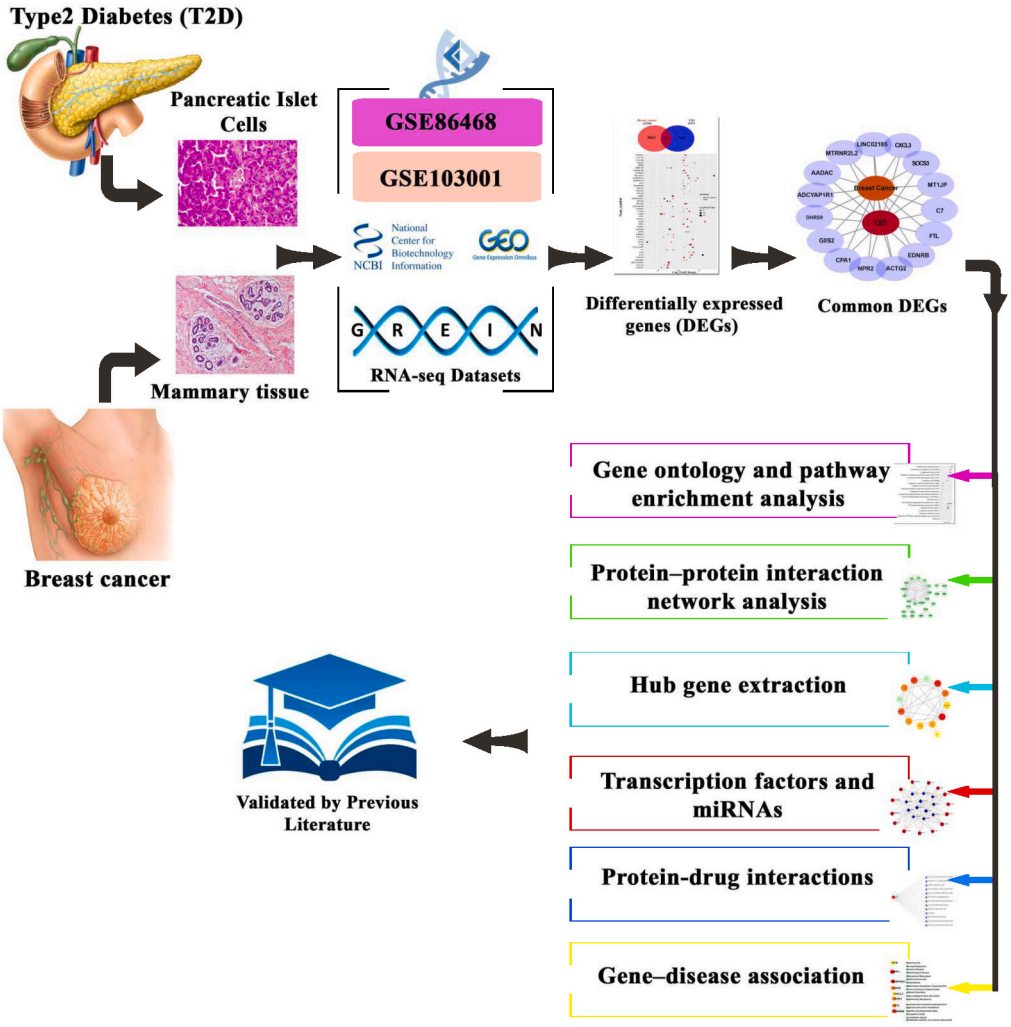

Fig. 1. Graphical representation of the underlying procedure of this research.

T2D, this study also offers significant information for drug development and design. Fig. 1 illustrates the sequential workflow of our study.

2. Materials and methods

2.1. Datasets used for this research

We examined datasets from the Gene Expression Omnibus (GEO) of the National Centre for Biotechnology Information (NCBI) [23] and GREIN. GREIN is a website-based interactive tool that enables quick entrance to GEO RNA-seq data. Several research used GREIN to carry out differential gene expression inspection on the respective RNA-seq data [24]. Several datasets have been rejected because they had fewer than six samples, no control and case requirements, replicated datasets, undesirable formatting, unrelated experimental emphasis, or non-human RNAseq data. Each of the six samples must contain a minimum of 3 case samples and healthy control. We searched for a dataset having an adequate number of controls and cases for our research. This approach identified two datasets pertinent to T2D and breast cancer and were suitable for our inquiry. Datasets on genomic expression bearing identification numbers GSE103001 [25] and GSE86468 [26] were evaluated (having a healthy control group and case group) for our research. The breast cancer dataset (GSE103001) contained mammary tissue-derived gene expression information. Strand-specific RNA fingerprinting was conducted on an Illumina HiSeq 2000 and the dataset consisted of 22 initial breast carcinoma exhibiting estrogen and its corresponding normal tissues. Analyzing the ER+ breast cancer dataset (GSE103001) can provide valuable insights into the biology of this disease and guide the development of new treatments. Second, the dataset may have been collected with specific research questions in mind, and therefore may contain more specific and relevant information than larger, more general

datasets like TCGA or METABRIC. The T2D dataset (GSE86468) included gene expression data from human pancreatic islets. The dataset contained RNA-seq profiles on a Illumina NextSeq 500 of 24 human pancreatic islet cell samples (6 cases,6 control).

2.2. Raw counts preprocessing and differentially expression analysis

Differentially expressed at the transcriptional level refers to genes that display statistically significant differences in expression between multiple experimental conditions [27]. The NCBI Gene Expression Omnibus provided the gene expression microarray datasets used in our research. These datasets were created by identifying differentially expressed genes (DEGs) linked to each pathology by contrasting healthy tissue with diseased tissue. The primary objective of this analysis was to extract DEGs from the datasets GSE86468 and GSE103001. According to two criteria (absolute log fold change value higher than or equivalent to 1 and adjusted p-value 0.05), we calculated the range of significant genes from Differentially Expressed Genes (DEGs). Then, a VENN analysis program named Jvenn [28] was used to acquire the mutual DEGs of GSE86468 and GSE103001.

2.3. Analysis of gene ontology and pathway enrichment

Expression inquiry of gene sets is a remarkable logical exertion to categorize similar molecular findings, including physiological procedures or chromosome sites related to several interconnected disorders [29]. Utilizing EnrichR [30], a complete gene array enhancement web tool studies on gene ontology, functional augmentation (biological methods, cellular constituents, and molecular functions) and pathway embellishment, were carried out to describe the molecular mechanisms and signal transduction way of common DEGs. Gene ontology provides enormous computational information resources by considering the components and functions of genes. An ontological concept specifies a collection of information within a specific circumstance [31]. We looked at five databases at this time: Bio Carta [32], Elsevier Pathway, KEGG [33], Reactome [34], and WikiPathway [35]. These databases of genomic information aim to provide a global perspective on the molecular functions and interactions of genes and metabolic pathways in various organisms. Then, we used shared DEGs between T2D and breast cancer to search for highly enriched pathways. Then, for analysis, these pathways were separated into functional categories. On the other hand, the GO technique classified them according to operational divisions using biological process (BP), cellular component (CC), and molecular function (MF) [36].

2.4. Analysis of protein–protein interacting networks

Protein-protein infrastructure (that forms at the end of a protein's journey within a cell) signifies the protein processes by forming similar protein affiliations. The evaluation and study of the PPI structure and its functionalities is the core and primary objective of comprehension and gaining insights into the intracellular component functions [37][38][39]. To depict the functional and physical relationships between breast cancer and T2D, we built the protein PPI connection obtained from common DEGs Utilizing the STRING (version 11.5) repository [37]. Highly interacting proteins were found utilizing topological features (degree above 15°) from PPI analysis. STRING aims to provide expanded insights on PPI in the context of various category confidence ratings (low, moderate, and high). We uploaded our PPI network in Cytoscape (version 3.7.0) for visual illustration and additional PPI connection experiments [40]. Cytoscape (v.3.7.0), a widely accessible network representation platform, consolidates several datasets to produce better results for multiple interactions, including PPIs, DNA interconnections, DNA-protein couplings, and more.

2.5. Identification of hub genes and submodule analysis

Hub genes are genes with a strong relationship to module possibilities [41]. The characteristics and progression of the disease are significantly influenced by these genes [42]. A Cytoscape plugin termed Cytohubba was used to rank and retrieve important or potential target biological network elements depending on different network properties. Cytohubba provides 11 ways to analyze networks from distinct perspectives [43]. Using the BottleNeck and DMNC method of Cytohubba, we determined the best 15 hub genes within the PPI connection.

2.6. Involvement of transcription factors and microRNAs in common DEGs

Transcription factors are proteins that play an essential part in the transformation of DNA into RNA, also known as transcription; therefore, it is essential for molecular insights [44]. To express a specific and distinct combination of proteins and RNA molecules in each type of cell in our body, transcription factors (TFs) are essential proteins that decode the information in our genome [45]. We examined the relationships between DEGs and transcription factors (TFs) as well as DEGs and microRNAs (miRNAs) to determine the governing molecules (i.e., TFs and miRNAs) that control DEGs of importance after or during transcription. We employed the NetworkAnalyst system to identify TFs from the JASPAR [46] and Chip-X [47] databases which seek to connect to our common DEGs. Tarbase [48] and mirTarbase [49] are the leading experimental validity data for miRNAs–targeting gene relations. The TFs were sorted depending on the degree greater or equal to 20 from the DEGs-TFs connection. Based on the DEGs-miRNAs network degree above or equivalent to 15, we have selected the miRNAs. Network Analyzer in Cytoscape [40] and Network Analyst [50] utilized for the topological study. Cytoscape enables researchers to select prime miRNAs having significant degrees, identify biological activities and characteristics, and provide a viable biological hypothesis.

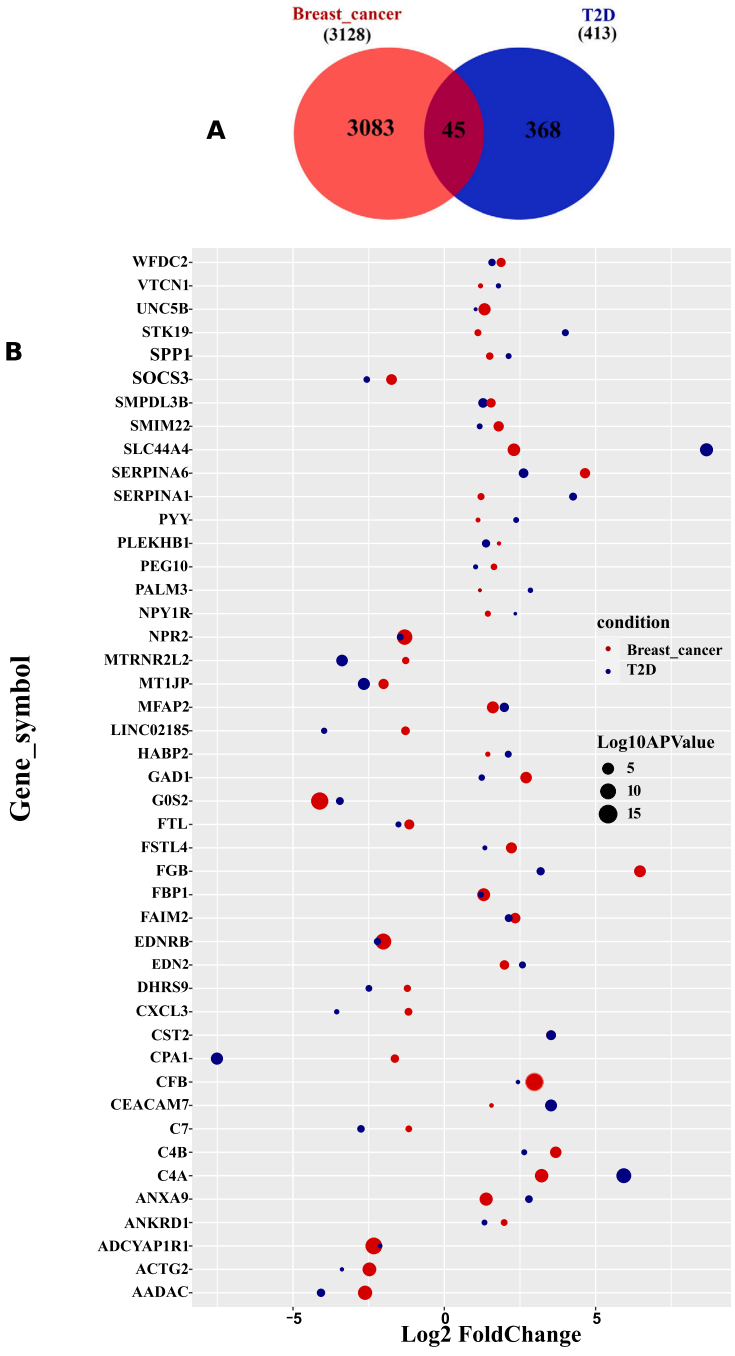

Fig. 2. The diagram provides an overview of the transcriptome analysis. A. Utilizing a venn diagram, identify the common indicator genes between breast cancer and T2D B. Bubble graph of common genes along with their corresponding log10APValue(log10 adjusted p-value) and Log2 fold change. Log2 fold change depicts the variation between the expression levels of each gene and log10APValue indicates the significance level of each gene.

2.7. Evaluation of applicant drugs

The vital element of this investigation was the forecast of protein-drug interactions (PDI) or the classification of medicinal compounds. A drug must bind to a receptor protein or an enzyme before it can act on a receptor or metabolize by an enzyme [51]. We have utilized The NetworkAnalyst platform to determine the potential for protein-drug interactions. We used the DrugBank database in the construction of protein-drug interactions [52]. Information about the protein and pharmacological targets was obtained from the DrugBank database (Version 5.0). PDI was generated for all connected, prevalent, and causative genes of the disorders targeted.

Table 1
The summary of the transcriptomic data and analysis. It comprises the dataset identification code, the sampling location, the actual number of the raw gene, the specimen value, and the notable genes.

Disease Name	GEO Platform	Tissues/Cells	GEO Accession	RAW Genes	Case Samples	Control Samples	Significant	Up Reg. Genes	Down Reg. Genes
Type-2 Diabetes (T2D)	Illumina NextSeq 500 (Homo sapiens)	Pancreatic Islet	GSE-86468	21144	6	6	413	217	196
Breast cancer	Illumina HiSeq 2000 (Homo sapiens)	Mammary tissue	GSE-103001	23588	22	22	3128	1465	1663

2.8. Gene–disease association analysis

DisGeNET is a network that combines information from different archives about human variant-disease associations (VDAs) and gene-disease associations (GDAs) (such as inherited, complicated, and ecologic diseases) [53]. DisGeNET is one of the largest collections of its kind presently offered, with 117000 genomic variants, 17000 genes and 24000 disorders [54]. DisGeNET data may be examined via the DisGeNET Cytoscape plugin, Enrichr-Ma'ayan Laboratory tool, and NetworkAnalyst platform. We have utilized the NetworkAnalyst platform to detect gene-disease correlations and complications with common DEGs.

3. Results

3.1. Exploration of transcriptomic data for gene expression

According to our methodological requirements, Table 1 represents the statistical data for female T2D and breast cancer. Table 1 includes the following information: GEO identification names, cellular characteristics, the quantity of control and case specimens, source genes for every dataset, significant genes, upregulated genes, and downregulated genes. To evaluate the impacts of gene expression on T2D patients and breast cancer patients, we obtained RNA-Seq information using Grein or NCBI. According to the study, the "Illumina Next Seq 500 (Homo sapiens)" GEO platform supplied T2DM information, and the "Illumina HiSeq 2000 (Homo sapiens)" forum provided breast cancer information. A total of 12 breast tumor samples 96 for the case and 6 for the control) were used in the T2D study. Analysis of differential expression (commonly referred to as signature data extraction) revealed a total of 21144 genes for T2D in females. In analyzing the signature data, we considered two conditions: "Log2 Fold-Change" and "Adjusted P-Value". After applying these criteria, we found 413 significant genes with adjusted p-values lower than 0.05 and abs(LogFC) larger than or similar to 1.0. Considering significant genes, we identified 217 up-regulated genes and 196 down-regulated genes. On the other hand, the dataset for breast cancer included 44 samples in total, 22 for the case and 22 for the control. We found 23588 genes after generating signature data. Then we used the same criteria, and we found 3128 significant genes. Among them, 1465 were up-regulated genes and 1663 were down-regulated genes. Then, we compared the T2D up-regulated genes and the breast cancer up-regulated genes. We also compared the genes that were down-regulated in both circumstances. We found 30 common genes that were up-regulated and 15 common genes that were down-regulated between T2D and breast cancer. The most essential common up-regulated genes were *CFB, C4A, ANXA9, FBP1, SLC44A4, UNC5B, MFAP2, FGB, GAD1, C4B, FSTL4, FAIM2, SMIM22, SERPINA6, SMPDL3B, EDN2, WFDC2, CST2, PLEKHB1, SPP1, SERPINA1, STK19, PEG10, NPY1R, ANKRD1, HABP2, VTCN1, PYY, CEACAM7, PALM3*. Also, the most important common down-regulated genes between female breast cancer and T2D were *G0S2, ADCYAP1R1, EDNRB, NPR2, AADAC, ACTG2, SOCS3, MT1JP, FTL, LINC02185, CPA1, CXCL3, MTRNR2L2, DHRS9, C7*. These DEGs were identified based on their association with the estrogen receptor (ER) signaling pathway, which play a critical role in the development and progression of ER+ breast cancer. The 45 DEGs identified in this study may not be unique to ER+ breast cancer and may be involved in other types of cancer as well. Nevertheless, their association with ER signaling and other cancer-related pathways suggests that they may play a substantial role in the development and progression of ER+ breast cancer.

3.2. Exploration of pathway and GO functional correlation

We employed five distinct databases for pathway analysis: Bio Carta [32], Elsevier Pathway, KEGG [33], Reactome [34], and WikiPathway [35]. Then, we investigated highly enriched pathways employing DEGs common to breast cancer and T2D. Then, to facilitate analysis, these pathways were divided into functional groups. We identified 479 signaling pathways associated with both disorders. We manually selected lesser pathways in consideration of adjusted p-values less than 0.05. Ultimately, we obtained the top 31 signaling pathways. Lastly, we organized the pathways according to the adjusted p-value in ascending order. Fig. 4 shows the top 20 pathways connected to T2D and breast cancer. The GO technique divided them into distinct classes using biological process (BP), molecular function (MF), and cellular component (CC). We analyzed the BP, CC, and MF database of GO terminologies. The most significant GO terms adjusted p-value under 0.05. Table 2 outlines the ten leading entries in the categories of biological methods, cellular constituents, and molecular processes.

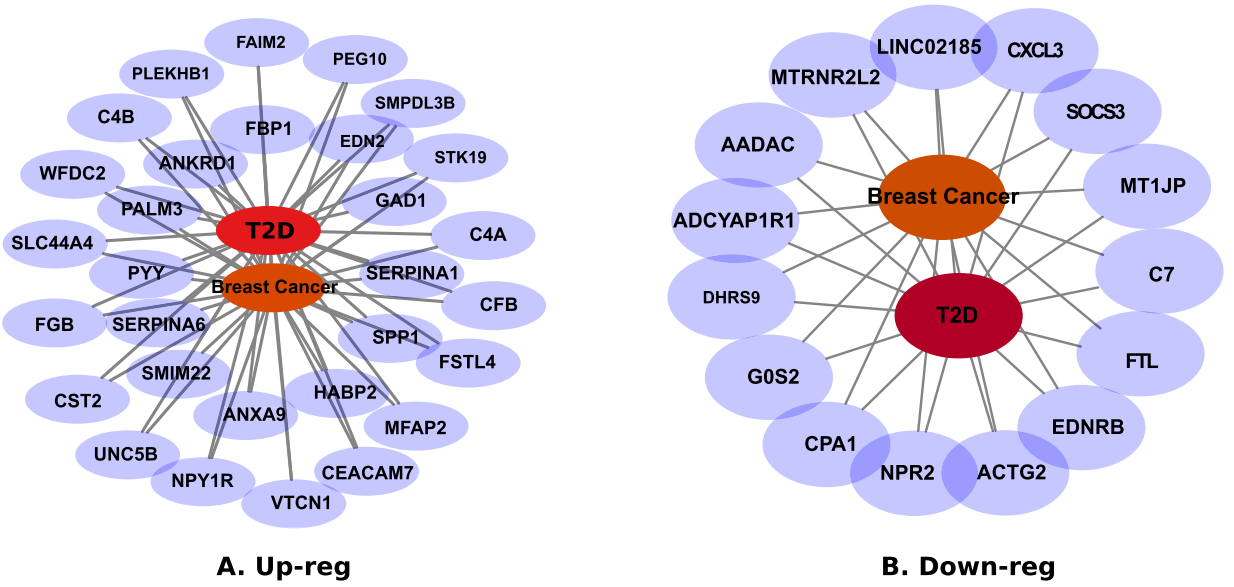

Fig. 3. Up regulated and Down regulated genes between T2D patients and breast cancer are represented separately.

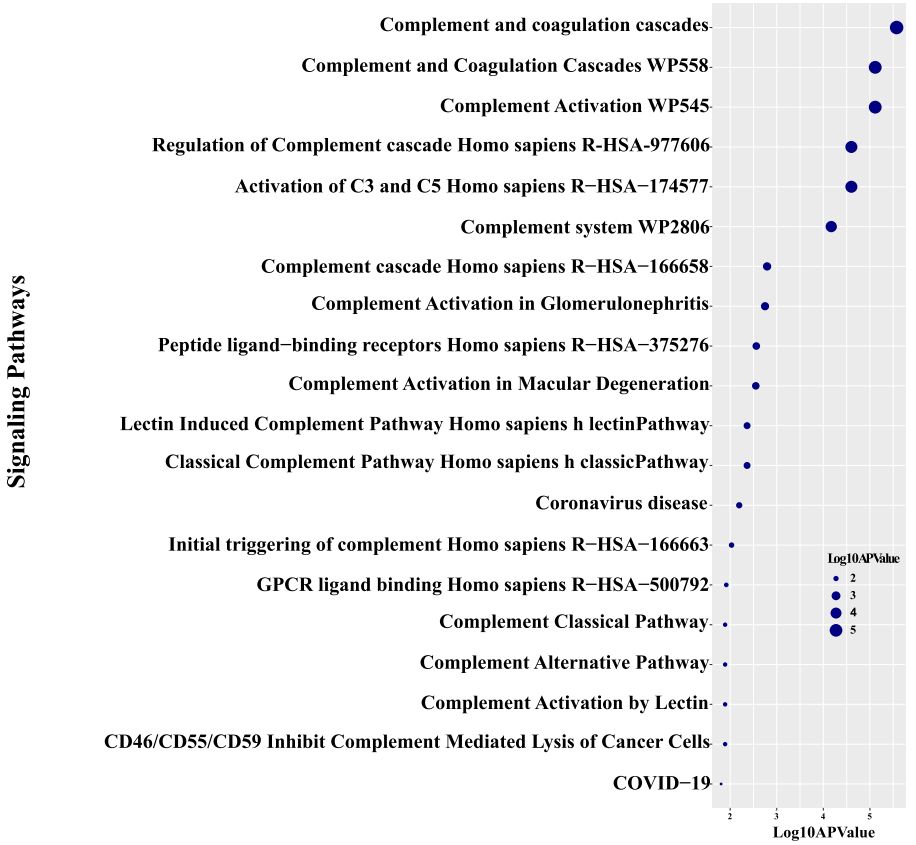

Fig. 4. 20 leading pathways related to both diseases are depicted using a bubble graph based on a transcriptomic study. The log10APValue indicates the significance level of each pathway.

Table 2
Exploration of DEGs common by breast tumors and type 2 diabetes from an ontological perspective.

Category	GO ID	Term	P-values	Genes
GO Biological Process	GO:0030449	Regulation of complement activation	4.77E-06	C4B-C4A-C7-CFB
	GO:0002697	Regulation of immune effector process	6.04E-06	C4B-C4A-C7-CFB
	GO:0002920	Regulation of humoral immune response	6.51E-06	C4B-C4A-C7-CFB
	GO:2000427	Positive regulation of apoptotic cell clearance	0.0001374	C4B-C4A
	GO:2000425	Regulation of apoptotic cell clearance	0.000137	C4B-C4A
	GO:0014829	Vascular associated smooth muscle contraction	0.000137	EDN2-EDNRB
	GO:0042310	Vasoconstriction	0.000268	EDN2-EDNRB
	GO:0048246	Macrophage chemotaxis	0.000322	EDN2-EDNRB
	GO:0014821	Phasic smooth muscle contraction	0.000322	EDN2-EDNRB
	GO:1905517	Macrophage migration	0.000510	EDN2-EDNRB
GO Cellular Component	GO:0099512	Supramolecular fiber	0.000826	MFAP2-ACTG2
	GO:0031093	Platelet alpha granule lumen	0.009974	FGB-SERPINA1
	GO:0005579	Membrane attack complex	0.013425	C7
	GO:0032982	Myosin filament	0.013425	ACTG2
	GO:0031982	Vesicle	0.014255	FGB-GAD1-AN-XA9
	GO:0031091	Platelet alpha granule	0.017488	FGB-SERPINA1
	GO:0061200	Clathrin sculpted gamma aminobutyric acid transport vesicle	0.017861	GAD1
	GO:0061202	Clathrin sculpted gamma aminobutyric acid transport vesicle membrane	0.017861	GAD1
	GO:0044754	Autolysosome	0.022278	FTL
	GO:0001527	Microfibril	0.024479	MFAP2
GO Molecular Function	GO:0004867	Serine type endopeptidase inhibitor activity	0.000332	SERPINA1-SERPINA6-WFDC2
	GO:0004866	Endopeptidase inhibitor activity	0.002204	SERPINA1-SERPINA6-WFDC2
	GO:0030283	Testosterone dehydrogenase NADP activity	0.011200	DHRS9
	GO:0015220	Choline transmembrane transporter activity	0.011200	SLC44A4
	GO:0004983	Neuropeptide Y receptor activity	0.011200	NPY1R
	GO:0005179	Hormone activity	0.013337	PYY-EDN2
	GO:0008528	G-protein coupled peptide receptor activity	0.014666	ADCYAP1R1-EDNRB
	GO:0019828	Aspartic type endopeptidase inhibitor activity	0.015646	WFDC2
	GO:0008199	Ferric-iron binding	0.015646	NA
	GO:0018455	Alcohol dehydrogenase NADP activity	0.015646	DHRS9

3.3. Classification of hub proteins and submodule

We created a PPI network on the STRING based on proteins expressed by DEGs common to breast cancer and T2D. Fig. 5 displays the outcomes of this preparation and evaluation in Cytoscape. The PPI network has 25 nodes and 76 edges. Protein subnetworks associated with two or more diseases are known to be related. Although further research is required to comprehend their roles, these recently discovered hub proteins may prove useful as therapeutic targets. Fig. 6(A, B) shows the identified hub proteins using the algorithms BottleNeck and DMNC. We identified 23 hub genes using the BottleNeck algorithm, with the top 10 highly significant genes highlighted by the colors red, orange, and yellow. Additionally, utilizing the DMNC algorithm, we discovered 12 hub genes, with 10 of them being the most prominent (represented by red, yellow, and orange colors) as evident in Table 3. The most significant genes are hub genes with a high degree of interconnection between potential modules. A high degree of connection indicates a connection ranking among the top 10%. For instance, if the module size was 1000, the top 100 genes were considered hub genes. From both algorithms, we identified the top 15 DEGs as the most significant genes-*FGA, SOCS3, UNC5B, CFB, ITGAV, EDN2, NPY1R, C4A, SERPINA6, HABP2, SERPINA1, C3, C7, FGB*, and *C4B*. In addition to the most important genes, both methods shared five genes-*FGA, CFB, C4A, SERPINA6*, and *HABP2*. We created two submodule networks to understand hub genes' interconnection and closeness (Fig. 6). These core genes may serve as possible biomarkers, resulting in novel therapeutic approaches for examined disorders. Table 3 represents the summary of hub genes.

3.4. Determination of transcriptional and post-transcriptional regulators of the differentially expressed genes

Transcription factors (TFs) are proteins responsible for the manifestation of significant genes [55]. Transcription is the process by which genes are translated into RNA or protein [56]. Gene expression is regulated by transcription factors, which are present in all alive organisms [57]. TF genes control a wide range of biological processes, making them extremely important [58]. MicroRNAs are a type of non-coding RNAs that influence gene manifestation in significant ways [59]. MiRNA contents (enriched miRNA) may significantly impact physiological functions, transcription, and message transmission [60][61][62]. MiRNAs regulate breast tumors cell proliferation and cycle advancement by engaging with cyclins, protein kinases, and growth promoters or suppressors [63]. Numerous studies have linked miRNA dysregulation to T2DM pathogenesis including insulin secretion, glucose metabolism, insulin deficiency, lipogenesis, islet maturation, beta-cell and adipocyte proliferation, as well as diabetes management [64][65]. We performed the TFs-DEGs interaction results illustrated in Fig. 7 and the miRNAs-DEGs reaction study depicted in Fig. 8 to explore the post-transcriptional and transcriptional moderators of the DEGs. Fig. 7 displays the TF-Gene connections. The governing genes are: *PLEKHB1, FTL, SMPDL3B, GAD1, SPP1, PEG10, FAIM2, FSTL4, SOCS3, EDN2, FBP1, UNC5B, G0S2, ANKRD1, SERPINA1, SLC44A4,*

Table 3
Summary of breast cancer and type 2 diabetes (T2D) hub genes acquired by study of protein-protein interactions.

Gene Symbol	Description	Feature
FGA	Fibrinogen alpha chain	Signaling receptor binding and protein macromolecule adaptor activity
SOCS3	Suppressor Of Cytokine Signaling 3	Protein kinase inhibitor activity
UNC5B	Unc-5 Netrin Receptor B	Programmed Cell Death
CFB	Complement Factor B	Serine-type endopeptidase activity and complement binding
ITGAV	Integrin Subunit Alpha V	Protease binding and voltage-gated calcium channel activity
EDN2	Endothelin 2	Hormone activity
NPY1R	Neuropeptide Y Receptor Y1	G protein-coupled receptor activity and peptide YY receptor activity
C4A	Complement C4A	Endopeptidase inhibitor activity and complement component C1q complex binding
SERPINA6	Serpin Family A Member 6	Serine-type endopeptidase inhibitor activity and steroid binding
HABP2	Hyaluronan Binding Protein 2	Calcium ion binding and glycosaminoglycan binding
SERPINA1	Serpin Family A Member 1	Identical protein binding and protease binding.
C3	Complement C3	Signaling receptor binding and C5L2 anaphylatoxin chemotactic receptor binding
C7	Complement C7	Complement cascade and Initial triggering of complement
FGB	Fibrinogen Beta Chain	Signaling receptor binding and chaperone binding
C4B	Complement C4B	Carbohydrate binding and complement binding

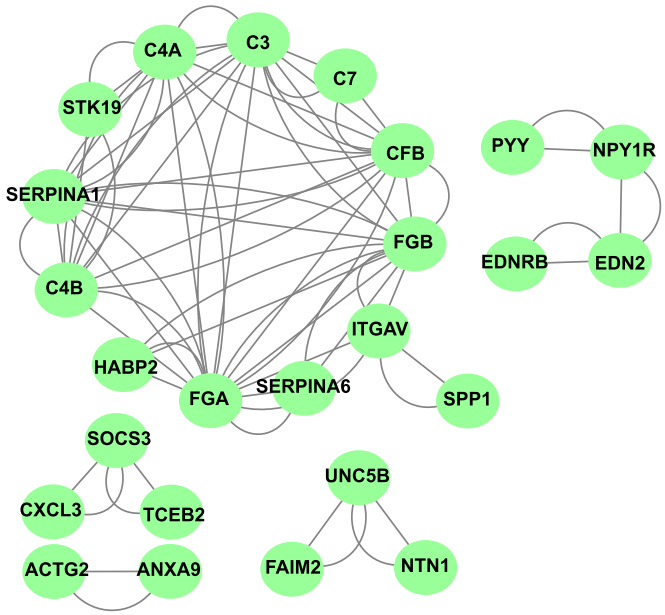

Fig. 5. The figure describes the PPI network of DEGs between T2D and breast cancer. Circular nodes indicate DEGs and edges denote node relations. The PPI network consists of 25 nodes and 76 edges. Employing String, the PPI network was constructed and displayed in Cytoscape.

AADAC, CFB, ADCYAP1R1, EDNRB, MFAP2, C4A, ACTG2, HABP2, FGB, TCF3, ZNF281, SPI1, RAD21, GATA2, PPARG, TRIM28, TP63, HNF4A, SOX2, POU5F1, NANOG, USF2, FOXC1, NFKB1, FOXL1, E2F1, TFAP2A, SRF, YY1, NFIC. Fig. 8 graphically shows miRNA's genes relation: mir-4784, mir-193b-3p, mir-765, mir-3150b-3p, mir-484, mir-5096, mir-2114-5p, mir-98-5p, mir-335-5p, mir-204-5p, mir-30b-5p, mir-30a-5p, mir-30c-5p, mir-30e-5p, mir-221-3p, mir-30d-5p, let-7f-5p, mir-7-5p, mir-146a-5p, mir-16-5p, mir-129-2-3p, mir-1-3p, mir-27a-3p, mir-374a-5p, mir-34a-5p, mir-124-3p, mir-1343-3p, let-7b-5p, mir-195-5p. Summary of post-transcriptional regulatory biomolecules of T2D and breast cancer is shown in Table 4.

3.5. Detection of potential drugs

Protein-drug interactions are essential for understanding ligand affinity [66]. It is essential to evaluate protein–drug associations to comprehend the structural characteristics indicated for receptor accessibility [67][68]. We identified 13 potential therapeutic compounds for common DEGs as suspected drug targets in T2D and breast cancer. Fig. 9 represents Protein Drug Interactions for T2D and breast cancer complicacy and the common gene is *CPA1*. We considered other genes as well but found that CPA1 had the strongest connections to drug targets and therefore focused on it as a potential therapeutic target.

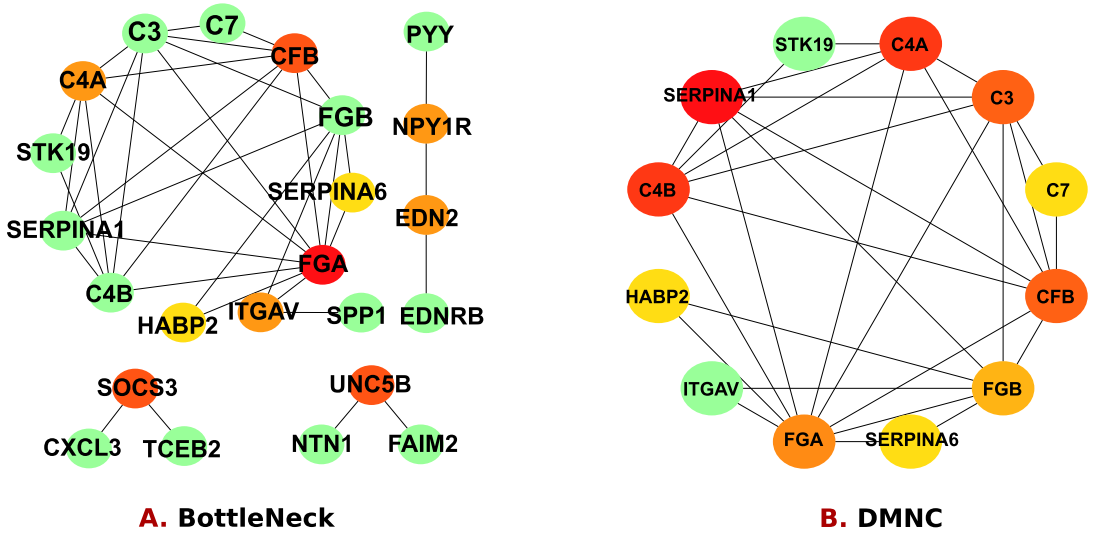

Fig. 6. Utilizing the Cytohubba addon in Cytoscape to identify hub genes from the PPI network. We used the most recent BottleNeck and DMNC Cytohubba plugin methods to obtain hub genes. The BottleNeck network contains 23 nodes, 37 edges, and DMNC nodes and edges totaling 12 and 29. Node color indicates the degree of interaction (red means high level, orange/yellow means moderate level, and green means low level).

Table 4
Summary of post-transcriptional regulatory biomolecules of T2D and breast cancer.

Symbol	Description	Feature
mir-193b-3p	MicroRNA-193b	Anti metastatic role in the breast cancer
mir-765	MicroRNA-765	Tumor suppressor and eliminates lipids in clear cell renal cell carcinoma.
mir-3150b-3p	MicroRNA-3150b	Inhibits the proliferation and invasion of cervical cancer, colorectal cancer, hepatocellular carcinoma.
mir-484	MicroRNA-484	Effects on inflammation, apoptosis, and mitochondrial function.
mir-5094	MicroRNA-5094	Suppresses human peripheral blood T cell proliferation.
mir-2114-5p	MicroRNA-2114	Involved in osteosarcoma metastasis.
mir-98-5p	MicroRNA-98	Anti-carcinogenic functions. Regulates the proliferation and metastasis of breast cancer cells.
mir-335-5p	MicroRNA-335	Inhibits the growth, chemo-sensitivity, and metastasis of human breast cancer cells. Overexpression impairs insulin secretion.
mir-204-5p	MicroRNA-204	Inhibits the proliferation, migration, and invasion of various carcinoma.
mir-30b-5p	MicroRNA-30b	Promotes proliferation, migration, and invasion of breast cancer cells. Involved in the pathological process of diabetes mellitus.
mir-30a-5p	MicroRNA-30a	Significantly promotes cell apoptosis, and acts as a tumor suppressor. Associated with dysglycaemia.
mir-30c-5p	MicroRNA-30c	Suppressed pancreatic cancer cell proliferation. Protects against diabetic cardiomyopathy.
mir-30e-5p	MicroRNA-30e	Correlates with the clinical stage of breast cancer. Suppression of hepatocellular carcinoma.
mir-221-3p	MicroRNA-221	Promotes breast cancer resistance to adriamycin. Key factors of insulin/insulin-like signaling pathway.
mir-30d-5p	MicroRNA-30d	Induces insulin transcription factor and insulin production. Responsible for tumor development and progression
let-7f-5p	MicroRNA-Let-7f	Inhibits tumor invasion and metastasis. Promotes angiogenesis.
mir-7-5p	MicroRNA-7	Inhibit the invasion and metastasis of breast cancer cells. Regulates pancreatic beta cell function.
mir-146a-5p	MicroRNA-146a	Act as a modulator of the innate immune response. Role in the proliferation of breast cancer cells.
mir-16-5p	MicroRNA-16	Inhibit gastric cancer, breast cancer, and chordoma. Affected insulin-mediated glucose handling.
mir-129-2-3p	MicroRNA-129-2	Suppresses proliferation and migration of carcinoma cell .
mir-1-3p	MicroRNA-1	Function as a tumor suppressor for breast cancer. Involved in the pathophysiology of diabetes.
mir-27a-3p	MicroRNA-27a	Role in oncogenesis, cell-proliferation, tumor cell metabolism and chemotherapy-resistance.
mir-374a-5p	MicroRNA-374a	Suppressor of oncogenic signaling in breast cancer. Regulates inflammatory response in diabetic nephropathy.
mir-34a-5p	MicroRNA-34a	Affecting numerous oncogenes and cancer pathways.
mir-124-3p	MicroRNA-124	Role in CNS development, such as neuronal differentiation, maturation, and survival.
mir-1343-3p	MicroRNA-1343	As a tumor suppressor for many types of cancer.
let-7b-5p	MicroRNA-let-7b	An important role in cell maturation. Inhibition of breast cancer tumor growth. Contributes to Diabetic Retinopathy.
mir-195-5p	MicroRNA-195	Promotes pancreatic beta-cell dedifferentiation. Inhibits proliferation, invasion, and metastasis in breast cancer.

3.6. Identification of disease association

Cooperatively interacting and disease genes play roles in the progression of complicated disorders [69]. Targeting disease-associated functions or pathways provides insight into disease mechanisms [70]. Disorder-specific therapeutic design techniques begin with identifying the connection between genes and diseases [71]. Based on NetworkAnalyst's evaluation of the gene-disease connection, we observed that mammary Neoplasms and autosomal recessive predisposition were most precisely synchronized to our identified hub genes. Other includes - diabetes mellitus, pneumonia, liver cirrhosis, pulmonary fibrosis, neoplasm metastasis, adenocarcinoma, asbestosis, liver Cirrhosis, mood Disorders, neurodegenerative disorders, mental and motor retardation, global developmental delay, cognitive delay and cerebellar Ataxia. Fig. 10 shows the connection between genetic disease.

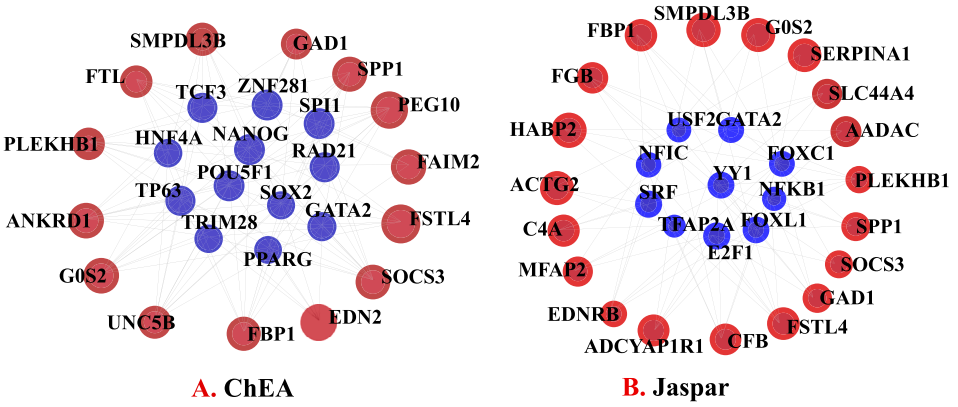

Fig. 7. Two different methods, ChEA and Jaspar, demonstrated TF-Gene correlations to highlight the relationship between breast cancer and T2D. The circles of blue demonstrate TF-genes that interact with the DEGs gene (circular blue color).

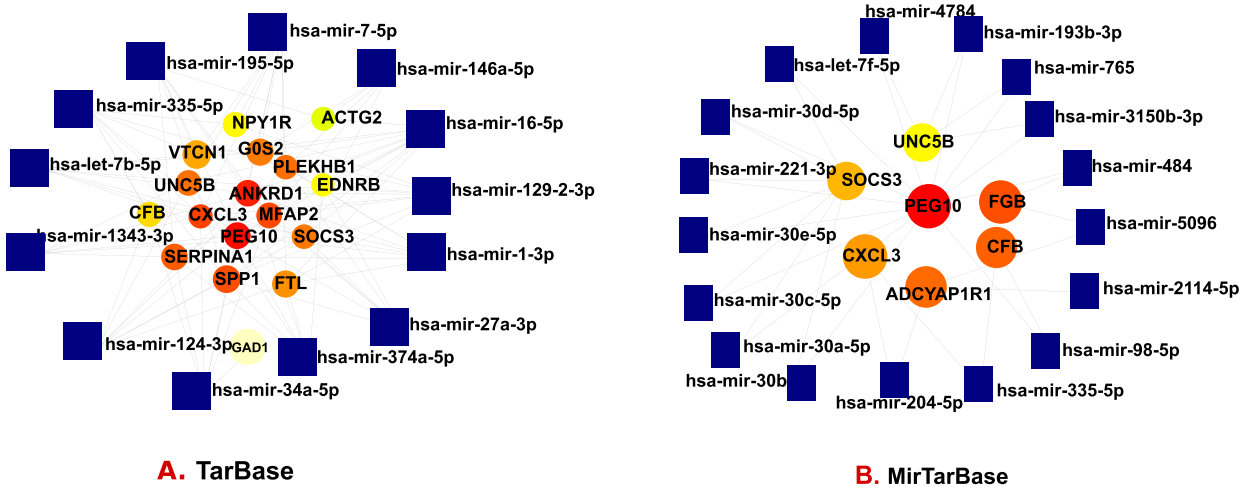

Fig. 8. Applying two algorithms named TarBase and MirTarBase, the miRNA gene between breast cancer and T2D has been established. The circular red, orange, and yellow colors (red means high level, orange means moderate level, and yellow means low level) denote DEGs genes interacting with the mi-RNA gene (blue color square box).

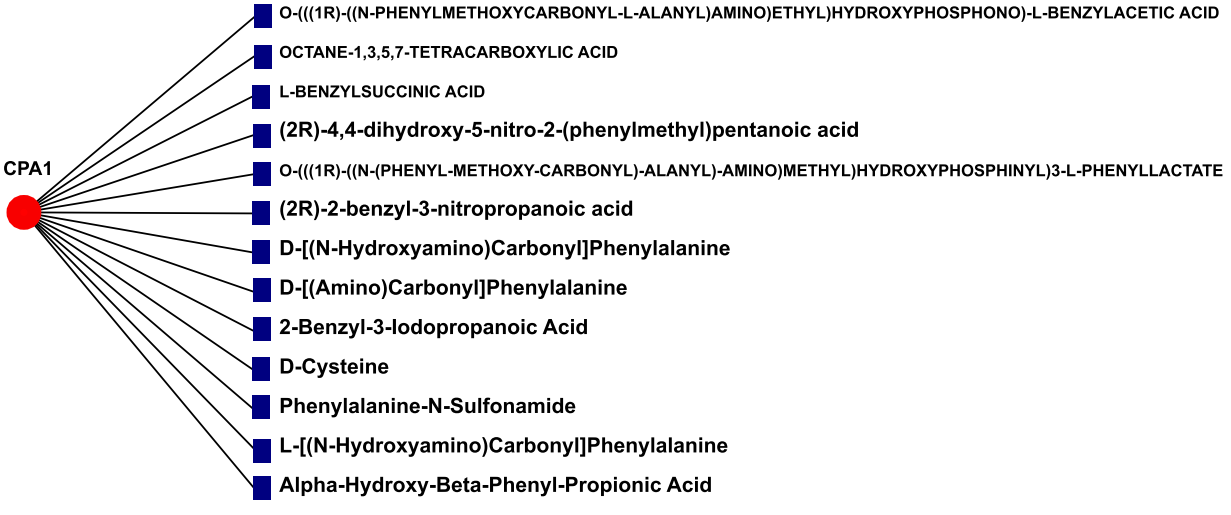

Fig. 9. Using the Protein Drug Interaction method, the diagram depicts the Protein Drug combination between breast cancer and T2D.

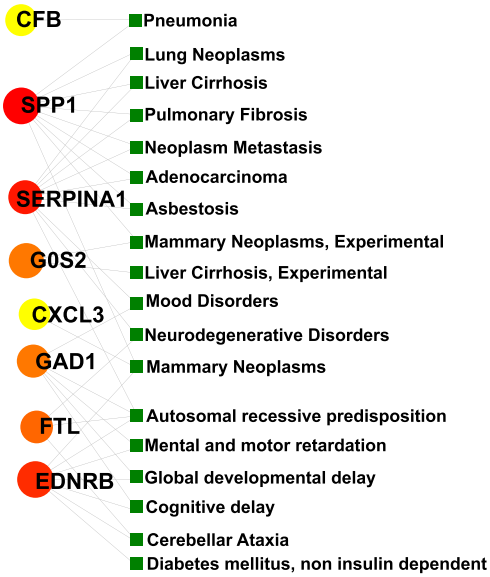

Fig. 10. The gene-disease correlation network depicts illnesses related to reciprocal DEGs.

4. Discussion

Diabetic women are 23% more likely than non-diabetic women to develop breast carcinoma [72]. A history of diabetes is associated with a 37% mortality risk in breast cancer patients [73][74]. Additionally, diabetes can exacerbate the symptoms of breast cancer, and the symptoms of diabetes in women may resemble those of breast cancer. Concentrating on the effect of comorbid conditions on long-term results is crucial due to the prolonged lifespan of breast cancer survivors. Diabetes and breast cancer have common symptoms that impair quality of life, including distress, fatigue, diminished physical function, anxiety, and sexual dysfunction [75][76]. Both disorders also share central nervous system complications, such as peripheral neuropathy, impaired cognitive function, and attention dysfunction [77][78][79][80]. Even though diabetes and breast cancer have different etiologies for these symptoms, cohabitation can worsen their conditions. Type 2 diabetic women have marginally more likely to develop breast cancer. Due to insulin resistance, any organ with a greater concentration of estrogen receptors, including the ovaries, breast, and endometrium, has an increased risk of developing cancer [81]. Different cytokines, IGFBP-3, and IGF-1 levels can raise the risk of breast cancer development [82][22]. Some T2D medications have both positive and negative affect on the development of different types of breast cancer in females. Metformin, a drug used to reduce blood glucose levels in individuals with T2D, may lower the likelihood that women will develop estrogen-positive (ER-positive) breast cancer. According to a study that appeared in Annals of Oncology, diabetic women who had used the medicine metformin for at least ten years had a 38% lower risk of developing ER-positive breast carcinoma than non-diabetic women. In fact, compared to non-diabetic women, it was associated with a 25% elevated risk of ER-negative breast carcinoma and a 74% elevated chance of triple-negative breast carcinoma [83][84]. After one year of use, sitagliptin may lower the incidence of breast carcinoma in female patients suffering from T2D. However, it is not always possible to pinpoint the exact mechanisms or biological pathways that contribute to the elevated risk of breast carcinoma among T2D sufferers. Because T2D and breast cancer have such significant therapeutic implications, we have tried to take advantage of the complexity of these two diseases by examining pathways and interactions between them. T2D promotes breast cancer, but how is unknown, making hypothesis-driven biology or endocrinological study challenging. It is essential to utilize well-established bioinformatics techniques and analysis methods to investigate functional disorders interacting with genes and pathways to Determine new factors that contribute to comorbidity interactions. We analyzed GEO RNA-seq datasets from publicly available archives. To gain insight into how these comorbidities interact, we discovered DEGs shared by T2D and breast cancer and built networks using the DEGs. These DEGs helped to find linked GO keywords and dysregulated biological pathways. After applying an adjusted p-value level of 0.05, there was a specific technical challenge, and manual collection significantly reduced the number of pathways and GO classes. In addition to pathways and GO categories, we investigated previously unstudied DEGs-TFs and DEGs-miRNA interactions, protein-protein interaction (PPI) study, and hub protein identity. We searched the STRING database for PPIs corresponding to the products of our targeted genes. This study discussed the recognized protein-protein interactions that may be the primary agents of pathogenesis for several illnesses. We didn't evaluate projected PPIs for this reason; instead, we only looked at experimentally verified PPI data. With the help of topological parameters, we created the PPI utilizing the DEGs shared by T2D and breast cancer. Many existing breast cancer datasets, including GSE29431, GSE32641, GSE61304, GSE70947, and GSE86374, contained the observed DEGs, which lends credence to the robustness and reproducibility of the findings [85][86].

Differential Expression Analysis (DEA) was performed to evaluate any substantial dysregulation. DEA is proceeded by the detection of ubiquitous genes, including up- or down-regulated genes (expressed in Fig. 2 and Fig. 3, consecutively), pathways (shown

in Fig. 4), GO (displayed in Table 2), PPIs (shown in Fig. 5), hub–protein interactions (expressed in Fig. 6 and Table 3), TFs gene interactions (displayed in Fig. 7), gene–miRNA interactions (shown in Fig. 8 and Table 4), PDIs (represented in Fig. 9) and GDAs (displayed in Fig. 10). The T2D RNA-seq dataset was obtained from patients with type 2 diabetes and non-diabetes, and the breast cancer dataset was collected from the healthy control group and case group (shown in Table 1). In addition, we compared our findings to earlier research published in several journals. In Fig. 1, the diagram of our strategy has been graphically illustrated and effectively oriented.

Gene Ontology (GO) enrichment analysis creates hypotheses on the innate physiological facts of research and evaluates high throughput molecular data [87]. Initially, we concentrated on the leading GO biological process: the regulation of complement activation (4 genes) and the regulation of immune effector process (4 genes). The cellular complement system is an important mediator of the innate immune response, contributing to the control of several phases of an inflammatory response [88]. In T2DM, an inflammatory reaction arises due to an autoimmune response to high blood glucose and inflammatory mediators generated by adipocytes and macrophages in adipose tissue. This mild and persistent inflammation leads to damage to the beta cells of the pancreas and leads to an inadequate production of insulin [89]. Immunosurveillance is a mechanism of the inherent immune system that identifies and destroys cancerous cells in the earlier stages of cancer development [90]. In cancer immunosurveillance, humoral immunity exerts pressure on a growing tumor, eliminating malignant cells before a tumor develops [91]. This same immunological pressure influences tumor growth and selects for specific mutations, generating immune-evasive breast cancer. Secondly, we focused on cellular components; the top GO terms are Supramolecular fiber, platelet alpha granule, membrane attack complex, myosin filament, vesicle, clathrin sculpted gamma-aminobutyric acid transport vesicle, autolysosome, and microfibril. P-selectin is a protein found in the alpha granules of platelets that promote the proliferation of tumor cells in breast tissues [92]. In patients with T2D and angiopathy, a higher number of activated large platelets circulates [93]. In T2DM, the membrane attack complex serves as a marker for complement overexpression [94]. Protectin on breast cancer cell membranes prevents complement membrane attack. Neutralizing protectin on cancer cells enhances their complement lysis sensitivity [95]. T2DM causes alterations in cardiac myosin that reduce the heart's contraction rate [96] and an invasive breast tumor cell line, myosin-interacting guanine nucleotide exchange factor (MyoGEF) regulates the expression of RhoA and RhoC as well as cell polarity and invasion [97]. Abnormal extracellular vesicles (EVs) can cause insulin resistance and lead to T2DM complications [98]. EVs transport proteins and nucleic acids that promote cancer growth and metastasis [99]. Finally, we focused on the top two GO molecular functions: endopeptidase inhibitor activity (3 genes) and serine type endopeptidase inhibitor activity (3 genes). Trypsin, a digestive serine protease that promotes invasion, proliferation, and metastasis, has also been associated with several malignancies. Protease engagement in carcinoma promotes the use of endopeptidase inhibitors as anti-cancer agents [100]. Endopeptidase activity facilitates inflammatory cell invasion of pancreatic islets and inhibits beta-cell multiplication. The role of endopeptidase in cancer proposes the use of proteolytic enzyme inhibitors as antineoplastic drugs [101].

Pathway analysis is the means of representing an organism's responses via internal alterations. Almost all pathways are associated with the complement system. The complement system referred to as the complement cascade, is a fundamental component of our innate immune system and contains above 50 soluble protein regulators and receptors [102]. In malignant tumors, complement protein expression is elevated, and complement activation in the tumor microenvironment develops carcinogenesis and expansion. Multiple studies have demonstrated that Complement C1q, Complement C3a, and Complement C5a are highly expressed in breast cancer and promote carcinogenesis, tumor development, proliferation, and metastasis [103]. Several studies have shown the relationship between the activation of the complement system and the onset and development of metabolic diseases such as insulin resistance and T2DM. Complement C3, complement C4, factor D, and complement CD59 has been involved in the diagnostics of diabetes [104][105]. Possible treatments for insulin resistance include suppression of the Complement C1 and Complement C3 complement activation as well as its regulators [106]. Complement C3a and C5a's impacts on adipose tissue and their receptors are crucial to both the pathophysiology of adipose tissue and the onset of metabolic disorders [107]. Elevated complement levels are associated with diabetic complications such as neuropathy, nephropathy, and retinopathy [108]. G protein-coupled receptors (GPCRs) are the broadest class of receptor molecules associated with the genesis and development of numerous cancers, notably breast cancer [109]. For the treatment of Type 2 diabetes mellitus, GPCRs have been examined as prospective therapeutics [110]. In human breast cancer, the complement-regulating proteins CD46 and CD59 have elevated levels of expression. CD46, CD59, and CD55 are connected to diabetic retinopathy.

Hub genes have a substantial impact on the features and course of the disease. Fibrinogen alpha chain (FGA) exhibits significant expression in breast cancer. Fibrinogen is pro-tumorigenic in breast cancer due to its ability to bind to growth factors, such as vascular endothelial growth factor, promote angiogenesis, interact with other matrix proteins that serve as a physiological defensive barrier, and act as a scaffold for cellular proliferation via integrin binding [111]. FGA polymorphisms may play a significant role in the progression of endothelial dysfunction, ultimately resulting in type 2 diabetes and insulin resistance. FGA polymorphisms increase plasma fibrinogen levels, levels of lipids, and glucose [112]. Suppressor Of Cytokine Signaling 3 (SOCS3) is a significant factor in insulin resistance in muscle and abnormalities in glycaemic control, indicating that suppressing SOCS3 in muscle may be an effective method for recovering insulin action in T2D patients [113]. Suppression of SOCS1 and SOCS3 genes may facilitate the neoplastic transition of epithelial tissues, according to a theory based on the DNA-hypermethylation of SOCS genes in breast cancer [114]. By triggering PI3K/Akt signaling, Unc-5 Netrin Receptor B (UNC5B) plays a significant role in the progression of breast tumor cell proliferation and metastasis, and targeting UNC5B is a possible technique for customized breast carcinoma treatment [115]. Expression of UNC5B and netrin-1 is associated with diabetic nephropathy and angiogenesis [116]. Changes in Complement Factor B (CFB) expression as a potential cause of insulin resistance and hypertension, and CFB is higher in human cohorts with T2D and cardiovascular disease [117]. Breast cancer molecular subtypes, especially the LA subtype, have a high correlation with CFB [118]. Suppression of Integrin Subunit Alpha V (ITGAV) decreased cell proliferation, migration, and self-renewal in breast cancer cell lines

through modifying BCL2 and PXN expression, and metastatic breast cancers with elevated ITGAV expression may be treated by targeting ITGAV [119]. In human breast cancer, excessive EDN2 expression is often inhibited by promoter methylation, presumably causing abnormal regulation of the ET-axis, which may exacerbate this disease [120]. In type 2DM, increased Neuropeptide Y Receptor Y1 (NPY1R) mRNA expression in human islets impaired insulin production by enhancing Y1 receptor activation [121]. In breast cancer, NPY1R overexpression is linked with disease stage and lymph node metastases. Compared to no NPY1R expression, patients with higher NPY1R expression exhibited decreased tumor-specific mortality [122]. Breast cancer chemotherapy resistance is associated with Serpin Family A Member 6(SERPINA6) [123]. By interacting with hyaluronic acid (HA), the Hyaluronan Binding Protein 2 (HABP2) gene generates an extracellular serine protease that promotes the breakdown of the extracellular matrix, causing rupture of the endothelium, ultimately enhancing tumor metastasis and angiogenesis [124]. Serpin Family A Member 1 (SERPINA1) is associated with cardiovascular complexity in T2D [125]. SERPINA1 expression targets ER and HER2-regulated genes and is associated with HER2-positive and estrogen receptor-positive (ER+) breast cancer and predicts survival in ER+ patients, it may have diagnostic potential [126]. Elevated Complement C3 expression levels are related to metastasis in breast cancer patients, as determined by ELISA test of serum samples from breast cancer patients [127]. C3 is related with the progression of diabetes after adjusting for possible confounders like BMI, insulin, and other inflammatory markers [128]. In the initial phases of diabetes, Complement C7 and its gene product are expressed more in early diabetic nephropathy kidneys (EDN) due to complement deposition [129]. C7 activates MAPK, VEGF, or JAK/STAT signaling pathways, which promotes the development of breast cancer [130]. Fibrinogen b-gene(FGB) polymorphism connects with cardiovascular risk in type 2DM and glycemic control rigorously affects fibrinogen plasma concentrations [131]. Plasma fibrinogen levels may predict postoperative metastasis and mortality in advanced breast cancer patients undergoing neoadjuvant chemotherapy [132]. C4B-binding protein (C4BP) influences human islet amyloid polypeptide (IAPP)-mediated inflammation, amyloid accumulation in pancreatic islets, and IAPP-mediated beta cell damage [133]. In breast cancer, the complement inhibitor CSMD1 acts as a tumor suppressor by suppressing Complement C4B [134]. These identified hub genes have the potential to serve as biomarkers, and it will be significant to eradicate mutations in these genes to achieve recovery. We still have not identified the relationship between ITGAV, EDN2, and SERPINA6 genes with T2D and HABP2 in breast cancer. However, these genes may serve as prospective therapeutic targets, opening the door for deeper analysis and research.

We also examined the TF-gene and miRNA interactions to identify the transcriptional and post-transcriptional regulators of the common DEGs. Transcription factors control the rate of transcription, and the post-transcriptional gene is regulated by miRNAs. To comprehend the onset of disease, TFs, and miRNAs are significant. In this way, the common DEGs, TFs, and miRNAs were connected to our analysis. We discovered several TF genes, including *FTL, SPP1, PEG10, SOCS3, EDN2, FBP1, UNC5B, and SERPINA1*. *FTL* (Ferritin levels) is linked to the breast cancer development [135]. Among Tamoxifen-treated women with ER+ breast cancer, high levels of secretory phosphoprotein 1 (SPP1) gene expression is linked to an elevated risk of distant recurrence [136]. By binding to ER, estrogen acts as a growth factor in malignant tumors, promoting cell division and development. Tamoxifen inhibits the effects of estrogen by inhibiting DNA synthesis by binding to estrogen receptors on tumors and other tissue targets. SPP1 is significantly more abundant in breast carcinoma tissue, making it a potential option for immunotherapy and a novel prognostic biomarker [136][137]. The overexpression of PEG10 (Paternally Expressed Imprinted Gene 10) in breast cancer cells promotes migration and invasion as well as cell cycle, clone formation and cell proliferation [138]. BMI and breast cancer are linked to the expression of the SOCS3 gene, which suppresses cytokine signaling. Inhibiting cytokine signaling and promoting protein degradation, SOCS proteins are negative JAK/STAT pathway regulators [139]. Proliferation and anchorage-independent growth are both decreased when SOCS3 gene expression is overexpressed. Several cytokines, including prolactin, GH, ILs, and insulin, promote SOCS3 gene expression, while glucocorticoids decrease it. Cytokines and SOCS proteins influence the mammary gland's development and functionality [139]. In the mammary gland of humans, EDN2 (Endothelin2) acts as a tumor suppressor [120]. Breast cancer is also linked to fructose-1, 6-bisphosphates (FBP1). Compared to individuals with luminal breast cancer and patients with basal-like breast cancer have lower FBP1 expression levels in tumor tissue. FBP1 may have anticancer effects in breast cancer cells because of the potential decrease in HIF-1 alpha overexpression [140]. Breast cancer patients have a worse prognosis when UNC5B is highly expressed because it promotes tumor growth and metastasis [115]. Breast cancer cell growth and metastasis are decreased by UNC5B knockdown, which also affects PI3K/Akt signaling activity. A potential approach for customized breast cancer treatment is to target UNC5B, a biomarker with diagnostic and prognostic potential. SERPINA1 is a direct estrogen receptor target gene and a survival predictor. Additionally, it has been linked to a number of tumors, including breast, lung, and colon cancer. Breast cancer patients have increased SERPINA1 expression, which results in E2-independent ER binding [141]. The identified miRNAs, such as mir-193b-3p,mir-98-5p,mir204-5p,mir-30e-5p, and mir-146a-5p are associated with regulating the proliferation and metastasis of breast cancer cells. We also predicted mir-335-5p, mir-30b-5p, mir-7-5p, mir-16-5p, mir-13p, mir-374a-5p, mir-let-7b-5p, mir-195-5p that are associated with different genes of breast cancer and T2D. Among them, overexpression of mir-335-5p impairs insulin secretion but inhibits the growth, chemoresistance, and metastasis of human breast cancer cells; mir-30b-5p, mir-7-5p,mir-13p, mir-16-5p, mir-195-5p inhibit the migration, proliferation, and invasion of breast cancer cells and also involved in the pathological process of T2D. Further, mir-374a-5p is a breast cancer suppressor of oncogenic signaling and also regulates inflammatory response in diabetic nephropathy. Furthermore, mir-let-7b-5p inhibits breast cancer tumor growth and also contributes to diabetic retinopathy. The multiplication and apoptosis of pancreatic beta-cells are regulated by miR-16-5p-targeted genes [142]. MiR-16-5p can suppress proliferation of breast cancer cells through down-regulating expression of ANLN [143].

From gene-disease analysis, we may predict numerous diseases are associated with T2D and breast cancer, including respiratory, cerebral, cardiovascular, liver, congenital, and some types of carcinoma. T2D is associated with an increase in pneumonia-related mortality and the development of pneumonia following breast cancer radiotherapy [144][145]. Previously existing T2D may raise the risk of lung cancer and women with breast carcinoma have an elevated chance of developing subsequent lung cancer, presumably

due to smoking interaction and radiotherapy [146][147]. Increased estrogen levels are linked to liver cirrhosis, which may have a causative relationship with breast cancer and approximately 30% of those who have cirrhosis also have T2D [148][149]. The lung is one of the targeted organs of diabetes, manifesting as pulmonary fibrosis [150]. Prediabetes may enhance tumor development and metastasis by triggering mitogenic signaling [13]. Women exposed to asbestos have greater incidences of breast cancer, according to several reports [151]. The main symptoms of major depressive disorder are mood disorders and psychomotor retardation. Depression impairs the prognosis of diabetes in diabetic patients, and depression may increase the likelihood of acquiring T2D by 60% [152]. Depression is common but underappreciated and untreated among breast cancer patients, causing amplification of somatic symptoms, impaired functioning, and poor treatment compliance [153]. Insulin resistance, typically linked with T2DM, may result in a lack of insulin actions in the CNS, leading to neurodegenerative disorders (cognitive delay, cerebellar ataxia) and global developmental delay [154]. Compared to type 2 diabetes, autosomal recessive predisposition has a significant impact on the development of breast cancer in women [155].

We discovered 13 potential therapeutic substances (Fig. 9) linked to the carboxypeptidase A1 (CPA1) gene. The zinc metalloprotease carboxypeptidase A1 (CPA1) is generated by the acinar cells of the pancreas and is responsible for cleaving the C-terminal branched-chain and aromatic amino acids from ingested proteins [156]. In chronic pancreatitis, incorrect CPA1 mutations were discovered, and mutation induced CPA1 misfolding induces endoplasmic reticulum(ER) stress [157]. Chronic pancreatitis causes 25-80% of diabetes cases and increases the risk of pancreatic cancer [158]. Pancreatic cancer is closely linked with an about 10% elevated risk of breast and ovarian cancer occurrence and death [159]. High blood sugar and dyslipidemia associated with type II diabetes alter ER homeostasis, leading to irreversible activation of the unfolded protein response (UPR). The activity of various UPRs involved in breast cancer carcinogenesis [160]. No pharmacological data exists about these 13 potential medicinal compounds. Their application in medicine has not been investigated yet. These compounds may prevent mutation of this carboxypeptidase A1 (CPA1) gene. Thus, the mechanisms mentioned earlier are stopped, and the progression of chronic pancreatitis-induced diabetes, breast cancer, and ER stress-induced diabetes, breast cancer is halted.

It is significant to acknowledge that our study had certain limitations. The datasets employed contained a restricted number of samples, and we did not collect any other types of diabetic datasets. Furthermore, we did not consider age, sex, race, and other pertinent characteristics in our analysis. Thus, additional validation is required to thoroughly analyze the biological relevance associated with the study's findings.

5. Conclusions

The current study used a multi-omics technique to examine transcriptomics datasets of T2D and breast cancer to decode the shared genes expressed in both conditions. Incorporating the overlapped DEGs with other biomolecular interaction networks led to the identification of 15 hub proteins from protein-protein interactions, regulatory TFs from DEG-TF interactions research, and miRNAs from DEGs-miRNA interactions study. The TFs and miRNAs were also novel as the hub genes and pathways, meaning they weren't previously confirmed to play a significant role in the relationships between T2D and breast cancer. In this method, the current work showed biomarkers at the protein, RNA, pathway, and GO levels, but further research is required to identify them officially as biomarkers. These findings point to T2D genes that are differentially expressed and may be significant in the development of breast cancer. As a result, we used transcriptome analysis to identify shared molecular pathways and biomarkers in patients with T2D and breast cancer to understand how the two diseases are related. They may also provide new information about these diseases. Researchers and medical professionals may also use it as a crucial technique to identify the underlying pathophysiology and nature of the specific fundamental disease to develop more effective and reliable treatments, possibly within a highly individualized and customized pharmacotherapeutic structure.

Author contribution statement

Md Sumon Sarkar: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Md Misor Mia: Conceived and designed the experiments; Performed the experiments; Wrote the paper.

Md Al Amin: Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Md Sojib Hossain: Contributed reagents, materials, analysis tools or data.

Md Zahidul Islam: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data associated with this study has been deposited at https://www.ncbi.nlm.nih.gov/gds under the accession number GSE86468 and GSE103001.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e16151.

References

[1] Y.-S. Sun, Z. Zhao, Z.-N. Yang, F. Xu, H.-J. Lu, Z.-Y. Zhu, W. Shi, J. Jiang, P.-P. Yao, H.-P. Zhu, Risk factors and preventions of breast cancer, Int. J. Biol. Sci. 13 (11) (2017) 1387.
[2] M. Akram, M. Iqbal, M. Daniyal, A.U. Khan, Awareness and current knowledge of breast cancer, Biol. Res. 50 (1) (2017) 1–23.
[3] M. Cejuela, B. Martin-Castillo, J.A. Menendez, S. Pernas, Metformin and breast cancer: where are we now?, Int. J. Mol. Sci. 23 (5) (2022) 2705.
[4] Breast cancer-scope of the problem, https://www.who.int/news-room/fact-sheets/detail/breast-cancer. (Accessed 26 May 2022).
[5] Z. Kleibl, V.N. Kristensen, Women at high risk of breast cancer: molecular characteristics, clinical presentation and management, Breast 28 (2016) 136–144.
[6] F. James, S. Wootton, A. Jackson, M. Wiseman, E. Copson, R. Cutress, Obesity in breast cancer–what is the risk factor?, Eur. J. Cancer 51 (6) (2015) 705–720.
[7] A.E. Uloko, B.M. Musa, M.A. Ramalan, I.D. Gezawa, F.H. Puepet, A.T. Uloko, M.M. Borodo, K.B. Sada, Prevalence and risk factors for diabetes mellitus in Nigeria: a systematic review and meta-analysis, Diabetes Ther. 9 (3) (2018) 1307–1316.
[8] Y. Zheng, S.H. Ley, F.B. Hu, Global aetiology and epidemiology of type 2 diabetes mellitus and its complications, Nat. Rev. Endocrinol. 14 (2) (2018) 88–98.
[9] L. Chen, D.J. Magliano, P.Z. Zimmet, The worldwide epidemiology of type 2 diabetes mellitus—present and future perspectives, Nat. Rev. Endocrinol. 8 (4) (2012) 228–236.
[10] A.B. Olokoba, O.A. Obateru, L.B. Olokoba, Type 2 diabetes mellitus: a review of current trends, Oman Med. J. 27 (4) (2012) 269.
[11] M.A. Zeru, E. Tesfa, A.A. Mitiku, A. Seyoum, T.A. Bokoro, Prevalence and risk factors of type-2 diabetes mellitus in Ethiopia: systematic review and meta-analysis, Sci. Rep. 11 (1) (2021) 1–15.
[12] K. Shikata, T. Ninomiya, Y. Kiyohara, Diabetes mellitus and cancer risk: review of the epidemiological evidence, Cancer Sci. 104 (1) (2013) 9–14.
[13] G. Shlomai, B. Neel, D. LeRoith, E.J. Gallagher, Type 2 diabetes mellitus and cancer: the role of pharmacotherapy, J. Clin. Oncol. 34 (35) (2016) 4261.
[14] L. Szablewski, Diabetes mellitus: influences on cancer risk, Diabetes/Metabolism Res. Rev. 30 (7) (2014) 543–553.
[15] S. Liao, J. Li, W. Wei, L. Wang, Y. Zhang, J. Li, C. Wang, S. Sun, Association between diabetes mellitus and breast cancer risk: a meta-analysis of the literature, Asian Pac. J. Cancer Prev. 12 (4) (2011) 1061–1065.
[16] P.F. Christopoulos, P. Msaouel, M. Koutsilieris, The role of the insulin-like growth factor-1 system in breast cancer, Mol. Cancer 14 (1) (2015) 1–14.
[17] F. Gu, F.R. Schumacher, F. Canzian, N.E. Allen, D. Albanes, C.D. Berg, S.I. Berndt, H. Boeing, H.B. Bueno-de Mesquita, J.E. Buring, et al., Eighteen insulin-like growth factor pathway genes, circulating levels of igf-i and its binding protein, and risk of prostate and breast cancer, Cancer Epidemiol. Biomark. Prev. 19 (11) (2010) 2877–2887.
[18] P.J. Wysocki, B. Wierusz-Wysocka, Obesity, hyperinsulinemia and breast cancer: novel targets and a novel role for metformin, Expert Rev. Mol. Diagn. 10 (4) (2010) 509–519.
[19] M.F. Atoum, F. Alzoughool, H. Al-Hourani, Linkage between obesity leptin and breast cancer, Breast Cancer, Basic Clin. Res. 14 (2020) 1178223419898458.
[20] S.C. Larsson, C.S. Mantzoros, A. Wolk, Diabetes mellitus and risk of breast cancer: a meta-analysis, Int. J. Cancer 121 (4) (2007) 856–862.
[21] G. Maskarinec, A. Fontaine, J.E. Torfadottir, L.L. Lipscombe, I.C. Lega, J. Figueroa, S. Wild, The relation of type 2 diabetes and breast cancer incidence in Asian, Hispanic and African American populations—a review, Can. J. Diabetes 42 (1) (2018) 100–105.
[22] P. Ferroni, S. Riondino, O. Buonomo, R. Palmirotta, F. Guadagni, M. Roselli, Type 2 diabetes and breast cancer: the interplay between impaired glucose metabolism and oxidant stress, Oxidative Medicine Cell. Longevity 2015 (2015).
[23] T. Barrett, S.E. Wilhite, P. Ledoux, C. Evangelista, I.F. Kim, M. Tomashevsky, K.A. Marshall, K.H. Phillippy, P.M. Sherman, M. Holko, et al., Ncbi geo: archive for functional genomics data sets—update, Nucleic Acids Res. 41 (D1) (2012) D991–D995.
[24] N.A. Mahi, M.F. Najafabadi, M. Pilarczyk, M. Kouril, M. Medvedovic Grein, An interactive web platform for re-analyzing geo rna-seq data, Sci. Rep. 9 (1) (2019) 1–9.
[25] S. Wenric, S. ElGuendi, J.-H. Caberg, W. Bezzaou, C. Fasquelle, B. Charloteaux, L. Karim, B. Hennuy, P. Frères, J. Collignon, et al., Transcriptome-wide analysis of natural antisense transcripts shows their potential role in breast cancer, Sci. Rep. 7 (1) (2017) 1–12.
[26] N. Lawlor, J. George, M. Bolisetty, R. Kursawe, L. Sun, V. Sivakamasundari, I. Kycia, P. Robson, M.L. Stitzel, Single-cell transcriptomes identify human islet cell signatures and reveal cell-type–specific expression changes in type 2 diabetes, Genome Res. 27 (2) (2017) 208–222.
[27] A. Anjum, S. Jaggi, E. Varghese, S. Lall, A. Bhowmik, A. Rai, Identification of differentially expressed genes in rna-seq data of arabidopsis thaliana: a compound distribution approach, J. Comput. Biol. 23 (4) (2016) 239–247.
[28] P. Bardou, J. Mariette, F. Escudié, C. Djemiel, C. Klopp, jvenn: an interactive Venn diagram viewer, BMC Bioinform. 15 (1) (2014) 1–7.
[29] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, et al., Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc. Natl. Acad. Sci. 102 (43) (2005) 15545–15550.
[30] M.V. Kuleshov, M.R. Jones, A.D. Rouillard, N.F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S.L. Jenkins, K.M. Jagodnik, A. Lachmann, et al., Enrichr: a comprehensive gene set enrichment analysis web server 2016 update, Nucleic Acids Res. 44 (W1) (2016) W90–W97.
[31] G.O. Consortium, The gene ontology (go) database and informatics resource, Nucleic Acids Res. 32 (suppl_1) (2004) D258–D261.
[32] D. Nishimura, Biocarta, Biotech Softw. Internet Rep., Comput. Softw. J. Scient. 2 (3) (2001) 117–120.
[33] M. Kanehisa, S. Goto, Kegg: Kyoto encyclopedia of genes and genomes, Nucleic Acids Res. 28 (1) (2000) 27–30.
[34] G. Joshi-Tope, M. Gillespie, I. Vastrik, P. D'Eustachio, E. Schmidt, B. de Bono, B. Jassal, G. Gopinath, G. Wu, L. Matthews, et al., Reactome: a knowledgebase of biological pathways, Nucleic Acids Res. 33 (suppl_1) (2005) D428–D432.
[35] D.N. Slenter, M. Kutmon, K. Hanspers, A. Riutta, J. Windsor, N. Nunes, J. Mélius, E. Cirillo, S.L. Coort, D. Digles, et al., Wikipathways: a multifaceted pathway database bridging metabolomics to other omics research, Nucleic Acids Res. 46 (D1) (2018) D661–D667.
[36] G.O. Consortium, et al., Creating the gene ontology resource: design and implementation, Genome Res. 11 (8) (2001) 1425–1433.
[37] D. Szklarczyk, A.L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N.T. Doncheva, J.H. Morris, P. Bork, et al., String v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets, Nucleic Acids Res. 47 (D1) (2019) D607–D613.
[38] M.K. Islam, M.H. Rahman, M.R. Islam, M.Z. Islam, M.M.I. Mamun, A. Azad, M.A. Moni, Network based systems biology approach to identify diseasome and comorbidity associations of systemic sclerosis with cancers, Heliyon 8 (2) (2022) e08892.
[39] A. Ben-Hur, W.S. Noble, Kernel methods for predicting protein–protein interactions, Bioinformatics 21 (suppl_1) (2005) i38–i46.
[40] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, Genome Res. 13 (11) (2003) 2498–2504.
[41] Y. Liu, H.-Y. Gu, J. Zhu, Y.-M. Niu, C. Zhang, G.-L. Guo, Identification of hub genes and key pathways associated with bipolar disorder based on weighted gene co-expression network analysis, Front. Physiol. 10 (2019) 1081.
[42] T. Li, X. Gao, L. Han, J. Yu, H. Li, Identification of hub genes with prognostic values in gastric cancer by bioinformatics analysis, World J. Surg. Oncol. 16 (1) (2018) 1–12.

[43] C.-H. Chin, S.-H. Chen, H.-H. Wu, C.-W. Ho, M.-T. Ko, C.-Y. Lin, cytohubba: identifying hub objects and sub-networks from complex interactome, BMC Syst. Biol. 8 (4) (2014) 1–7.
[44] A.G. Papavassiliou, Transcription factors, N. Engl. J. Med. 332 (1) (1995) 45–47.
[45] G. Caramori, K. Ito, I.M. Adcock, Transcription factors in asthma and copd, Idrugs, Invest. Drugs J. 7 (8) (2004) 764–770.
[46] A. Sandelin, W. Alkema, P. Engström, W.W. Wasserman, B. Lenhard, Jaspar: an open-access database for eukaryotic transcription factor binding profiles, Nucleic Acids Res. 32 (suppl_1) (2004) D91–D94.
[47] A. Lachmann, H. Xu, J. Krishnan, S.I. Berger, A.R. Mazloom, A. Ma'ayan, Chea: transcription factor regulation inferred from integrating genome-wide chip-x experiments, Bioinformatics 26 (19) (2010) 2438–2444.
[48] P. Sethupathy, B. Corda, A.G. Hatzigeorgiou, Tarbase: a comprehensive database of experimentally supported animal microrna targets, RNA 12 (2) (2006) 192–197.
[49] H.-Y. Huang, Y.-C.-D. Lin, J. Li, K.-Y. Huang, S. Shrestha, H.-C. Hong, Y. Tang, Y.-G. Chen, C.-N. Jin, Y. Yu, et al., mirtarbase 2020: updates to the experimentally validated microrna–target interaction database, Nucleic Acids Res. 48 (D1) (2020) D148–D154.
[50] J. Xia, E.E. Gill, R.E. Hancock, Networkanalyst for statistical, visual and network-based meta-analysis of gene expression data, Nat. Protoc. 10 (6) (2015) 823–844.
[51] X.-X. Yang, Z.-P. Hu, S.Y. Chan, S.-F. Zhou, Monitoring drug–protein interaction, Clin. Chim. Acta 365 (1–2) (2006) 9–29.
[52] D.S. Wishart, Y.D. Feunang, A.C. Guo, E.J. Lo, A. Marcu, J.R. Grant, T. Sajed, D. Johnson, C. Li, Z. Sayeeda, et al., Drugbank 5.0: a major update to the drugbank database for 2018, Nucleic Acids Res. 46 (D1) (2018) D1074–D1082.
[53] J. Piñero, À. Bravo, N. Queralt-Rosinach, A. Gutiérrez-Sacristán, J. Deu-Pons, E. Centeno, J. García-García, F. Sanz, L.I. Furlong, Disgenet: a comprehensive platform integrating information on human disease-associated genes and variants, Nucleic Acids Res. (2016) gkw943.
[54] J. Piñero, J.M. Ramírez-Anguita, J. Saüch-Pitarch, F. Ronzano, E. Centeno, F. Sanz, L.I. Furlong, The disgenet knowledge platform for disease genomics: 2019 update, Nucleic Acids Res. 48 (D1) (2020) D845–D855.
[55] T. Mitsis, A. Efthimiadou, F. Bacopoulou, D. Vlachakis, G.P. Chrousos, E. Eliopoulos, Transcription factors and evolution: an integral part of gene expression, World Acad. Sci. J. 2 (1) (2020) 3–8.
[56] T. Glisovic, J.L. Bachorik, J. Yong, G. Dreyfuss, Rna-binding proteins and post-transcriptional gene regulation, FEBS Lett. 582 (14) (2008) 1977–1986.
[57] O. Hobert, Gene regulation by transcription factors and microrna, Science 319 (5871) (2008) 1785–1786.
[58] J.M. Vaquerizas, S.K. Kummerfeld, S.A. Teichmann, N.M. Luscombe, A census of human transcription factors: function, expression and evolution, Nat. Rev. Genet. 10 (4) (2009) 252–263.
[59] J.-W. Wei, K. Huang, C. Yang, C.-S. Kang, Non-coding rnas as regulators in epigenetics, Oncol. Rep. 37 (1) (2017) 3–9.
[60] A.F. Samad, M. Sajad, N. Nazaruddin, I.A. Fauzi, A.M. Murad, Z. Zainal, I. Ismail, Microrna and transcription factor: key players in plant regulatory network, Front. Plant Sci. 8 (2017) 565.
[61] M. Inui, G. Martello, S. Piccolo, Microrna control of signal transduction, Nat. Rev. Mol. Cell Biol. 11 (4) (2010) 252–263.
[62] B. Chen, H. Li, X. Zeng, P. Yang, X. Liu, X. Zhao, S. Liang, Roles of microrna on cancer cell metabolism, J. Transl. Med. 10 (1) (2012) 1–12.
[63] R.-u. Takahashi, H. Miyazaki, T. Ochiya, The roles of microrna in breast cancer, Cancers 7 (2) (2015) 598–616.
[64] M.P. LaPierre, M. Stoffel, Microrna as stress regulators in pancreatic beta cells and diabetes, Mol. Metab. 6 (9) (2017) 1010–1023.
[65] S. Calderari, M.R. Diawara, A. Garaud, D. Gauguier, Biological roles of microrna in the control of insulin secretion and action, Physiol. Genomics 49 (1) (2017) 1–10.
[66] J. de Azevedo, F. Walter, R.A. Caceres, I. Pauli, L.F.S. Timmers, G.B. Barcellos, K.B. Rocha, M.B. Soares, Protein-drug interaction studies for development of drugs against plasmodium falciparum, Curr. Drug Targets 10 (3) (2009) 271–278.
[67] M. Al-Mustanjid, S.H. Mahmud, M.R.I. Royel, M.H. Rahman, T. Islam, M.R. Rahman, M.A. Moni, Detection of molecular signatures and pathways shared in inflammatory bowel disease and colorectal cancer: a bioinformatics and systems biology approach, Genomics 112 (5) (2020) 3416–3426.
[68] S.H. Mahmud, W. Chen, H. Meng, H. Jahan, Y. Liu, S.M. Hasan, Prediction of drug-target interaction based on protein features using undersampling and feature selection techniques with boosting, Anal. Biochem. 589 (2020) 113507.
[69] Y.-Q. Qiu, S. Zhang, X.-S. Zhang, L. Chen, Detecting disease associated modules and prioritizing active genes based on high throughput data, BMC Bioinform. 11 (1) (2010) 1–12.
[70] Y. Liu, M. Koyutürk, J.S. Barnholtz-Sloan, M.R. Chance, Gene interaction enrichment and network analysis to identify dysregulated pathways and their interactions in complex diseases, BMC Syst. Biol. 6 (1) (2012) 1–14.
[71] Z. Ahmed, S. Zeeshan, D. Mendhe, X. Dong, Human gene and disease associations for clinical-genomics and precision medicine research, Clin. Transl. Med. 10 (1) (2020) 297–318.
[72] K. De Bruijn, L. Arends, B. Hansen, S. Leeflang, R. Ruiter, C. Van Eijck, Systematic review and meta-analysis of the association between diabetes mellitus and incidence and mortality in breast and colorectal cancer, J. Br. Surg. 100 (11) (2013) 1421–1429.
[73] S. Shao, A.A. Gill, S.H. Zahm, I. Jatoi, C.D. Shriver, K.A. McGlynn, K. Zhu, Diabetes and overall survival among breast cancer patients in the US military health system, Cancer Epidemiol. Biomark. Prev. 27 (1) (2018) 50–57.
[74] Y. Zhou, X. Zhang, C. Gu, J. Xia, Influence of diabetes mellitus on mortality in breast cancer patients, ANZ J. Surg. 85 (12) (2015) 972–978.
[75] S.-H. Doong, A. Dhruva, L.B. Dunn, C. West, S.M. Paul, B.A. Cooper, C. Elboim, G. Abrams, J.D. Merriman, D.J. Langford, et al., Associations between cytokine genes and a symptom cluster of pain, fatigue, sleep disturbance, and depression in patients prior to breast cancer surgery, Biol. Res. Nursing 17 (3) (2015) 237–247.
[76] D.S. Hershey, S.J. Pierce, Examining patterns of multivariate, longitudinal symptom experiences among older adults with type 2 diabetes and cancer via cluster analysis, Eur. J. Oncol. Nursing 19 (6) (2015) 716–723.
[77] D. Von Ah, B. Habermann, J.S. Carpenter, B.L. Schneider, Impact of perceived cognitive impairment in breast cancer survivors, Eur. J. Oncol. Nursing 17 (2) (2013) 236–241.
[78] D.S. Hershey, J. Tipton, B. Given, E. Davis, Perceived impact of cancer treatment on diabetes self-management, Diabetes Educator 38 (6) (2012) 779–790.
[79] Z. Tang, J. Wang, H. Zhang, L. Sun, F. Tang, Q. Deng, J. Yu, Associations between diabetes and quality of life among breast cancer survivors, PLoS ONE 11 (6) (2016) e0157791.
[80] P.A. Vissers, C.R. Cardwell, L.V. van de Poll-Franse, I.S. Young, F. Pouwer, L.J. Murray, The association between glucose-lowering drug use and mortality among breast cancer patients with type 2 diabetes, Breast Cancer Res. Treat. 150 (2) (2015) 427–437.
[81] A.O. Eketunde, Diabetes as a risk factor for breast cancer, Cureus 12 (5) (2020).
[82] E.J. Gallagher, D. LeRoith, Epidemiology and molecular mechanisms tying obesity, diabetes, and the metabolic syndrome with cancer, Diabetes Care 36 (Supplement_2) (2013) S233–S239.
[83] Y.-M. Park, D. Bookwalter, K. O'Brien, C. Jackson, C. Weinberg, D. Sandler, A prospective study of type 2 diabetes, metformin use, and risk of breast cancer, Ann. Oncol. 32 (3) (2021) 351–359.
[84] A. Lohmann, P. Goodwin, Diabetes, metformin and breast cancer: a tangled web, Ann. Oncol. 32 (3) (2021) 285–286.
[85] J. Li, G. Huang, C. Ren, N. Wang, S. Sui, Z. Zhao, M. Li, Identification of differentially expressed genes-related prognostic risk model for survival prediction in breast carcinoma patients, Aging (Albany NY) 13 (12) (2021) 16577.
[86] J. Kim, In silico analysis of differentially expressed genesets in metastatic breast cancer identifies potential prognostic biomarkers, World J. Surg. Oncol. 19 (1) (2021) 188.

[87] A. Tomczak, J.M. Mortensen, R. Winnenburg, C. Liu, D.T. Alessi, V. Swamy, F. Vallania, S. Lofgren, W. Haynes, N.H. Shah, et al., Interpretation of biological experiments changes with evolution of the gene ontology and its annotations, Sci. Rep. 8 (1) (2018) 1–10.
[88] M.M. Markiewski, J.D. Lambris, The role of complement in inflammatory diseases from behind the scenes into the spotlight, Am. J. Pathol. 171 (3) (2007) 715–727.
[89] A. Berbudi, N. Rahmadika, A.I. Tjahjadi, R. Ruslami, Type 2 diabetes and its impact on the immune system, Curr. Diabetes Rev. 16 (5) (2020) 442.
[90] M.J. Smyth, G.P. Dunn, R.D. Schreiber, Cancer immunosurveillance and immunoediting: the roles of immunity in suppressing tumor development and shaping tumor immunogenicity, Adv. Immunol. 90 (2006) 1–50.
[91] G.P. Dunn, A.T. Bruce, H. Ikeda, L.J. Old, R.D. Schreiber, Cancer immunoediting: from immunosurveillance to tumor escape, Nat. Immunol. 3 (11) (2002) 991–998.
[92] G. Mendoza-Almanza, L. Burciaga-Hernández, V. Maldonado, J. Melendez-Zajgla, J. Olmos, Role of platelets and breast cancer stem cells in metastasis, World J. Stem Cells 12 (11) (2020) 1237.
[93] D. Tschoepe, P. Roesen, J. Esser, B. Schwippert, H. Nieuwenhuis, B. Kehrel, F. Gries, Large platelets circulate in an activated state in diabetes mellitus, in: Seminars in Thrombosis and Hemostasis, vol. 17, 1991, pp. 433–438, Copyright© 1991 by Thieme Medical Publishers, Inc.
[94] E.A. Arakelova, M.R. Ovsepyan, A.S. Boyadzhyan, A.A. Arakelyan, A.A. Gevorkyan, A.A. Mamikonyan, The membrane attack complex as an indicator of complement hyperactivation in type 2 diabetes mellitus, Diabetes Mellitus 14 (3) (2011) 17–20.
[95] J. Hakulinen, S. Meri, Expression and function of the complement membrane attack complex inhibitor protectin (cd59) on human breast cancer cells, Lab. Invest., J. Techn. Methods Pathol. 71 (6) (1994) 820–827.
[96] A. Malhotra, V. Sanghi, Regulation of contractile proteins in diabetic heart, Cardiovasc. Res. 34 (1) (1997) 34–40.
[97] D. Wu, M. Asiedu, Q. Wei, Myosin-interacting guanine exchange factor (myogef) regulates the invasion activity of mda-mb-231 breast cancer cells through activation of rhoa and rhoc, Oncogene 28 (22) (2009) 2219–2230.
[98] Y. Xiao, L. Zheng, X. Zou, J. Wang, J. Zhong, T. Zhong, Extracellular vesicles in type 2 diabetes mellitus: key roles in pathogenesis, complications, and therapy, J. Extracellular Vesicles 8 (1) (2019) 1625677.
[99] Q.B. Zha, Y.F. Yao, Z.J. Ren, X.J. Li, J.H. Tang, Extracellular vesicles: an overview of biogenesis, function, and role in breast cancer, Tumor Biol. 39 (2) (2017) 1010428317691182.
[100] J.K. Woodward, I. Holen, R.E. Coleman, D.J. Buttle, The roles of proteolytic enzymes in the development of tumour-induced bone disease in breast and prostate cancer, Bone 41 (6) (2007) 912–927.
[101] R. Syed, R. Farukh, R. Shaista, M. Akbar, A. Shajrul, Role of proteases in cancer: a review, Biotechn. Mol. Biol. Rev. 7 (4) (2012) 90–101.
[102] C.A. Janeway Jr, P. Travers, M. Walport, M.J. Shlomchik, The complement system and innate immunity, in: Immunobiology: The Immune System in Health and Disease, 5th edition, Garland Science, 2001.
[103] R. Zhang, Q. Liu, T. Li, Q. Liao, Y. Zhao, Role of the complement system in the tumor microenvironment, Cancer Cell Int. 19 (1) (2019) 1–12.
[104] D. McMillan, Elevation of complement components in diabetes mellitus, Diabete Metabolisme 6 (4) (1980) 265–270.
[105] B.C. King, A.M. Blom, Non-traditional roles of complement in type 2 diabetes: metabolism, insulin secretion and homeostasis, Mol. Immunol. 84 (2017) 34–42.
[106] M. Copenhaver, C.-Y. Yu, R.P. Hoffman, Complement components, c3 and c4, and the metabolic syndrome, Curr. Diabetes Rev. 15 (1) (2019) 44–48.
[107] Y. Mamane, C. Chung Chan, G. Lavallee, N. Morin, L.-J. Xu, J. Huang, R. Gordon, W. Thomas, J. Lamb, E.E. Schadt, et al., The c3a anaphylatoxin receptor is a key mediator of insulin resistance and functions by modulating adipose tissue macrophage infiltration and activation, Diabetes 58 (9) (2009) 2006–2017.
[108] K.L. Rasmussen, B.G. Nordestgaard, S.F. Nielsen, Complement c3 and risk of diabetic microvascular disease: a cohort study of 95202 individuals from the general population, Clin. Chem. 64 (7) (2018) 1113–1124.
[109] R. Lappano, Y. Jacquot, M. Maggiolini, Gpcr modulation in breast cancer, Int. J. Mol. Sci. 19 (12) (2018) 3840.
[110] K.W. Sloop, P.J. Emmerson, M.A. Statnick, F.S. Willard, The current state of gpcr-based drug discovery to treat metabolic disease, Br. J. Pharmacol. 175 (21) (2018) 4060–4071.
[111] J. Yang, X. Xiong, S. Liu, J. Zhu, M. Luo, L. Liu, L. Zhao, Y. Qin, T. Song, C. Huang, Identification of novel serum peptides biomarkers for female breast cancer patients in western China, Proteomics 16 (6) (2016) 925–934.
[112] J.-Y. Hwang, M.-H. Ryu, M.-J. Go, B.-S. Oh, Y.-S. Cho, Association between single nucleotide polymorphisms of the fibrinogen alpha chain (fga) gene and type 2 diabetes mellitus in the Korean population, Genomics Inform. 7 (2) (2009) 57–64.
[113] S.B. Jorgensen, H.M. O'Neill, L. Sylow, J. Honeyman, K.A. Hewitt, R. Palanivel, M.D. Fullerton, L. Öberg, A. Balendran, S. Galic, et al., Deletion of skeletal muscle socs3 prevents insulin resistance in obesity, Diabetes 62 (1) (2013) 56–64.
[114] M. Evans, C. Yu, A. Lohani, R. Mahdi, X. Liu, A. Trzeciak, C. Egwuagu, Expression of socs1 and socs3 genes is differentially regulated in breast cancer cells in response to proinflammatory cytokine and growth factor signals, Oncogene 26 (13) (2007) 1941–1948.
[115] S. Wu, X. Guo, J. Zhou, X. Zhu, H. Chen, K. Zhang, Y. Lu, Y. Chen, High expression of unc5b enhances tumor proliferation, increases metastasis, and worsens prognosis in breast cancer, Aging (Albany NY) 12 (17) (2020) 17079.
[116] X. Jiao, D. Zhang, Q. Hong, L. Yan, Q. Han, F. Shao, G. Cai, X. Chen, H. Zhu, Netrin-1 works with unc5b to regulate angiogenesis in diabetic kidney disease, Front. Med. 14 (3) (2020) 293–304.
[117] P.M. Coan, M. Barrier, N. Alfazema, R.N. Carter, S. Marion de Procé, X.C. Dopico, A. Garcia Diaz, A. Thomson, L.H. Jackson-Jones, B. Moyon, et al., Complement factor b is a determinant of both metabolic and cardiovascular features of metabolic syndrome, Hypertension 70 (3) (2017) 624–633.
[118] C. He, Y. Li, R. Zhang, J. Chen, X. Feng, Y. Duan, Low cfb expression is independently associated with poor overall and disease-free survival in patients with lung adenocarcinoma, Oncol. Lett. 21 (6) (2021) 1–10.
[119] I.W.-Y. Cheuk, M.T. Siu, J.C.-W. Ho, J. Chen, V.Y. Shin, A. Kwong, Itgav targeting as a therapeutic approach for treatment of metastatic breast cancer, Am. J. Cancer Res. 10 (1) (2020) 211.
[120] F. Wiesmann, J. Veeck, O. Galm, A. Hartmann, M. Esteller, R. Knüchel, E. Dahl, Frequent loss of endothelin-3 (edn3) expression due to epigenetic inactivation in human breast cancer, Breast Cancer Res. 11 (3) (2009) 1–18.
[121] C.-H. Yang, D. Ann-Onda, X. Lin, S. Fynch, S. Nadarajah, E.G. Pappas, X. Liu, J.W. Scott, J.S. Oakhill, S. Galic, et al., Neuropeptide y1 receptor antagonism protects β-cells and improves glycemic control in type 2 diabetes, Mol. Metab. 55 (2022) 101413.
[122] L. Liu, Q. Xu, L. Cheng, C. Ma, L. Xiao, D. Xu, Y. Gao, J. Wang, H. Song, Npy1r is a novel peripheral blood marker predictive of metastasis and prognosis in breast cancer patients, Oncol. Lett. 9 (2) (2015) 891–896.
[123] J.J. De Ronde, E.H. Lips, L. Mulder, A.D. Vincent, J. Wesseling, M. Nieuwland, R. Kerkhoven, M.-J.T. Vrancken Peeters, G.S. Sonke, S. Rodenhuis, et al., Serpina6, bex1, agtr1, slc26a3, and laptm4b are markers of resistance to neoadjuvant chemotherapy in her2-negative breast cancer, Breast Cancer Res. Treat. 137 (1) (2013) 213–223.
[124] J.S. Brand, K. Humphreys, J. Li, R. Karlsson, P. Hall, K. Czene, Common genetic variation and novel loci associated with volumetric mammographic density, Breast Cancer Res. 20 (1) (2018) 1–11.
[125] Y. He, G.E. Ronsein, C. Tang, G.P. Jarvik, W.S. Davidson, V. Kothari, H.D. Song, J.P. Segrest, K.E. Bornfeldt, J.W. Heinecke, Diabetes impairs cellular cholesterol efflux from abca1 to small hdl particles, Circ. Res. 127 (9) (2020) 1198–1210.
[126] H.J. Chan, H. Li, Z. Liu, Y.-C. Yuan, J. Mortimer, S. Chen, Serpina1 is a direct estrogen receptor target gene and a predictor of survival in breast cancer patients, Oncotarget 6 (28) (2015) 25815.
[127] Z. Zheng, Y.-n. Li, S. Jia, M. Zhu, L. Cao, M. Tao, J. Jiang, S. Zhan, Y. Chen, P.-J. Gao, et al., Lung mesenchymal stromal cells influenced by th2 cytokines mobilize neutrophils and facilitate metastasis by producing complement c3, Nat. Commun. 12 (1) (2021) 1–15.

[128] G. Engstrom, B. Hedblad, K.-F. Eriksson, L. Janzon, F. Lindgärde, Complement c 3 is a risk factor for the development of diabetes: a population-based cohort study, Diabetes 54 (2) (2005) 570–575.
[129] M. Sircar, I.A. Rosales, M.K. Selig, D. Xu, Z.K. Zsengeller, I.E. Stillman, T.A. Libermann, S.A. Karumanchi, R.I. Thadhani, Complement 7 is up-regulated in human early diabetic kidney disease, Am. J. Pathol. 188 (10) (2018) 2147–2154.
[130] H. Zhang, Y. Zhao, X. Liu, L. Fu, F. Gu, Y. Ma, High expression of complement component c7 indicates poor prognosis of breast cancer and is insensitive to taxane-anthracycline chemotherapy, Front. Oncol. 11 (2021).
[131] A. Ceriello, F. Mercuri, D. Fabbro, R. Giacomello, G. Stel, C. Taboga, L. Tonutti, E. Motz, G. Damante, Effect of intensive glycaemic control on fibrinogen plasma concentrations in patients with type ii diabetes mellitus. Relation with β-fibrinogen genotype, Diabetologia 41 (11) (1998) 1270–1273.
[132] Y. Mei, H. Liu, X. Sun, X. Li, S. Zhao, R. Ma, Plasma fibrinogen level may be a possible marker for the clinical response and prognosis of patients with breast cancer receiving neoadjuvant chemotherapy, Tumor Biol. 39 (6) (2017) 1010428317700002.
[133] K. Kulak, G.T. Westermark, N. Papac-Milicevic, E. Renström, A.M. Blom, B.C. King, The human serum protein c4b-binding protein inhibits pancreatic iapp-induced inflammasome activation, Diabetologia 60 (8) (2017) 1522–1533.
[134] A. Blom, The role of complement inhibitors beyond controlling inflammation, J. Intern. Med. 282 (2) (2017) 116–128.
[135] S.I. Shpyleva, V.P. Tryndyak, O. Kovalchuk, A. Starlard-Davenport, V. Chekhun, F.A. Beland, I.P. Pogribny, Role of ferritin alterations in human breast cancer cells, Breast Cancer Res. Treat. 126 (1) (2011) 63–71.
[136] A. Göthlin Eremo, K. Lagergren, L. Othman, S. Montgomery, G. Andersson, E. Tina, Evaluation of spp1/osteopontin expression as predictor of recurrence in tamoxifen treated breast cancer, Sci. Rep. 10 (1) (2020) 1–9.
[137] Y. Song, M. Lu, L. Feng, Q. Chen, H. Huang, Q. Lin, Identification of potential immunotherapy biomarkers for breast cancer by bioinformatics analysis, Biosci. Rep. 42 (2) (2022) BSR20212035.
[138] X. Li, R. Xiao, K. Tembo, L. Hao, M. Xiong, S. Pan, X. Yang, W. Yuan, J. Xiong, Q. Zhang, Peg10 promotes human breast cancer cell proliferation, migration and invasion, Int. J. Oncol. 48 (5) (2016) 1933–1942.
[139] J.G. Santillán-Benítez, H. Mendieta-Zerón, L.M. Gómez-Oliván, A.O. Quiroz, J.J. Torres-Juárez, J.M. González-Bañales, Jak2, stat3 and socs3 gene expression in women with and without breast cancer, Gene 547 (1) (2014) 70–76.
[140] L. Shi, C. Zhao, H. Pu, Q. Zhang, Fbp1 expression is associated with basal-like breast carcinoma, Oncol. Lett. 13 (5) (2017) 3046–3056.
[141] Z. Zhong, W. Jiang, J. Zhang, Z. Li, F. Fan, Identification and validation of a novel 16-gene prognostic signature for patients with breast cancer, Sci. Rep. 12 (1) (2022) 1–10.
[142] A.E. Sørensen, M.N. van Poppel, G. Desoye, P. Damm, D. Simmons, D.M. Jensen, L.T. Dalgaard, D.C.I. Group, The predictive value of mir-16,-29a and-134 for early identification of gestational diabetes: a nested analysis of the dali cohort, Cells 10 (1) (2021) 170.
[143] S. Ghafouri-Fard, T. Khoshbakht, B.M. Hussen, S.T. Abdullah, M. Taheri, M. Samadian, A review on the role of mir-16-5p in the carcinogenesis, Cancer Cell Int. 22 (1) (2022) 342.
[144] J.B. Kornum, R.W. Thomsen, A. Riis, H.-H. Lervang, H.C. Schønheyder, H.T. Sørensen, Type 2 diabetes and pneumonia outcomes: a population-based cohort study, Diabetes Care 30 (9) (2007) 2251–2257.
[145] V. Cottin, R. Frognier, H. Monnot, A. Levy, P. DeVuyst, J.-F. Cordier, et al., Chronic eosinophilic pneumonia after radiation therapy for breast cancer, Eur. Respir. J. 23 (1) (2004) 9–13.
[146] M. Prochazka, F. Granath, A. Ekbom, P.G. Shields, P. Hall, Lung cancer risks in women with previous breast cancer, Eur. J. Cancer 38 (11) (2002) 1520–1525.
[147] J.-Y. Lee, I. Jeon, J.M. Lee, J.-M. Yoon, S.M. Park, Diabetes mellitus as an independent risk factor for lung cancer: a meta-analysis of observational studies, Eur. J. Cancer 49 (10) (2013) 2411–2423.
[148] D. Garcia-Compean, J.O. Jaquez-Quintana, J.A. Gonzalez-Gonzalez, H. Maldonado-Garza, Liver cirrhosis and diabetes: risk factors, pathophysiology, clinical implications and management, World J. Gastroenterol. 15 (3) (2009) 280.
[149] H.T. Sørensen, S. Friis, J.H. Olsen, A.M. Thulstrup, L. Mellemkjaer, M. Linet, D. Trichopoulos, H. Vilstrup, J. Olsen, Risk of breast cancer in men with liver cirrhosis, Am. J. Gastroenterol. 93 (2) (1998) 231–233.
[150] D. Wang, Y. Ma, X. Tong, Y. Zhang, H. Fan, Diabetes mellitus contributes to idiopathic pulmonary fibrosis: a review from clinical appearance to possible pathogenesis, Front. Public Health 8 (2020) 196.
[151] S.B. Dalsgaard, E.T. Würtz, J. Hansen, O.D. Røe, Ø. Omland, A cohort study on cancer incidence among women exposed to environmental asbestos in childhood with a focus on female cancers, including breast cancer, Int. J. Environ. Res. Public Health 19 (4) (2022) 2086.
[152] S. Bădescu, C. Tătaru, L. Kobylinska, E. Georgescu, D. Zahiu, A. Zăgrean, L. Zăgrean, The association between diabetes mellitus and depression, J. Med. Life 9 (2) (2016) 120.
[153] J.R. Fann, A.M. Thomas-Rich, W.J. Katon, D. Cowley, M. Pepping, B.A. McGregor, J. Gralow, Major depression after breast cancer: a review of epidemiology and treatment, Gen. Hosp. Psychiatry 30 (2) (2008) 112–126.
[154] M. Ristow, Neurodegenerative disorders associated with diabetes mellitus, J. Mol. Med. 82 (8) (2004) 510–529.
[155] S. Shiovitz, L.A. Korde, Genetics of breast cancer: a topic in evolution, Ann. Oncol. 26 (7) (2015) 1291–1299.
[156] R. Uhlig, H. Contreras, S. Weidemann, N. Gorbokon, A. Menz, F. Büscheck, A.M. Luebke, M. Kluth, C. Hube-Magg, A. Hinsch, et al., Carboxypeptidase a1 (cpa1) immunohistochemistry is highly sensitive and specific for acinar cell carcinoma (acc) of the pancreas, Am. J. Surg. Pathol. 46 (1) (2022) 97.
[157] R.K. Yadav, S.-W. Chae, H.-R. Kim, H.J. Chae, Endoplasmic reticulum stress and cancer, J. Cancer Prev. 19 (2) (2014) 75.
[158] M.D. Bellin, M. Lowe, M.B. Zimmerman, M. Wilschanski, S. Werlin, D.M. Troendle, U. Shah, S.J. Schwarzenberg, J.F. Pohl, E. Perito, et al., Diabetes mellitus in children with acute recurrent and chronic pancreatitis: data from the insppire cohort, J. Pediatr. Gastroenterol. Nutr. 69 (5) (2019) 599.
[159] D. Yadav, A.B. Lowenfels, The epidemiology of pancreatitis and pancreatic cancer, Gastroenterology 144 (6) (2013) 1252–1261.
[160] S. Mustapha, M. Mohammed, A.K. Azemi, A.I. Jatau, A. Shehu, L. Mustapha, I.M. Aliyu, R.N. Danraka, A. Amin, A.A. Bala, et al., Current status of endoplasmic reticulum stress in type ii diabetes, Molecules 26 (14) (2021) 4362.