Mathematical Biology and Bioinformatics
2024. V. 19. № 2. P.276–292. doi: 10.17537/2024.19.276

========== **BIOINFORMATICS** ==========

Unveiling Potential Therapeutic Targets for Breast Cancer Recurrence: Differentially Expressed Genes and Pathways in Post-Surgery Patients

Hamad Ali Hamad[1], Hammad Khalaf Saeed[2], Talib Hammad Hussein[2], Abdalwahab Bdewi Hussain[1], Nur Fariesha Md Hashim[*3]

[1]*College of Applied Sciences, University of Fallujah, Fallujah, Anbar, Iraq*
[2]*Ramadi General Hospital, Al-Anbar Health Directorate, Ministry of Health, Anbar, Iraq*
[3]*Faculty of Medicine and Health Sciences, University Putra Malaysia, Serdang, Malaysia*

Abstract. Various intrinsic and extrinsic factors, including genetic changes and environmental factors, have been reported to contribute to tumor recurrence. However, insufficient information about the significantly dysregulated genes and pathways responsible for cancer recurrence, even after surgical removal of tumors and chemotherapy. The aim of this research is to find out the fundamental genes linked with progression of cancer that may play a critical role in breast cancer recurrence.

To achieve this, a microarray dataset of Affymetrix Human Genome U133 Plus 2.0 Array platform was used to identify downregulated and upregulated genes that associated with tumor recurrence in post-surgery patients. The study includes 20 specimen, 10 samples extracted at the time of diagnosis and 10 samples taken 30 minutes post-surgery and chemotherapy. Genes that stand out from the rest in their level of expression were further subjected to subsequent functional enrichment analysis and hub genes identification to pinpoint the key genes associated with recurrence.

Results revealed that significantly overexpressed genes were found to be enriched in cancer progression-associated signaling pathways, for example, Wnt pathway and proteoglycans in cancer. Moreover, the identified key hub genes (COL1A1, IGF1, COL1A2, DCN, LUM, MMP2, JUN, CXCL12, THBS2, and LOX) majorly found to play a role in gene expression regulation, dysregulated immune system, epithelial-to-mesenchymal transition, and extracellular matrix remodeling thus promoting the development of cancer and increasing the chances of recurrence after surgery and chemotherapy.

The findings have uncovered key therapeutic targets associated with tumor recurrence through potential ECM-related genes whose overexpression may significantly contribute to tumorigenesis in breast cancer survivors by epithelial-to-mesenchymal transition and targeting them may improve the chances of better survival breast cancer patients and increase the quality of life by reducing the chances of recurrence. However, the study is solely bioinformatics-based; therefore, future study will be experimental validations to bring forth these key genes as potential therapeutic targets.

Key words: *breast cancer, microarray dataset, overexpressed genes, ECM-related genes, cancer progression-associated signaling pathways, Wnt pathway.*

INTRODUCTION

Breast cancer remains one of the largely frequent and significant causes of death in women in every single country [1]. Breast cancer, being malignant, can metastasize to other organs such as bones, brain, liver and lungs, which reports its incurable characteristics [2]. In general, cancer disease is considered to be propelled by continuous inheritable abnormalities, which involve alterations in tumor suppressor genes, chromosomal abnormalities, and oncogenes [3]. Some

* nurfariesha@upm.edu.my

studies revealed that breast cancer is also driven by mutations in epigenetics, which do not exert influence on the primary sequence of DNA [4].

According to the GLOBOCAN 2020 world statistics, incidences of approximately 19.3 million population along with 10 million deaths were reported due to cancers [5]. Among the instances indicated above, female breast cancer exhibited the highest prevalence of all malignancies, with 2.3 million reported cases [6]. It was observed in a 2020 report that there was a 30% increase in female breast cancer, with increased novel cases of approximately 276,480 and estimated deaths of more than 42,000. Each year, the demographics rise for breast cancers to be 1.3 million women are diagnosed with breast cancer globally [2].

Based on the order of prevalence along with the progesterone expression receptor (PR), Estrogen expression receptor (ER), and human epidermal growth factor receptor 2 (HER2), breast cancer is classified into four main subtypes, including Luminal A (HR+, PR+, and HER2-), luminal B (HR +, PR+ AND HER2+), HER2 (ER-, PR-, HER2+), and triple-negative (TNBC) (ER-, PR- and HER2-). Luminal A comprises 60% of breast cancer, whereas luminal B represents 30% of breast cancer [7, 8]. HER2 consists of 10% of breast cancer. Lastly, TNBC consists of 15-20% of breast cancer, which is considered to be the most aggressive cancer with a worse prognosis as compared with other subtypes and frequently occurs in younger women [8].

Malignant cancer of the breast considers as a highly diverse, heterogeneous disease, and unique factors accord to its occurrence [9]. There are two types of risk factors, one of which is intrinsic or inherent, including age, race, family history, gene mutations, reproductive history, sex, and obesity [10]. The second one is the extrinsic factors, including diet, radiation exposure history, alcohol and tobacco consumption, post-menopausal hormone therapy, and physical inactivity [11, 12]. Despite these risk factors, women with unknown family histories are also at increased risk of having breast cancer [11]. The fact that most women diagnosed with breast cancer show no identified risk [13].

Moreover, the signs and symptoms are the crucial yet essential phase in the recognition of breast cancer. At an early stage of breast cancer, symptoms are undiagnosable in most people [11]. Different symptoms appear at advanced stages of cancer; including a breast thickening or a lump, breast appearance changes, pain in breasts, change in nipple appearance or fluid discharge and nipple or skin edema, retraction, erythema linked with malignancies [11, 13]. The most common symptom includes mass accumulation in breasts, and detection of which is the most common symptom in females. Mastalgia may or may not be associated with breast cancer [13].

The utmost step before treatment is to diagnose the cancer through mammography, ultrasonography, and biopsy, along with other clinical examinations [14]. The three biopsy techniques are currently being used for the diagnosis of breast cancer. Fine needle aspiration technique (FNA), the second is the core needle biopsy, and the third is the open (surgical) biopsy [15, 16]. The FNA is done to check the solid lump or cyst by placing a small needle inside the breast without any cut. The core needle biopsy also helps in diagnosing the lump that indicates breast cancer [17]. No cut is required either in this type of biopsy. The open (surgical) biopsy, also commonly known as excisional biopsy, requires a cut to be made and then diagnoses the lump or area of concern [13, 15].

Breast cancer treatments depend on the subtypes and stages of cancer and proliferation in the other parts of the body [11]. The different therapies to reduce cancer include surgery, radiotherapy, and chemotherapy. The surgery is performed to eradicate cancer of breast, radiotherapy is conducted to minimize recurrence risk, and chemotherapy is required to destroy cancer cells, along with preventing distant metastasis [11].

Breast surgery is only considered to be performed if cancer is symptomatic [18]. Surgery as a treatment for breast cancer, over the past years, has significantly evolved [14]. Different types of surgeries are performed, including lumpectomy, in which only cancer-causing tissues are removed with safety margin, and mastectomy, which involves whole breast removal [19]. Along with this, it also involves the removal of lymph nodes to evaluate the ability of cancer proliferation [18].

Chemotherapy is the process widely used in treating breast cancer that involves using drugs or natural agents to inhibit the proliferation of cancer cells particularly breast. Chemotherapy aims to target the ER because most of breast cancers are ER+, and it contributes to 60% of cancers [2]. There are two types of chemotherapies; adjuvant chemotherapy, which is administered to patients after surgery to eradicate remaining cancer cells to reduce the recurrence chances, and the second one is neoadjuvant, which is administered before surgery to stop the spread of cancer and make the surgery more effective [20,21].

Breast tumorigenesis recurrence is the resurgence of cancer cells after initial treatments, appears to be one of the prime causes of death, and is a severe medicinal manifestation [22]. Studies revealed that a higher risk of recurrence is linked with ER-positive breast cancers [23]. The implication of cancer stem cells (CSCs) has been observed in the recurrence process of the tumor, and regulation of various signaling pathways results in the self-renewal ability of CSCs [24].

Different pathways, including transforming growth factor (TGF-β), wingless related integration site (Wnt/β-catenin), Hedgehog, Nuclear factor kappa B (NF-κB), Notch, β1 integrins, and Hippo have been involved in cancer and normal cells' process of self-renewal [24]. Numerous signaling pathways in the breast CSCs show a pivotal role in cancer recurrence, chemotherapy, and resistance to radiation treatment, including NF-κB, EMT, and heat shock protein 27 (hsp27) signaling even after the surgery [23]. Tumor recurrence and drug resistance are related closely, and for cancer to reoccur, it is obliged to reduce the cytotoxic effects of drugs by cancer cells. Therefore, CSCs play a key role in drug resistance and tumor recurrence [25].

Dysregulation of TGF-β signaling triggers cancer recurrence and initiates the tumor recurrence when its immunosuppressive action becomes significant [26]. Activation of the Notch signaling is believed to regulate the tumor recurrence and enhance the invasiveness of cancer along with this ligand Jagged-1 in its active form also contributes to the recurrence of the tumor [25]. The β1 integrin, along with different features such as epidermal growth factor receptor (EGFR), focal adhesion kinase (FAK), and extracellular signal-regulated kinase (ERK), impacts the breast tumor recurrence and progression [27]. Similarly, some studies point out strong evidence to support the correlation between the hedgehog signaling pathway and breast cancer recurrence. Based on the chemotherapy resistance and recurrence of breast cancer, upregulation of Gli1 and smoothened in tamoxifen-resistance derivatives other than tamoxifen-sensitive MCF-7 and T47D breast tumor cells were reported [27]. These signaling pathways and dysregulated genes mentioned above disclose their effect on recurrence of breast cancer after surgery and necessitate thorough transcriptomic analysis of breast cancer patients' data who undergo surgery and may have greater chances of tumor recurrence.

The study employed differential gene expression analysis to analyze dysregulated genes that may contribute to breast cancer recurrence following surgical intervention and chemotherapy. Moreover, the pathway enrichment analysis helped us understand the biological pathways and processes that may display essential function in tumor progression. Furthermore, protein-protein interaction (PPI) evaluation was implemented to analyze network patterns of dysregulated genes, highlighting their interactions within the network. Subsequently, the top ten hub genes, elucidating the highest interaction from these networks, were identified. The identification of these hub genes provided valuable insights as potential key regulators underlying breast cancer recurrence following surgical intervention and chemotherapy.

METHODOLOGY

Overview of the study

This study primarily focused on unveiling the key genes that might play a crucial role in the recurrence of breast cancer post-tumor removal following chemotherapy. The study aims to identify the differentially expressed genes (DEGs) associated with pre- and post-chemotherapy and tumor removal surgery to prevent cancer recurrence and understand the potential therapeutic

interventions. The differential expression analysis was performed on the surgical specimen microarray data, extracted at the time of diagnosis and following the tumor removal from female patients, to detect genes exhibiting significant expression differences amongst samples.

Furthermore, functional enrichment evaluation was carried out on the identified DEGs to understand the related biological processes and pathways that might lead to breast cancer recurrence. Moreover, the PPI was conducted, identifying the top ten upregulated and downregulated genes. This analysis highlighted the significance of genes that may like to breast cancer recurrence following surgical intervention and chemotherapy.

Data collection

The microarray dataset present under accession ID: E-GEOD-28583 was retrieved from the ArrayExpress (https://www.ebi.ac.uk/biostudies/arrayexpress) repository and was selected relevant to breast cancer surgical specimens at diagnosis and post-tumor removal samples following the chemotherapy. ArrayExpress is one of the major public repositories for functional genomics datasets, primarily containing genome-wide gene expression data from microarray or next-generation sequencing (NGS) platforms [28, 29]. The study incorporates the microarray dataset generated using the Affymetrix Human Genome U133 Plus 2.0 Array platform. A total of 20 specimens, comprising 10 samples at diagnosis and 10 samples taken 30 minutes post-tumor removal following chemotherapy, were obtained from female patients with 54.2 years (mean age). This dataset utilized for the identification of dysregulated genes that might be crucial in the recurrence of breast cancer.

Microarray dataset analysis

R version 4.3.1 was used to analyze the microarray dataset. The sample data relationship format (SDRF) data, having essential sample evidence, was administered via giving row names with the equivalent column present in the data file array. Hence, a marked data frame was formed for additional analysis. Columns representing specific sample phenotypes were taken out from the pData data frame. The expression data underwent a log2 transformation, and principal component analysis (pca) was used to determine the ratio of the first two principal components (PCs). The rowMedians function was used to determine row medians following summarization and background correction without normalization. Transcripts with low intensity were investigated using predetermined threshold of four. Genes with median intensities below the cutoff were not included, whereas genes with intensities above it were kept if show the minimum sample requirement. Gene annotation was conducted using the hgu133plus2.db package, probe IDs were excluding that corresponded to variety of genes. Moreover, the limma package was used to apply experimental Bayes conducted t-statistics using the eBayes function, allowing the identification of DEGs. The subset function was employed to extract DEGs with a p-value less than 0.05 and log fold change (FC) values greater than 1 for upregulated and less than -1 for downregulated genes, focusing on genes that are both biologically significant and statistically significant [28].

Functional enrichment analysis

Using the online tool GeneCODis 4, the functional enrichment analysis of the dysregulated genes was carried out. This involved identifying the Kyoto Encyclopedia of Genes & Genomes (KEGG) pathways associated with these genes as well as the Gene Ontology (GO) terms, such as biological processes (BP), molecular function (MF), and cellular components (CC). GeneCodis 4 (https://genecodis.genyo.es/) is an API and web tool used for the analysis of gene lists, protein lists, and regulatory elements such as miRNAs, transcription factors (TFs), and CpGs, focusing on the Modular Enrichment & Singular Enrichment Analysis (MEA & SEA) [30]. Moreover, these findings helped identify overrepresented pathways in a gene list, uncovering the altered molecular mechanisms that underlie the dysregulated genes.

Protein-protein interaction and identification of Hub genes

The search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database was used to conduct the PPI analysis (https://string-db.org/), a PPI database with deciphering and visualizing interactions [31]. The DEGs found through the microarray data analysis were used as input into STRING database by using the default settings. The resulting PPI genes were then imported into Cytoscape, a commonly used software tool for visualizing and analyzing biological networks. Furthermore, the top 10 upregulated and downregulated hub genes from the PPI network. Were identified using Cytoscape's CytoHubba module. By showing the connections (edges) that each gene has with other genes in the network, the degree of centrality of each gene within the network was determined. The top 10 upregulated and downregulated hub genes were identified based on the centrality of the genes.

RESULTS

Identification of differentially expressed genes via microarray data analysis

A differential gene expression analysis was carried out on the dataset, which included samples from patients with breast cancer, using the limma package. In the dataset, 320 genes showed upregulation with FC > 1 and a p-value < 0.05, while 28 genes exhibited downregulation with a FC < -1 and a p-value < 0.05. These results are depicted in Figure 1.

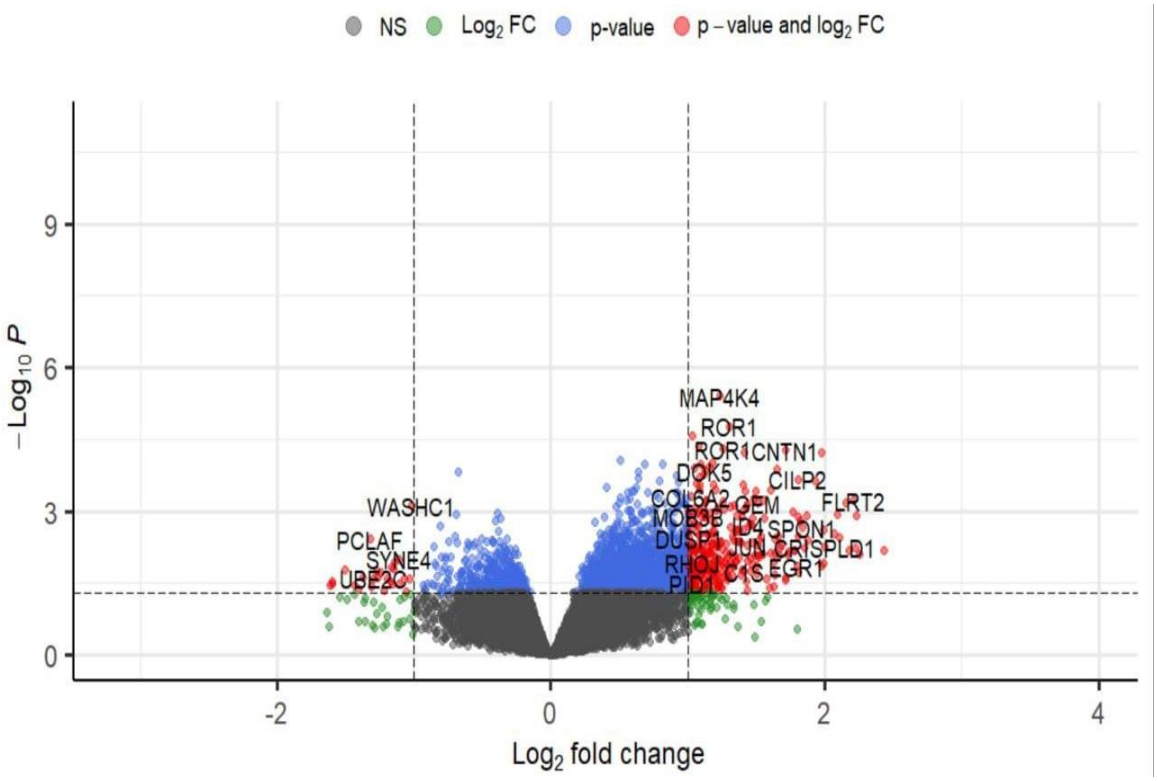

Fig. 1: Enhanced volcano plots representing DEGs in the microarray dataset. Post-tumor removal vs. pre-tumor-removal. Total 29799 variables.

In the microarray dataset, the top 10 genes exhibiting overexpression were COL14A1, OGN, SFRP4, FOS, FLRT2, CCN1, FBLN1, CCN3, CRISPLD1, and CCDC80, while PBK, FSIP1, IL20, NEK2, MSMB, ASPM, ANLN, PCLAF, BIRC5, and UBE2C genes demonstrated under-expression. The top 10 upregulated and downregulated genes are listed in Table 1, Figure 2 and Table 2, Figure 3 respectively.

Table 1. Top 10 overexpressed genes

Gene Names	p-value	logFC
COL14A1	0.006675562	2.432428916
OGN	0.00815681	2.254641431
SFRP4	0.00126765	2.234314057
FOS	0.005780726	2.225645101
FLRT2	0.000606222	2.206055053
CCN1	0.000685949	2.15790777
FBLN1	0.003506652	2.106942977
CCN3	0.001153155	2.089315774
CRISPLD1	0.005581805	1.996647629
CCDC80	0.012107191	1.995166665

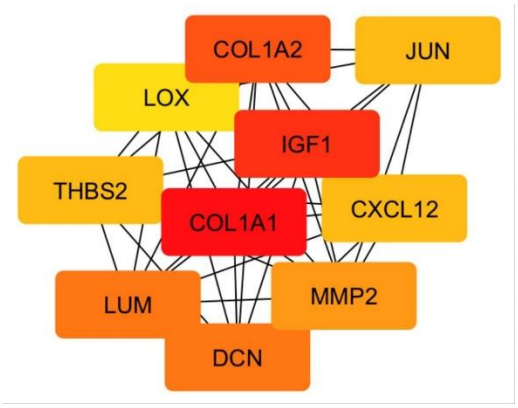

Fig. 2: Visual representation of top 10 upregulated hub genes.

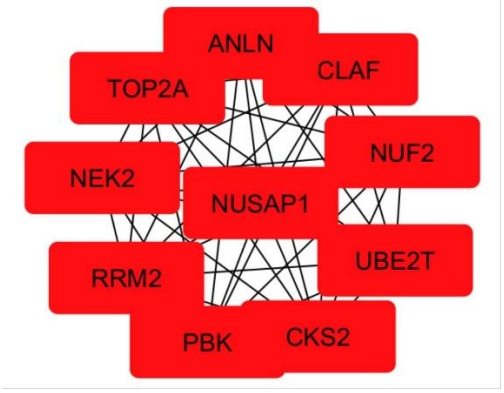

Fig. 3: Visual representation of top 10 downregulated hub genes.

Table 2. Top 10 underexpressed genes

Gene Names	p-value	logFC
PBK	0.034884	-1.61449
FSIP1	0.03283	-1.5988
IL20	0.02882	-1.59826
NEK2	0.017217	-1.5019
MSMB	0.035411	-1.43741
ASPM	0.043178	-1.40103
ANLN	0.024503	-1.37175
PCLAF	0.003854	-1.32132
BIRC5	0.037319	-1.31232
UBE2C	0.025185	-1.29386

Gene ontology term analysis of dysregulated genes

The GO terms linked to the dysregulated genes that were found using GeneCodis4. Analysis revealed various BP, CC, and MF influenced by the dysregulated genes after chemotherapy. The BP analysis revealed that the upregulation of genes was enhanced in extracellular matrix organization, cell adhesion, collagen fibril organization, angiogenesis, positive regulation of cell-substrate adhesion, osteoblast differentiation, intramembranous ossification, Wnt signaling pathway, response to mechanical stimulus, and negative regulation of canonical Wnt signaling pathway (Supplementary Figure S1). However, downregulated genes were enriched in cell division, cell cycle, chromosome segregation, positive regulation of exit from mitosis, mitotic cytokinesis, blastocyst development, protein modification by small protein conjugation, protein K11-linked ubiquitination, mitotic cell cycle, and mitotic sister chromatid segregation (Supplementary Figure S2).

Furthermore, the MF investigation revealed that the overexpression genes show fundamental role in extracellular matrix structural element, fibronectin binding, integrin binding, heparin-binding, extracellular matrix structural constituent conferring tensile strength, collagen binding, extracellular matrix structural constituent conferring compression resistance, insulin-like growth factor binding, platelet-derived growth factor binding, and Wnt-protein binding (Supplementary Figure S3). Nonetheless, the downregulated genes were involved in ubiquitin-like protein transferase activity, chromatin binding, interleukin-20 receptor binding, ubiquitin-conjugating enzyme activity, cobalt ion binding, DNA topoisomerase type II (double strand cut, ATP-hydrolyzing) activity, ribonucleoside-diphosphate reductase activity (thioredoxin disulfide as acceptor), interleukin-22 receptor binding, hemoglobin alpha binding, and ubiquitin-binding (Supplementary Figure S4).

Additionally, the CC analysis of the upregulated genes revealed that the genes were primary found in the extracellular space, collagen trimer, collagen-containing extracellular matrix, extracellular region, extracellular matrix, basement membrane, endoplasmic reticulum lumen, elastic fiber, collagen type I trimer, and interstitial matrix (Supplementary Figure S5). However, the silencing genes were found to be in chromosome, centromeric region, midbody, kinetochore, microtubule, centriole, nucleus, Integral component of nuclear outer membrane, DNA

topoisomerase type II (double strand cut, ATP-hydrolyzing) complex, survivin complex, and spindle (Supplementary Figure S6).

Pathway enrichment analysis for identified dysregulated genes

S&Using the GeneCodis4 tool, pathway enrichment analysis was carried out to identify dysregulated genes implicated in different biological pathways that may be responsible for the recurrence of breast cancer after chemotherapy treatment. The KEGG pathways in which the genes showed dysregulation were found using the enrichment analysis. The analysis indicated the overexpression of the genes involved in breast cancer, absorption of protein, Wnt signaling pathway, proteoglycans in cancer, complement, and coagulation cascades, fluid shear stress and atherosclerosis, african trypanosomiasis, AGE-RAGE signaling pathway in diabetic complications, and relaxin signaling. However, the under-expression of genes was observed in platinum drug resistance, p53 signaling pathway, purine metabolism, apoptosis, hepatitis B, pathways in cancer, drug metabolism - other enzymes, pyrimidine metabolism, colorectal cancer, and glutathione metabolism. Enrichment analysis pathways of the dysregulated genes is illustrated in Figure 4 and Figure 5.

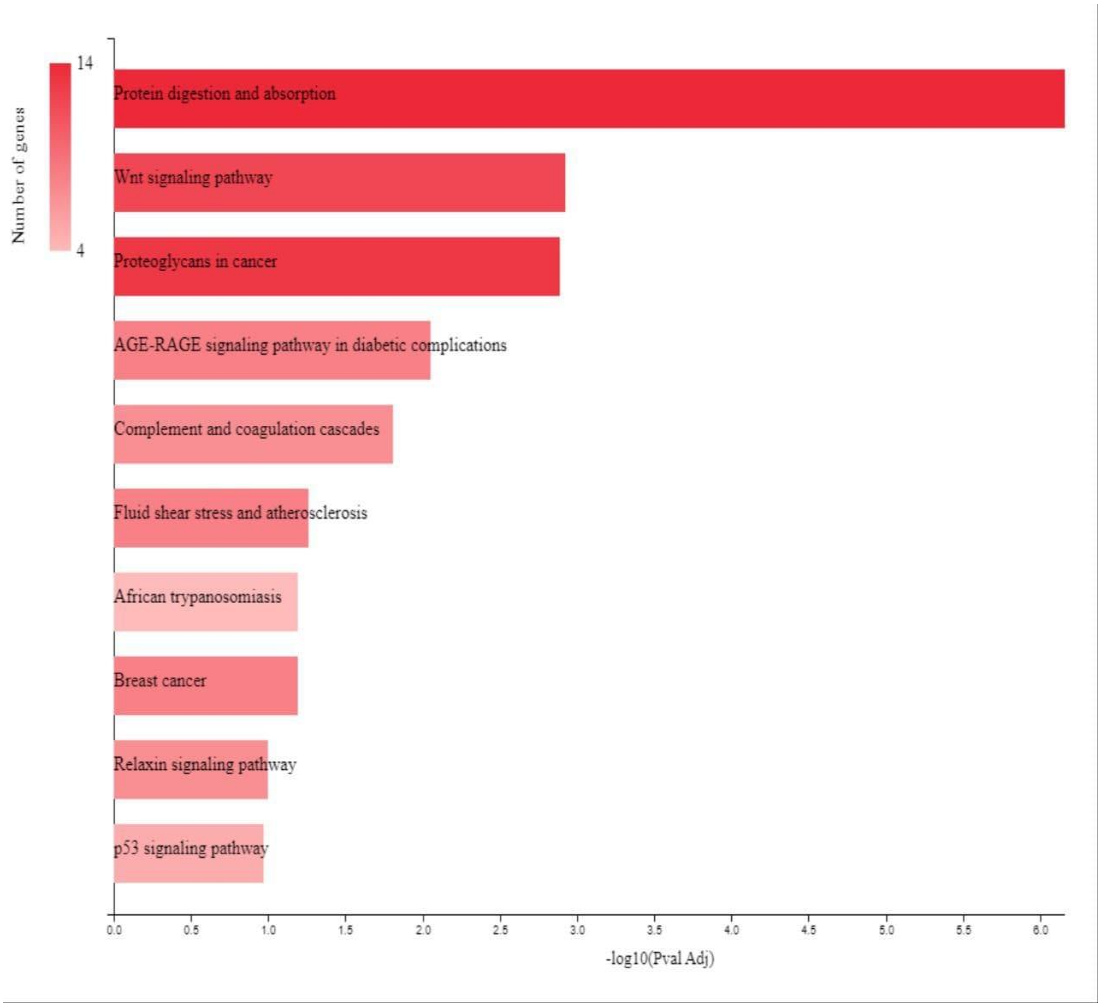

Fig. 4: Pathway enrichment analysis of upregulated genes.

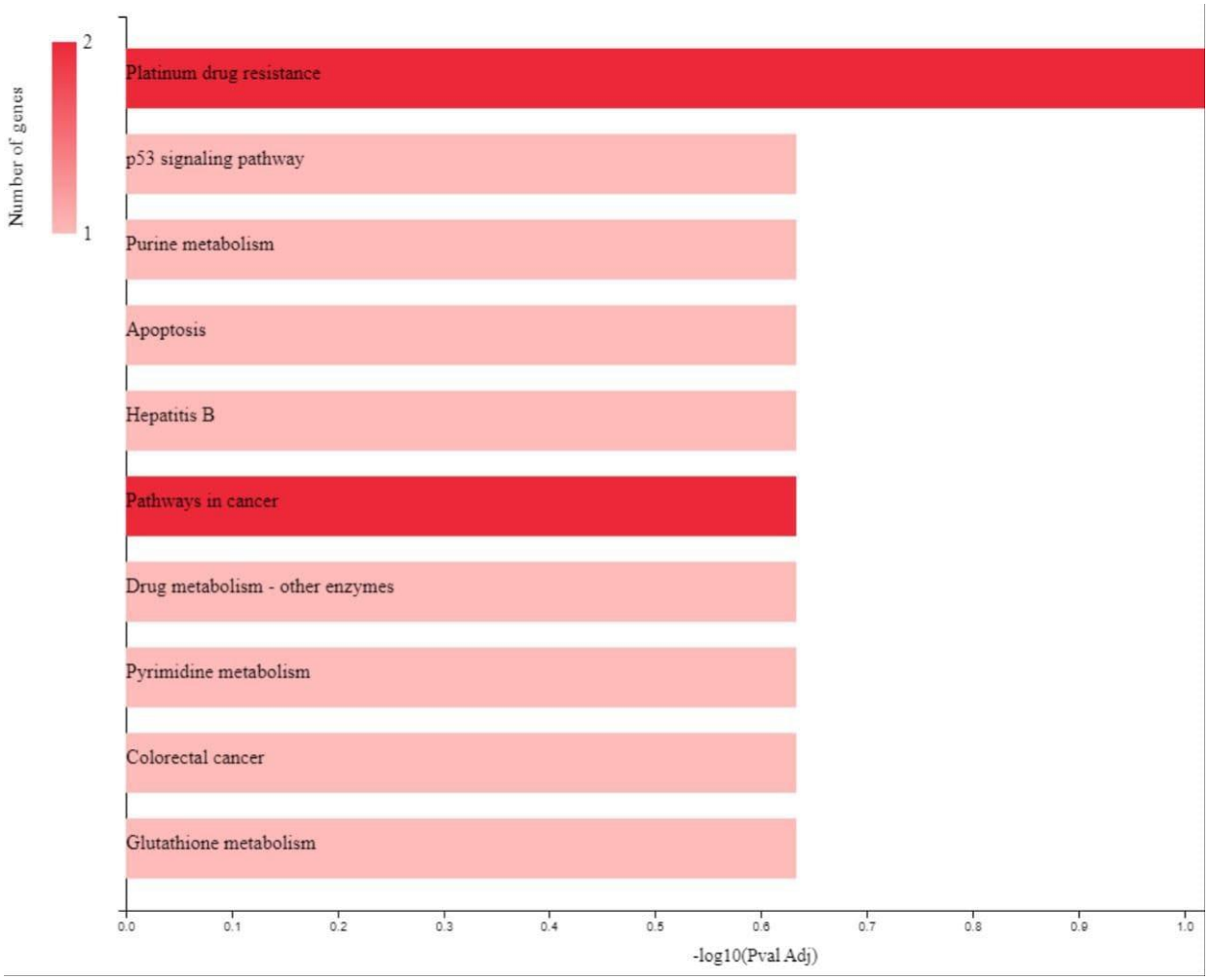

Fig. 5: Pathway enrichment analysis of downregulated genes.

Protein-protein interactions and hub genes identification

The STRING database was used for the PPI analysis of the dysregulated genes. As a result, a network of genes was produced that is shown in (Figure 6 and Figure 7) and has important functional relationships as well as possible interactions among the dysregulated genes. The hup genes in the network were found using the cytohubba plug-in in Cytoscape. The top 10 upregulated and downregulated genes and their respective scores are listed in Table 3 and Table 4. Moreover, these top 10 dysregulated hub genes are illustrated in Figure 3.

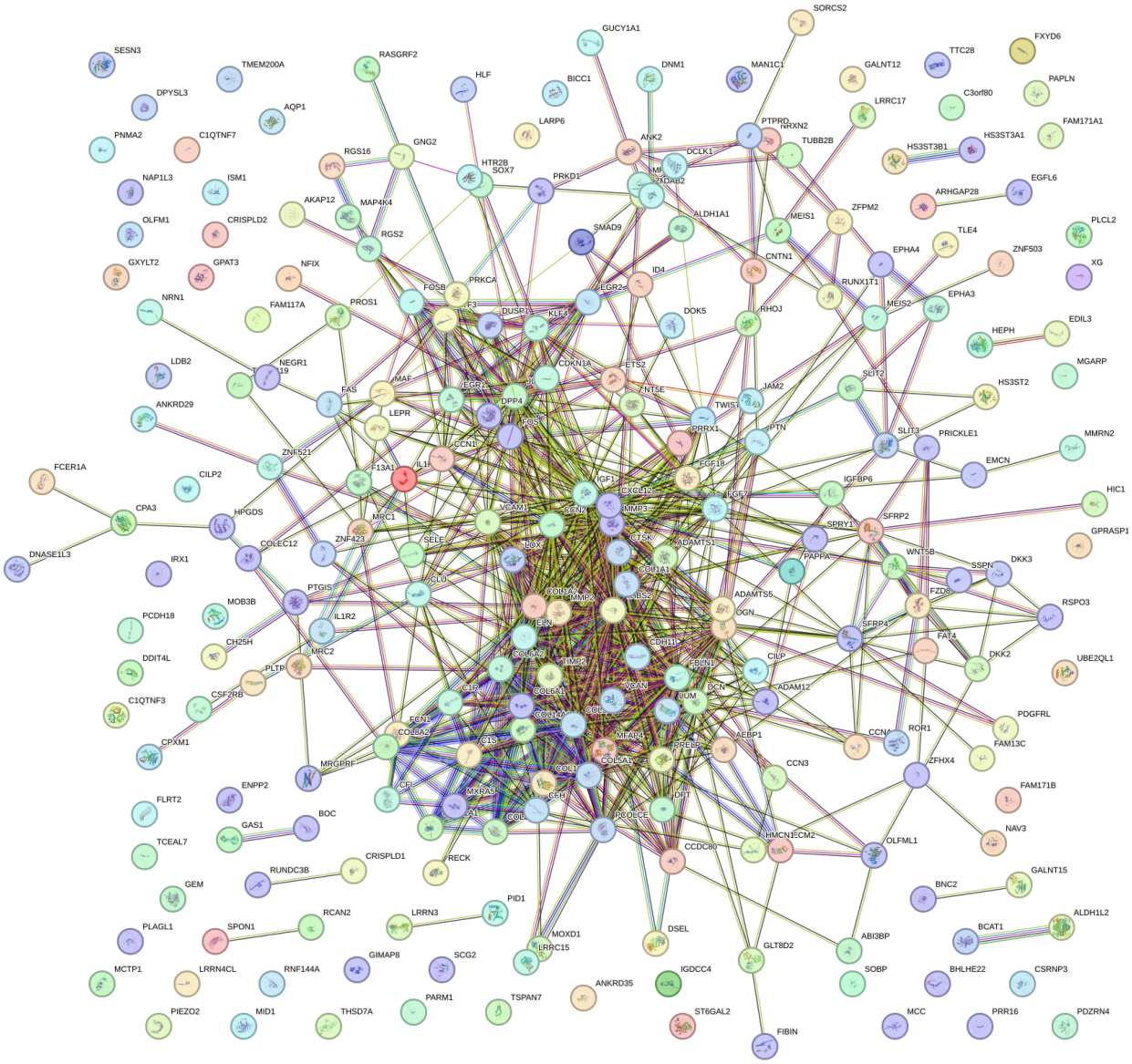

Fig. 6: Protein-Protein Interaction network of upregulated genes identified by STRING. Full size image

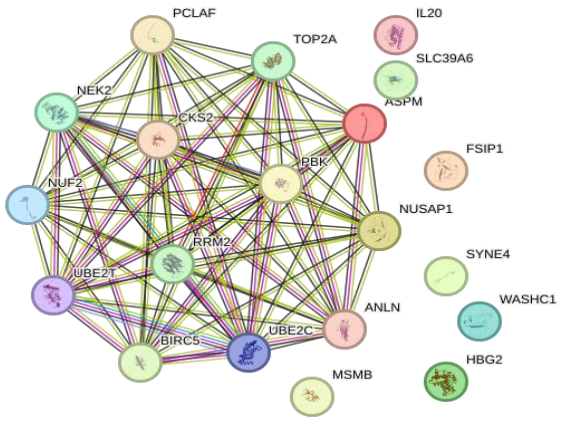

Fig. 7: Protein-Protein Interaction network of downregulated genes identified by STRING. Full size image

Table 3. Top 10 upregulated hub genes and their scores ranked by degree method

Name	Score
COL1A1	54
IGF1	45
COL1A2	44
DCN	42
LUM	42
MMP2	39
JUN	38
CXCL12	38
THBS2	38
LOX	37

Table 4. Top 10 downregulated hub genes and their scores ranked by degree method

Name	Score
RRM2	12
NEK2	12
CKS2	12
PBK	12
UBE2T	12
TOP2A	12
NUSAP1	12
ANLN	12
NUF2	12
PCLAF	12

DISCUSSION

Breast cancer is one of the most prevalent forms of female cancer, with a high mortality rate worldwide. Intriguingly, at the time of diagnosis, around 75% of the primary breast cancers have already spread, leading to early metastasis at the near or distant site, giving rise to micro-metastases [32]. These micro-metastases can easily survive as dormant tumors, making them unaltered during chemotherapy even after surgery. Over time, the alterations in immune cells, growth factors, and cytokines within the tumor microenvironment (TME) eventually lead to tumor growth by cessation of cancer dormancy [33, 34].

In this present study, the longitudinal transcriptomic analysis was performed on breast cancer patients' data for which the samples were collected at the time of cancer diagnosis and after they underwent surgery to pinpoint the dysregulated genes whose expression may significantly lead to

tumor recurrence and tumor progression even after surgical removal of the tumor and during chemotherapy. Differential gene expression analysis of this dataset resulted in a significantly higher number of upregulated genes compared to downregulated genes. The top 10 upregulated genes were observed to be COL14A1, OGN, SFRP4, FOS, FLRT2, CCN1, FBLN1, CCN3, CRISPLD1, and CCDC80.

The upregulated expression of these genes is associated with tumor aggressiveness, influencing crucial cellular processes including proliferation, metastasis, apoptosis, and drug resistance [35]. The COL14A1 and FOS genes are reported to play crucial role in cell proliferation in breast cancer [36, 37]. SFRP4 has been found to play a significant role in the progression of breast cancer cells, showing promise as both prognostic biomarkers and potential therapeutic targets [38]. However, the overexpression of OGN and FLRT2 genes have been found to have a significant inhibitory effect on cell proliferation and migration/invasion in breast cancer cells [39, 40]. Similarly, the FBLN1, CCN1 and CCN3 genes inhibit adhesion and motility in the breast cancer cells through their expression [41, 42]. Lastly, the high expression of CRISPLD1 is reported in top 10 genes with inverse DNA methylation relation in breast ductal carcinoma. However, the role of this gene in breast cancer is under research [43]. Despite the CCDC80 gene regulated the apoptotic pathways, its higher expression is yet not reported in breast cancer [44].

Multiple factors are known to be involved in tumor recurrence, including overexpression of certain genes related to molecular pathways such as Wnt-signaling, PI3K/AKT pathway, and other growth factors-related pathways activated via cross-talk signaling of related genes [45–46]. Functional enrichment analysis revealed significantly enriched pathways, including protein digestion and absorption, Wnt signaling pathway, proteoglycans in cancer, AGE-RAGE signaling pathway in diabetic complications, complement and coagulation cascades, fluid shear stress and atherosclerosis, African trypanosomiasis, breast cancer, relaxin signaling, and p53 signaling pathways. Among these pathways, the Wnt signaling pathway was found to be the significant key player that may have a role in tumor recurrence in breast cancer. Wnt signaling pathway upregulation has been reported to be associated with tumor progression and development and has been found in a significant proportion of cases due to high levels of beta-catenin expression, which further regulates the expression of multiple downstream genes involved in tumor progression [46].

Additionally, the PPI analysis of dysregulated genes and further hub genes identification among those networks of the upregulated and downregulated genes revealed the most interacting ones among them that may have a role in promoting tumor growth and development. Intriguingly, the majority of the identified hub genes (COL1A1, IGF1, COL1A2, DCN, LUM, MMP2, JUN, CXCL12, THBS2, and LOX) were mainly found to be associated with extracellular matrix (ECM) remodeling and downstream gene expression regulation, and their over-expression exhibits the potential role in tumor development even when the tumor has been removed after surgery.

The ECM proteins play a significant role in the development and progression of tumors after surgical removal of tumors in breast cancer. The dynamic communication of cancer cells with TME comprised of ECM and stromal cells is crucial to stimulate cancer heterogeneity, leading to cancer progression and recurrence after surgery. These cellular crosstalk and cell-ECM interactions are majorly responsible for ECM remodeling, eventually creating a favorable environment for tumorigenesis and leading to tumor recurrence [47]. Moreover, it has been reported that even after surgical removal of the tumors and exposure to therapy, the stromal tissues within TME facilitate refuge to cancer cells, rendering them chemo-resistant and making the tumor recur [48].

In normal breast tissues, ECM is essential for the structural support of the cells by regulating multiple cellular processes such as proliferation, growth, homeostasis, differentiation, and morphogenesis [49]. Dysregulation of these normal cellular processes promotes increased proliferation and cancer progression via over-expression of collagen proteins and other ECM proteins, including Lysyl Oxidases (LOX) and Matrix Metalloproteinases (MMPs) [49]. Over-expression of LOX enhances the cross-linking of collagen and increases ECM stiffness; hence, its

increased activity has been reported to promote cancer progression and metastasis [50, 51]. Moreover, the over-expression of collagen genes found in our study, including Collagen type 1 alpha 1 (COL1A1) and Collagen type 1 alpha 2 (COL1A2), are reportedly known to be involved in playing a crucial role in ECM remodeling to promote tumorigenesis in breast tissues [52].

These collagen proteins are known to be the key modulators of epithelial-to-mesenchymal transition, which is a crucial hallmark of cancer development. COL1A1, highly expressed in multiple cancers, also regulates cellular processes, leading to enhanced cell proliferation, migration, and invasion, thus playing an oncogenic role in various pathways [53].

It is therefore suggested to target these identified ECM proteins right after the surgical procedure along with the chemotherapy to mitigate the incidence of tumor recurrence. Targeting these collagen proteins along with other ECM proteins identified in this study may significantly help mitigate the chances of tumor recurrence after surgery, eventually leading to better survival of breast cancer patients. However, these promising findings are solely bioinformatics-based, which should be tested and validated through laboratory experiments to be utilized as therapeutic strategies to mitigate the chances of breast cancer recurrence.

CONCLUSIONS

Breast cancer recurrence, a pressing concern affecting women globally, remains a significant challenge despite the development of breast surgical procedures and current chemotherapy. This longitudinal transcriptomic analysis has unveiled potential genes associated with the extracellular matrix (ECM), shedding light on their role in promoting tumorigenesis through epithelial-to-mesenchymal transition in breast cancer survivors. Notably, collagen proteins COL1A1 and COL1A2 emerge as key players in ECM remodeling, while LOX and MMP2 contribute to increased ECM stiffness and collagen cross-linking, thereby facilitating epithelial-to-mesenchymal transition. Beyond ECM modulation, the study highlights the pivotal role of the oncogenic transcription factor JUN. As a key regulator of numerous downstream oncogenic genes, JUN proves instrumental in fostering tumor growth in breast cancer, ultimately leading to recurrence post-surgery. These findings have paved the way to develop therapeutic strategies to mitigate the chances of tumor recurrence in breast cancer patients after they undergo surgical removal of tumors.

REFERENCES

1. Anayyat U., Ahad F., Muluh T.A., Zaidi S.A.A., Usmani F., Yang H., Li M., Hassan H.A., Wang X. Immunotherapy: Constructive Approach for Breast Cancer Treatment. *Breast Cancer: Targets and Therapy*. 2023. V. 15. P. 925–951 doi: 10.2147/BCTT.S424624
2. Momenimovahed Z., Salehiniya H. Epidemiological characteristics of and risk factors for breast cancer in the world. *Breast Cancer: Targets and Therapy*. 2019. V. 11. P.151–164. doi: 10.2147/BCTT.S176070
3. Pasculli B., Barbano R., Parrella P. Epigenetics of breast cancer: Biology and clinical implication in the era of precision medicine. *Semin. Cancer Biol*. 2018. V. 51. P. 22–35. doi: 10.1016/j.semcancer.2018.01.007
4. Rahman M.M., Brane A.C., Tollefsbol T.O. MicroRNAs and Epigenetics Strategies to Reverse Breast Cancer. *Cells*. 2019. V. 8. Article No. 1214. doi: 10.3390/cells8101214
5. Bray F., Laversanne M., Sung H., Ferlay J., Siegel R.L., Soerjomataram I., Jemal A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J. Clin*. 2024. V. 74. No.3. P. 229–263. doi: 10.3322/caac.21834
6. Arnold M., Morgan E., Rumgay H., Mafra A., Singh D., Laversanne M., Vignat J., Gralow J.R., Cardoso F., Siesling S., et al. Current and future burden of breast cancer: Global statistics for 2020 and 2040. *The Breast*. 2022. V. 66. No. 15–23. doi: 10.1016/j.breast.2022.08.010

7. Hwang K.T., Kim J., Jung J., Chang J.H., Chai Y.J., Oh S.W., Oh S., Kim Y.A., Park S.B., Hwang K.R. Impact of breast cancer subtypes on prognosis of women with operable invasive breast cancer: A population-based study using SEER database. *Clin. Cancer Res*. 2019. V. 25. No. 6. P. 1970–1979. doi: 10.1158/1078-0432.CCR-18-2782
8. Tsang JY, Gary MT. Molecular classification of breast cancer. *Advances in anatomic pathology*. 2020. V. 1. No. 27. P. 27–35. doi: 10.1097/PAP.0000000000000232
9. Prat A., Pineda E., Adamo B., Galván P., Fernández A., Gaba L., Díez M., Viladot M., Arance A., Muñoz M. Clinical implications of the intrinsic molecular subtypes of breast cancer. *The Breast*. 2015. V. 1. No. 24. P. S26–S35. doi: 10.1016/j.breast.2015.07.008
10. Kamińska M., Ciszewski T., Łopacka-Szatan K., Miotła P., Starosławska E. Breast cancer risk factors. *Menopause Rev*. 2015. V. 14. No. 3. P. 196–202. doi: 10.5114/pm.2015.54346
11. Zhao P., Xia N., Zhang H., Deng T. The Metabolic Syndrome Is a Risk Factor for Breast Cancer: A Systematic Review and Meta-Analysis. *Obes. Facts*. 2020. V. 13. No. 4. P. 384–396. doi: 10.1159/000507554
12. Azadnajafabad S., Moghaddam S.S., Keykhaei M., Shobeiri P., Rezaei N., Ghasemi E., Mohammadi E., Ahmadi N., Ghamari A., Shahin S., et al. Expansion of the quality of care index on breast cancer and its risk factors using the global burden of disease study 2019. *Cancer Med*. 2023. V. 12. No. 2. P. 1729–1743. doi: 10.1002/cam4.4951
13. Schünemann H.J., Lerda D., Quinn C., Follmann M., Alonso-Coello P., Rossi P.G., Lebeau A., Nyström L., Broeders M., Ioannidou-Mouzaka L., et al. Breast cancer screening and diagnosis: A synopsis of the European Breast Guidelines. *Ann. Intern. Med*. 2020. V. 172. No. 1. P. 46–56. doi: 10.7326/M19-2125
14. Waks AG, Winer EP. Breast cancer treatment: A review. *JAMA*. 2019. V. 321. No.3. P. 288–300. doi: 10.1001/jama.2018.19323
15. Bick U., Trimboli R.M., Athanasiou A., Balleyguier C., Baltzer P.A., Bernathova M., Borbély K., Brkljacic B., Carbonaro L.A., Clauser P., et al. Image-guided breast biopsy and localisation: Recommendations for information to women and referring physicians by the European Society of Breast Imaging. *Insights into Imaging*. 2020. V. 11. Article No. 12. doi: 10.1186/s13244-019-0803-x
16. Louro J., Román M., Posso M., Comerma L., Vidal C., Saladié F., Alcantara R., Sanchez M., Quintana M.J., Del Riego J. et al. Differences in breast cancer risk after benign breast disease by type of screening diagnosis. *The Breast*. 2020. V. 54. No. 1. P. 343–348. doi: 10.1016/j.breast.2020.09.005
17. Tripathi K., Yadav R., Maurya S.K. A comparative study between fine-needle aspiration cytology and core needle biopsy in diagnosing clinically palpable breast lumps. *Cureus*. 2022. V. 14. Article No. e27709. doi: 10.7759/cureus.27709
18. Riis M. Modern surgical treatment of breast cancer. *Ann. Med. Surg*. 2020. V. 56. P. 95–107. doi: 10.1016/j.amsu.2020.06.016
19. Kaidar-Person O., Offersen B.V., Boersma L.J., de Ruysscher D., Tramm T., Kühn T., Gentilini O., Mátrai Z., Poortmans P. A multidisciplinary view of mastectomy and breast reconstruction: Understanding the challenges. *The Breast*. 2021. V. 56. P. 42–52. doi: 10.1016/j.breast.2021.02.004
20. Labrosse J., Osdoit M., Hamy A.S., Coussy F., Pierga J.Y., Reyal F., Laas E. Adjuvant chemotherapy for breast cancer after preoperative chemotherapy: A propensity score matched analysis. *PLoS ONE*. 2020. V. 15. No. 6. Article No. e0234173. doi: 10.1371/journal.pone.0234173
21. Buzdar AU. Preoperative chemotherapy treatment of breast cancer—a review. *Cancer: Interdisciplinary International Journal of the American Cancer Society*. 2007. V. 110. No. 11. P. 2394–2407. doi: 10.1002/cncr.23083
22. Chen X., Liu L., Wang Y., Liu B., Zeng D., Jin Q., Li M., Zhang D., Liu Q., Xie H. Identification of breast cancer recurrence risk factors based on functional pathways in tumor

and normal tissues. *Oncotarget*. 2016. V. 8. No. 13. P. 20679–20694. doi: 10.18632/oncotarget.11557

23. Rueda O.M., Sammut S.J., Seoane J.A., Chin S.F., Caswell-Jin J.L., Callari M., Batra R., Pereira B., Bruna A., Ali H.R., et al. Dynamics of breast-cancer relapse reveal late-recurring ER-positive genomic subgroups. *Nature*. 2019. V. 567. No. 7748. P. 399–404. doi: 10.1038/s41586-019-1007-8
24. Ordaz-Ramos A., Tellez-Jimenez O., Vazquez-Santillan K. Signaling pathways governing the maintenance of breast cancer stem cells and their therapeutic implications. *Front. Cell Dev. Biol*. 2023. V. 11. Article No. 1221175. doi: 10.3389/fcell.2023.1221175
25. Hamad H.A., Gopalsamy B., Kqueen C.Y., Hashim N.F.M., Potential Ability of Phytochemical in Inhibition of Invadopodia Formation and HIF-1α in Cancer Metastasis. *Malaysian Journal of Medicine & Health Sciences*. 2019. V. 15. P. 71–80.
26. Kamdje A.H.N., Etet P.F.S., Vecchio L., Muller J.M., Krampera M., Lukong K.E. Signaling pathways in breast cancer: therapeutic targeting of the microenvironment. *Cell Signal*. 2014. V. 26. No. 12. P. 2843–2856. doi: 10.1016/j.cellsig.2014.07.034
27. Ahmad A. Pathways to breast cancer recurrence. *International Scholarly Research Notices*. 2013. V. 2013. No. 1. Article No. 290568. doi: 10.1155/2013/290568
28. Jabbar FA., AlChalabi R., AL-Tarboolee AY., Shaban SA., Suleiman AA. Alternative Splicing in Pancreatic Ductal Adenocarcinoma Leads to Dysregulated Immune System. *Mathematical Biology and Bioinformatics*. 2024. V. 19. No. 1. P. 15–35 doi: 10.17537/2024.19.15
29. Alaimo S., Di Maria A., Shasha D., Ferro A., Pulvirenti A. TACITuS: transcriptomic data collector, integrator, and selector on big data platform. *BMC Bioinformatics*. 2019. V. 20. No. 9. Article No. 366. doi: 10.1186/s12859-019-2912-4
30. Garcia-Moreno A., López-Domínguez R., Villatoro-García J.A., Ramirez-Mena A., Aparicio-Puerta E., Hackenberg M., Pascual-Montano A., Carmona-Saez P., Functional Enrichment Analysis of Regulatory Elements. *Biomedicines*. 2022. V. 10. Article No. 590. doi: 10.3390/biomedicines10030590
31. Szklarczyk D., Kirsch R., Koutrouli M., Nastou K., Mehryary F., Hachilif R., Gable A.L., Fang T., Doncheva N.T., Pyysalo S,. et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res*. 2023. V. 51. No. D1. P. D638–D646. doi: 10.1093/nar/gkac1000
32. Riggio A.I., Varley K.E., Welm A.L. The lingering mysteries of metastatic recurrence in breast cancer. *Br. J. Cancer*. 2021. V. 124. P. 13–26. doi: 10.1038/s41416-020-01161-4
33. Baghban R., Roshangar L., Jahanban-Esfahlan R., Seidi K., Ebrahimi-Kalan A., Jaymand M., Kolahian S., Javaheri T., Zare P. Tumor microenvironment complexity and therapeutic implications at a glance. *Cell Commun. Signal*. 2020. V. 18. No. 1. Article No. 59. doi: 10.1186/s12964-020-0530-4
34. Khalaf K., Hana D., Chou J.T.T., Singh C., Mackiewicz A., Kaczmarek, M. Aspects of the Tumor Microenvironment Involved in Immune Resistance and Drug Resistance. *Front. Immunol*. 2021. V. 12. Article No. 656364. doi: 10.3389/fimmu.2021.656364
35. Mares-Quiñones M.D., Galán-Vásquez E., Pérez-Rueda E., Pérez-Ishiwara D.G., Medel-Flores M.O., Gómez-García M.D.C. Identification of modules and key genes associated with breast cancer subtypes through network analysis. *Sci. Rep*. 2024. V. 14. Article No. 12350. doi: 10.1038/s41598-024-61908-4
36. Jiang Y., Chen F., Ren X., Yang Y., Luo J., Yuan J., Yuan J., Tong, Q. RNA-Binding Protein COL14A1, TNS1, NUSAP1 and YWHAE Are Valid Biomarkers to Predict Peritoneal Metastasis in Gastric Cancer. *Front. Oncol*. 2022. V. 12. Article No. 830688. doi: 10.3389/fonc.2022.830688
37. Wu Z.H., Zhang Y.J., Yue J.X., Zhou T. Comprehensive Analysis of the Expression and Prognosis for SFRPs in Breast Carcinoma. *Cell Transplant*. 2020. Article No. 29. doi: 10.1177/0963689720962479

38. Motrich R.D., Castro G.M., Caputto, B.L. Old Players with a Newly Defined Function: Fra-1 and c-Fos Support Growth of Human Malignant Breast Tumors by Activating Membrane Biogenesis at the Cytoplasm. *PLoS ONE*. 2013. V. 8, No. 1 Article No. e53211. doi: 10.1371/journal.pone.0053211
39. Nulali J., Zhan M., Zhang K., Tu P., Liu Y., Song H. Osteoglycin: An ECM Factor Regulating Fibrosis and Tumorigenesis. *Biomolecules*. 2022. V. 12. No. 11. Article No. 1674. doi: 10.3390/biom12111674
40. Bae H., Kim B., Lee H., Lee S., Kang H.S., Kim, S.J. Epigenetically regulated Fibronectin leucine rich transmembrane protein 2 (FLRT2) shows tumor suppressor activity in breast cancer cells. *Sci. Rep*. 2017. V. 7. Article No. 272. doi: 10.1038/s41598-017-00424-0
41. Ahmed K.A., Hasib T.A., Paul S.K., Saddam M., Mimi A., Saikat A.S.M., Faruque H.A., Rahman M.A., Uddin M.J., Kim, B. Potential Role of CCN Proteins in Breast Cancer: Therapeutic Advances and Perspectives. *Curr. Oncol*. 2021. V. 28. No. 6. P. 4972–4985. doi: 10.3390/curroncol28060417
42. Harikrishnan K., Joshi O., Madangirikar S., Balasubramanian N. Cell Derived Matrix Fibulin-1 Associates With Epidermal Growth Factor Receptor to Inhibit Its Activation, Localization and Function in Lung Cancer Calu-1 Cells. *Front. Cell Dev. Biol*. 2020. V. 8. Article No. 522. doi: 10.3389/fcell.2020.00522
43. Shan N.L., Minden A., Furmanski P., Bak M.J., Cai L., Wernyj R., Sargsyan D., Cheng D., Wu R., Kuo H.C.D., et al. Analysis of the Transcriptome: Regulation of Cancer Stemness in Breast Ductal Carcinoma In Situ by Vitamin D Compounds. *Cancer Prev. Res.* 2020. V. 13. P. 673–686. doi: 10.1158/1940-6207.CAPR-19-0566
44. Pei G., Lan Y., Lu W., Ji L., Hua Z.C. The function of FAK/CCDC80/E-cadherin pathway in the regulation of B16F10 cell migration. *Oncol. Lett*. 2018. V. 16. No. 4. P. 4761–4767. doi: 10.3892/ol.2018.9159
45. Miricescu D., Totan A., Stanescu-Spinu I.I., Badoiu S.C., Stefani C., Greabu, M. PI3K/AKT/mTOR Signaling Pathway in Breast Cancer: From Molecular Landscape to Clinical Aspects. *Int. J. Mol. Sci*. 2021. V. 22. Article No. 173. doi: 10.3390/ijms22010173
46. Abreu de Oliveira W.A., El Laithy Y., Bruna A., Annibali D., Lluis F. Wnt Signaling in the Breast: From Development to Disease. *Front. Cell Dev. Biol.* 2022. V. 10. Article No. 884467. doi: 10.3389/fcell.2022.884467
47. You M., Xie Z., Zhang N., Zhang Y., Xiao D., Liu S., Zhuang W., Li L., Tao Y. Signaling pathways in cancer metabolism: mechanisms and therapeutic targets. *Signal Transduct. Target. Ther*. 2023. V. 8. Article No. 196. doi: 10.1038/s41392-023-01442-3
48. Hamad H.A., Enezei H.H., Alrawas A., Zakuan N.M., Abdullah N.A., Cheah Y.K., Hashim N.F.M. Identification of Potential Chemical Substrates as Fuel for Hypoxic Tumors That May Be Linked to Invadopodium Formation in Hypoxia-Induced MDA-MB-231 Breast-Cancer Cell Line. *Molecules*. 2020. V. 25. No. 17. Article No. 3876. doi: 10.3390/molecules25173876
49. Popova N.V., Jücker M. The Functional Role of Extracellular Matrix Proteins in Cancer. *Cancers*. 2022. V. 14. No. 1. Article No. 238. doi: 10.3390/cancers14010238
50. Wang W., Wang X., Yao F., Huang C. Lysyl Oxidase Family Proteins: Prospective Therapeutic Targets in Cancer. *Int. J. Mol. Sci*. 2022. V. 23. No. 20. Article No. 12270. doi: 10.3390/ijms232012270
51. Meysami M., Rahaie M., Ebrahimi A., Samiee, F. Four Matrix Metalloproteinase genes involved in murine breast cancer affected by ginger extract. *Gene Rep*. 2021. V. 25. Article No. 101332. doi: 10.1016/j.genrep.2021.101332
52. Jiang Y., Wang B., Li J.K., Li S.Y., Niu R.L., Fu N.Q., Zheng J.J., Liu G., Wang Z.L. Collagen fiber features and COL1A1: are they associated with elastic parameters in breast lesions, and can COL1A1 predict axillary lymph node metastasis? *BMC Cancer*. 2022. V. 22. No. 1. Article No. 1004. doi: 10.1186/s12885-022-10092-7

53. Li X., Sun X., Kan C., Chen B., Qu N., Hou N., Liu Y., Han F. COL1A1: A novel oncogenic gene and therapeutic target in malignancies. *Pathol. Res. Pract.* 2022. V. 236. Article No. 154013. doi: 10.1016/j.prp.2022.154013

Received 12.04.2024.
Revised 17.06.2024.
Published 24.08.2024.