Eurasian Journal of Medicine and Oncology

ORIGINAL RESEARCH ARTICLE

Differential gene expression and gene ontology associated with breast cancer development and progression: A meta-analysis study

Najia El Aboudi[1], Faissal Ouardi[1], Mouna Ababou[2], Laraqui Abdelilah[3], Oubaida Elbiad[3], and Bouabid Badaoui[4]*

[1]Department of Informatics, Faculty of Science, Mohammed V University, Rabat, Morocco

[2]Laboratory of Human Pathologies, Genomic Center of Human Pathologies, Faculty of Sciences, Mohammed V University, Rabat, Morocco

[3]Biosafety Level 3 and Research Laboratory, Mohammed V Military Teaching Hospital, Rabat, Morocco

[4]Center of Plant and Microbial Biotechnology, Biodiversity, and Environment, Faculty of Science, Mohammed V University, Rabat, Morocco

***Corresponding author:**
Bouabid Badaoui
(b.badaoui@um5r.ac.ma)

Citation: El Aboudi N, Ouardi F, Ababou M, Abdelilah L, Elbiad O, Badaoui B. Differential gene expression and gene ontology associated with breast cancer development and progression: A meta-analysis study. *Eurasian J Med Oncol*. 2025;9(4):296-310. doi: 10.36922/EJMO025060025

Received: February 7, 2025

Revised: April 27, 2025

Accepted: May 6, 2025

Published online: November 10, 2025

Copyright: © 2025 Author(s). This is an Open Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abstract

Introduction: Breast cancer (BC) is heterogeneous and remains a major health priority. Robust biomarkers are needed to improve early detection and guide therapy.
Objective: This study investigates the molecular mechanisms underlying BC development (normal versus cancer) and progression (early versus late).
Methods: A meta-analysis of 51 microarray studies comparing BC and normal cells, as well as early- and late-stage BC cells, was conducted using microarray data obtained from the Gene Expression Omnibus and ArrayExpress databases. Five meta-analysis methods, each based on different statistical approaches, were applied, and the overlapping results were identified.
Results: A total of 3,362 and 95 differentially expressed genes (DEGs) associated with BC development and progression were identified. Among these DEGs, the upregulation of *COL10A1*, *COL11A1*, *TOP2A*, *CDK1*, *MMP11*, and *S100P* and the downregulation of *ADH1B*, *SFRP1*, *DST*, *LEP*, *ADIPOQ*, and *CHRDL1* were the most significantly associated with BC development. DEGs such as *DHTKD1* and *CBX3* (upregulated) and *MAP3K20* (downregulated) were found to be among the most significantly differentially expressed in BC progression. The top gene ontology terms enriched in BC development included regulation of signaling receptor activity and cytokine-mediated signaling pathway. In addition, the cellular macromolecule biosynthetic process and response to organic substances were significantly enriched in BC progression.
Conclusion: Many of the DEGs may serve as potential therapeutic targets for BC.

Keywords: Breast cancer; Meta-analysis; Microarray data; Differentially expressed genes; Early-stage cancer; Late-stage cancer; Gene ontology

1. Introduction

Breast cancer (BC) is the most frequently diagnosed cancer among women worldwide, corresponding to 23.8% of all cancers, affecting 2.3 million women each year, among

whom 665,684 succumb to death.[1] BC remains a major health problem and currently represents a top biomedical research priority. New diagnostic and therapeutic tools are required to improve the personalized management of each patient.

BC is a complex and heterogeneous disease[2] with varied molecular features, tumor characteristics, expression patterns, and responses to therapy. Perou *et al.*[3] established the mammary tumor classification system based on the analysis of gene expression. They classified BC into four molecular classes: luminal-like, basal-like, normal-like, and human epidermal growth factor receptor 2-positive BC.[3]

Early detection of BC is of considerable clinical importance, and it can be used to make treatment decisions while the tumor burden is still low and when patients are most likely to respond to adjuvant therapy.[4] BC markers that showed evidence of clinical utility and were recommended for use in practice include carcinoembryonic antigen, estrogen receptor, progesterone receptor, human epidermal growth factor receptor 2, urokinase plasminogen activator, plasminogen activator inhibitor 1, the tumor suppressor *p53*, cathepsin D, and multiparameter assays for gene expression.[5]

Although some genes were found to be responsible for the development of BC, none of the mentioned markers can diagnose early- or late-stage BC. Thus, there is a crucial need to identify new biomarkers that enable the detection of early molecular phases in BC carcinogenesis and facilitate targeted treatments to prevent relapses. This necessity has recently driven intense research efforts to discover robust clinical and pathological prognostic and predictive factors.

Recent microarray-based expression profiling studies have provided evidence that breast tumors of different grades have distinct genomic, transcriptomic, and immunohistochemical profiles.[6] DNA microarrays can assess the expression of thousands of genes simultaneously.[7] Specifically, gene-expression profiling can potentially aid in the identification of prognostic markers for different types of cancers and new therapeutic targets. It has also been established that tumors that remain localized and primary tumors that metastasize have different gene expression profiles.[8]

Many microarray studies have resulted in a significant increase in publicly available databases, such as the Gene Expression Omnibus (GEO) from the National Center of Biotechnology Information (NCBI) and the ArrayExpress database from the European Bioinformatics Institute.[9,10] From these databases, numerous studies related to relevant diseases or biological processes (BP) can be accessed, and the number of available studies continues to grow rapidly. The most widely known and frequently used platforms include the Affymetrix GeneChips (*in situ* synthesized oligonucleotide microarray) and the Illumina high-density bead arrays.[11]

Combining transcriptomic studies, through meta-analysis, for a disease or condition could increase sensitivity and validate conclusions. Meta-analysis is most commonly applied to identify differentially expressed genes (DEGs),[12] which may serve as biomarkers or be used as features in classification models or classifiers to further refine a clinically useful gene signature.[13]

Several meta-analysis studies have been conducted to identify DEGs involved in BC and other subtypes of BC. Karn *et al.*[14] combined 3,030 Affymetrix HG-U133A microarrays from primary BC samples, including 238 of their own samples and 2,792 samples from 22 different publicly available datasets.[14] Sotiriou *et al.*[15] analyzed microarray data from 189 invasive BC samples (GSE2990) and three published gene expression datasets of breast carcinomas.[15] In addition, a meta-analysis was conducted on six triple-negative BC (TNBC) microarray expression data sets collected from the NCBI GEO database to identify robust DEGs of TNBC.[16]

Other studies using a few microarrays have also highlighted DEGs in BC.[17,18] R programming (http://www.r-project.org) and the associated Bioconductor project have become one of the most widespread tools for bioinformatics and biostatistics in recent years. Many newly developed meta-analysis methods have been published, including the Bayesian approach,[19] Fisher inverse chi-square,[20] probability of expression,[21] meta differential expression through distance synthesis,[22] and RankProd.[23] Furthermore, gene ontology (GO) analysis has become a commonly used approach for functional studies of large-scale genomic or transcriptomic data.

To date, published meta-analyses have employed different statistical approaches, often leading to differing conclusions. In addition, many of these studies relied on preprocessed data, which may hinder the accuracy of the analysis. Hence, in the present study, we performed a large-scale meta-analysis combining different statistical approaches to compare the expression profiles in BC and normal breast tissues, and in early- and late-stage BC tissues using microarray data accessible through GEO and ArrayExpress databases.

2. Methods

2.1. Data collection

We used "breast cancer" as the keyword to search for genome-wide expression studies in the NCBI-GEO

(http://www.ncbi.nlm.nih.gov/geo/) and European Molecular Biology Laboratory-European Bioinformatics Institute ArrayExpress databases (https://www.ebi.ac.uk/arrayexpress/). We used the GEOquery and ArrayExpress R packages in the R software (version R 4.1.3) to download cell intensity files. Only studies utilizing the Affymetrix platform were selected to analyze DEGs in BC, because various gene-expression profiling platforms measure the same gene with differing precision, relative scales, and dynamic ranges.[24] Only datasets that met the inclusion criteria were included in the analysis. The inclusion criteria were studies employing expression profiling by array and studies with cell intensity files. Studies involving cell lines were excluded from the analysis. The following information was also extracted from each identified study: accession number, country, year, status (cancer or normal), stage, platform, number of probes, ethnic group, and author. All publicly available microarray datasets that corresponded to BC studies using the Affymetrix platform published before March 4, 2019, were included. A total of 51 microarray datasets that corresponded to 10,959 chips, including 5,497 cases and 5,462 controls, were considered in this study (Table S1).

2.2. Preliminary preprocessing of datasets

We used the arrayQualityMetrics R package to reanalyze the quality of each dataset, which was scored on the basis of spatial, boxplot, heatmap, and run-length encoding metrics. We removed any array that failed in at least two metrics, as performed in a previous study.[25] Normalization for each dataset was performed independently using the robust multi-array average (RMA) expression measure[26] given in the affy package. RMA uses a model-based background correction, quantile normalization, and probe-level summarization to standardize signal intensities and reduce technical noise.

Subsequently, we extracted the common genes across all datasets using the R merge function; the data where genes do not overlap were excluded. To perform the meta-analysis, each dataset must consist of a control group; hence, for datasets lacking a control group, we considered a common reference control group from the other datasets included in this study. Although standardized preprocessing and normalization procedures were applied, the integration of data from multiple sources may introduce batch effects and other forms of variability.

2.3. Meta-analysis of the microarray data

The aim of this study was twofold: (i) to identify the set of DEGs between BC and control tissues using the complete dataset and (ii) to identify the DEGs associated with early- versus late-stage BC. Early-stage BC was defined as arrays corresponding to stage T1, node-negative (N0), grade 1, and early or primary BC. Late stage included all arrays in stage T4, node-positive (N1), metastatic (M1), and grade 3 BC cases. Accordingly, we highlighted 21 datasets harboring 1,438 chips, of which 1,153 were early-stage BC and 127 were late-stage BC.

We used the R Meta-Analysis of MicroArray package (R-Bioconductor 3.7 in R version 3.5.2) to perform meta-analysis of microarray, focusing on the overlapping outputs among five different methods: (i) *p*-value combination method (*p*val), (ii) effect size combination method (Es), (iii) meta-analysis pattern matches (MAP-Matches), (iv) meta-analysis of ranked discovery datasets (METRADISC), and (v) *z*-statistic method.

The p-value combination methods applied in this study included Fisher's method,[27] Stouffer's method,[28] the minimum p-value approach,[29] and the maximum p-value approach.[30] These methods aggregate significance levels obtained from individual studies. Other methods combine the effect size from each dataset. Depending on the assumption of homogeneity, either a fixed-effect model (FEM) or a random-effect model (REM) was adopted. FEM model combines the effect size across all studies by assuming a simple linear model with an underlying true effect size plus a random error in each study, while REM[31] allows random effects for the inter-study heterogeneity in the model. MAP-Matches involves computing *t*-statistics for each pairwise class comparison in individual datasets, followed by the construction of a *t*-statistics matrix for pattern matching. METRADISC[32] is a rank-based method that estimates heterogeneity across datasets. The *z*-statistic method utilizes the posterior mean differential expression, weighted by variance. The method assumes a standard normal distribution due to the classic Bayesian probability calculation and derives a prior distribution of differential expression to assess significance.

The datasets were transformed to a special object class from the MetaArray package, where separate slots were designated for gene expression data matrices (expression profiles), clinical sample characteristics (BC [1] and normal [0]), and dataset identifiers.

Five meta-analysis approaches were employed to identify the DEGs, after which the overlapping genes were considered as the final output of the meta-analysis. The top 100 DEGs genes in BC development and progression were considered for further analysis.

2.4. Gene ontology analysis

Gene ontology analysis was performed for the two meta-analyses using the topGO R package to identify the most represented biological functions in the two datasets. The

topGO package was designed to work with various test statistics and algorithms, taking GO dependencies into account.[33] The DEGs were analyzed using the topGO package, which uses Fisher's test to identify significantly enriched GO terms.[34,35]

3. Results

3.1. DEGs in BC development and progression

A total of 22,268 probes present across all datasets were considered for the identification of DEGs. Five different meta-analysis methods (*p*val, Es, MAP-Matches, METRADISC, and *z*-statistic) were performed, resulting in the identification of 8,031, 11,765, 21,205, 20,304, and 9,102 DEGs, respectively ($p<0.05$). The obtained DEGs from each method were plotted into a Venn diagram (Figure 1A), which highlighted 3,362 common DEGs in BC across all meta-analyses.

From a total of 10,959 microarray datasets used in BC, a subset of 12,80 was associated with early- (1,153 arrays) and late-stage (127 arrays) BC and was used for the meta-analysis of BC progression. The DEGs identified through the meta-analysis were plotted into a Venn diagram (Figure 1B), resulting in 95 common DEGs among all the meta-analysis methods. The top DEGs in BC development (100 DEGs) and progression (95 DEGs) are presented in Tables S2 and S3, respectively.

Among these DEGs, Table 1 shows the highly differentially expressed Oncogenes and Tumor suppressor genes in BC development and progression.

3.2. Gene ontology functional enrichment analysis

DEGs from BC development were subjected to enrichment analysis using the topGO R package. The most enriched BP were regulation of signaling receptor activity, cytokine-mediated signaling pathway, response to drug, negative regulation of cell proliferation, leukocyte migration, organic hydroxy compound transport, positive regulation of cell proliferation, G-protein coupled receptor signaling pathway, receptor-mediated endocytosis, and response to glucocorticoid (Table 2).

The most enriched GO terms in BC progression included cellular macromolecule biosynthetic process, response to organic substance, response to endogenous stimulus, cell development, cellular protein localization, cellular macromolecule localization, nucleic acid metabolic process, cellular response to organic substance, transcription by RNA polymerase II, and organonitrogen compound biosynthetic process (Table 3).

We found 17 and eight DEGs highly implicated in GO terms corresponding to BC development and progression, respectively. Among those genes, some were involved in more than one GO term (Table 4). These genes are more likely to play a fundamental role in cancer development and progression, and particular attention should be given to them.

4. Discussion

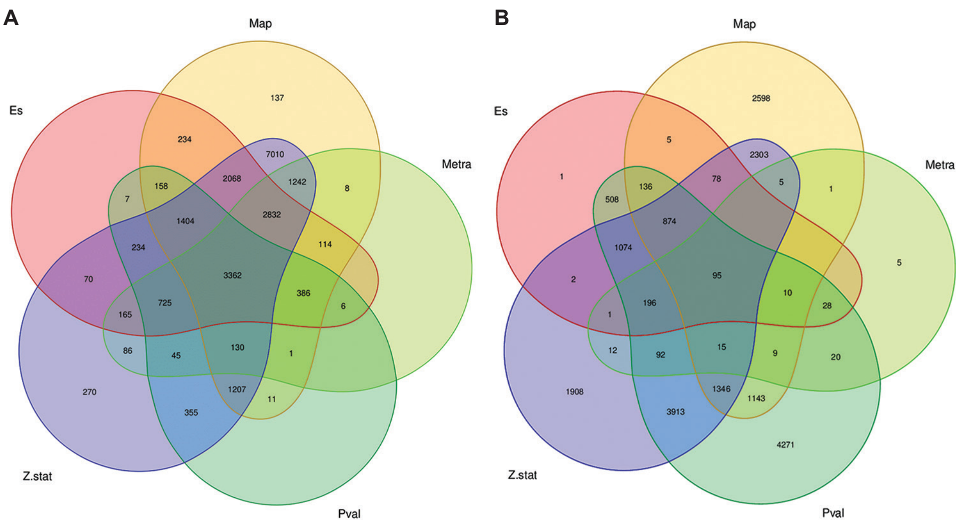


Figure 1. Venn diagrams of differentially expressed genes. (A) Venn diagram illustrating the differentially expressed genes in BC development using five statistical approaches. A total of 3,362 genes were common between the five meta-analysis methods. (B) Venn diagram illustrating the differentially expressed genes in BC progression using five statistical approaches. A total of 95 genes were common between the five meta-analysis methods.
Abbreviations: Es: Effect size combination method; Map: Meta-analysis pattern matches; Metra: Meta-analysis of ranked discovery datasets; Pval: *p*-value combination method; Z.stat: *z*-statistic method.

Table 1. Top differentially expressed oncogenes and tumor suppressor genes associated with breast cancer development and progression

Cases	Tumor suppressor genes	Oncogenes
Normal versus breast cancer	*CHL1*; *ABCA8*; *SFRP1*; *MME*	*COL10A1*; *COL11A1*; *ID4*; *CDK1*; *MELK*; *ATAD2*; *CEP55*; *HMMR*; *OXTR*
Early- versus late-stage breast cancer		*EVI5*; *SLC35A2*

Table 2. Gene ontology terms for breast cancer development

Gene ontology ID	Term	Annotated	Significant	Expected	Rank in classic Fisher	Classic Fisher	Classic KS	Elim KS
GO: 0010469	Regulation of signaling receptor activity	98	98	98	1	1	2.9e-10	2.9e-10
GO: 0019221	Cytokine-mediated signaling pathway	227	227	227	2	1	1.7e-10	7.7e-09
GO: 0042493	Response to drug	234	234	234	3	1	4.1e-09	9.6e-07
GO: 0008285	Negative regulation of cell proliferation	147	147	147	4	1	2.1e-08	1.7e-06
GO: 0050900	Leukocyte migration	130	130	130	5	1	1.7e-07	2.5e-06
GO: 0015850	Organic hydroxy compound transport	60	60	60	6	1	3.2e-06	3.2e-06
GO: 0008284	Positive regulation of cell proliferation	217	217	217	7	1	8.2e-14	4.9e-06
GO: 0007186	G protein-coupled receptor signaling pathway	179	179	179	8	1	2.7e-09	5.5e-06
GO: 0006898	Receptor-mediated endocytosis	97	97	97	9	1	5.9e-06	5.9e-06
GO: 0051384	Response to glucocorticoid	37	37	37	10	1	6.2e-06	6.2e-06

Note: Classic indicates that each gene ontology (GO) term was tested independently, not taking the GO hierarchy into account. Elim processes the GO terms by traversing the GO hierarchy from bottom to top.
Abbreviation: KS: Kolmogorov–Smirnov.

Table 3. Gene ontology terms for breast cancer progression

Gene ontology ID	Term	Annotated	Significant	Expected	Rank in classic Fisher	Classic Fisher	Classic KS	Elim KS
GO: 0034645	Cellular macromolecule biosynthetic process	37	37	37	1	1	0.0036	0.0036
GO: 0010033	Response to organic substance	21	21	21	2	1	0.0049	0.0049
GO: 0009719	Response to endogenous stimulus	12	12	12	3	1	0.0084	0.0084
GO: 0048468	Cell development	11	11	11	4	1	0.0155	0.0155
GO: 0034613	Cellular protein localization	21	21	21	5	1	0.0158	0.0158
GO: 0070727	Cellular macromolecule localization	21	21	21	6	1	0.0158	0.0158
GO: 0090304	Nucleic acid metabolic process	39	39	39	7	1	0.0174	0.0174
GO: 0071310	Cellular response to organic substance	18	18	18	8	1	0.0192	0.0192
GO: 0006366	Transcription by RNA polymerase II	18	18	18	9	1	0.0212	0.0212
GO: 1901566	Organonitrogen compound biosynthetic process	16	16	16	10	1	0.0306	0.0306

Note: Classic indicates that each gene ontology (GO) term was tested independently, not taking the GO hierarchy into account. Elim processes the GO terms by traversing the GO hierarchy from bottom to top.
Abbreviation: KS: Kolmogorov–Smirnov.

Table 4. Highly involved genes in the enriched gene ontology terms in breast cancer development and progression

Cases	Upregulated genes	Downregulated genes
Normal versus breast cancer	*INHBA*[a]; *CXCL10*[a]; *SPP1*[a]; *FN1*[a]; *CDK1*[a]; *ECT2*; *ARF1*; *RAB31*; *SDC1*	*LEP*[a]; *ADIPOQ*[a]; *SFRP1*[a]; *VEGFD*; *CXCL2*; *CD36*; *ANGPT1*; *TF*
Early- versus late-stage breast cancer	*CBX3*[a]; *DPM1*; *SHARPIN*; *RAC1*; *RNF10*; *YWHAZ*[a]; *XPO1*[a]; *DHTKD1*	-

Note: [a]genes implicated in more than one gene ontology term.

BC remains the most aggressive cancer, resulting in substantial mortality. The fight against BC represents a major challenge as its heterogeneity renders it difficult to understand. Failure to account for this variability can limit the generalizability and specificity of the findings. Gene expression studies combined with bioinformatics would permit the identification of biological features and target genes associated with BC, which could improve the understanding of molecular mechanisms shaping cancer development and progression.[36,37]

This study is a meta-analysis of publicly available transcriptomic datasets rather than an original clinical investigation. While this design enhances statistical power and generalizability by integrating data across diverse populations and platforms, it also limits the depth of clinical interpretation. The lack of patient-level clinical metadata restricts our ability to establish causal or prognostic relationships. As a result, further validation in prospective clinical studies is warranted.

Control samples were obtained from datasets different from the case samples. Integrating data from multiple sources introduces the potential for batch effects, despite normalization efforts. To our knowledge, we report herein the largest combined clinical and gene expression meta-analysis of BC using raw data and different statistical approaches to obtain the most pertinent results. Our meta-analysis integrated 51 BC microarray datasets and led to the identification of 3,362 and 95 DEGs for BC development and progression, respectively. These findings are discussed in the context of BP represented by the DEGs.

This study did not include external validation of the identified gene expression changes. The findings were derived from a single analytical pipeline without confirmation using experimental techniques. As a result, the reliability and reproducibility of the results are limited. Future studies should prioritize external validation to strengthen the biological and clinical relevance of the identified genes.

4.1. BC development

Regulation of signaling receptor activity is defined as any process that modulates the frequency, rate, or extent of a signaling receptor activity. Receptor activity is manifested when a molecule combines with an extracellular or intracellular messenger to initiate a change in cell activity.[38] The *INHBA* gene, upregulated and involved in this biological process (Figure 2), encodes for inhibin subunit beta A, members of the transforming growth factor β superfamily,[39] where its activation results in cell growth, proliferation, differentiation, metabolism, homeostasis, apoptosis, and carcinogenesis.[40] This gene has been previously studied in BC, where its mRNA expression was higher than in normal breast tissues.[41] Therefore, *INHBA* may serve as a prognostic indicator for BC patients.[42] In addition, it has been reported that *INHBA* signaling promotes BC metastasis by regulating *IL13Rα2* expression.[43]

LEP gene, significantly downregulated in this biological process (Figure 2), encodes for leptin, which may promote mammary tumor growth through multiple mechanisms such as modulation of the extracellular environment, downregulation of apoptosis, and/or upregulation of anti-apoptotic genes.[44] Leptin level plays a role in BC and has the potential to be developed as a diagnostic tool.[45] In premenopausal women, researchers found reduced BC risk associated with elevated leptin levels.[46] Ishikawa *et al.*[47] showed that leptin may play a role in the carcinogenesis and metastasis of BC, possibly in an autocrine manner. In a contradictory result, Niu *et al.*[45] showed that the mean serum leptin level of case groups was significantly higher than that of control groups, indicating that higher leptin may be associated with increased incidence and development of BC. Several studies have identified the associations of several polymorphisms in *LEP* with BC risk,[48,49] whereas others suggested null associations.[49,50]

Cytokine-mediated signaling pathway regulates many downstream cellular processes.[38] *FN1* gene, upregulated in this BP, has been found to be upregulated in both cancer epithelium and stroma relative to normal epithelium and stroma, respectively.[51] *FN1* encodes for a heterodimeric glycoprotein form at the cell surface and extracellular matrix that binds to interleukin[52] and is found to be involved in cell adhesion and cell metastasis.[53] The *FN1* gene was found to be a key regulator in BC development and a potential target in cancer treatment.[54] Moreover, *FN1* expression is directly regulated by microRNA-206, which has been demonstrated to be associated with metastatic cancer types, including BC.[55,56]

Response to drug is defined as any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a drug stimulus.[38] *CDK1* and *SFRP1* are associated in this BP. Specifically, *CDK1*, implicated in mammalian mitosis,[57,58] was involved in the G2/M phase arrest in many tumor types[59] as well as the initiation and development of cancer. *SFRP1* acts as a tumor suppressor gene,[60] and *SFRP1* promoters are often hypermethylated in tumor cells.[61] *SFRP1* was overexpressed in TNBC compared to other BC subtypes.[62] Knockdown of *SFRP1* increased the migration and invasion potential of tumor cells as well as reduced apoptotic events. Lack of

Figure 2. Means and standard error of the differentially expressed genes in breast cancer development, implicated in the regulation of signaling receptor activity. Cancer is represented as blue, and the control is in red.

SFRP1 protein expression is associated with poor overall survival in patients with early BC.[63]

Basal-like subtype is associated with low methylation levels of *SFRP1*, suggesting that the methylation levels of specific BC genes may potentially serve as epigenetic biomarkers and prognostic factors.[64] In BC, hypermethylation of the *SFRP1* promoter has been correlated to poor prognosis, presumably due to elevated levels of Wnt signaling,[61] indicating that the level of *SFRP1* may be associated with BC.

Negative regulation of cell proliferation and leukocyte migration are processes that control cell proliferation.[38] *INHBA* and *SFRP1* are involved in the former, while *FN*1 and *LEP* are associated with the latter.

COL10A1, *COL11A1*, *TOP2A*, *CDK1*, *MMP11*, and *S100P* were the most significantly upregulated and *ADH1B*, *SFRP1*, *DST*, *LEP*, *ADIPOQ*, and *CHRDL1* were most significantly downregulated genes in BC development. *COL10A1* was overexpressed in BC and was involved in tumor vasculature through the regulation of cell proliferation, migration, and invasion.[65-69] Protein expression of the *COL10A1* gene in the plasma could be a potential biomarker for early BC.[70] In addition, *COL11A1* was overexpressed in invasive ductal and lobular carcinomas[71] and has been reported to be a fundamental gene in late-stage BC.[72] *TOP2A* provokes tumorigenesis through chromosome instability[73,74] and DNA damage[75] and is considered a biomarker of cancer proliferation[76,77]

as it was highly overexpressed in proliferative subtypes of BC, like human epidermal growth factor receptor 2-enriched BC, luminal BC, and TNBC.[78-81] *MMP-11* was overexpressed in BC cells and is associated with poor survival rates in patients with BC.[82] It is considered a potential biomarker for early BC diagnosis and prognosis,[83] metastatic BC,[84] and immunotherapeutic strategies.[85] *S100P* was upregulated in breast and many other cancers and controlled many carcinogenic processes.[86-91]

ADH1B was underexpressed in BC cells, which may induce metabolic pathway disorders leading to carcinogenesis.[92] Moreover, overexpression of *ADH1B* correlates with a decrease in mortality in patients with gastric cancer.[93]

DST may promote BC progression and behave as a tumor suppressor in breast epithelial cells.[94] *ADIPOQ* has been reported to increase BC risk and was correlated with poor outcomes in patients with early BC.[95] *Chrdl1* has a tumor suppressor activity in BC, and its overexpression is correlated with longer disease-free and distant metastasis-free survival times.[96] Thus, *Chrdl1* upregulation might be a potential biomarker for improved outcomes in patients with BC.

4.2. BC progression

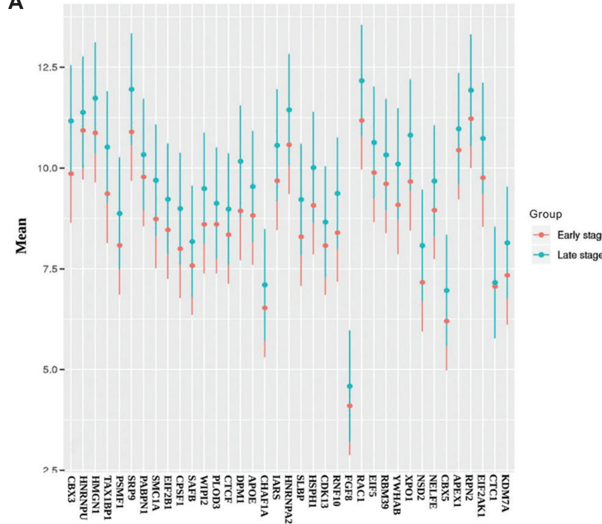
CBX3 was upregulated in the BP of cellular macromolecule biosynthesis (Figure 3A). *CBX3* silences gene expression,[97-100] and *CBX3/HP1γ* can affect transcriptional activation, alternative RNA splicing, DNA damage responses, transcription elongation, and cell growth and differentiation.[101-108] Overexpression of *CBX3/HP1γ* negatively impacts prognosis by promoting cancer development and growth in BC.[109]

The response to organic substances involving *SHARPIN* plays an important regulatory function in inflammation control.[110-113] Numerous studies have reported that *SHARPIN* is upregulated in various cancers, including BC,[114-116] and its expression correlates with poor prognosis in BC patients.[115,117,118]

Another enriched BP in BC progression is the response to endogenous stimulus, which might be linked to cancer in terms of gene expression, secretion, and cell movement.

Cell development concerns cell progression over time.[38] *RAC*, upregulated in this BP, is highly implicated in cell migration and survival by regulating cytoskeleton reorganization.[119] It has been reported to be overexpressed in BC as well as many other tumors[120] and is an indicator of cancer therapy resistance.[121-123] Furthermore, radioresistance of BC cells depends on *Rac1* signaling.[124]

Cellular protein localization harbored *YWHAZ* (tyrosine 3-monooxygenase/tryptophan 5-monooxygenase



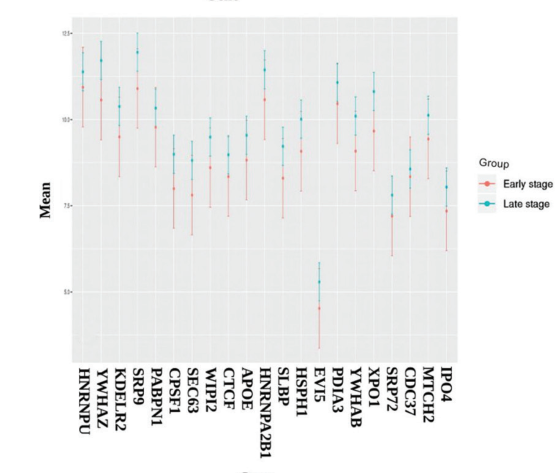


Figure 3. Means and standard error of the differentially expressed genes in breast cancer progression implicated in (A) cellular macromolecule biosynthetic process and (B) cellular protein localization. The late-stage cancer is represented in blue, and the early-stage cancer is in red.

activation protein zeta) gene that encodes for anti-apoptotic proteins.[125] This gene was overexpressed in cancer tissues (Figure 3B), and its experimental downregulation inhibited cell migration.[126] In addition, *YWHAZ* confers cancer cells apoptosis resistance and has been associated with low survival rates and drug resistance in BC patients.[127-129]

In BC progression, at least three genes were highly modulated: *DHTKD1* and *CBX3* were upregulated, and *MAP3K20* was downregulated. *DHTKD1* has been implicated in reducing cell apoptosis and increasing cell growth through mitochondrial biogenesis and function maintenance.[130] It was reported that mitochondrial dysfunction is associated with several diseases, including cancers.[131-133] Furthermore, the *DHTKD1*-encoded protein

was reported as a protein biomarker of late-stage BC.[134] *CBX3* is involved in cell growth and differentiation and DNA damage responses.[101-108] Specifically, overexpression of the *CBX3* gene is an indicator of poor prognosis in BC.[135]

The DEG discussed herein constitutes potential biomarkers for BC development and progression. Patients could be evaluated by dosing the expression of one or a set of biomarkers. Some biomarkers could be highly useful for the choice of treatment strategies. Among the robust DEGs, the *CDK* gene corresponds to known or potential drug targets. *CDK* genes, particularly *CDK4* and *CDK6*, are already established therapeutic targets in BC, with inhibitors such as palbociclib, ribociclib, and abemaciclib widely used in clinical practice. However, this study did not associate the identified DEGs identified to approved drugs, drug development pipelines, or databases.

Several genes were found to be upregulated; however, the study did not investigate whether these genes are associated with patient outcomes such as overall survival, progression-free survival, or response to specific treatments. Without linking gene expression changes to clinical endpoints, the functional and prognostic significance of these findings remains unclear. Future work should integrate survival analysis or treatment response data to identify genes with potential clinical utility.

A major limitation of this study is the absence of adjustment for critical clinical and demographic factors such as age, treatment history, hormone receptor status, and ethnicity. These variables are well-documented confounders in gene expression analyses and can significantly influence transcriptional profiles. The lack of adjustment may mask true biological signals, thereby limiting the interpretability and generalizability of our findings.

5. Conclusion

In this study, we combined many statistical methods to perform a stringent meta-analysis for BC development and progression. This approach highlighted 3,362 and 95 DEGs associated with BC development and progression, respectively. The top BP of the GO term with the associated genes were scrutinized to gain a deeper insight into BC development and progression.

Acknowledgments

None.

Funding

This study was supported by the National Centre for Scientific and Technical Research (CNRST) through the Programme de Bourses d'Excellence de Recherche (PBER).

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Bouabid Badaoui
Data curation: Najia El Aboudi, Faissal Ouardi
Formal analysis: Najia El Aboudi, Bouabid Badaoui
Investigation: Mouna Ababou, Laraqui Abdelilah, Oubaida Elbiad
Writing–original draft: Najia El Aboudi
Writing–review & editing: Bouabid Badaoui

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

All data analyzed have been presented in the paper.

References

1. Bray F, Laversanne M, Sung H, *et al*. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2024;74(3):229-263.

   doi: 10.3322/caac.21834

2. Einbeigi Z, Bergman A, Kindblom LG, *et al*. A founder mutation of the BRCA1 gene in Western Sweden associated with a high incidence of breast and ovarian cancer. *Eur J Cancer*. 2001;37(15):1904-1909.

   doi: 10.1016/S0959-8049(01)00223-4

3. Perou CM, Sorlie T, Eisen MB, *et al*. Molecular portraits of human breast tumours. *Nature*. 2000;406(6797):747-752.

   doi: 10.1038/35021093

4. Shah R, Rosso K, Nathanson SD. Pathogenesis, prevention, diagnosis and treatment of breast cancer. *World J Clin Oncol*. 2014;5(3):283-298.

   doi: 10.5306/wjco.v5.i3.283

5. Donepudi MS, Kondapalli K, Amos SJ, Venkanteshan P. Breast cancer statistics and markers. *J Cancer Res Ther*. 2014;10(3):506-511.

   doi: 10.4103/0973-1482.137927

6. Rakha EA, Reis-Filho JS, Baehner F, *et al*. Breast cancer prognostic classification in the molecular era: The role of histological grade. *Breast Cancer Res*. 2010;12(4):207.

   doi: 10.1186/bcr2607

7. Schena M, Shalon D, Davis RW, Brown PO. Quantitative

monitoring of gene expression patterns with a complementary DNA microarray. *Science*. 1995;270(5235):467-470.

doi: 10.1126/science.270.5235.467

8. Montel V, Huang TY, Mose E, Pestonjamasp K, Tarin D. Expression profiling of primary tumors and matched lymphatic and lung metastases in a xenogeneic breast cancer model. *Am J Pathol.* 2005;166(5):1565-1579.

   doi: 10.1016/S0002-9440(10)62372-3

9. Clough E, Barrett T. The gene expression omnibus database. *Methods Mol Biol*. 2016;1418:93-110.

   doi: 10.1007/978-1-4939-3578-9_5

10. Brazma A, Parkinson H, Sarkans U, *et al*. ArrayExpress-a public repository for microarray gene expression data at the EBI. *Nucleic Acids Res*. 2005;33(Database issue):D553-D555.

    doi: 10.1093/nar/gki056

11. Miller MB, Tang YW. Basic concepts of microarrays and potential applications in clinical microbiology. *Clin Microbiol Rev*. 2009;22(4):611-633.

    doi: 10.1128/CMR.00019-09

12. Tseng GC, Ghosh D, Feingold E. Comprehensive literature review and statistical considerations for microarray meta-analysis. *Nucleic Acids Res*. 2012;40(9):3785-3799.

    doi: 10.1093/nar/gkr1265

13. Hu P, Wang X, Haitsma JJ, *et al*. Microarray meta-analysis identifies acute lung injury biomarkers in donor lungs that predict development of primary graft failure in recipients. *PLoS One*. 2012;7(10):e45506.

    doi: 10.1371/journal.pone.0045506

14. Karn T, Metzler D, Ruckhäberle E, *et al*. Data driven derivation of cutoffs from a pool of 3,030 Affymetrix arrays to stratify distinct clinical types of breast cancer. *Breast Cancer Res Treat*. 2010;120(3):567-579.

    doi: 10.1007/s10549-009-0416-z

15. Sotiriou C, Wirapati P, Loi S, *et al*. Gene expression profiling in breast cancer: Understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst*. 2006;98(4):262-272.

    doi: 10.1093/jnci/djj052

16. Naorem LD, Muthaiyan M, Venkatesan A. Integrated network analysis and machine learning approach for the identification of key genes of triple-negative breast cancer. *J Cell Biochem*. 2019;120(4):6154-6167.

    doi: 10.1002/jcb.27903

17. Li Y, Bai W, Zhang L. The overexpression of CD80 and ISG15 are associated with the progression and metastasis of breast cancer by a meta-analysis integrating three microarray datasets. *Pathol Oncol Res.* 2020;26(1):443-452.

    doi: 10.1007/s12253-018-0478-5

18. Meng L, Xu Y, Xu C, Zhang W. Biomarker discovery to improve prediction of breast cancer survival: Using gene expression profiling, meta-analysis, and tissue validation. *Onco Targets Ther.* 2016;9:6177-6185.

    doi: 10.2147/OTT.S113855

19. DuMouchel WH, Harris JE. Bayes methods for combining the results of cancer studies in humans and other species. *J Am Stat Assoc*. 1983;78:293-315.

    doi: 10.1080/01621459.1983.10477968

20. Fisher RA. *Statistical Methods for Research Workers*. Edinburgh: Oliver & Boyd; 1950. p. 11.

21. Parmigiani G, Garrett ES, Anbazhagan R, Gabrielson E. A statistical framework for expression-based molecular classification in cancer. *J R Stat Soc Ser B Stat Methodol*. 2002;64:717-736.

    doi: 10.1111/1467-9868.00358

22. Campain A, Yang YH. Comparison study of microarray meta-analysis methods. *BMC Bioinformatics.* 2010;11:408.

    doi: 10.1186/1471-2105-11-408

23. Breitling R, Armengaud P, Amtmann A, Herzyk P. Rank products: A simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett.* 2004;573:83-92.

    doi: 10.1016/j.febslet.2004.07.055

24. Tan PK, Downey TJ, Spitznagel EL Jr., *et al.* Evaluation of gene expression measurements from commercial microarray platforms. *Nucleic Acids Res*. 2003;31(19):5676-5684.

    doi: 10.1093/nar/gkg763

25. Mabbott NA, Baillie JK, Hume DA, Freeman TC. Meta-analysis of lineage-specific gene expression signatures in mouse leukocyte populations. *Immunobiology*. 2010;215(9-10):724-736.

    doi: 10.1016/j.imbio.2010.05.012

26. Irizarry RA, Hobbs B, Collin F, *et al*. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics.* 2003;4(2):249-264.

    doi: 10.1093/biostatistics/4.2.249

27. Fisher RA. *Statistical Methods for Research Workers.* Edinburgh: Genesis Publishing; 1925.

28. Stouffer SA. A study of attitudes. *Sci Am*. 1949;180(5):11-15.

    doi: 10.1038/scientificamerican0549-11

29. Wishart J. The methods of statistics: An introduction mainly for workers in the biological sciences; the combination of observation. *Nature*. 1932;130:491-492.

    doi: 10.1038/130491a0

30. Wilkinson B. A statistical consideration in psychological

research. *Psychol Bull*. 1951;48(3):156-168.

doi: 10.1037/h0059111

31. Choi JK, Yu U, Kim S, Yoo OJ. Combining multiple microarray studies and modeling interstudy variation. *Bioinformatics*. 2003;19(suppl 1):i84-i90.

doi: 10.1093/bioinformatics/btg1010

32. Zintzaras E, Ioannidis JPA. Meta-analysis for ranked discovery datasets: Theoretical framework and empirical demonstration for microarrays. *Comput Biol Chem*. 2008;32(1):38-46.

doi: 10.1016/j.compbiolchem.2007.09.003

33. Alexa A, Rahnenführer J. topGO: Enrichment Analysis for Gene Ontology. R Package Version 2.46.0. Bioconductor; 2020. Available from: https://bioconductor.org/packages/topGO [Last accessed on 2025 May 01].

34. Rivals I, Personnaz L, Taing L, Potier MC. Enrichment or depletion of a GO category within a class of genes: Which test? *Bioinformatics*. 2007;23(4):401-407.

doi: 10.1093/bioinformatics/btl633

35. Alexa A, Rahnenführer J, Lengauer T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*. 2006;22(13):1600-1607.

doi: 10.1093/bioinformatics/btl140

36. Kumar A, Rajendran V, Sethumadhavan R, Purohit R. Identifying novel oncogenes: A machine learning approach. *Interdiscip Sci*. 2013;5(4):241-246.

doi: 10.1007/s12539-013-0151-3

37. Rajendran V, Sethumadhavan R. Drug resistance mechanism of PncA in *Mycobacterium tuberculosis. J Biomol Struct Dyn*. 2014;32(2):209-221.

doi: 10.1080/07391102.2012.759885

38. Binns D, Dimmer E, Huntley R, Barrell D, O'Donovan C, Apweiler R. QuickGO: A web-based tool for Gene Ontology searching. *Bioinformatics*. 2009;25:3045-3046.

doi: 10.1093/bioinformatics/btp536

39. Gaddy-Kurten D, Tsuchida K, Vale W. Activins and the receptor serine kinase super-family. *Recent Prog Horm Res*. 1995;50:109-129.

doi: 10.1016/b978-0-12-571150-0.50010-x

40. Chen YG, Wang Q, Lin SL, Chang CD, Chuang J, Ying SY. Activin signaling and its role in regulation of cell proliferation, apoptosis, and carcinogenesis. *Exp Biol Med* (*Maywood*). 2006;231(5):534-544.

doi: 10.1177/153537020623100507

41. Reis FM, Cobellis L, Tameirão LC, *et al*. Serum and tissue expression of activin A in postmenopausal women with breast cancer. *J Clin Endocrinol Metab*. 2002;87(5):2277-2282.

doi: 10.1210/jcem.87.5.8512

42. Wang J, Jiang H, Huang Z, *et al*. The expression status of INHBA as a prognostic marker for human breast cancer. *Int J Clin Exp Pathol*. 2016;9(11):11334-11342.

43. Kalli M, Mpekris F, Wong CK, *et al*. Activin A signaling regulates IL13Rα2 expression to promote breast cancer metastasis. *Front Oncol*. 2019;9:32.

doi: 10.3389/fonc.2019.00032

44. Jarde T, Perrier SM, Vasson P, Caldefie-Chezet F. Molecular mechanisms of leptin and adiponectin in breast cancer. *Eur J Cancer*. 2011;47(1):33-43.

doi: 10.1016/j.ejca.2010.09.005

45. Niu J, Jiang L, Guo W, Shao L, Liu Y, Wang L. The association between leptin level and breast cancer: A meta-analysis. *PLoS One*. 2013;8(6):e67349.

doi: 10.1371/journal.pone.0067349

46. Cleveland R, Gammon MD, Long CM, *et al*. Common genetic variations in the LEP and LEPR genes, obesity and breast cancer incidence and survival. *Breast Cancer Res Treat*. 2010;120(3):745-752.

doi: 10.1007/s10549-009-0503-1

47. Ishikawa M, Kitayama J, Nagawa H. Enhanced expression of leptin and leptin receptor (OB-R) in human breast cancer. *Clin Cancer Res*. 2004;10(13):4325-4331.

doi: 10.1158/1078-0432.CCR-03-0749

48. Snoussi K, Strosberg AD, Bouaouina N, Ben Ahmed S, Helal AN, Chouchane L. Leptin and leptin receptor polymorphisms are associated with increased risk and poor prognosis of breast carcinoma. *BMC Cancer*. 2006;6:38.

doi: 10.1186/1471-2407-6-38

49. Teras LR, Goodman M, Patel AV, *et al*. No association between polymorphisms in LEP, LEPR, ADIPOQ, ADIPOR1, or ADIPOR2 and postmenopausal breast cancer risk. *Cancer Epidemiol Biomarkers Prev*. 2009;18(9):2553-2557.

doi: 10.1158/1055-9965.EPI-09-0542

50. Woo HY, Park H, Ki CS, Park YL, Bae WG. Relationships among serum leptin, leptin receptor gene polymorphisms, and breast cancer in Korea. *Cancer Lett*. 2006;237(1):137-142.

doi: 10.1016/j.canlet.2005.05.041

51. Casey T, Bond J, Tighe S, *et al*. Molecular signatures suggest a major role for stromal cells in development of invasive breast cancer. *Breast Cancer Res Treat*. 2009;114(1):47-62.

doi: 10.1007/s10549-008-9982-8

52. Pankov R, Yamada KM. Fibronectin at a glance. *J Cell Sci*. 2002;115(Pt 20):3861-3863.

doi: 10.1242/jcs.00059

53. Wang J, Du Q, Li C. Bioinformatics analysis of gene

expression profiles to identify causal genes in luminal B2 breast cancer. *Oncol Lett.* 2017;14(6):7880-7888.

doi: 10.3892/ol.2017.7256

54. Liu X, Ma Y, Yang W, Wu X, Jiang L, Chen X. Identification of therapeutic targets for breast cancer using biological informatics methods. *Mol Med Rep.* 2015;12(2):1789-1795.

doi: 10.3892/mmr.2015.3565

55. Kondo N, Toyama T, Sugiura H, Fujii Y, Yamashita H. miR-206 expression is down-regulated in estrogen receptor alpha-positive human breast cancer. *Cancer Res.* 2008;68(13):5004-5008.

doi: 10.1158/0008-5472.CAN-08-0180

56. Adams BD, Claffey KP, White BA. Argonaute-2 expression is regulated by epidermal growth factor receptor and MAP kinase signaling and correlates with a transformed phenotype in breast cancer cells. *Endocrinology*. 2009;150(1):14-23.

doi: 10.1210/en.2008-0984

57. Malumbres M, Barbacid M. Cell cycle, CDKs and cancer: A changing paradigm. *Nat Rev Cancer.* 2009;9:153-166.

doi: 10.1038/nrc2602

58. Morgan DO. Principles of CDK regulation. *Nature*. 1995;374:131-134.

doi: 10.1038/374131a0

59. Vassilev LT, Tovar C, Chen S, *et al*. Selective small-molecule inhibitor reveals critical mitotic functions of human CDK1. *Proc Natl Acad Sci U S A*. 2006;103(28):10660-10665.

doi: 10.1073/pnas.0600447103

60. Finch PW, He X, Kelley MJ, *et al*. Purification and molecular cloning of a secreted, frizzled-related antagonist of Wnt action. *Proc Natl Acad Sci U S A*. 1997;94(13):6770-6775.

doi: 10.1073/pnas.94.13.6770

61. Veeck J, Niederacher D, An H, *et al*. Aberrant methylation of the Wnt antagonist SFRP1 in breast cancer is associated with unfavourable prognosis. *Oncogene*. 2006;25(24):3479-3488.

doi: 10.1038/sj.onc.1209386

62. Bernemann C, Hülsewig C, Ruckert C, *et al.* Influence of secreted frizzled receptor protein 1 (SFRP1) on neoadjuvant chemotherapy in triple-negative breast cancer does not rely on WNT signaling. *Mol Cancer.* 2014;13:174.

doi: 10.1186/1476-4598-13-174

63. Klopocki E, Kristiansen G, Wild PJ, *et al.* Loss of SFRP1 is associated with breast cancer progression and poor prognosis in early stage tumors. *Int J Oncol*. 2004;25(3):641-649.

doi: 10.3892/ijo.25.3.641

64. Jeong YJ, Jeong HY, Bong JG, Park SH, Oh HK. Low methylation levels of the SFRP1 gene are associated with the basal-like subtype of breast cancer. *Oncol Rep.* 2013;29(5):1946-1954.

doi: 10.3892/or.2013.2335

65. Zhang M, Chen H, Wang M, Bai F, Wu K. Bioinformatics analysis of prognostic significance of COL10A1 in breast cancer. *Biosci Rep.* 2020;40(2):BSR20193286.

doi: 10.1042/BSR20193286

66. Brodsky AS, Xiong J, Yang D, *et al*. Identification of stromal COLXα1 and tumor-infiltrating lymphocytes as putative predictive markers of neoadjuvant therapy in ER+/HER2+ breast cancer. *BMC Cancer.* 2016;16:274.

doi: 10.1186/s12885-016-2302-5

67. Chapman KB, Prendes MJ, Sternberg H, *et al*. COL10A1 expression is elevated in diverse solid tumor types and is associated with tumor vasculature. *Future Oncol*. 2012;8(8):1031-1040.

doi: 10.2217/fon.12.79

68. Huang H, Li T, Ye G, *et al*. High expression of COL10A1 is associated with poor prognosis in colorectal cancer. *Onco Targets Ther.* 2018;11:1571-1581.

doi: 10.2147/OTT.S160196

69. Li T, Huang H, Shi G, *et al.* TGF-β1–SOX9 axis–inducible COL10A1 promotes invasion and metastasis in gastric cancer via epithelial–mesenchymal transition. *Cell Death Dis*. 2018;9(9):849.

doi: 10.1038/s41419-018-0877-2

70. Giussani M, Landoni E, Merlino G, *et al*. Extracellular matrix proteins as diagnostic markers of breast carcinoma. *J Cell Physiol.* 2018;233(8):6280-6290.

doi: 10.1002/jcp.26513

71. Turashvili G, Bouchal J, Baumforth K, *et al*. Novel markers for differentiation of lobular and ductal invasive breast carcinomas by laser microdissection and microarray analysis. *BMC Cancer*. 2007;7:55.

doi: 10.1186/1471-2407-7-55

72. Kim H, Watkinson J, Varadan V, Anastassiou D. Multi-cancer computational analysis reveals invasion-associated variant of desmoplastic reaction involving INHBA, THBS2 and COL11A1. *BMC Med Genomics*. 2010;3:51.

doi: 10.1186/1755-8794-3-51

73. Wang JC. Cellular roles of DNA topoisomerases: A molecular perspective. *Nat Rev Mol Cell Biol.* 2002;3(6):430-440.

doi: 10.1038/nrm831

74. Chen T, Sun Y, Ji P, Kopetz S, Zhang W. Topoisomerase IIα in chromosome instability and personalized cancer therapy. *Oncogene*. 2015;34(31):4019-4031.

doi: 10.1038/onc.2014.332

75. Tewey KM, Rowe TC, Yang L, Halligan BD, Liu LF.

Adriamycin-induced DNA damage mediated by mammalian DNA topoisomerase II. *Science*. 1984;226(4673):466-468.

doi: 10.1126/science.6093249

76. Mueller RE, Parkes RK, Andrulis I, O'Malley FP. Amplification of the TOP2A gene does not predict high levels of topoisomerase II alpha protein in human breast tumor samples. *Genes Chromosomes Cancer.* 2004;39(4):288-297.

    doi: 10.1002/gcc.20008

77. Ali Y, Abd Hamid S. Human topoisomerase II alpha as a prognostic biomarker in cancer chemotherapy. *Tumour Biol.* 2016;37(1):47-55.

    doi: 10.1007/s13277-015-4270-9

78. Martin M, Romero A, Cheang MC, *et al*. Genomic predictors of response to doxorubicin versus docetaxel in primary breast cancer. *Breast Cancer Res Treat*. 2011;128(1):127-136.

    doi: 10.1007/s10549-011-1461-y_xffff_

79. Nakagawa M, Bando Y, Nagao T, *et al*. Expression of p53, Ki-67, E-cadherin, N-cadherin and TOP2A in triple-negative breast cancer. *Anticancer Res*. 2011;31(6):2389-2393.

80. Rody A, Karn T, Ruckhäberle E, *et al.* Gene expression of topoisomerase II alpha (TOP2A) by microarray analysis is highly prognostic in estrogen receptor-positive breast cancer. *Breast Cancer Res Treat.* 2009;113(3):457-466.

    doi: 10.1007/s10549-008-9964-x

81. Hua W, Sa KD, Zhang X, *et al*. MicroRNA-139 suppresses proliferation in luminal-type breast cancer cells by targeting topoisomerase II alpha. *Biochem Biophys Res Commun*. 2015;463(4):1077-1083.

    doi: 10.1016/j.bbrc.2015.06.061

82. Naghshvar F, Torabizadeh Z, Charati JY, Akbarnezhad M. Relationship between MMP-11 expression in invasive ductal breast. *Middle East J Cancer.* 2017;8(2):69-75.

83. Zhang X, Huang S, Guo J, *et al*. Insights into the distinct roles of MMP-11 in tumor biology and future therapeutics. *Int J Oncol*. 2016;48(5):1783-1793.

    doi: 10.3892/ijo.2016.3400

84. González de Vega R, Clases D, Fernández-Sánchez ML, *et al*. MMP-11 as a biomarker for metastatic breast cancer by immunohistochemical-assisted imaging mass spectrometry. *Anal Bioanal Chem*. 2019;411(3):639-646.

    doi: 10.1007/s00216-018-1365-3

85. Roscilli G, Cappelletti M, De Vitis C, *et al*. Circulating MMP-11 and specific antibody immune response in breast and prostate cancer patients. *J Transl Med*. 2014;12:54.

    doi: 10.1186/1479-5876-12-54

86. Donato R. S100: A multigenic family of calciummodulated proteins of the EFhand type with intracellular and extracellular functional roles. *Int J Biochem Cell Biol.* 2001;33(7):637-668.

    doi: 10.1016/S1357-2725(01)00046-2

87. Arumugam T, Simeone DM, Van Golen K, Logsdon CD. S100P promotes pancreatic cancer growth, survival, and invasion. *Clin Cancer Res*. 2005;11(15):5356-5364.

    doi: 10.1158/1078-0432.CCR-05-0178

88. Gibadulinova A, Barathova M, Kopacek J, *et al*. Expression of S100P protein correlates with and contributes to the tumorigenic capacity of HeLa cervical carcinoma cells. *Oncol Rep.* 2005;14(2):575-582.

    doi: 10.3892/or.14.2.575

89. Schor AP, Carvalho FM, Kemp C, Silva ID, Russo J. S100P calcium-binding protein expression is associated with high-risk proliferative lesions of the breast. *Oncol Rep.* 2006;15(1):3-6.

    doi: 10.3892/or.15.1.3

90. Guerreiro da Silva D, Hu YF, Russo IH, *et al*. S100P calciumbinding protein overexpression is associated with immortalization of human breast epithelial cells *in vitro* and early stages of breast cancer development *in vivo*. *Int J Oncol.* 2000;16(2):231-240.

    doi: 10.3892/ijo.16.2.231

91. Yang R, Stöcker S, Schott S, *et al*. The association between breast cancer and S100P methylation in peripheral blood by multicenter case-control studies. *Carcinogenesis*. 2017;38(3):312-320.

    doi: 10.1093/carcin/bgx004

92. Jelski W, Chrostek L, Szmitkowski M, Markiewicz W. The activity of class I, II, III and IV alcohol dehydrogenase isoenzymes and aldehyde dehydrogenase in breast cancer. *Clin Exp Med.* 2006;6(2):89-93.

    doi: 10.1007/s10238-006-0101-z

93. Guo E, Wei H, Liao X, Xu Y, Li S, Zeng X. Prognostic value of alcohol dehydrogenase mRNA expression in gastric cancer. *Oncol Lett*. 2018;15(4):5505-5516.

    doi: 10.3892/ol.2018.8007

94. Jain PB, Guerreiro PS, Canato S, Janody F. The spectraplakin Dystonin antagonizes YAP activity and suppresses tumourigenesis. *Sci Rep.* 2019;9(1):19843.

    doi: 10.1038/s41598-019-56296-z

95. Goodwin PJ, Ennis M, Pritchard KI, *et al*. Fasting insulin and outcome in early-stage breast cancer: Results of a prospective cohort study. *J Clin Oncol*. 2002;20(1):42-51.

    doi: 10.1200/JCO.2002.20.1.42

96. Cyr-Depauw C, Northey JJ, Tabariès S, *et al*. Chordin-like 1 suppresses bone morphogenetic protein 4-induced breast cancer cell migration and invasion. *Mol Cell Biol*. 2016;36(10):1509-1525.

doi: 10.1128/MCB.00600-15

97. Mishima Y, Jayasinghe CD, Lu K, *et al*. Nucleosome compaction facilitates HP1γ binding to methylated H3K9. *Nucleic Acids Res*. 2015;43(21):10200-10212.

    doi: 10.1093/nar/gkv841

98. Singh PB. HP1 proteins-what is the essential interaction? *Genetika*. 2010;46(10):1424-1429.

    doi: 10.1134/S1022795410100297

99. Jones DO, Mattei MG, Horsley D, Cowell IG, Singh PB. The gene and pseudogenes of Cbx3/mHP1γ. *DNA Seq*. 2001;12(3):147-160.

    doi: 10.3109/10425170109080769

100. Minc E, Courvalin JC, Buendia B. HP1γ associates with euchromatin and heterochromatin in mammalian nuclei and chromosomes. *Cytogenet Cell Genet*. 2000;90(3-4):279-284.

    doi: 10.1159/000056789

101. Huang C, Su T, Xue Y, *et al*. Cbx3 maintains lineage specificity during neural differentiation. *Genes Dev*. 2017;31(3):241-246.

    doi: 10.1101/gad.292169.116

102. Bot C, Pfeiffer A, Giordano F, Manjeera DE, Dantuma NP, Ström L. Independent mechanisms recruit the cohesin loader protein NIPBL to sites of DNA damage. *J Cell Sci*. 2017;130(6):1134-1146.

    doi: 10.1242/jcs.197236

103. Velez G, Lin M, Christensen T, Faubion WA, Lomberk G, Urrutia R. Evidence supporting a critical contribution of intrinsically disordered regions to the biochemical behavior of full-length human HP1γ. *J Mol Model*. 2016;22(1):12.

    doi: 10.1007/s00894-015-2874-z

104. Morikawa K, Ikeda N, Hisatome I, Shirayoshi Y. Heterochromatin protein 1γ overexpression in P19 embryonal carcinoma cells elicits spontaneous differentiation into the three germ layers. *Biochem Biophys Res Commun*. 2013;431(2):225-231.

    doi: 10.1016/j.bbrc.2012.12.128

105. Thorne JL, Ouboussad L, Lefevre PF. Heterochromatin protein 1 gamma and IκB kinase alpha interdependence during tumour necrosis factor gene transcription elongation in activated macrophages. *Nucleic Acids Res*. 2012;40(16):7676-7689.

    doi: 10.1093/nar/gks509

106. Smallwood A, Hon GC, Jin F, Henry RE, Espinosa JM, Ren B. CBX3 regulates efficient RNA processing genomewide. *Genome Res*. 2012;22(8):1426-1436.

    doi: 10.1101/gr.124818.111

107. Ruan J, Ouyang H, Amaya MF, *et al*. Structural basis of the chromodomain of Cbx3 bound to methylated peptides from histone H1 and G9a. *PLoS One*. 2012;7(4):e35376.

    doi: 10.1371/journal.pone.0035376

108. Abe K, Naruse C, Kato T, Nishiuchi T, Saitou M, Asano M. Loss of heterochromatin protein 1 gamma reduces the number of primordial germ cells via impaired cell cycle progression in mice. *Biol Reprod*. 2011;85(5):1013-1024.

    doi: 10.1095/biolreprod.111.091512

109. Liang YK, Lin HY, Chen CF, Zeng D. Prognostic values of distinct CBX family members in breast cancer. *Oncotarget*. 2017;8(54):92375-92387.

    doi: 10.18632/oncotarget.21325

110. Wang Z, Potter CS, Sundberg JP, Hogenesch H. SHARPIN is a key regulator of immune and inflammatory responses. *J Cell Mol Med*. 2012;16(10):2271-2279.

    doi: 10.1111/j.1582-4934.2012.01574.x

111. Tokunaga F, Iwai K. LUBAC, a novel ubiquitin ligase for linear ubiquitination, is crucial for inflammation and immune responses. *Microbes Infect*. 2012;14(7-8):563-572.

    doi: 10.1016/j.micinf.2012.01.011

112. Kasirer-Friede A, Tjahjono W, Eto K, Shattil SJ. SHARPIN at the nexus of integrin, immune, and inflammatory signaling in human platelets. *Proc Natl Acad Sci U S A*. 2019;116(11):4983-4988.

    doi: 10.1073/pnas.1819156116

113. McGowan HW, Schuijers JA, Grills BL, *et al*. Sharpin is a key regulator of skeletal homeostasis in a TNF dependent manner. *J Musculoskelet Neuronal Interact*. 2014;14(4):454-463.

114. Zhuang T, Yu S, Zhang L, *et al*. SHARPIN stabilizes estrogen receptor alpha and promotes breast cancer cell proliferation. *Oncotarget*. 2017;8(44):77137-77151.

    doi: 10.18632/oncotarget.20368

115. Yang H, Yu S, Wang W, *et al*. SHARPIN facilitates p53 degradation in breast cancer cells. *Neoplasia*. 2017;19(2):84-92.

    doi: 10.1016/j.neo.2016.12.002

116. De Melo J, Tang D. Elevation of SIPL1 (SHARPIN) increases breast cancer risk. *PLoS One*. 2015;10(5):e0127546.

    doi: 10.1371/journal.pone.0127546

117. Bii VM, Rae DT, Trobridge GD. A novel gammaretroviral shuttle vector insertional mutagenesis screen identifies SHARPIN as a breast cancer metastasis gene and prognostic biomarker. *Oncotarget*. 2015;6(37):39507-39520.

    doi: 10.18632/oncotarget.6232

118. Kao KJ, Chang KM, Hsu HC, Huang AT. Correlation of microarray-based breast cancer molecular subtypes and clinical outcomes: Implications for treatment optimization. *BMC Cancer*. 2011;11:143.

    doi: 10.1186/1471-2407-11-143

119. Bosco EE, Mulloy JC, Zheng Y. Rac1 GTPase: A “Rac” of all

trades. *Cell Mol Life Sci*. 2009;66(3):370-394.

doi: 10.1007/s00018-008-8552-x

120. Schnelzer A, Prechtel D, Knaus U, *et al*. Rac1 in human breast cancer: Overexpression, mutation analysis, and characterization of a new isoform, Rac1b. *Oncogene*. 2000;19(26):3013-3020.

doi: 10.1038/sj.onc.1203621

121. Bid HK, Roberts RD, Manchanda PK, Houghton PJ. RAC1: An emerging therapeutic option for targeting cancer angiogenesis and metastasis. *Mol Cancer Ther.* 2013;12(10):1925-1934.

doi: 10.1158/1535-7163.MCT-13-0164

122. Wetterskog D, Shiu KK, Chong I, *et al*. Identification of novel determinants of resistance to lapatinib in ERBB2-amplified cancers. *Oncogene*. 2014;33(8):966-976.

doi: 10.1038/onc.2013.41

123. Skvortsov S, Dudás J, Eichberger P, *et al*. Rac1 as a potential therapeutic target for chemoradioresistant head and neck squamous cell carcinomas. *Br J Cancer*. 2014;110(11):2677-2687.

doi: 10.1038/bjc.2014.221

124. Hein AL, Post CM, Sheinin YM, *et al*. RAC1 GTPase promotes the survival of breast cancer cells in response to hyperfractionated radiation treatment. *Oncogene*. 2016;35(49):6319-6329.

doi: 10.1038/onc.2016.163

125. Li FQ, Mofunanya A, Harris K, Takemaru KI. Chibby cooperates with 1433 to regulate βcatenin subcellular distribution and signaling activity. *J Cell Biol.* 2008;181(7):1141-1154.

doi: 10.1083/jcb.200709091

126. Wang W, Zhang L, Wang Y, *et al*. Involvement of miR451 in resistance to paclitaxel by regulating YWHAZ in breast cancer. *Cell Death Dis*. 2017;8(10):e3071.

doi: 10.1038/cddis.2017.460

127. Neal CL, Yao J, Yang WT, Zhou XY, Nguyen NT, Lu J. 1433ζ overexpression defines high risk for breast cancer recurrence and promotes cancer cell survival. *Cancer Res*. 2009;69(8):3425-3432.

doi: 10.1158/0008-5472.CAN-08-2765

128. Bergamaschi A, Katzenellenbogen BS. Tamoxifen downregulation of miR451 increases 1433ζ and promotes breast cancer cell survival and endocrine resistance. *Oncogene*. 2012;31(1):39-47.

doi: 10.1038/onc.2011.223

129. Fang Y, Shen H, Cao Y, *et al*. Involvement of miR30c in resistance to doxorubicin by regulating YWHAZ in breast cancer cells. *Braz J Med Biol Res*. 2014;47(1):60-69.

doi: 10.1590/1414-431X20133324

130. Xu W, Zhu H, Gu M, *et al*. DHTKD1 is essential for mitochondrial biogenesis and function maintenance. *FEBS Lett.* 2013;587(21):3587-3592.

doi: 10.1016/j.febslet.2013.08.047

131. Stark R, Roden M. ESCI award 2006. Mitochondrial function and endocrine diseases. *Eur J Clin Invest.* 2007;37(4):236-248.

doi: 10.1111/j.1365-2362.2007.01773.x

132. DiMauro S. A history of mitochondrial diseases. *J Inherit Metab Dis.* 2011;34(2):261-276.

doi: 10.1007/s10545-010-9082-x

133. Ježek P, PlecitáHlavatá L, Smolkova K, Rossignol R. Distinctions and similarities of cell bioenergetics and the role of mitochondria in hypoxia, cancer, and embryonic development. *Int J Biochem Cell Biol*. 2010;42(5):604-622.

doi: 10.1016/j.biocel.2009.11.008

134. Dong H, Zhang S, Wei Y, *et al.* Bioinformatic analysis of differential expression and core genes in breast cancer. *Int J Clin Exp Pathol*. 2018;11(3):1146-1156.

135. Zhong X, Kan A, Zhang W, *et al.* CBX3/HP1γ promotes tumor proliferation and predicts poor survival in hepatocellular carcinoma. *Aging* (*Albany NY*). 2019;11(15):5483-5497.

doi: 10.18632/aging.102132