IJBCS
ISSN: 2147-1428

Original Article

International Journal of Basic and Clinical Studies (IJBCS)

2022; 11(2): 67-74 Elgun T. and Sarac T.

Example of Bioinformatic Analysis in Breast Cancer: Comparison of Invasive Ductal Carcinoma and Invasive Lobular Carcinoma for The Definition of New Biomarkers

Tuğba ELGÜN[1] , Tuba SARAÇ[2]

[1] Biruni University, Medical School, Department of Medical Biyology, Istanbul, Turkey
[2]Biruni University, Istanbul, Türkiye

Abstract

In our study, it was aimed to determine the gene/genes that can be evaluated as biomarkers in early diagnosis with bioinformatic analysis in Invasive ductal carcinoma (IDC) and Invasive lobular carcinoma (ILC), which are frequently encountered breast cancer types.

IDC is the most common subtype, 80% of all breast cancers. Invasive Ductal means that cancer has invaded or spread to the surrounding breast tissues from the ductal wall. ILC is around 10% of all breast cancers. In this type, the cancer is surrounding the milk duct, but their shapes are not deformed. This is why it is very difficult to identify in mamammographyGenetic tests and their analysis is very important for the early detection of cancer. Physicians can estimate bad prognoses and can develop better treatment plans. 17 ILC and 83 IDC patients' data were gathered from online databases. GEO, Array express, SRA. We developed software using R Bioconductor for differential expression analysis of the data. Our analysis picked TFF3, MMP9, DUSP1, SCGB2A2, CTHRC1, APOD, TGFB3, NMU, IGFBP1, CD34 genes as differentially expressed. The relations among genes are demonstrated using the online STRING tool. The threshold for the expression levels is selected in terms of Log-fold change (LogFC). The down limit is LogFC -0.9; the down limit is logFC 1.693. According to the STRING analysis, 4 genes were strongly linked; MM9, CD34, IGFBP1 and CTHRC1. CD34 is upregulated in ILC (logFC 1.693, $p<0.05$) and downregulated in IDC (logFC -0.9, $p<0.05$). As a result of our study, we suggest that CD34 gene expression should be primarily evaluated in distinguishing between ILC and IDC.

Keywords: Invasive ductal carcinoma, Invasive lobular carcinoma, bioinformatic analysis, biomarker

Correspondence: Dr. Tuğba ELGÜN, Biruni University, Department of Medical Biyology, İstanbul, Türkiye telgun@biruni.edu.tr

IJBCS
ISSN: 2147-1428

Original Article

International Journal of Basic and Clinical Studies (IJBCS)

2022; 11(2): 67-74 Elgun T. and Sarac T.

Introduction

Every day, many new data, studies and articles take their place in scientific databases. Thus, a very large and very large database is formed (1). Bioinformatics, which is an interdisciplinary science, provides the opportunity to organize and analyze existing data by developing techniques for storing biological data and obtaining information from the database (2).

Breast cancer is a disease process that occurs with the uncontrolled and rapid proliferation of the cells lining the milk glands and milk ducts in the breast due to various pathological reasons, as well as the spread (metastasis) to the surrounding tissues (invasion) and organs in other parts of the body. Invasive Ductal means that cancer has invaded or spread to the surrounding breast tissues from the ductal wall. In the ILC type, the cancer is surrounding the milk duct, but their shapes are not deformed (3). This is why it is very difficult to identify in mammography. In early-stage cancers, genetic testing is very important. Physicians can estimate bad prognoses and can develop better treatment plans (4).

Breast cancers have a rate of approximately 30% among all cancers. It is divided into invasive and non-invasive. The ratio of invasive cancers to all breast cancers is approximately 80%. There are 2 types of breast cancer: Lobular cancer (ILC) which develops from the milk-secreting part and ductal cancer (IDC) which develops from the milk ducts (4,5). Nearly 85% of invasive cancers are IDC, and close to 10-14% are ILC. ILC differs from IDC in terms of its pathological and clinical features (6).

Genetic diagnosis is important in identifying and categorizing breast cancer at an early stage. After genetic diagnosis, bioinformatic analysis and evaluation, and clinical interpretation are of great importance in the treatment process (7). In our study, it was aimed to determine the gene/genes that can be evaluated as biomarkers in early diagnosis with bioinformatics analysis in Invasive ductal carcinoma (IDC) and Invasive lobular carcinoma (ILC), which are frequently encountered breast cancer types (8).

IJBCS
ISSN: 2147-1428

Original Article

International Journal of Basic and Clinical Studies (IJBCS)

2022; 11(2): 67-74 Elgun T. and Sarac T.

Method

In our study, 17 ILC and 83 IDC patients' data was gathered from online databases: GEO, Arrayexpress, and SRA. We developed Differential Expression Analysis software using R Bioconductor for the identification of up-and down-regulated and differentially expressed gegenesens from our gathered array data. Our analysis successfully identified TFF3, MMP9, DUSP1, SCGB2A2, CTHRC1, APOD, TGFB3, NMU, IGFBP1, CD34 genes as differentially expressed. The relations among these genes are investigated and demonstrated using the STRING: functional protein association networks online database tool.

According to the literature annotations and the gene expression data analyzed in our framework study, the direct or indirect effects were revealed. For a successful classification, the a statistically significant threshold value for the expression levels was picked in terms of Log-fold change (LogFC). Log-fold change (LogFC) values were taken into account while determining the exponential expression levels. Negative FC values mean down regulations and positive FC values mean up regulations of the selected genes. LogFC values lower than -0.9 and higher than 1.693 were assigned.

Results

According to the STRING analysis, 4 genes were strongly linked, MM9, CD34, IGFBP1 and CTHRC1 (Figure 1). In the bioinformatics analysis we performed to distinguish between ILC and IDC, it was determined that TFF3, MMP9, DUSP1, SCGB2A2, CTHRC1, APOD, TGFB3, NMU, IGFBP1 genes did not make a significant difference in terms of expression. Unlike these genes, CD34 gene expression increased in ILC (logFC 1.693, $p<0.05$); Expression was decreased in IDC (logFC −0.9, $p<0.05$).

IJBCS
ISSN: 2147-1428

Original Article

International Journal of Basic and Clinical Studies (IJBCS)

2022; 11(2): 67-74 **Elgun T. and Sarac T.**

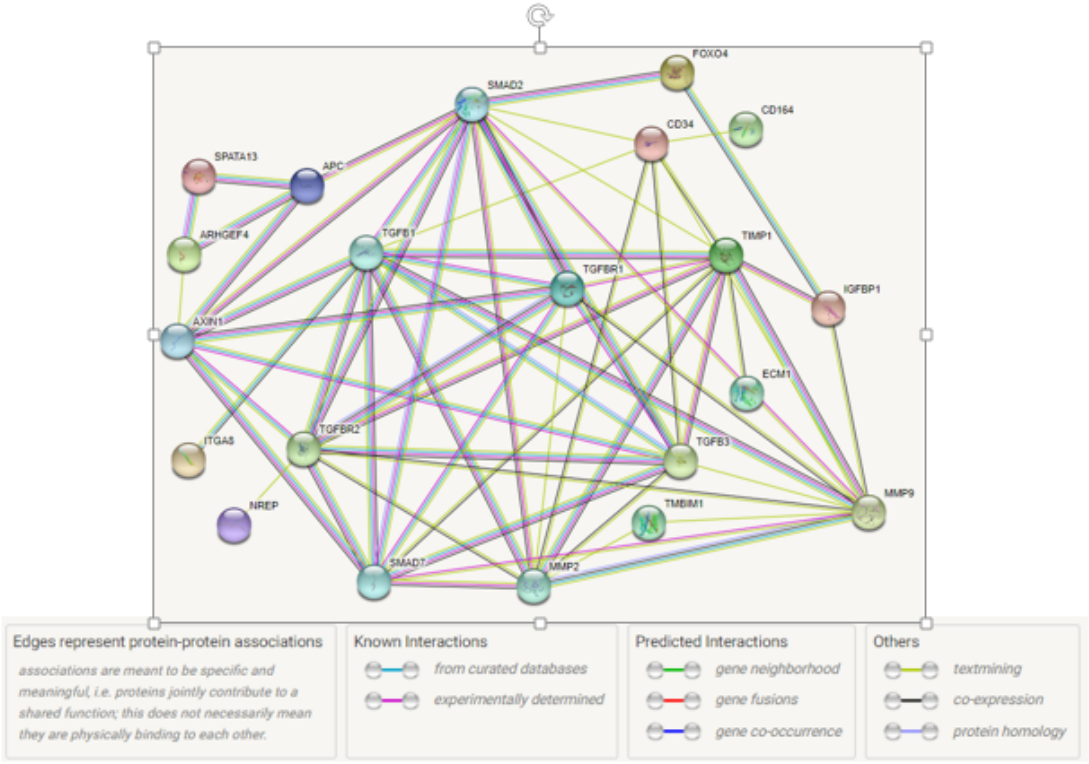

Figure 1: Analysis of Protein-Protein Relationship With String Gene

As a result of our study, we suggest that CD34 gene expression should be primarily evaluated in distinguishing between ILC and IDC (Table 1) (Figüre 2).

IJBCS
ISSN: 2147-1428

Original Article

International Journal of Basic and Clinical Studies (IJBCS)

2022; 11(2): 67-74 Elgun T. and Sarac T.

Table 1: Evaluation of gene regulation of ILC cancers with IDC.

GenBank Access	GenBank Access	Gene	Gene Name
ILC	**IDC**		
NM_003226	**NM_003226**	TFF3	Trefoil Factor 3
NM_004994	**NM_004994**	MMP9	Matrix Metallopeptidase 9
NM_004417	**NM_004417**	DUSP1	Dual Specificity Phosphatase 1
NM_002411	**NM_002411**	SCGB2A2	Secretoglobin Family 2A Member 2
BC021025	**BC021025**	CTHRC1	Collagen Triple Helix Repeat Containing 1
NM_001647	**NM_001647**	APOD	Apolipoprotein D
NM_003239	**NM_003239**	TGFB3	Transforming Growth Factor Beta 3
NM_006681	NM_006681	NMU	Neuromedin U
NM_000596	NM_000596	IGFBP1	Insulin Like Growth Factor Binding Protein 1
BX640941	BX640941	**CD34***	**CD34 molecule**
UP-REGULATION DOWN REGULATION			
*p<0.05			

IJBCS
ISSN: 2147-1428

Original Article

International Journal of Basic and Clinical Studies (IJBCS)

2022; 11(2): 67-74 Elgun T. and Sarac T.

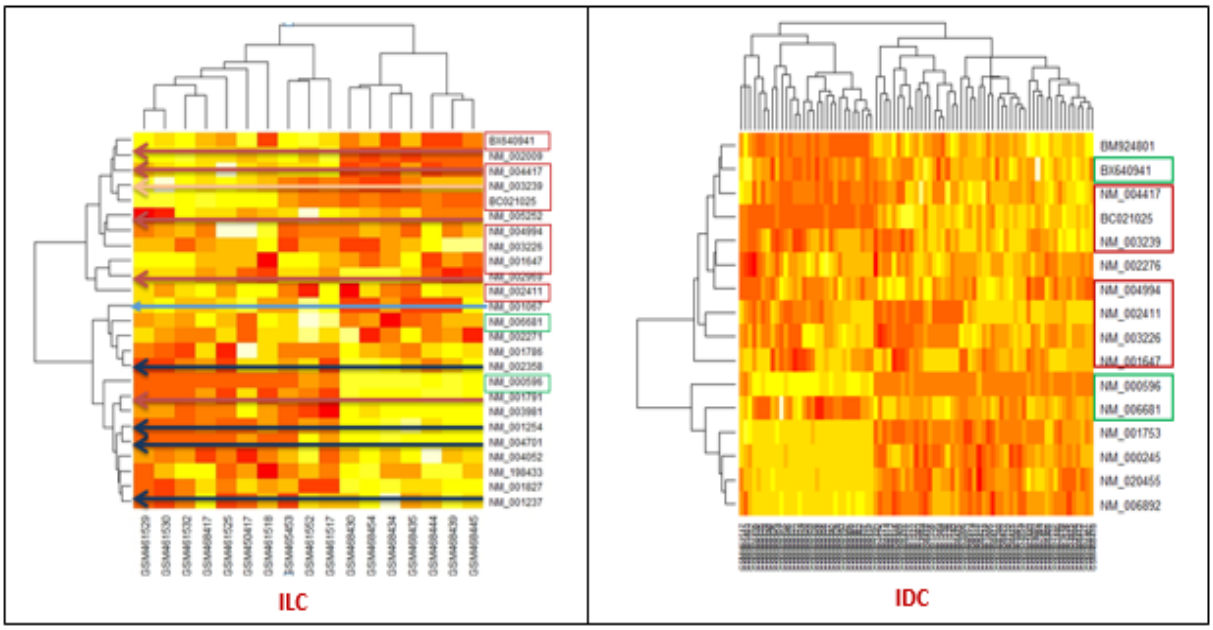

Figure 2: Associating genes with Heat-Map

Discussion

Significant differences between ILC and IDC are shown. E-cadherin (CDH1) protein loss, which allows cells to adhere to each other, occurs in almost all ILC cases (9). Because of the problem of interconnection between cells, tumors are arranged as strings, whereas in IDC, a rounded tumor is formed due to the tightening of cells (10). Another difference shown is that ILC tumor cells do not use as much sugar as IDC tumors. Therefore, the proliferation of ILC cells is slower. In addition, almost all of the cells in the ILC are ER-2 positive and thus they respond better to hormonal treatments. However, It has been shown that the recurrence rate of ILC is higher (8).

CD34 is a transmembrane glycoprotein involved in the modulation of cell adhesion and signal transduction and is thought to be expressed by mesenchymal cells at various sites, including the normal breast stroma. Most benign breast lesions with invasive carcinoma of the breast express large numbers of CD34-positive stromal cells in the stroma. Matrix metalloproteinases (MMPs) are a family of zinc-linked endopeptidases that degrade the ECM. These proteases cleave collagen from the epithelium and vascular basement membrane and cause the migration of tumor cells. For IDC and ILC, a significant effect of the MMP9 gene on the expression level

was not demonstrated in our study, as stated in the literature. In our study, no significant difference was found in the expression levels of MMP9 and CTHRC1 genes for IDC and ILC. As a result of the evaluation, CD34 is frequently used as a biomarker in breast cancers, supporting the literature on marker research related to IDC and ILC. Lack of CD34 gene expression in the cell causes myofibrosarcoma in the breast (11).

Again, in similar studies in the literature, differences in the expression of GATA3, FOXA1, and PTEN genes were observed between ILC and IDC. At the same time, increased expression of the IGF-1R receptor gene was detected in ILC cells. High expression of the FGFR4 gene has been observed in ILC resistant to hormonal therapy. Studies have shown that mutations of Her2/Her3 genes are also higher in ILC (12).

References

1. R.L. Siegel, K.D. Miller, H.E. Fuchs and A. Jemal, Can553 cer statistics, 2021, CA: A Cancer Journal for Clinicians 71 554 (2021), 7–33. 555 [2]
2. H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjo556 mataram, A. Jemal and F. Bray, Global cancer statistics 2020: 557 GLOBOCAN estimates of incidence and mortality worldwide 558 for 36 cancers in 185 countries, CA: A Cancer Journal for 559 Clinicians 71 (2021), 209–249.
3. A.G. Waks and E.P. Winer, Breast cancer treatment: A review, 568 Jama 321 (2019), 288–300.
4. Kaçar A , Paker İ , Akbıyık F , Arıkök AT , Mambet A; Cd117 and Cd34 Staining Patterns in Childhood Benign Mammary Lesions; Vol. 28, No. 1, 2012; Page 31-37; doi: 10.5146/tjpath.2012.01094.
5. Yayoi Adachi et al. Comparison of clinical outcomes between luminal invasive ductal carcinoma and luminal invasive lobular carcinoma. BMC Cancer (2016) 16:248 DOI 10.1186/s12885-016-2275-4.
6. Romualdo Barroso-Sousa and Otto Metzger-Filho, Differences between invasive lobular and invasive ductal carcinoma of the breast: results and therapeutic implications; Ther Adv Med Oncol. 2016 Jul; 8(4): 261–266.
7. Oh CD, Maity SN, Lu JF, Zhang J, Liang S, Coustry F, de Crombrugghe B, Yasuda H PLoS One. Identification of SOX9 interaction sites in the genome of chondrocytes 5(4):e10113 (2010).
8. Rastegar F et al. Mesenchymal stem cells: Molecular characteristics and clinical applications; World J Stem Cells. 2(4):67-80 (2010).

9. Bajrami I et al. E-cadherin/ROS1 inhibitor synthetic lethality in breast cancer. Jul 2018. Cancer Research 78 (13). DOI10.1158/2159-8290.CD-17-0603.

10. Siegel R, Ma J, Zou Z, Jemal A; Cancer statistics, 2014; CA: a Cancer Journal for Clinicians, 07 Jan 2014, 64(1):9-29, DOI: 10.3322/caac.21208, PMID: 24399786

11. Ilirjana Bajrami et al. 2018; E-Cadherin/ROS1 Inhibitor Synthetic Lethality in Breast Cancer; American Association for Cancer Research; DOI: 10.1158/2159-8290.CD-17-0603.

12. Nagle, AM et al. Loss of E-cadherin Enhances IGF1-IGF1R Pathway Activation and Sensitizes Breast Cancers to Anti-IGF1R/InsR Inhibitors. Clin Cancer Res. 2018 Oct 15; 24(20): 5165–5177. doi: 10.1158/1078-0432.CCR-18-0279