Review Article

A Review on Identification of Novel Biomarkers to Identify Genes in Breast Cancer

Nimra Munir[1], Nabila Ghulam Rasool[1], Yasmeen Batool[2], Rubab Yousaf[1*], Taiyyiaba Basharat[3], Sabahat Irfan[1], Hiba Sharif[1], Fazeela Zaka[4]

[1]Department of Biochemistry, University of Agriculture, Faisalabad, Pakistan
[2]Department of Zoology University of Agriculture, Faisalabad, Pakistan
[3]Department of Plant Breeding and Genetics, University of Sargodha, Pakistan
[4]Department of Plant Breeding and Genetics, University of Agriculture, Faisalabad, Pakistan

Abstract

Breast cancer is the most common occurring cancer among women worldwide. Risk factors included drinking lot of alcohol, lack of physical exercise, family history, ionizing radiations and hormones replacement therapy during menupuase. BRAC1 gene play important role in biochemical processes occurring in cells of the breast cancer. It is highly abundant in tumor cells as compared to the normal cells of human body. It makes it strong and ideal candidate for biochemical and functional characterization at molecular level. The molecular pathways associated with tumor formation and progression of breast cancer not completely discovered yet. There is need to design such potential target or drugs that can target the cells of breast cancer and regulate the epithelial mesenchymal transition as well as reduces the number of cases due to breast cancer. Information regarding genes involved molecular pathways remains unclear. There is need to identify genes involved inthe development of breast cancer as well as prognosis. GEO database of NCBI especially used for breast cancer datasets searching.Cytohubba is used to identify the DEGs. Functional enrichment software that is used for classification and characterization of genes at functional level. Biological database such as STRING that designed for to construct PPI networks of DEGs. MCODE in Cytoscape-Intergerated plug in experimentally applied to check the interactions in PPI network. Cytohubba also used to check out the either hub genes functional or target newly drug for the breast cancer treatment. Kaplan–Meier is biological software that is used to check out the survival of individual hub genes subjected to breast cancer. X2K (Expression2Kinases) used to check out the expression of hub genes. The between hub gene and drug especially be designed in database of drug bank.
Keywords: Ctyphuba, X2K, PPI, MCODE, STRING, DEGs, GEO, Hub genes.

INTRODUCTION

Breast cancer is the most common type of malignant tumour among females. In United States, breast cancer is the major cause of death due to genetic mutations in breast and ultimately leads to cancer. Number of cases due to breast cancer increasing day by day has become a global health issue and. In 2017, 268,600 new cases of breast cancer reported and death rate 41,710 in United States [1]. Risk factors included drinking lot of alcohol, lack of physical exercise, family history, ionizing radiations and hormones replacement therapy during menupuase.

Breast cancer is the most common type of malignant tumor. Breast cancer has high rate of prevalence among women due to progression of metastasis. Epithelial cell adhesion molecule (EpCAM) acts as a oncogene which ultimately increases the formation of tumors in breast cells [2]. Role of N-linked glycosylation of epithelial cell adhesion molecule in regulation of epithelial mesenchymal transition (EMT) during pathogenesis of breast cancer remains unclear.

Breast cancer is classified into several types due to differences in molecular characteristics, histological features and outcome clinical features. Molecular classification of breast cancer divided into six major subgroups such as normal like, basal like, HER2-poisitive, claud in low and luminal type A and B. Claud in low characterized due to lack of estrogen receptor and progesterone receptor. HER 2 belong to triple negative breast cancer that has high possibility to

occur again in visceral tissues and cause visceral metastases [3].

Breast cancer is the most common type of malignant tumor among females. Breast cancer has become a global health issue and number of cases due to breast cancer increasing day by day. About 5-10% cases of breast cancer arise due to mutations acquired from their parents. Risk factors included drinking of alcohol, lack of physical exercise, family history, ionizing radiations and hormones replacement therapy during menopause [4].

Breast cancer has high rate of prevalence among women due to progression of metastasis. Biochemical role of epithelial cell adhesion molecule in regulation of epithelial mesenchymal transition during pathogenesis of breast cancer not completely described and remains unclear. There is need to design such potential target or drugs that can target the cells of breast cancer and regulate the EMT as well as reduces the number of cases due to breast cancer [10].

There are certain evidences for the potential roles of BRAC1 genes at molecular level such as RNA synthesis and nuclear exportation but molecular mechanisms remain unclear how this BRAC1 gene is involved. BRAC1 genes also act as nuclear protein that involved in regulation of elongation of transcription. BRAC1 genes biochemically interact with C-terminal of RNA polymerase II FCP1, SPT5 which is nuclear transcription factor, and POLDIP3 which is a Polymerase delta interacting proteins that help in localization of the nuclear regions such as nuclear speckles and SNRPD3 plays a significant role in splicing of mRNA. BRAC1 genes also show interaction with CIZI which is present in the replication fork also playing significant role in DNA replication [6].

The purpose of this review is to study the bioinformatics approach to identify potential crucial genes and molecular pathways of prognosis and tumoerigenis of breast cancer. Association of hub genes with formation of tumors and prognosis of breast cancer.

Diagnostic Biomarkers in Breast Cancer
Epithelial cell adhesion molecule (EpCAM) acts as a oncogene which ultimately increases the formation tumors in breast cell. EpCAM significantly express in epithelial cells as well as in tumor cells [5]. EpCAM acts as a glycoprotein that contains important three sites for N-glycosylation reaction which are experimentally proved the biochemical effects on apoptosis as well as adhesion in cells of breast [6].

Tumor microenvironment (TME) is a huge mass of tumors in which tumor causing cells combine together and play significant role progression of cancers

[7]. Hypoxia leads to metastasis of breast cancer that dependent on of tumor microenvironment [8]. Hypoxia also increases the rate of breast cancer due to invasions of tumor which untimely leads to more formation tumor formation in cell line of breast. There is need to design such potential target or drugs that can target the cells of breast cancer and regulate the EMT as well as reduces the number of cases due to breast cancer [9].

BRAC1 gene play important role in biochemical processes occurring in cells of the breast cancer. Intact, BRAC1 genes composed of amino acid sequence identical in mouse and human and different from other species that is encoded by single amino acid. The biological functions of BRAC1 genes remain unclear and not completely described yet. This BRAC1 gene is highly abundant in tumor cells as compared to the normal cells of human body. It makes it strong and ideal candidate for biochemical and functional characterization at molecular level [8, 9].

BRAC1 gene play important role in biochemical processes occurring in cells of the breast cancer. Intact, BRAC1 genes composed of amino acid sequence identical in mouse and human and different from other species that is encoded by single amino acid. The biological functions of BRAC1 genes remain unclear and not completely described yet. This BRAC1 gene is highly abundant in tumor cells as compared to the normal cells of human body. It makes it strong and ideal candidate for biochemical and functional characterization at molecular level. If there is severe defect found in BRAC1 genes, then CENP-E mRNA incorrectly splicing occurs. It also leads to further defects in genetic defects, as result cell cycle arrested and occurring of other genetic defects. It bind to Mmi1 in EMC to form Erh1-Mmi1 complex [2, 5]. EMC localizes as well as scattered to the nuclear region by interacting with two complexes. One complex is MTREC for Mtl1-Red1 that significantly cooperates with Mmi1 for biological degradation. The second complex is CCR4-NOT that is involved in regulation of heterchromatin also its function. Both these complexes such as MTREC and CCR4-NOT are highly conserved in the cells of breast also interact with these two complexes. It is highly abundant in tumor cells as compared to the normal cells of human body. It makes it strong and ideal candidate for biochemical and functional characterization at molecular level [11, 12].

Breast cancer has rate of prevalence among women due to progression of metastasis. Mechanisms associated with occurrence and progression of breast cancer at molecular level remains unclear. It is critical to correctly identify genes associated with tumor formation and progression of breast cancer at molecular level for purpose of help for identifying new diagnostic biomarkers as well as therapeutic target [12].
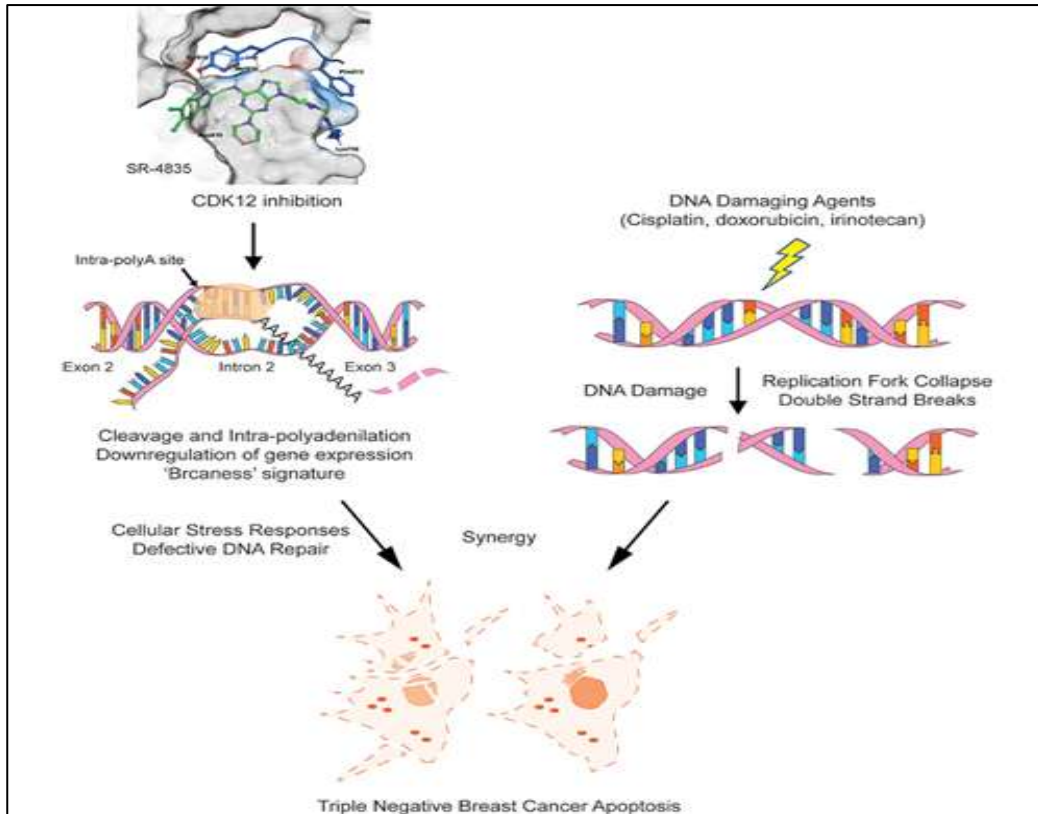
Fig-1: Shows the apoptosis process involved in the breast cancer

Data Retrieval

Triple negative breast cancer classified into basal like immune activated, basal like immune suppressed, luminal androgen receptor and mesenchymal tumour. These subgroups have common activated, repressed genes and signaling pathways. These genes and signaling pathways are involved in formation of tumours and ultimately developing of breast cancer [11].

GEO database of NCBI is used for breast cancer datasets searching [13]. Functional enrichment, detection of hub genes, Identification of DEGs, PPI Network Construction and Module Analysis, Survival Analysis of Hub Genes, Expression Analysis of Hub Genes, hub gene-drug interaction network and statistical analysis especiallybe carried out into following sections. DEGs identify by using Cytohubba. DEGs also characterize using LIMMA by adjusting the p-value for the purpose of reduction in errors bars [14].

Table-1: Risk Factors for the metastasis of the breast cancer

Risk factor	Relative risk	High risk group
Age	>10	Elderly
Reproductive risk factors		
Age at menarche	3	Menarche before age 11
Age at menopause	2	Menopause after age 54
Age at first pregnancy	3	Nulliparous or first child in early 40s
Lifestyle factors		
Diet	1.5	High intake of saturated fat
Body weight (postmenopausal)	2	Body mass index > 35
Alcohol	1.3	Excessive intake
Hormonal status		
Oral contraceptives	1.24	Current use
Hormone replacement therapy	1.35	Use for > 10 years
Radiation	3	Abnormal exposure after age 10
Family history	≥ 2	Breast cancer in first degree relative when young

Functional Enrichment Software

Functional enrichment is software that used for classification and characterization of genes at functional level [20]. Gene Ontology function and Kyoto Encyclopedia of Genes and Genomes pathways enrichment analyses of the DEGs experimentally perform by using the principles of functional enrichment analysis [15]. Biological database such as STRING that designed for to construct PPI networks of DEGs due to known PPIs and analysis especially be performed to check interactions among proteins [16].

MCODE in Cytoscape-Intergerated plug in applied to check the interactions in PPI network. Cytohubbathat used to check out the either hub genes functional or target newly drug for the breast cancer treatment [25]. Kaplan–Meier is a reliable tool that used to check out the survival of individual hub genes in relation to breast cancer. The expression analysis of the hub genes performed to analyze cancer transcriptome data X2K (Expression2Kinases) used to check out the expression of hub [17]. The between hub gene and drug designed in database of drug bank. It is helpful in decreasing or increasing expression of protein expression of the hub genes [18].
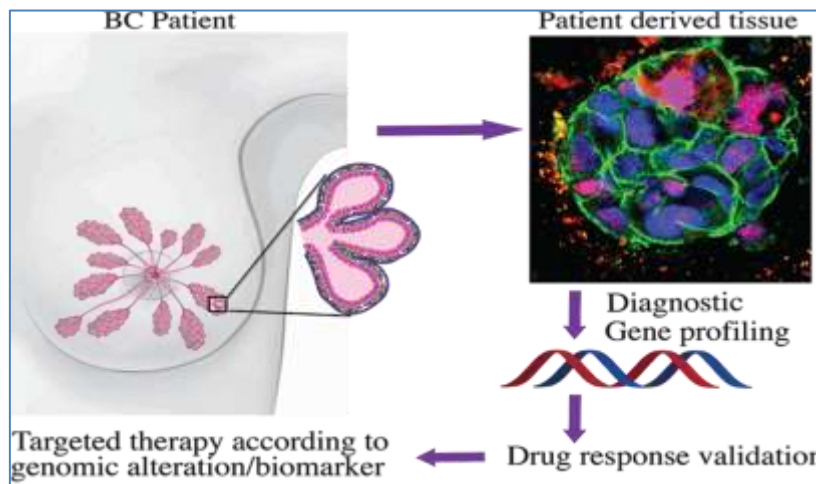

Fig-2: Shows Novel biomarkers for the Prognosis of the breast cancer

Immunoblotting

Immunoblotting analysis that is used to identify the novel biomarkers in the breast cancer genes. This method also determines the presence of the cells that are the cancerous as compared to the other cells [24]. The standard steps in the previous studies remains unclear and new methods protols added. Cells washed with phosphate buffer saline and concentration of proteins determined by using the Coomassie Brilliant Blue G 250. Sodium dodecyl sulfate polyacrylamide gel used to separate the protein and transferred to nitrocellulose membrane. The results finally detected by using the chemiluminescence assays [19].
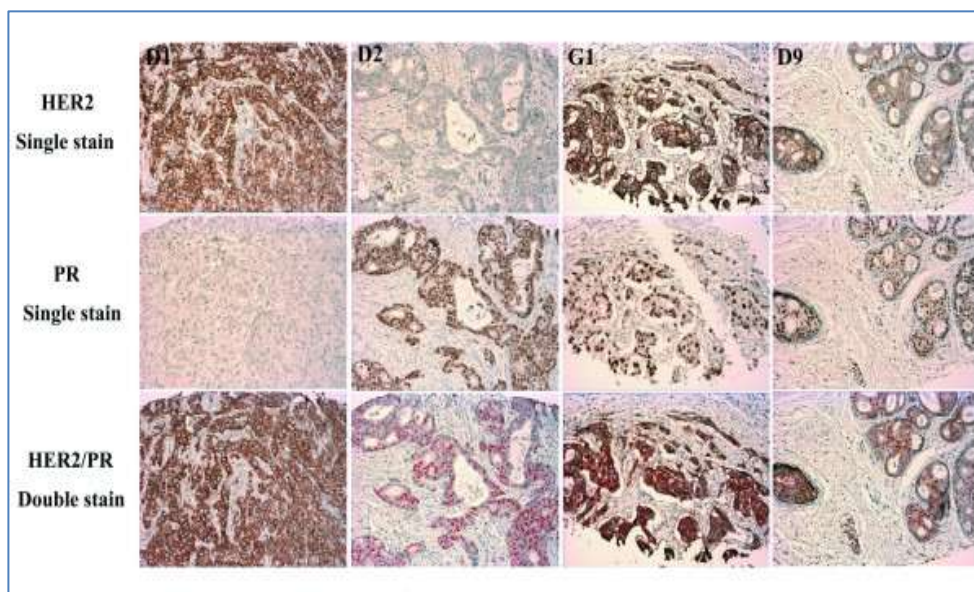

Fig-3: Shows various images Immunoblotting to targeting cells of breast cancer

Immunofluorescence staining that is used to identify the novel biomarkers in the breast cancer genes. This type of staining used now for the discovery of the new biomarkers in the novel cells. Glass coverslips used to cover the cells. 5% formaldehyde especiallybe used to fix the cells for about 25minutes at 37°C. Cells incubated with SOX2 and OCT4 antibodies for a period of 24 hours at 5°C and finally incubating with conjugated secondary antibodies that fluorescence antibodies and staining with using DAPI. Fluorescence microscopy that is specially used measurement of images [20, 24].

Sphere Formation

Sphere Formation Assay is used to identify the novel biomarkers in the breast cancer genes [22, 23]. This method has gained certain advantages over the other methods. Single cell inserted into six well plates and treated with 250μM CoCl2 with sequence of plasmid such as EpCAM-over and EpCAM-mutant[23]. Inverted microscope that is used to determine the number of spheroids and size. It is important during experiment that only single cell be inserted to each well [21].

CONCLUSION

This study especiallybe helpful in better for clear identification of genes associated with tumor formation and progression of breast cancer at molecular level also for indentifying new diagnostic biomarkers as well as therapeutic target. These strategies helpful for the prognosis of breast cancer patiens. This research especiallybe helpful for evaluation on the ability of the EpCAM acting as a therapeutic target for breast cancer in future studies. This study especiallly be also helpful in regulating of expression of epithelial mesenchymal transition at genetic level in future studies.

REFERENCES

1. Bader, G.D., & Hogue, C.W. (2003). An automated method for finding molecular complexes in large protein interaction networks. BMC bioinformatics, 4: 2.
2. Chandrashekar, D.S., B. Bashel, S.A.H. Balasubramanya, C.J. Creighton, I. Ponce-Rodriguez,
3. Chakravarthi, B.V., & Varambally, S. (2017). UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia, 19: 649-658.
4. Davis, A.P., Wiegers, T.C., Johnson, R.J., Lay, J.M., Lennon-Hopkins, K., Saraceni-Richards, C., Sciaky, D., Murphy, C.G., & Mattingly, C.J. (2013). Text mining effectively scores and ranks the literature for improving chemical-gene-disease curation at the comparative toxicogenomics database. PLoS One. 8.
5. Dey, N., De, P., & Leyland-Jones, B. (2017). PI3K-AKT-mTOR inhibitors in breast cancers: From tumor cell signaling to clinical trials. Pharmacology & therapeutics, 175: 91-106.
6. Diboun, I., Wernisch, L., Orengo, C.A., & Koltzenburg, M. (2006). Microarray analysis after RNA amplification can detect pronounced differences in gene expression using limma. BMC genomics.7: 252.
7. Kast, K., Link, T., Friedrich, K., Petzold, A., Niedostatek, A., Schoffer, O., Werner, C., Klug, S.J., Werner, A., & Gatzweiler A. (2015). Impact of breast cancer subtypes and patterns of metastasis on outcome.Breast cancer research and treatment, 150: 621-629.
8. Lacny, S., Wilson, T., Clement, F., Roberts, D.J., Faris, P., Ghali W.A., & Marshall, D.A. (2018). Kaplan–Meier survival analysis overestimates cumulative incidence of health-related events in competing risk settings: a meta-analysis. Journal of clinical epidemiology, 93: 25-35.
9. Liberati, A., Altman, D., Tetzlaff, J., Mulrow, C., Gøtzsche, P., & Ioannidis, J.C. M., Devereaux, P.J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. PLoS Medicine. 6.
10. Pathan, M., Keerthikumar, S., Chisanga, D., Alessandro, R., Ang, C.S., Askenase, P., Batagov, A.O., Benito-Martin. A., Camussi, G., & Clayton A. (2017). A novel community driven software for functional enrichment analysis of extracellular vesicles data.Journal of extracellular vesicles, 6: 1321455.
11. Siegel, R.L., Miller K.D., & Jemal, A. (2019). Cancer statistics. CA: a cancer journal for clinicians, 69: 7-34.
12. Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., & Kuhn, M. (2015). STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic acids research*, *43*(D1), D447-D452.
13. Asgarova, A., Asgarov, K., Godet, Y., Peixoto, P., Nadaradjane, A., Boyer-Guittaut, M., & Pallandre, J. R. (2018). PD-L1 expression is regulated by both DNA methylation and NF-kB during EMT signaling in non-small cell lung carcinoma. *Oncoimmunology*, *7*(5), e1423170.
14. Cai, L., Gu, Z., Zhong, J., Wen, D., Chen, G., He, L., & Gu, Z. (2018). Advances in glycosylation-mediated cancer-targeted drug delivery. *Drug Discovery Today*, *23*(5), 1126-1138.
15. Campton, D. E., Ramirez, A. B., Nordberg, J. J., Drovetto, N., Clein, A. C., Varshavskaya, P., ... & Blau, S. (2015). High-recovery visual identification and single-cell retrieval of circulating tumor cells for genomic analysis using a dual-technology platform integrated with automated immunofluorescence staining. *BMC cancer*, *15*(1), 360.

16. Cong, W., Zhou, A., Liu, Z., Shen, J., Zhou, X., Ye, W., & Jin, L. (2015). Highly sensitive method for specific, brief, and economical detection of glycoproteins in sodium dodecyl sulfate-polyacrylamide gel electrophoresis by the synthesis of a new hydrazide derivative. *Analytical chemistry*, *87*(3), 1462-1465.
17. Gadalla, S. E., Öjemalm, K., Vasquez, P. L., Nilsson, I., Ericsson, C., Zhao, J., & Nistér, M. (2013). EpCAM associates with endoplasmic reticulum aminopeptidase 2 (ERAP2) in breast cancer cells. *Biochemical and biophysical research communications*, *439*(2), 203-208.
18. Liu, X., Gao, J., Sun, Y., Zhang, D., Liu, T., Yan, Q., & Yang, X. (2017). Mutation of N-linked glycosylation in EpCAM affected cell adhesion in breast cancer cells. *Biological Chemistry*, *398*(10), 1119-1126.
19. Liu, X., Yang, L., Zhang, D., Liu, T., Yan, Q., & Yang, X. (2019). Deglycosylation of epithelial cell adhesion molecule affects epithelial to mesenchymal transition in breast cancer cells. *Journal of cellular physiology*, *234*(4), 4504-4514.
20. Nguyen, T. H., Kandel, M. E., Rubessa, M., Wheeler, M. B., & Popescu, G. (2017). Gradient light interference microscopy for 3D imaging of unlabeled specimens. *Nature communications*, *8*(1), 1-9.
21. Naeem, M., Ashraf, A., Safdar, H. M. Z., Khan, M. Q., Rehman, S. U., Iqbal, R., ... & Ahmad, G. Biochemical changes in patients with chronic kidney failure in relation to complete blood count and anemia.
22. Zhang, D., Liu, X., Gao, J., Sun, Y., Liu, T., Yan, Q., & Yang, X. (2017). The role of epithelial cell adhesion molecule N-glycosylation on apoptosis in breast cancer cells. *Tumor Biology*, *39*(3), 1010428317695973.
23. Naeem, M., Ali, J., Hassan, M. Z., Arshad, B., Rao, M. H. I., Sarmad, M. S. K., ... & Hussain, M. U. Novel Approach towards DNA Barcoding as a Tool in Molecular Biology and Biological Activities of Cyclotides with Particular Emphasizes at Molecular Level.
24. Naeem, A., Saddique, S., & Chand, S. A. (2019). Advancement and Future Directions towards Herbal Treatment for Various Diseases.