# INVESTIGATION OF DIFFERENTIAL GENE EXPRESSION OF ER NEGATIVE AND ER POSITIVE BREAST CANCER CELL

Atishya Mahesh Jain
Karunya Institute of Technology
and Sciences
Coimbatore, India

*Abstract*— **This paper presents about differential gene expression of ER negative and ER positive breast cancer cells. Cancer is a major problem in many countries, including India. Cancer is basically a disorder of autonomic cell proliferation and is of a very specific nature that makes them difficult to treat. In all types of cancer, the incidence and mortality rate of breast cancer are higher due to late diagnosis of the disease, in the progression of normal cell in the cancerous cell, the cell has to go through various obstacles and has to overcome to divide continuously. And spread (Hanahan and Weinberg, 2000). Despite recent improvements in breast cancer mortality, many patients recover from the onset of the disease after an initial response to traditional therapies such as chemotherapy, radiotherapy, and endocrine therapies (Creighton et al. 2009). According to a report by the World Health Organization, the survival of breast cancer can be enhanced by preliminary investigations that are similar to the basics of breast cancer regulations. Various modern therapies have been suggested for the treatment of breast cancer. The basic treatments and drugs prescribed for people at risk of breast cancer are enterogens such as tamoxifen and raloxifen which are thought to prevent the occurrence of breast cancer and those who are more likely to develop it (Peng et al., 2009). In addition to medications and treatments, there are many other techniques such as surgery, where the tumor can be physically removed. Despite these advances, the opposite nature of the disease has made it challenging for breast cancer cells to exhibit and treat.**

## I. INTRODUCTION

Cancer – To classify breast cancer into different types based on their molecular profile and also to understand the biomarkers present in the cell along with their role in the process of tumorogenesis (Poliak, 2011). When it comes to breast cancer, there are many risk factors that lead to the onset of the disease. Some of the major factors that lead to breast cancer risk are divided into five different categories. They are:

1) Aging: It is the first and leading risk factor for the onset of breast cancer as studies show that the incidence of breast cancer and age is directly proportional. Thus, screening testing after a person reaches the age of 40 is really important to avoid the onset of HBC (Siegel, 2017).

2) Family history: Family history plays a major role in the incidence of breast cancer. A woman with a family background is more likely to have one or more breast cancers. In addition, the risk is 2 times higher in a woman if two or more of her relatives are suffering from breast cancer. This is mainly due to mutated genes like BRCA1, BRCA2, etc. (Brewer, 2017)

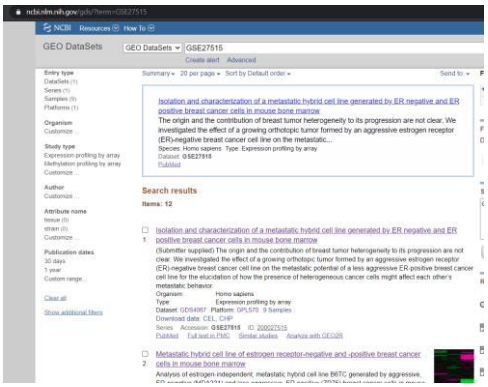
3)) Fertility factors: This is another criterion. Various reproductive factors such as late pregnancy, late menopause, etc., play an important role in increasing the risk of early onset of breast cancer (W. Shabruk 2006; Horn et al., 2017; Dale et al., 2017).

4)) Estrogen: Estrogen- Both exogenous and endogenous are highly associated with the risk of breast cancer. The former is basically produced by the ovaries (Endogenous, 2013) while the latter is acquired by external sources such as oral contraceptives and HRT.

5)) Lifestyle: Contemporary lifestyle includes consumption of elevated alcohol and excess fat through diet which results in increased risk of breast cancer. This is because both alcohol and fat intake increase the level of estrogen in the blood. In addition, smoking during adolescence increases the risk of breast cancer outbreaks. (Mackenzie et al., 2013; Ketsburg et al., 2015; Gaudet et al., 2017; Kispert et al., 2017).

## II. METHODS AND MATERIALS

***A.Microaaray Dataset Retrival:*** The microarray dataset is derived from gene expression n minibus (GEO) which is a web-based tool as shown in Figure 5.1. Initially, numerous microarray datasets were analyzed based on their experimental setups and the sample shown in Figure 5.2. Next, we selected microarray datasets GSE27515 for differential gene expression and pathway breeding studies. The specimens used in this dataset are estrogen-independent cell lines with estrogen-positive (control) and estrogen-negative (sample) cell lines. The platform file, as well as the category matrix file, have been downloaded for a detailed study about the genes present in that cell line.



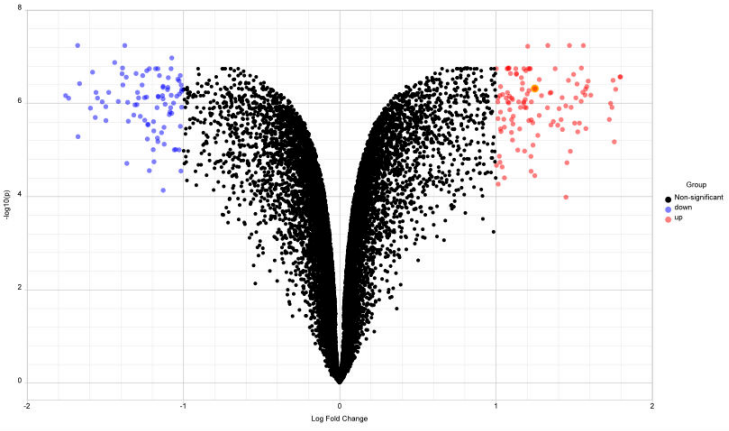

***B.Gene Enrichment Analysis:***

**BOX PLOT :** Box plot is a graphical method for portraying groups of numerical data through quartiles. They do not depend on a specific parameter. They display the variations present in the sample within a statistical population.

**COUNT SUM PLOT:** Count Sum plot is one among the quality check plots which gives the sum of the read counts from all the features of each sample along with the control.

***C. Differential Experssion Analysis:***

Differential expression analysis is used to identify genes whose expression differs in different situations. This is a statistical phenomenon that occurs when thousands are compared for a small number of samples. The selection of differentially expressed genes (DEGs) is a basic technique for comparing expression profiles between different conditions. One of the main applications of microarray analysis is the classical class comparison experiment comparative analysis of gene expression profiles between 2 or more different conditions or phenotypes.




To understand the biological effect, we find out which genes are up-regulated or down-regulated (decreased in regulation) in a specific physical or pathological condition, such as drug-treated specimens, compared to the reference (control) condition. Various methods and statistical testing methods. Statistical testing is used to determine if, for a given gene, the observed difference in the reading calculation is significant.

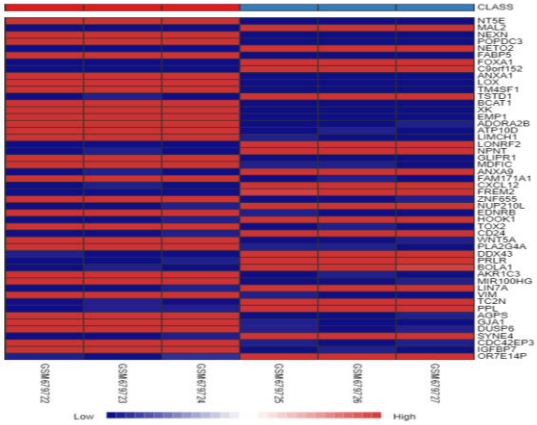
***D.Gene Functional Analysis***

Gene functional analysis provides association between the gene product and the function in order to annotate genes based on existing biological knowledge. It is one of the functional enrichment techniques which play a role in classifying the genes based on their functional characteristics and attributes. The genes from GSE27515 are uploaded and are analyzed using the PANTHER database.

***E.Pathway Enrichment Analysis***

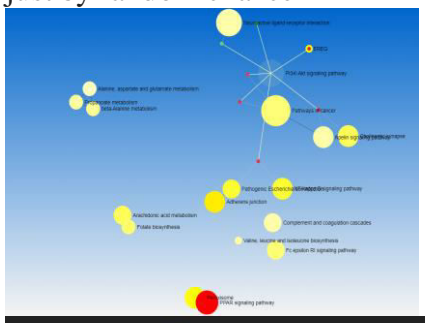
**Over Representation Analysis- Heatmap Clustering:** Over Representation Analysis (ORA) measures the percentage of in the pathway that has the Differential Gene Expression. ORA aim to obtain the important and relevant pathway based on the p-value this ORA result shows the relationship between the pathway of interest to various other pathways also the relationship of a particular gene of interest with other pathways and genes in the same pathway or different pathway. The ORA heatmaps are interactive, allowing users to easily visualize, perform enrichment analysis, and define gene signatures using groups of genes from the heatmap. These are used in molecular biology to represent the expression level of many genes across a number of comparable samples as they are obtained from DNA

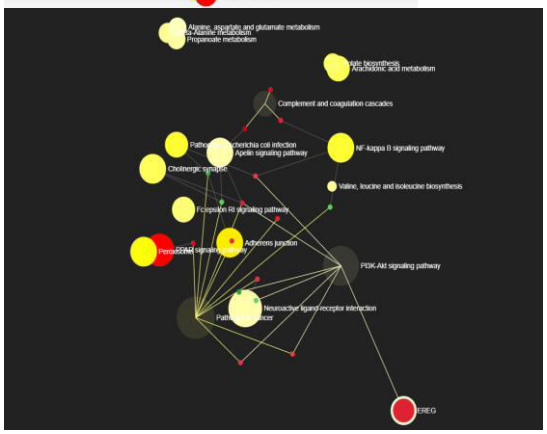
microarrays. The red coloured genes are upregulated ones and EREG gene is selected for further study.




## ENRICHMENT NETWORK

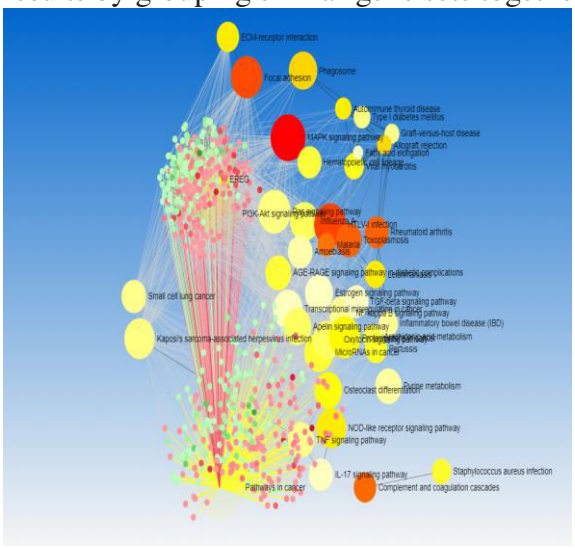
Over Representation Analysis (ORA) is a statistical technique that identifies gene sets or pathways that have significant overlap with the selected gene of interest. These algorithms compare the proportion of genes within deferentially expressed genes, which are associated with a particular functional gene ontology term with the proportion of gene obtained just by random chance.






## GENE SET ENRICHMENT ANALYSIS

A common approach for interpreting gene expression data is gene set enrichment analysis based on the functional annotation of the differentially expressed genes. This is useful to find out if the differentially expressed gene are associated with a certain biological process or molecular function. The Gene Ontology, containing the annotation of gene products, is commonly used. It works by comparing the frequency of individual annotations in the gene list with a reference list. Enrichment of biological pathways supplied by KEGG, or Reactome. Each significantly enriched gene set from GSEA is represented as a Node. Gene sets with overlapping genes are connected with an Edge. The network visualization simplifies the interpretation of GSEA results by grouping similar gene sets together.




### III. RESULT AND ANALYSIS

One of the main causes of cancer is copy number change (CNA) in DNA, and microarrays play an important role in identifying it. Genesis analysis is a tool for gene interpretation. We can interpret and analyze large gene lists using prior knowledge (Nam and Kim, 2008). PI3K / ACT / MTOR is a major intracellular signaling pathway that responds to stimulation of nutrients, hormones, and growth factor. Is and it is well established. To play a very significant role in the growth and proliferation of tumor cells (Cantley, 2002). Activation of this pathway is one of the main causes of cancer cell resistance to antitumor therapy. PI3K / ACT / MTOR indicates the critical object of the study to understand the development and progression of this disease. Thus, the potential of this pathway in breast cancer patients may play a role in therapeutic targeting, as well as prognosis and diagnostic value. Despite the existence of selective PI3K / Akt / mTOR pathway inhibitors and existing clinical trials, cellular mechanisms are not yet known. , 4-triphosphate or PIP which, in turn, leads to phosphorylation of ACT, which is serine / threonine kinase, which affects cancer cell cycling, survival and growth (Zhao and

Vogt, 2008). Dysregulation of this pathway is associated with a variety of cancer hallmarks, including uncontrolled proliferation, genomic instability, and metabolic reprogramming in tumor cells.

## IV. CONCLUSION

This paper provides a detailed overview of the MicroAr dataset based on its experimental setup and compressed the dataset GSE 27515 for which enrichment and pathway analysis has been performed. Specific gene expression studies have been performed to understand the expression and function of genes involved in the PI3K-AKT signaling pathway where after thorough analysis; A single gene EREG has been compressed for further study. The simultaneous analysis of gene- nantology studies for the functions of genes has been performed. Primers can be designed and characterized for the corresponding erasure gene by PCR standardization. No medication was developed to avoid blind targeting and so we can minimize chemotherapeutic side effects. Further studies will focus on standardizing the cloning process, standardizing the transaction process, and determining expression studies for gene EREG.

## REFERENCES

[1] *Akhtar M. Dasgupta S. Rangwala M. Triple negative breast cancer: an Indian perspective. Breast Cancer, (2015); 7: 239-43*

[2] *Akram M. Iabal M. Daniyal M and Khan AU. Awareness and current knowledge of breast cancer, Biol Res, (2017); 50:33. DOI 10.1186/s40659-017-0140-9*

[3] *Al-Hajj M. Wicha MS. Benito-Hernandez A. Morrison SJ. Clarke MF. Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci USA. (2003): 100(7):3983-3988. doi:10.1073/pnas.0530291100.*

[4] *Ali 1. Wani WA and Saleem K. Cancer scenario in India with future perspectives. Cancer Therapy, (2011); 8: 56-70.*

[5] *Alizadeh AA. Eisen MB. Davis RE. Ma C. Lossos IS and Rosenwald. A. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling.Nature, (2000); 403, 503-511*

[6] *Allard WJ. Matera J. Miller MC. Repollet M. Connelly MC, Rao C. Tibbe AG. Uhr JW and Terstappen LW. Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases. Clin Cancer Res, (2004); 10: 6897-6904.*

[7] Anuradha D and Lakshmi A. Mucinous carcinoma of breast with neuroendocrine differentiation: a rare case report with review of literature. Int J Res Med Sci. (2017); 2:1751-4.

[8] Arpino G. Bardou VJ. Clark GM and Elledge RM. Infiltrating lobular carcinoma of the breast: tumor characteristics and clinical outcome. Breast Cancer Res.(2004); 6:149-52

[9] Ashburner M. Ball CA. Blake JA. Botstein D and Butler H. Gene ontology: tool for the unification of biology. The Gene Ontology, Consortium. Nature genetics (2000); 25: 25-29

[10] Babu GR. Lakshmi SB. Thivagaraian JA. Epidemiological correlates of breast cancer in South India. Asian Pac J Cancer Prev, (2013); 14: 5077-83.

[11] Balasubramaniam S. Rotti S. Vivekanandam S. Risk factors of female breast carcinoma: a case control study at Puducherry. Indian J Cancer, (2013); 50: 65-70.

[12] Bardou VJ. Arpino G. Elledge RM. Osborne CK. Clark GM. Progesterone receptor status significantly improves outcome prediction over estrogen receptor status alone for adjuvant endocrine therapy in two large breast cancer databases. J Clin Oncol, (2003); 21(10):1973-1979.

[13] Bednarek A. Sahin A. Brenner A. Johnston D. Aldaz C. Analysis of telomerase activity levels in breast cancer positive detection at the in situ breast carcinoma stage. Clin Cancer Res, (1997); 3(1):11-6.

[14] Beniamin O. Anderson. M.D. UICC World Cancer Congress (2014). GLOBAL BREAST CANCER TRENDS

[15] Bensimon J. Altmeyer-Morel S. Benielloun H. Chevillard S. Lebeau J. CD24(-/low) stem-like breast cancer marker defines the radiation-resistant cells involved in memorization and transmission of radiation-induced genomic instability.Oncogene, (2013); 32(2):251-258. doi:10.1038/onc.2012.31

[16] Chlebowski RT. H S. Influence of estrogen plus progestin on breast cancer and mammography in healthy postmenopausal women: The women's health initiative randomized trial. JAMA, (2003); 289(24):3243-3253

[17] Clarke MF. Dick JE. Dirks PB. Cancer stem cells-perspectives on current status and future directions: AACR Workshop on cancer stem cells. Cancer Res, (2006):66:9339-44

[18] Ferlay J. Shin HR. Bray F. Forman D. Mathers C and Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer (2010) : 127: 2893-917

[19] Fidler IJ. Gersten DM and Hart IR. The biology of cancer invasion and metastasis.Adv Cancer Res. (1978); 28: 149-250.