

# ANALYSIS OF BREAST CANCER BASED ON THE GENE CLUSTER AND GENE EXPRESSION

<sup>1</sup>Mrs. Sandhya G, <sup>2</sup>Dr.G. T. Raju, <sup>3</sup>Mrs.D.Vasumathi  
1,Associate professor, Department of CS&E, RLJIT, Doddaballapur  
2, Professor, Dean & Head, Department of CS&E, RNSIT, Bangalore.  
3,Professor,Department of CS&E, JNTUH,Hyderabad

## Abstract:

**The main cancer deaths in women are due to the breast cancer. 1 out of 8 women's have a risk to develop breast cancer. In this paper we discuss about the factors and analysis of the breast cancer depending the gene expressions and gene clusters.**

## Introduction:

After skin cancer, breast cancer is the most frequently diagnosed cancer in women. It is second only to lung cancer in cancer-related deaths. Most breast cancer cases occur in women ages 40 or older. Gene expression profiling has been of great help in defining molecular phenotypes of many kinds of tumors. Stanford microarray center scientists with have studied genome-wide expression patterns in several cancers including lymphoma, breast, lung, liver, ovarian cancer and soft tissue sarcomas 1–5. Gene expression studies have made it clear that there is considerable diversity among breast tumors, both biologically and clinically 5-11. This is not a novel concept, as epidemiological studies previously had inferred the existence of two or more subpopulations of breast cancer 12.

Gene expression data derived from clinical cancer specimens provide an opportunity to characterize cancer-specific transcriptional programs. Here, we present an analysis delineating a correlation based gene expression landscape of breast cancer that identifies modules with strong associations to breast cancer specific and general tumor biology.

## Breast cancer

Breast cancer usually presents as solitary, painless, palpable lump which is detected quite often by self examination. Higher the age there will be more chances of breast lump turning out to be a cancer tissue. Currently the main concentration is on the early diagnosis by mammography. There are mainly two morphological types in breast carcinoma, they include

### 1. NON-INVASIVE (IN-SITU) CARCINOMA

It is characterized by presence of tumor cells within ducts or lobules without evidence of invasion. There are two types of carcinoma in situ- intraductal carcinoma which is confined with larger mammary ducts and lobular carcinoma which is not a palpable or grossly visible tumor.

### 2. INVASIVE CARCINOMA

It has various morphologic types. The most common is the infiltrating duct carcinoma which is the classic breast cancer and accounts for 70% cases of breast. Infiltrating lobular carcinoma

which comprises of 5% of all breast cancers. Medullary carcinoma is a variation of ductal carcinoma and comprises about 1% of all breast cancers.

### Factors causing the breast cancer

- **Family history:** First degree relatives like mother, sister, daughter of women with breast cancer has 2 to 6 fold high risk of development of breast cancer.
- **Genetic mutations:** Around 10% of breast cancers have found to have inherited mutations; the main is the BRCA gene. The most risk is due to BRCA 1 gene located at chromosome 17, BRCA 2 located at chromosome 13 and mutation in TP53 tumor suppressor gene on chromosome 17. The mutation of these genes causes the risk of developing breast cancer.
- **Breast density:** Breasts which is having more glandular and connective tissue and having less fat indicates that breast is having more density. This is more difficult to detect by mammography.
- **History of breast cancer:** The women who have been cured of a primary breast cancer are at increased risk of breast cancer the risk persists for 20 years or more.
- **Reproduction factors:** First menstrual cycle at the very early age, a late age at menopause, late age at the birth of the first child.
- **Postmenopausal Hormones:** The therapy with a combination of estrogen and progestin increases the risk.
- **Alcohol:** Consumption of more than 2 glasses of alcohol a day increases the risk.
- **Radiation therapy:** Women who have undergone radiation to the chest for any sort of treatment have an increased risk of breast cancer.
- **DES:** Diethylstilbestrol is a synthetic estrogen used to reduce the risk of miscarriages. According to the research the daughters born may also have risk of breast cancer.
- **Age and Gender:** There will be more risk of getting breast carcinoma as the age increases. Most advanced breast cancer cases are found in women over age 50. Compared to men, women's have high risk of getting breast cancer.

### Methods:

Repeated observation of breast tumor subtypes in independent gene expression data sets [Homo sapiens] This Super Series is composed of the following subset Series: GSE4335: Norway/Stanford Breast Tumors GSE4336: reported characteristic patterns of gene expression classifiers of breast tumors into clinically relevant subgroups. In this study, a total of 115 malignant breast tumors were analyzed by hierarchical clustering based on patterns of expression of 534 "intrinsic" genes and shown to subdivide into one basal-like, one ERBB2-overexpressing, two luminal-like, and one normal breast tissue-like subgroup. The genes used for classification were selected based on their similar expression levels between pairs of consecutive samples taken from the same tumor separated by 15 weeks of neo adjuvant treatment. Modules of highly connected genes were extracted from a gene co-expression network that was constructed based on Pearson correlation, and module activities were then calculated using a pathway activity score. Functional annotations of modules were experimentally validated with an siRNA cell spot microarray system using the KPL-4 breast cancer cell line, and by using gene expression data

from functional studies. Modules were derived using gene expression data representing 1,608 breast cancer samples and validated in data sets representing 971 independent breast cancer samples as well as 1,231 samples from other cancer forms.

## Results:

The initial co-expression network analysis resulted in the characterization of 8 tightly regulated gene modules. Cell cycle genes were divided into two transcriptional programs, and experimental validation using an siRNA screen showed different functional roles for these programs during proliferation. The division of the two programs was found to act as a marker for tumor protein p53 (*TP53*) status in luminal breast cancer, with the two programs being separated only in luminal tumors with functional protein 53 (p53).

Moreover, a module containing fibroblast and stroma related genes was highly expressed in fibroblasts, but was also up regulated by over expression of epithelial-mesenchymal transition factors such as transforming growth factor beta 1 (TGF-beta1) and Snail in immortalized human mammary epithelial cells. Strikingly, the stroma transcriptional program related to less malignant tumors for luminal disease and aggressive lymph node positive disease among basal-like tumors. Three representative samples from the ERBB2+ group and the Basal group were used for comparative pathway analysis. Out of the four clinically significant sample groups that were compared in the study, the two worst-prognosis categories, Basal and ErbB2 types, were similar in that adhesion, tissue and ECM-remodeling genes were overrepresented in these groups: ERK1 is down, but frizzled and SMAD genes were up-regulated in both; Estrogen receptor 1 is down; Brca1 gene is ON in both, (but more so in ERBB2)

## Conclusions:

By applying the modules to *TP53*-mutated samples we shed light on the biological consequences of non-functional p53 in otherwise low-proliferating luminal breast cancer. Furthermore, as in the case of the stroma module, we show that the biological and clinical interpretation of a set of co-regulated genes is subtype-dependent. Once the key genes were ,we were able to see the particular relationships within and between the reported cancer groups in the context of biological pathways. Comparing the two poor-prognosis groups up regulation of Extra cellular signaling genes involved in inflammatory and proteolytic processes in ERBB2+ type. This may account for the cell motility and ECM destruction responsible for the increased metastasis rate. On the other hand, proliferative and developmental processes displayed by activated cell growth and division genes may account for the worst survival prognosis in the Basal cancer group. Due to the significant degree of overlap between these and other classes, there is a clear incentive to pursue further molecular pathway analysis and classification of breast cancer.

## References:

→ Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G: Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000, 25:25-29.

→ Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP: Gene set enrichment analysis: a knowledge-based approach for interpreting genome wide expression profiles. *Proc Natl Acad Sci U S A* 2005, 102:15545-15550.

→ Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein Lonning P, Borresen-Dale AL: Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001, 98:10869-10874

→ van't Veer, L. J., Dai, H., van de Vijver, M. J., He, Y. D., Hart, A. A., Mao, M., Peterse, H. L., van der, K. K., Marton, M. J., Witteveen, A. T., et al. (2002) *Nature* 415, 530-536.[CrossRef][ISI][Medline]

→ West, M., Blanchette, C., Dressman, H., Huang, E., Ishida, S., Spang, R., Zuzan, H., Olson, J. A., Jr., Marks, J. R. & Nevins, J. R. (2001) *Proc. Natl. Acad. Sci. USA* 98, 11462-11467.[Abstract/Free Full Text]

IJOART