Identification Of Epidermal Growth Factor Receptor As Breast Cancer Master Regulator

Akhila Rachel James[1] , Sakthidasan Jayaprakash[2*]

[1]Akhila Rachel James: Research Scholar, Department of Biotechnology, Hindustan Institute of Technology and Science, Padur, Chennai, Tamilnadu, India

[2*]Sakthidasan Jayaprakash: Assistant Professor, Department of Biotechnology, Hindustan Institute of Technology and Science, Padur, Chennai, Tamilnadu, India

Abstract

Breast cancer emerged as the leading cause of cancer death in women. Computational approaches confined to survival analyses and fold change comparisons of gene expression between healthy and diseased profiles have revealed several differentially expressed genes in breast cancer. Identification of DEGs can help determine the master regulator that can be used as a potential drug target` for treating breast cancer. 587 profiles pertaining to breast cancer human tissue samples with 635 up-regulated genes were identified from Gene Expression Omnibus. Conduction of gene enrichment analysis revealed their role in different cancer pathways. Interaction networks of proteins and topological mathematical scoring functions proved the genes Epidermal Growth Factor Receptor (EGFR), Estrogen Receptor alpha (ESR1), Estrogen Receptor 2 (ESR2), KIT proto-oncogene (KIT), C-X-C Motif Chemokine Ligand 2 (CXCL2) and Dystrophin (DMD) as master regulators. A master regulator is a gene that stay at the top of hierarchy and regulates other cancer related genes. Human Protein Atlas determined the drug potency of these genes. Further analysis using FDA revealed EGFR, ESR1, ESR2 and KIT as the potential Food and Drug Administration approved master regulators. Based on ranks and scores in topological scoring models, EGFR was identified as a potential master regulator in treating breast cancer.

Keywords: Drug target, Gene enrichment, Gene expression, The Human Protein Atlas.

1.0 Introduction

Breast cancer is the most common cancer prominently occurring in women. More than 2 million new cases were diagnosed in the year 2021 accounting for the fifth leading cause of

cancer deaths worldwide. The incidence is escalating in developing countries because of adoption of western lifestyles, urbanization and increased life expectancy. Breast cancer is a heterogeneous disease with the involvement of innumerable genes in its progression, invasion and metastasis. Understanding the involvement of different genes and its pathways will help in managing the progression, treatment and survival rate of breast cancer (De Cicco et al., 2019).

Breast cancer and its complex multistep process alter numerous signalling pathways. The term master regulator, the gene that stays at the top of hierarchy and regulate other associated genes positively or negatively was coined by Ohno. These central regulatory genes efficiently remodel the transcript product thus, enhancing differentiation in cells leading to various diseases including cancer. Master regulators can be identified by construction of various mathematical network models. Targeting these genes can help to bring down breast cancer by increasing target specificity (Ohno, 1979, Balderas-Hernandez et al., 2013, Gevaert and Plevritis., 2013).

This study is to classify the differentially expressed genes in breast cancer by analysing the gene expression profiles of breast cancer patients. Gene Enrichment analysis, protein interaction network and topological scoring models helped in identification of master regulator that plays a major role in regulation of other cancer related genes driving breast cancer progression and metastasis.

2.0 Materials and Methods

2.1 Data Screening breast cancer DEGs

The expression profiles relevant to breast cancer were retrieved from Gene Expression Omnibus, a genomic data repository maintained by the National Centre for Biotechnology Information (NCBI) (Ariya et al., 2020). The profiles pertinent to Homo sapiens were procured. Among those, the profiles pertaining to cell lines, miRNA and gene sequence were eliminated. Eventually, datasets relevant to genes expressed in breast cancer in Homo sapiens were further analysed by transforming it to a logarithmic scale (Barrett et al., 2013).

The differentially expressed genes in breast cancer were identified by GEO2R. The criteria adopted were the false discovery rate, $\log_2 FC > 2$ and the p-value <1.0 (Benjamini et al., 2001).

2.2 Function and Pathway Enrichment Analysis

The enrichment analysis of the differentially expressed genes were applied for process, function and pathway enrichment analysis using the database GOrilla (**G**ene **O**ntology en **RI**chment ana**L**ysis and visua**LizA**tion tool). It helps to identify pathways and processes associated with factor regulating activity in cancer (Pan et al., 2019).

2.3 Protein-Interaction Network analysis

The signalling pathways and the interaction of genes were mapped and characterized by means of protein interaction network. Proteomics Standard Initiative Common QUery Interfa Ce (PSICQUIC), a web service which enables simultaneous computational access to numerous

molecular interaction data resources was used to construct the network on the Cytoscape platform developed by Shannon (Aranda et al., 2011, Shannon et al., 2003). The required information was retrieved from Reactome Fls and Biogrid database, and protein-protein interaction network was constructed based on the obtained results (Stark et al., 2006, Croft et al., 2011).

2.4 Topological Scoring Models from PPI

Based on different local and global mathematical model protocols, the topology of molecular interactions between the commonly up-regulated genes were analysed using the Cyto Hubba platform (Croft et al., 2011). Different algorithms were used to analyse the topological scoring functions namely Degree method (Deg), Maximum Neighbourhood Component (MNC), Density of Maximum Neighbourhood Component (DMNC), Maximal Clique Centrality (MCC), Closeness (Clo), Ec Centricity (EC), Radiality (Rad), Bottle Neck (BN), Stress (Str), Betweenness (BC) and Edge Percolated Component (EPC). The nodes and internodes were created (Chin et al., 2014).

2.6 Master regulators and potential drug target

The master regulator was determined based on the interaction network and survival analysis of the differentially expressed genes. The ranks, score and occurrence of gene in the network were considered. The potentiality of master regulators as a drug target were then evaluated using Human Protein Atlas (HPA) (Lindskog et al., 2016). HPA provides details about the gene and its acceptance by FDA as a potential drug target (Kadara et al., 2014).

3.0 Result

3.1 Identification of Differentially Expressed Genes

10 gene expression profiles pertaining to breast cancer were retrieved from Gene Expression Omnibus (GEO). A total of 587 samples were obtained from the profiles GSE86374, GSE72644, GSE22544, GSE71053, GSE16873, GSE17907, GSE17072, GSE15852, GSE9574 and GSE9309 of which 431 tumour and 142 normal samples were procured. Using GEO 2R, total DEGs expressed in the samples were identified. Based on the fold change and p-value, the up-regulated and the downregulated genes from the tumour and normal samples were recognized. This led to the identification of 1,640 up-regulated and 2,551 down-regulated genes from a total of 3,06,175 set of differentially expressed genes.

3.2 Screening of Breast cancer DEGs

Data mining to elucidate the top 250 significant DEGs were carried out based on the Benjamini and Hochberg method. A total of 635 up-regulated genes were selected for further analysis. The datasets were then compared to identify the common genes and this led to the identification 80 up-regulated genes that were common in more than one dataset. The details are given in table 2.

Table 2: Number of Differentially Expressed Gene from each Dataset

S.No	GEO Accession ID	No. of samples			DEGs			No. of significant top 250 genes	
		Total	Tumor	Normal	Total	Up-regulated	Down-regulated	Up-regulated	Down-regulated
1	GSE86374	159	124	35	33297	45	10	17	5
2	GSE72644	18	9	9	41094	77	150	13	18
3	GSE22544	20	16	4	14759	608	2068	217	2
4	GSE71053	18	12	6	54676	69	115	13	13
5	GSE16873	40	28	12	22283	36	10	78	25
6	GSE17907	55	51	4	24578	663	161	190	9
7	GSE17072	20	15	5	48803	5	2	1	0
8	GSE15852	86	43	43	22283	48	17	47	18
9	GSE9574	29	14	15	22283	13	4	12	3
10	GSE9309	142	133	9	22119	76	14	47	7
Total		587	431	142	306175	1640	2551	635	100

3.3 Gene Enrichment Analysis and Biological processes

Based on the GO biological process enrichment, the DEGs were mainly associated with cellular response to organic cyclic compound, detection of chemical stimulus involved in sensory perception, regulation of gliogenesis, regulation of lipid biosynthetic process, cellular response to steroid hormone stimulus and regulation of steroid metabolic process. Different genes involved are given in Table 3 and different biological functions are given in Figure 1.

Table 3: Gene Enrichment Analysis

GO term	Description	P-value	FDR q-value	Enrichment (N, B, n, b)
GO:0071407	cellular response to organic	2.17E-4	1E0	2.52 (360,20,107,15)

	cyclic compound			
GO:0050907	detection of chemical stimulus involved in sensory perception	3.1E-4	7.73E-1	72.00 (360,2,5,2)
GO:0014013	regulation of gliogenesis	5.04E-4	8.4E-1	3.50 (360,7,103,7)
GO:0046890	regulation of lipid biosynthetic process	7.11E-4	8.88E-1	2.65 (360,13,115,11)
GO:0071383	cellular response to steroid hormone stimulus	7.19E-4	7.18E-1	3.36 (360,7,107,7)
GO:0019218	regulation of steroid metabolic process	7.44E-4	6.19E-1	3.33 (360,9,96,8)
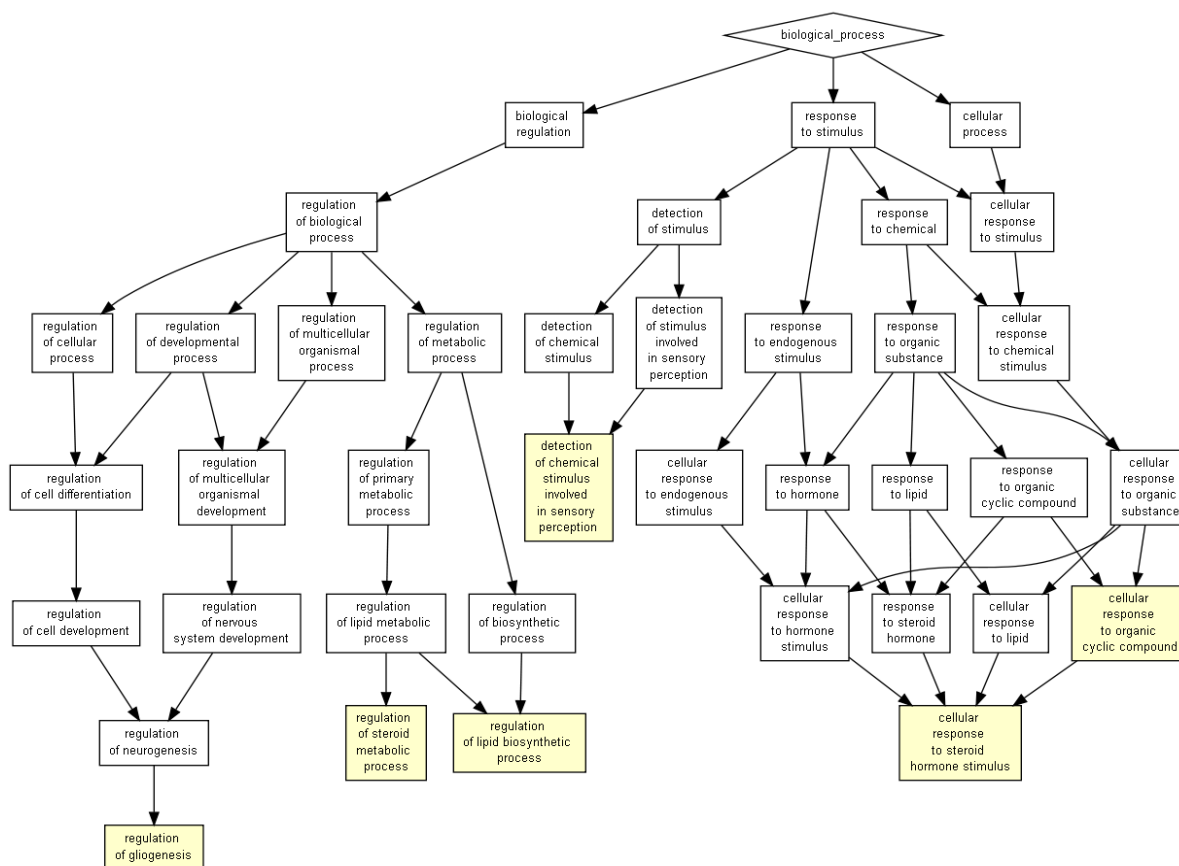
Figure 1: Gene Enrichment Biological processes

3.4 PPI Network Analysis

The genes work together with other genes and involve complex interactions to perform various biological function by formation of different pathways. The potential interaction was assessed among the genes involved in breast cancer. This was done using Reactome Fls and Biogrid database in the cytoscape platform. Results obtained were merged manually. The merged network had 1366 nodes and 1686 edges. The analysis revealed various common DEGs formed significant functional network, including regulating breast cancer related genes. Figure 2 shows the protein interaction network developed for commonly up-regulated genes.
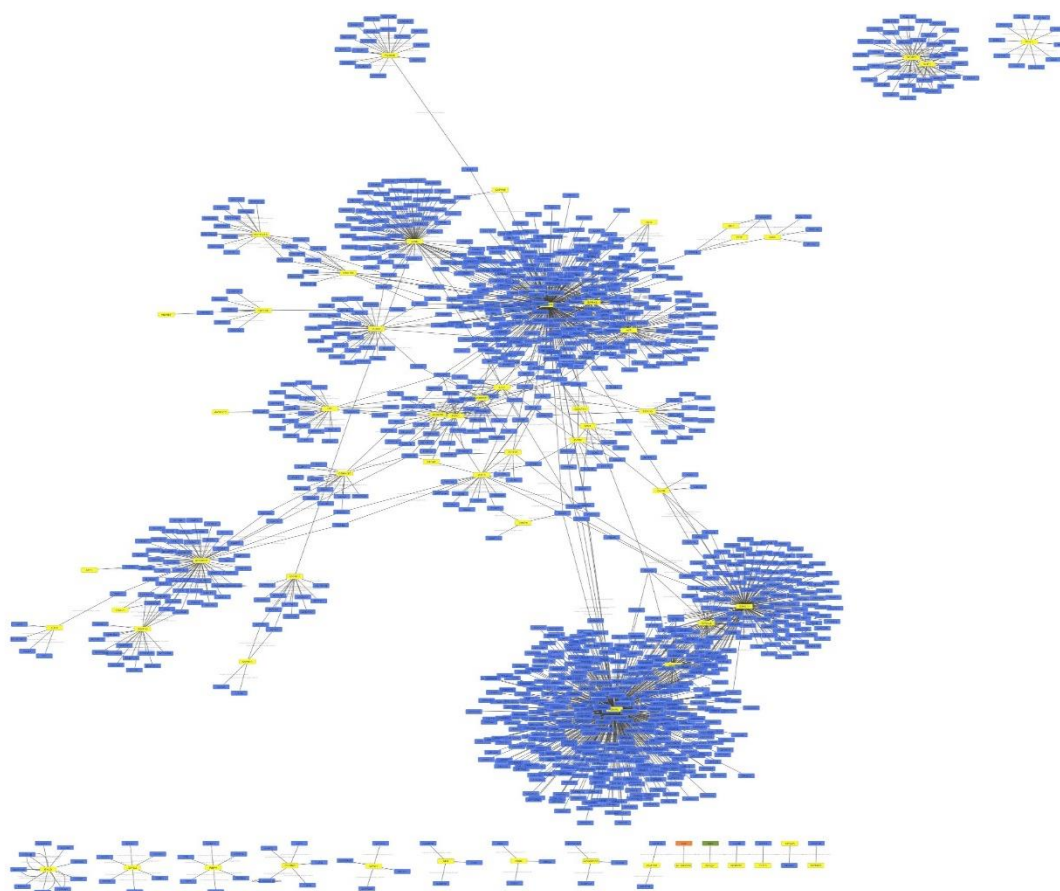
Figure 2: Merged protein interaction network of the up-regulated genes

3.5 Determining Topological Scoring Models

The topology of the DEGs were determined using Cyto Hubba platform. The merged network created from Reactome Fls and Biogrid database were given as the parent network from which the topological analysis was conducted. 12 different scoring methods were applied for the same. Based on the ranks, occurrence, score and impact in different scoring functions, the master regulators were identified. Figure 3A to 3L shows different topological scoring functions.
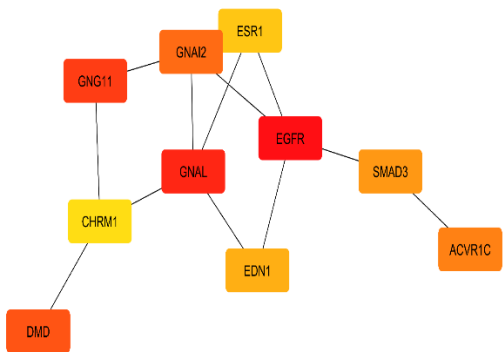
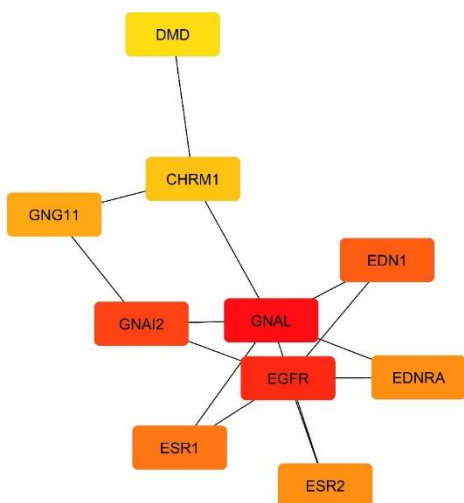**Figure 3A:** Betweenness
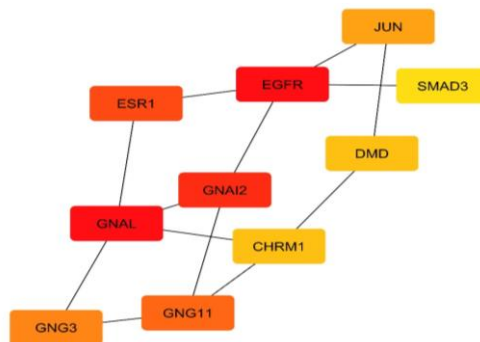
Figure 3B: Bottle Neck

Figure 3C: Closeness

Figure 3D: Clustering Coefficient

Figure 3E: Degree
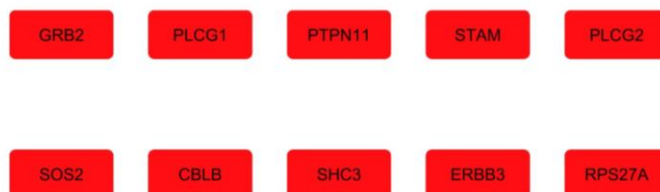
Figure 3F: DMNC
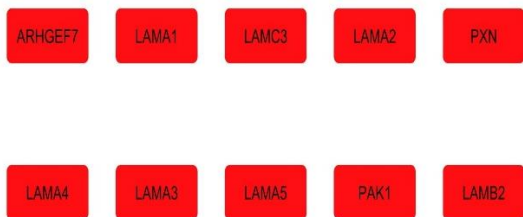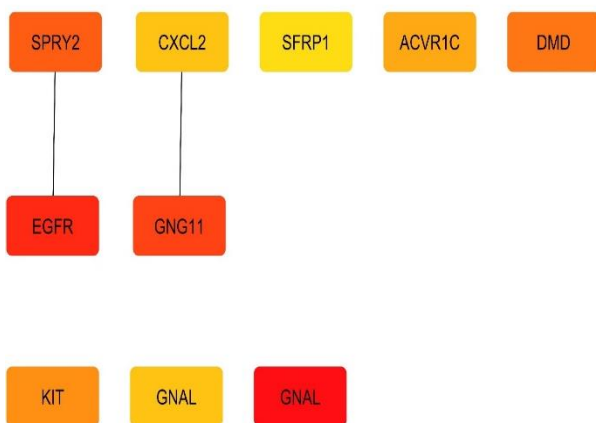
Figure 3G: Ec Centricity

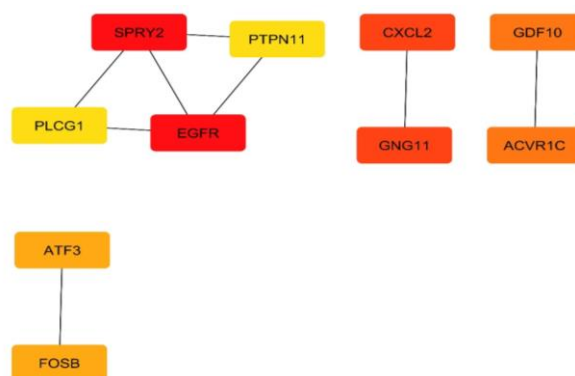**Figure 3H:** EPC

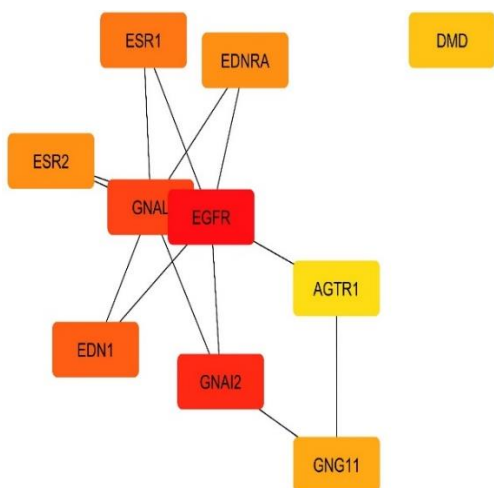**Figure 3I:** MCC

Figure 3J: MNC
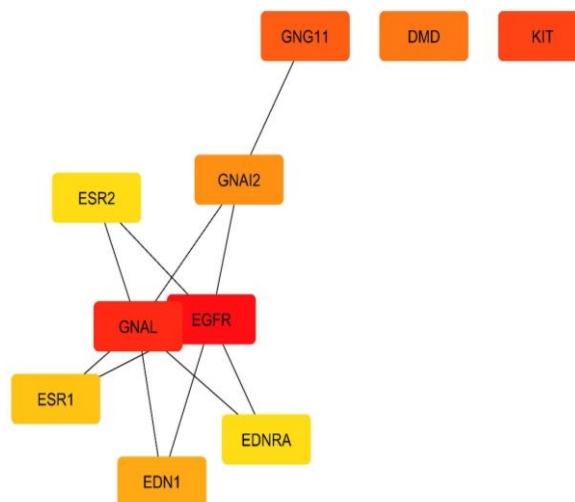

Figure 3K: Radiality

Figure 3L: Stress

3.6 Identification of Master Regulators

Topological analysis of the merged network of common up-regulated genes revealed the master regulators driving breast cancer. The genes GNG11, GNAI2, ESR1, GNAL, EGFR, CHRM1, SPRY2, DMD, CXCL2, EDNRA, ESR2 and KIT were found to have a higher impact on the topological scoring of network. Thus, several risk factors of breast cancer act upon these genes leading to cancer progression and metastasis.

3.7 Elucidation of Potential Drug Target

Among the 12 identified master regulators, the score for EGFR and GNG11 were higher. It was expressed in 9 different topological scoring functions. Cross-validation of the master regulators to be used as a drug target was carried out using Human Protein Atlas. These revealed the genes EGFR, ESR1, ESR2, KIT, CXCL2 and DMD as cancer related genes. Among these, EGFR, ESR1, ESR2 and KIT were found to be FDA approved drug target. Since, the ranking of EGFR was found to be higher in topological scoring models, it has been considered as a potential drug target for breast cancer patients.

4.0 Discussion

Breast malignancy is a leading cause of cancer death among women with poor prognosis rate. Non-specificity of the drug used and not targeting the master regulators of breast cancer have an adverse effect on tackling breast cancer. The underlying mechanism regulating breast cancer remains poorly understood. In this study, all the samples of human tissue of breast cancer deposited in Gene Expression Omnibus were analysed. The normal and tumour criteria were considered for the study. Through further analysis, master regulators and potential drug targets were identified.

The introduction of high throughput microarray analysis provided better information about the differentially expressed genes in tumour and normal tissues on a global level at a faster rate (Li et al., 2019). Through GO analysis, the enriched biological functions of different genes were analysed revealing the role of up-regulated genes in various cancer related functions. The cellular response to organic cyclic compounds is linked with viral carcinogenesis in breast (Changavi et al., 2015). Similarly, overexpression of genes involved in the generation of glial cells can lead to tumour. Alteration in lipids and steroids are found to be directly involved in triggering breast cancer. Thus, biological functions of these overexpressed genes can result in cancer progression.

The protein interaction network and topological interactions paved way for identification of the master genes involved in breast cancer. Identification of cancer related and FDA approved genes recognised Epidermal Growth Factor Receptor (EGFR), Estrogen Receptor 1 (ESR1), Estrogen Receptor 2 (ESR2) and KIT proto-oncogene as the hub genes involved in breast cancer. Based on the topological score and ranks, EGFR was identified as a potential drug target which is expressed in different cancers.

The Epidermal Growth Factor Receptor (EGFR) is a tyrosine kinase receptor in the HER family and is widely expressed in a number of tumours. Observed in about 15-45% of breast cancer, it plays a major role in cell proliferation. Phosphorylation of aberrant EGFR results in cell

proliferation, angiogenesis, invasion ad metastasis. Expression of EGFR induces resistance to chemotherapy and radiation treatment. It has been linked to many human epithelial cancers including breast cancer, non-small cell lung cancer, head and neck squamous cell carcinoma, pancreatic, colorectal and brain cancers (Masuda et al., 2012). EGFR promotes cancer growth and metastasis by Epithelial Mesenchymal Transmission. It is reported in 50% of TNBC (Zundelevich et al., 2020). The present study reveals EGFR as a master regulator that regulates other genes associated with cancer. Thus, targeting this using a drug which fulfils all the ADMET properties can help tackle EGFR aberrant breast cancer.

ESR1 or Estrogen Receptor 1 belong to the nuclear hormone receptor family and are found commonly in metastatic, endocrine therapy resistant cancers (10-50%). ESR1 mutation leads to severe progression of cancer and affects the overall survival of the patient. It functions as transcription unit that is indispensable in cellular multiplication and activation in normal cells. It works along with cyclins and CDKs to regulate cell cycle (Takeshima et al., 2020, Al-Eitan et al., 2019). Overexpression of ESR1 activates a large number of cyclins thus, increasing the replication rate of cell cycle making it tumorous.

Estrogen Receptor 2 (ESR2) is activated by estrogen and interacts with ESR1 in a dimeric manner. Studies have reported the association between breast cancer and ESR2 polymorphism (Borgquist et al., 2013). ESR2 polymorphisms are found to be cancer inducing in overweight women. Although ESR1 and ESR2 are interrelated, they play independent roles in tumour progression and metastasis (Harrel et al., 2017).

KIT is a receptor tyrosine kinase activated by its cognate ligand (KIT ligand) and is involved in the regulation of haematopoiesis (Jansson et al., 2014). Higher expression of KIT has been related to basal-like breast cancer and TNBC. The higher expression of mutated KITs has been reported in various neoplasms including breast cancer, leading to its impaired prognosis and aggressive phenotype (Takeshima et al., 2020).

These genes are found to be involved in cell cycle pathways. Cell cycle is an evolutionarily conserved procedure which is important for proper growth and development of an organism. Cell cycle dysfunction is a hallmark of cancer (Hanahan, 2022). The genes EGFR and KIT has been reported as proto-oncogenes.

5.0 Conclusion

The comparison of high throughput data of breast cancer patients paved way for the identification of up-regulated genes involved in breast cancer progression and metastasis. The gene enrichment analysis of the following genes proved their involvement in cancer. The protein interaction network and the topological scoring models revealed the master regulators of breast cancer. The genes EGFR, ESR1, ESR2, KIT, CXCL2 and DMD as the hubs where EGFR, ESR1, ESR2 and KIT were FDA approved targets. Thus, all the FDA genes reported in this study as the master regulators of breast cancer are found to be involved in the progression and metastasis of breast cancer.

None

Reference

1. Al-Eitan, L. N., Rababah, D. M., Alghamdi, M. A., and Khasawneh, R. H. (2019). Association between ESR1, ESR2, HER2, UGT1A4, and UGT2B7 polymorphisms and breast Cancer in Jordan: a case-control study. BMC cancer, 19(1), 1257. https://doi.org/10.1186/s12885-019-6490-7

2. Aranda, B., Blankenburg, H., Kerrien, S., Brinkman, F. S., Ceol, A., Chautard, E., Dana, J. M., De Las Rivas, J., Dumousseau, M., Galeota, E., Gaulton, A., Goll, J., Hancock, R. E., Isserlin, R., Jimenez, R. C., Kerssemakers, J., Khadake, J., Lynn, D. J., Michaut, M., O'Kelly, G., Hermjakob, H. (2011). PSICQUIC and PSISCORE: accessing and scoring molecular interactions. Nature methods, 8(7), 528–529. https://doi.org/10.1038/nmeth.1637

3. Ariya, S.S., James, A.R., Joseph, B. (2020). Computational Analysis of Oral Cancer Gene Expression Profile and Identification of MiRNAs and their Regulatory Hub Genes. J Complement Med Res, Vol 11, Issue 3

4. Balderas-Hernández, V. E., Alvarado-Rodríguez, M., and Fraire-Velázquez, S. (2013). Conserved versatile master regulators in signalling pathways in response to stress in plants. AoB PLANTS, 5, plt033. https://doi.org/10.1093/aobpla/plt033

5. Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., Marshall, K. A., Phillippy, K. H., Sherman, P. M., Holko, M., Yefanov, A., Lee, H., Zhang, N., Robertson, C. L., Serova, N., Davis, S., & Soboleva, A. (2013). NCBI GEO: archive for functional genomics data sets--update. Nucleic acids research, 41(Database issue), D991–D995. https://doi.org/10.1093/nar/gks1193

6. Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N., and Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. Behavioural brain research, 125(1-2), 279–284. https://doi.org/10.1016/s0166-4328(01)00297-2

7. Borgquist, S., Hjertberg, M., Henningson, M., Ingvar, C., Rose, C., and Jernström, H. (2013). Given breast cancer, is fat better than thin? Impact of the estrogen receptor beta gene polymorphisms. Breast cancer research and treatment, 137(3), 849–862. https://doi.org/10.1007/s10549-012-2367-z

8. Changavi, A. A., Shashikala, A., and Ramji, A. S. (2015). Epidermal Growth Factor Receptor Expression in Triple Negative and Nontriple Negative Breast Carcinomas. Journal of laboratory physicians, 7(2), 79–83. https://doi.org/10.4103/0974-2727.163129

9. Chin, C.H., Chen, S.H., Wu, H.H., Ho, C.W., Ko, M.T., and Lin, C.Y., (2014). cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst. Biol. 8 (4), 1–7.

10. Croft, D., Kelly, G. O., Wu, G., Haw, R., Gillespie, M., and Matthews, L. (2011). Reactome: a database of reactions, pathways and biological processes. Nucleic Acids Res. 39, 691–97.

11. De Cicco, P., Catani, M. V., Gasperi, V., Sibilano, M., Quaglietta, M., and Savini, I. (2019). Nutrition and Breast Cancer: A Literature Review on Prevention, Treatment and Recurrence. Nutrients, 11(7), 1514. https://doi.org/10.3390/nu11071514

12. Gevaert, O., and Plevritis, S. (2013). Identifying master regulators of cancer and their downstream targets by integrating genomic and epigenomic features. Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing, 123–134.

13. Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. Cell, 144(5), 646–674. https://doi.org/10.1016/j.cell.2011.02.013

14. Harrell, J. C., Shroka, T. M., and Jacobsen, B. M. (2017). Estrogen induces c-Kit and an aggressive phenotype in a model of invasive lobular breast cancer. Oncogenesis, 6(11), 396. https://doi.org/10.1038/s41389-017-0002-x

15. Jansson, S., Bendahl, P. O., Grabau, D. A., Falck, A. K., Fernö, M., Aaltonen, K., and Rydén, L. (2014). The three receptor tyrosine kinases c-KIT, VEGFR2 and PDGFRα, closely spaced at 4q12, show increased protein expression in triple-negative breast cancer. PloS one, 9(7), e102176. https://doi.org/10.1371/journal.pone.0102176

16. Kadara, H., Fujimoto, J., Yoo, S. Y., Maki, Y., Gower, A. C., Kabbout, M., Garcia, M. M., Chow, C. W., Chu, Z., Mendoza, G., Shen, L., Kalhor, N., Hong, W. K., Moran, C., Wang, J., Spira, A., Coombes, K. R., and Wistuba, I. I. (2014). Transcriptomic architecture of the adjacent airway field cancerization in non-small cell lung cancer. Journal of the National Cancer Institute, 106(3), dju004. https://doi.org/10.1093/jnci/dju004

17. Li, M., Guo, Y., Feng, Y. M., and Zhang, N. (2019). Identification of Triple-Negative Breast Cancer Genes and a Novel High-Risk Breast Cancer Prediction Model Development Based on PPI Data and Support Vector Machines. Frontiers in genetics, 10, 180. https://doi.org/10.3389/fgene.2019.00180

18. Lindskog C. (2016). The Human Protein Atlas - an important resource for basic and clinical research. Expert review of proteomics, 13(7), 627–629. https://doi.org/10.1080/14789450.2016.1199280

19. Masuda, H., Zhang, D., Bartholomeusz, C., Doihara, H., Hortobagyi, G. N., and Ueno, N. T. (2012). Role of epidermal growth factor receptor in breast cancer. Breast cancer research and treatment, 136(2), 331–345. https://doi.org/10.1007/s10549-012-2289-9

20. Ohno, S. (1979). Major sex-determining genes. Springer-Verlag.

21. Pan, Z., Li, L., Fang, Q., Qian, Y., Zhang, Y., Zhu, J., Ge, M., and Huang, P. (2019). Integrated Bioinformatics Analysis of Master Regulators in Anaplastic Thyroid Carcinoma. BioMed research international, 9734576. https://doi.org/10.1155/2019/9734576

22. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., and Ramage, D., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13 (11), 2498–2504.

23. Stark, C., Breitkreutz, B.J., Reguly, T., Boucher, L., Breitkreutz, A., and Tyers, M. (2006). Bio GRID: a general repository for interaction datasets. Nucleic Acids Res. 1 (34), 535–539.

24. Sehgal.P, Kumar.B, Sharma.M, Salameh A.A, Kumar.S, Asha.P (2022), Role of IoT In Transformation Of Marketing: A Quantitative Study Of Opportunities and Challenges, Webology, Vol. 18, no.3, pp 1-11

25. Kumar, S. (2020). Relevance of Buddhist Philosophy in Modern Management Theory. Psychology and Education, Vol. 58, no.2, pp. 2104–2111.

26. Roy, V., Shukla, P. K., Gupta, A. K., Goel, V., Shukla, P. K., & Shukla, S. (2021). Taxonomy on EEG Artifacts Removal Methods, Issues, and Healthcare Applications. Journal of Organizational and End User Computing (JOEUC), 33(1), 19-46. http://doi.org/10.4018/JOEUC.2021010102

27. Shukla Prashant Kumar, Sandhu Jasminder Kaur, Ahirwar Anamika, Ghai Deepika, Maheshwary Priti, Shukla Piyush Kumar (2021). Multi objective Genetic Algorithm and Convolutional Neural Network Based COVID-19 Identification in Chest X-Ray Images, Mathematical Problems in Engineering, vol. 2021, Article ID 7804540, 9 pages. https://doi.org/10.1155/2021/7804540

28. Takeshima, K., Hayashida, T., Maeda, H., Nakashoji, A., Yokoe, T., Seki, T., Takahashi, M., and Kitagawa, Y. (2020). Increased frequency of ESR1 mutation in metastatic breast cancer by dosing selective estrogen receptor modulator followed by aromatase inhibitor. Oncology letters, 20(2), 1231–1238. https://doi.org/10.3892/ol.2020.11669

29. Zundelevich, A., Dadiani, M., Kahana-Edwin, S., Itay, A., Sella, T., Gadot, M., Cesarkas, K., Farage-Barhom, S., Saar, E. G., Eyal, E., Kol, N., Pavlovski, A., Balint-Lahat, N., Dick-Necula, D., Barshack, I., Kaufman, B., and Gal-Yam, E. N. (2020). ESR1 mutations are frequent in newly diagnosed metastatic and loco-regional recurrence of endocrine-treated breast cancer and carry worse prognosis. Breast cancer research: BCR, 22(1), 16. https://doi.org/10.1186/s13058-020-1246-5