International Journal on Emerging Technologies 13(2): 01-05(2022)

ISSN No. (Print): 0975-8364
ISSN No. (Online): 2249-3255

Data Curation of Signalling Protein Molecules in Breast Cancer

Piyusha Sharma[1*] and Vachaspati Mishra[2]
[1]Assistant Professor, Department of Life Sciences,
Desh Bhagat University, Mandi Gobindgarh (Punjab), India.
[2]Professor, Department of Molecular Oncology,
Lovely professional University, Jalandhar (Punjab), India.

(Corresponding author: P. Sharma*)
(Received 15 May 2022, Revised 25 June 2022, Accepted 10 July 2022)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: Breast cancer is fundamentally a disease of regulation of tissue growth. Targeting the pathways that promote or sustain growth and invasion of carcinoma cells is critical to effective treatment of breast cancer. Data curation adopting a bioinformatics/ biological data basing approach is presently employed to identify the candidate protein molecules mediating breast cancer signalling with a promise for therapeutic targeting. Signalling proteins are highly specific and precise and act in a modular manner, which allows us to modulate such interactions specifically by the use of drugs or other molecule for a better drug target. With this aim in mind, we thus, are creating a catalogue of biomarkers of signalling proteins mediating breast cancer. Presently, we have manually curated signalling data on proteins mediating breast cancer on the basis of PubMed reports. Of the 150 entries made so far based on about 250 Pubmed reports, the receptor tyrosine kinase (RTK) cascade has been found to be implicated in a high proportion of Breast cancers and the major downstream effector in the process emerged out to be Activator protein-1 (AP-1), CDc42 and FOXO3 in these responses. Other key markers found out were CDc42 (a GTPase) and FOXO3 (a transcription factor). Receptor mainly involved is estrogen receptors which in particular have been well documented to play a critical role in the etiology and progression of breast cancer. This collection highlights many recent studies that catalogue these alterations and shed light on the mechanisms by which cancer genes function.

Keywords: Breast cancer, Oncology, Bioinformatics, AP-1, FOXO3, CDc42, PubMed.

I. INTRODUCTION

With 9.9 million deaths expected from the complicated disease in 2020, cancer is the leading cause of morbidity and mortality worldwide. In most nations, by the age of 70, it is also the leading cause of death. According to estimates, there will be 28.4 million new cases of cancer worldwide in 2040, which is a rise of almost 47% from the global average of 2020 [16]. With 2,261,419 new instances of breast cancer detected in 2020, accounting for 1 in 4 cancer cases, breast cancer remained the most common cancer among women worldwide and has since surpassed lung cancer in terms of incidence. In addition, female breast cancer has the highest mortality rate, accounting for 1 in 6 cancer fatalities worldwide in 2020 with 684,996 deaths [26]. Although there are regional variations in breast cancer incidence rates, Australia and New Zealand continue to have substantially higher rates (95.5 per 100,000), there have been noticeable increases in Africa and Asia. In recent years, there has been a historically low occurrence in Asian nations like Japan and Korea [17]. According to the expression of hormone and growth factor receptors, it is a highly heterogeneous disease. Patients with hormone-independent HER2 overexpressing subtypes of triple negative breast cancer (TNBC) and oestrogen, progesterone, and human epidermal growth factor receptor (HER2) negative TNBC all have relatively aggression, metastases, a poor prognosis, and medication resistance [11]. Receptors involved in breast cancer are estrogen receptor, oestrogen receptor, epidermal growth factor receptor etc. Research in the human breast field regarding the control of proliferation has stressed the functional implication of oestrogens, progesterone [2], epidermal growth factor [29], insuline-like growth factor [6], fibroblast growth factor, nerve growth factor, which are mitogenic for cancerous and not for normal cells [13]. Insulin-like growth factor (IGF)-II is a required intermediate for prolactin-induced up-regulation of cyclin D1 and proliferation in normal murine mammary epithelial cells in vivo and in vitro. Cross-talk between insulin-like growth factor 1 (IGF-1) and estrogen receptor alpha (ER) regulates gene expression in breast cancer cells [7]. The epidermal growth factor (EGF) and insulin-like growth factor (IGF) signaling pathways are critically involved in cancer development and progression [8]. In addition to the classical nuclear

estrogen receptors (ERs)-alpha and -beta, estrogen also signals through the seven-transmembrane G-protein-coupled receptor (GPR)-30 [2]. GPCR ligands bradykinins (BK) are inflammatory molecules that have been previously reported to contribute to the proliferation breast cancer cells [13]. The activating protein 1 (AP-1) transcription factor, a known regulator of processes essential for normal growth and development as well as carcinogenesis, is a potential site for cross-talk between these hormones in breast cancer cells [15]. The transcription factor AP-1, composed principally of Fos and Jun heterodimers, is required for cell invasion in vitro [1-5, 12], and in vivo [33]. In breast cells, the previous studies have suggested that growth factors and hormones, such as IGF, EGF, estrogens and retinoids, can modulate AP-1 transcriptional activity [24-4-28-19]. Other studies demonstrate that ER and AP-1 interact to regulate the expression of certain estrogen- and tamoxifen-regulated genes [20]. Activation of AP-1 may also contribute to tumor cell invasive capacity and to tamoxifen resistance [23-33-18-25]. One another signalling candidate molecule is Cdc42 (cell division cycle 42, GTP binding protein) is a Rho-GTPase that transduces extracellular signals from G-coupled protein receptors (GPCR) [10]. These small G-proteins, which include Rac1 and Cdc42, are well known for their ability to modulate and rearrange the actin cytoskeleton [22]. The FOXO classes of forkhead proteins are downstream targets of the phosphatidylinositol-3-kinase (PI3K)/Akt pathway [3]. The activity of FoxO3a is negatively regulated by Akt phosphorylation [21]. Each molecular subtype of breast cancer was examined for NOTCH-1, NOTCH-2, and NOTCH-3 expression. Comparing the basal/claudin-low subtype to other subtypes, NOTCH-1 levels are considerably greater in the basal/claudin-low subtype (P-value 0.000Another E3 ligase family that controls the Notch signalling pathway is called SKP1-CUL1-F-box [34]. With the rapidly increasing amount of molecular biological data available, the computer-based analysis of molecular interactions becomes more and more feasible [5]. Furthermore, sophisticated online cheminformatics tools are available for processing chemical structures, predicting properties, and generating 2D or 3D structure representations often required prior to more advanced analyses [9].The prediction of prognosis of tumors at molecular level is now possible due to the availability of data mining and other analysis tools and software's [31]. The anti-cancer treatment can be determined by using prognostic biomarkers as it provides accurate clinical estimation. The prognostic biomarkers have become an important part of precision medicine [32]. By data mining the information stored in the databases researchers can analyze the signalling between the protein molecules and look for prognostic biomarkers [30]. There are bioinformatics tools that are included in the LOGpc platform used for survival analysis of leiomyosarcoma, esophageal squamous cell carcinoma, kidney renal clear cell carcinoma, bladder cancer, cervical cancer, breast cancer, adrenocortical carcinoma, and uveal melanoma [27].

II. MATERIAL AND METHODS

Despite the fact that current research activities generate a vast amount of research data, a lack of knowledge about documentation, secure storage alternatives, and making research data easily accessible through research data repositories has led to the loss and/or rejection of priceless research data. The main goal of the current study is to showcase the top data curation techniques used in research data repositories. The study also provides a summary of breast cancer-related signalling molecules. On the basis of the scientific literature found in PubMed, information was carefully gathered about the signalling proteins that mediate breast cancer. A dataset is being created in which the information is being arranged under headings of signalling pathway involved, receptor involved, critical molecular interactions involved, induced/deinduced molecules involved and PubMed links to the relevant PubMed entries for reference. For studying candidate signalling proteins, critical responses were considered. The responses that were considered included upstream cascades and downstream molecules. A general protocol for data curation was followed which is given in Fig.1 and criteria followed for listing signalling molecules is given in Fig. 2. Based on the data curated cancer causing molecules enlisted in the literature were retrieved. Based on the repetition of their occurrence in maximum number of cancer research literature they were shortlisted and top three molecules were selected for further analysis.

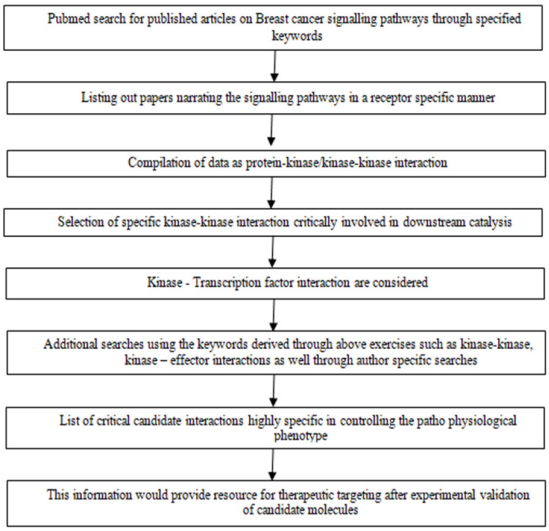

Fig. 1. Protocol for data curation for cataloguing the signalling pathway molecules.

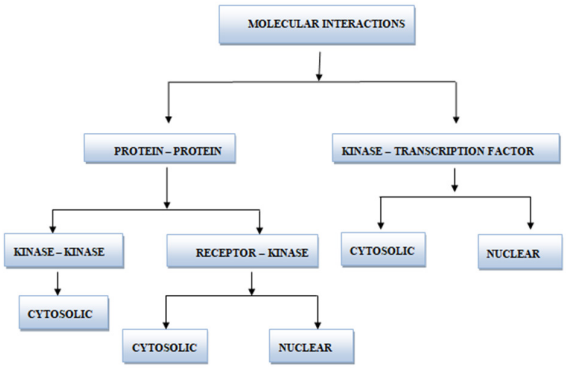

Fig. 2. Criteria followed for listing of signalling proteins associated with Breast Cancer.

III. RESULTS AND DISCUSSION

With a large number of studies of the scientific literature, a keyword search for "cell signalling in breast cancer" in PUBMED (http://www.ncbi.nlm.nih.gov/pubmed/) search engine, fetches 764 articles published. Therefore we firstly identified the relevant publications including cell signalling. Of the available 261 articles, about 250 articles were read and the currently curated dataset includes 150 entries derived from literature. Our finding presently envisages Receptor Tyrosine Kinase (RTK) cascade as the most common mediator of Breast cancers. The major downstream effector in the process emerged out to be Activator protein-1 (AP-1), cell division control protein 42 (CDc42), a GTPase and Forkhead box O3 (FOXO3a), transcription factor in these responses. .Receptor mainly involved is estrogen receptors which in particular have been well documented to play a critical role in the etiology and progression of breast cancer. Estrogen receptors are over-expressed in around 80% of breast cancer cases. At present, AP-1 seems to be a global biomarker in RTK-mediated response (55% direct cases including 10% indirect cases). Other key markers found out were CDc42 (35% cases) and FOXO3a (30% cases). Below is table: 1 showing this data followed by the graphical representation of the same.

Table 1: Data Annotated.

Sr. No.	Candidate molecule	No. of Cases
1.	AP-1	65%
2.	AP-1 directly involved	55%
3.	AP-1 indirectly through cross talk	10%
4.	CDc42	35%
5.	FOXO3a	30%

The breast is a challenging tissue to research. It is the main cause of death for women in many nations and is a multidisciplinary issue with no boundaries in terms of geography. There hasn't been a convenient forum for debating and settling ongoing disputes around breast cancer treatment. As a result, we are currently working on a project to compile a dataset describing important signalling molecules linked to breast cancer that may help us find new therapeutic targets and resistance mechanisms. Receptor Tyrosine Kinase (RTK) cascade is thought to be the most frequent mediator of breast cancer, according to our current conclusion, which is based on the study of studies available in PubMed. Forkhead box O3 (FOXO3a), a transcription factor in these reactions, and Activator protein-1 (AP-1), a cell division control protein 42 (CDc42), a GTPase, turned discovered to be the process' main downstream effectors. Estrogen receptors, in particular, have been well shown to play a vital role in the aetiology and progression of breast cancer. They are the principal receptor involved in the aforementioned cases. As a result, oestrogen receptors were chosen as the optimal target for the docking method in the current investigation. Cell signalling is a more effective method of understanding disease mechanism while trying to identify preventative factors. Structure-based drug design (SBDD) is an approach that is becoming more and more well-liked for finding new lead compounds. Various bioinformatical tools can be used to create drugs based on their structure.

IV. CONCLUSIONS

One of the most important issues in oncology is breast cancer, which is a challenging tissue to investigate. The dataset curated in the current study includes important signalling molecules linked to breast cancer. Receptor Tyrosine Kinase (RTK) cascade is currently thought to be the most frequent mediator of breast cancer. Activator protein-1 (AP-1), cell division control protein 42 (CDc42), a GTPase, and Forkhead box O3 (FOXO3a), a transcription factor, were found to be the process' main downstream effectors. Since it is clear that these prospective protein molecules have been shown to have a major role in the majority of cancer cases, a novel lead compound can be found through structure-based drug design employing a variety of bioinformatics tools. The identification of potential drug lead compounds for inhibiting the active site of protein molecules and preventing their involvement in the cancer pathway may be aided by In silico drug target discovery, drug design, docking or screening, drug metabolism prediction, interaction prediction, and general pharmaceutical education. The quick search for tiny compounds that might bind to biologically interesting targets is essential for the drug discovery process.

Author contributions. Dr. Vachaspati Mishra conceived, designed the analysis, Dr. Piyusha Sharma collected the data, contributed data or analysis tools, performed the analysis and wrote the paper.

Acknowledgement. I extend my sincere thanks to Dr. Vachaspati Mishra, Professor Molecular Oncology

Lovely Professional University and Dr. Deepinderpal Singh for giving me proper guidance throughout the course of study.
Conflict of Interest. None.

REFERENCES

[1]. AiPing, Lu., Zhang, F., Gupta, A. and Liu, J. (2002). Blockade of AP1 transactivation abrogates the abnormal expression of breast cancer-specific gene 1 in breast cancer cells. *Journal of Biological Chemistry*, *277*(35): 31364-31372.
[2]. Albanito, L., Sisci, D., Aquila, S., Brunelli, E., Vivacqua, A., Madeo, A. and Maggiolini, M. (2008). Epidermal growth factor induces G protein-coupled receptor 30 expression in estrogen receptor-negative breast cancer cells. *Endocrinology*, *149*(8): 3799-3808.
[3]. Andrew, S., Madureira, P. A., Pomeranz, K. M., Aubert, M., Brosens, J. J., Cook, S. J. and Lam, E. W. F. (2006). Paclitaxel-induced nuclear translocation of FOXO3a in breast cancer cells is mediated by c-Jun NH2-terminal kinase and Akt. *Cancer research*, *66*(1): 212-220.
[4]. Cai, D., Iyer, A., Felekkis, K. N., Near, R. I., Luo, Z., Chernoff, J. and Lerner, A. (2003). AND-34/BCAR3, a GDP exchange factor whose overexpression confers antiestrogen resistance, activates Rac, PAK1, and the cyclin D1 promoter. *Cancer research*, *63*(20): 6802-6808.
[5]. Arnott, C. H., Scott, K. A., Moore, R. J., Hewer, A., Phillips, D. H., Parker, P. and Owens, D. M. (2002). Tumour necrosis factor-α mediates tumour promotion via a PKCα-and AP-1-dependent pathway. *Oncogene*, *21*(31): 4728-4738.
[6]. Carver, K. C. and Schuler, L. A. (2008). Prolactin does not require insulin-like growth factor intermediates but synergizes with insulin-like growth factor I in human breast cancer cells. *Molecular Cancer Research*, *6*(4): 634-643.
[7]. Cascio, S., Bartella, V., Garofalo, C., Russo, A., Giordano, A. and Surmacz, E. (2007). Insulin-like growth factor 1 differentially regulates estrogen receptor-dependent transcription at estrogen response element and AP-1 sites in breast cancer cells. *Journal of Biological Chemistry*, *282*(6): 3498-3506.
[8]. Cui, X., Kim, H. J., Kuiatse, I., Kim, H., Brown, P. H. and Lee, A. V. (2006). Epidermal growth factor induces insulin receptor substrate-2 in breast cancer cells via c-Jun NH2-terminal kinase/activator protein-1 signaling to regulate cell migration. *Cancer research*, *66*(10): 5304-5313.
[9]. Ertl, P., Muehlbacher, J., Rohde, B. and Selzer, P. (2003). Web-based cheminformatics and molecular property prediction tools supporting drug design and development at Novartis. *SAR and QSAR in Environmental Research*, *14*(5-6): 321-328.
[10]. Etienne-Manneville, S. and Hall, A. (2002). Rho GTPases in cell biology. *Nature*, *420*(6916): 629-635.
[11]. Farghadani, R., and Naidu, R. (2021). Curcumin: modulator of key molecular signaling pathways in hormone-independent breast cancer. *Cancers*, *13*(14): 3427.
[12]. Gotoh, T., Cai, D., Tian, X., Feig, L. A.and Lerner, A. (2000). p130Cas regulates the activity of AND-34, a novel Ral, Rap1, and R-Ras guanine nucleotide exchange factor. *Journal of Biological Chemistry*, *275*(39): 30118-30123.
[13]. Greco, S., Elia, M. G., Muscella, A., Romano, S., Storelli, C. and Marsigliante, S. (2005). Bradykinin stimulates cell proliferation through an extracellular-regulated kinase 1 and 2-dependent mechanism in breast cancer cells in primary culture. *Journal of endocrinology*, *186*(2): 291-302.
[14]. Greco, S., Muscella, A., Elia, M. G., Romano, S., Storelli, C. and Marsigliante, S. (2004). Mitogenic signalling by B2 bradykinin receptor in epithelial breast cells. *Journal of cellular physiology*, *201*(1): 84-96.
[15]. Gutzman, J. H., Nikolai, S. E., Rugowski, D. E., Watters, J. J. and Schuler, L. A. (2005). Prolactin and estrogen enhance the activity of activating protein 1 in breast cancer cells: role of extracellularly regulated kinase 1/2-mediated signals to c-fos. *Molecular endocrinology*, *19*(7): 1765-1778.
[16]. Heer, E., Harper, A., Escandor, N., Sung, H., McCormack, V. and Fidler-Benaoudia, M. M. (2020). Global burden and trends in premenopausal and postmenopausal breast cancer: a population-based study. *The Lancet Global Health*, *8*(8): e1027-e1037.
[17]. Joko-Fru, W. Y., Jedy-Agba, E., Korir, A., Ogunbiyi, O., Dzamalala, C. P., Chokunonga, E. and Parkin, D. M. (2020). The evolving epidemic of breast cancer in sub-Saharan Africa: Results from the African Cancer Registry Network. *International Journal of Cancer*, *147*(8): 2131-2141.
[18]. Jorgensen, W. L. (2004). The many roles of computation in drug discovery. *Science*, *303*(5665): 1813-1818.
[19]. Lin, R., Bagrodia, S., Cerione, R. and Manor, D. (1997). A novel Cdc42Hs mutant induces cellular transformation. *Current Biology*, *7*(10): 794-797.
[20]. Morelli, C., Lanzino, M., Garofalo, C., Maris, P., Brunelli, E., Casaburi, I. and Ando, S. (2010). Akt2 inhibition enables the forkhead transcription factor FoxO3a to have a repressive role in estrogen receptor α transcriptional activity in breast cancer cells. *Molecular and cellular biology*, *30*(3): 857-870.
[21]. Patricia A. Madureira., Varshochi, R., Constantinidou, D., Francis, R. E., Coombes, R. C., Yao, K. M. and Lam, E. W. F. (2006). The Forkhead box M1 protein regulates the transcription of the estrogen receptor α in breast cancer cells. *Journal of Biological Chemistry*, *281*(35): 25167-25176.

[22]. Robert F. K., De Marco, P. C., Salasznyk, R. M., Ahuja, D., Hogg, M., Antoniotti, S. and Plopper, G. E. (2006). Apocynin derivatives interrupt intracellular signaling resulting in decreased migration in breast cancer cells. *Journal of Biomedicine and Biotechnology*, 1-10.
[23]. Sawhney, R. S., Liu, W. and Brattain, M. G. (2009). A novel role of ERK5 in integrin-mediated cell adhesion and motility in cancer cells via Faksignaling. *Journal of cellular physiology*, *219*(1): 152-161.
[24]. Shaulian, E. and Karin, M. (2001). AP-1 in cell proliferation and survival. *Oncogene*, *20*(19): 2390-2400.
[25]. Sobolev, V., Sorokine, A., Prilusky, J., Abola, E. E., and Edelman, M. (1999). Automated analysis of interatomic contacts in proteins. *Bioinformatics (Oxford, England)*, *15*(4): 327-332.
[26]. Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A. and Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, *71*(3): 209-249.
[27]. Wang, F., Wang, Q., Li, N., Ge, L., Yang, M., An, Y. and Guo, X. (2020). OSuvm: An interactive online consensus survival tool for uveal melanoma prognosis analysis. *Molecular Carcinogenesis*, *59*(1): 56-61.
[28]. Webb, P., Nguyen, P., Valentine, C., Lopez, G. N., Kwok, G. R., McInerney, E. and Kushner, P. J. (1999). The estrogen receptor enhances AP-1 activity by two distinct mechanisms with different requirements for receptor transactivation functions. *Molecular endocrinology*, *13*(10): 1672-1685.
[29]. Xiaoming Hu., Flaws, J. A., Sipes, I. G. and Hoyer, P. B. (2002). Activation of mitogen-activated protein kinases and AP-1 transcription factor in ovotoxicity induced by 4-vinylcyclohexene diepoxide in rats. *Biology of reproduction*, *67*(3): 718-724.
[30]. Xie, L., Dang, Y., Guo, J., Sun, X., Xie, T., Zhang, L. and Guo, X. (2019). High KRT8 expression independently predicts poor prognosis for lung adenocarcinoma patients. *Genes*, *10*(1): 36.
[31]. Xu, X. L., Gong, Y. and Zhao, D. P. (2018). Elevated PHD2 expression might serve as a valuable biomarker of poor prognosis in lung adenocarcinoma, but no lung squamous cell carcinoma. *Eur Rev Med Pharmacol Sci.*, *22*(24): 8731-8739.
[32]. Yang, J., Li, A., Li, Y., Guo, X. and Wang, M. (2019). A novel approach for drug response prediction in cancer cell lines via network representation learning. *Bioinformatics*, *35*(9): 1527-1535.
[33]. Yang, J. Y. and Hung, M. C. (2009). A new fork for clinical application: targeting forkhead transcription factors in cancer. *Clinical cancer research*, *15*(3): 752-757.
[34]. Yousefi, H., Bahramy, A., Zafari, N., Delavar, M. R., Nguyen, K., Haghi, A. and Babashah, S. (2022). Notch signaling pathway: a comprehensive prognostic and gene expression profile analysis in breast cancer. *BMC cancer*, *22*(1): 1282.

How to cite this article: Piyusha Sharma and Vachaspati Mishra (2022). Data Curation of Signalling Protein Molecules in Breast Cancer. *International Journal on Emerging Technologies*, *13*(2): 01–05.