ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Exploring the Molecular Link between Breast and Ovarian Cancer through Protein-Protein Interaction Network Analysis.

Pratiksha Bhoi, Anjali Nalawade, Gauri Shinde, Preeti Mate, and Sharddha Ranpise*.

Dr. D. Y. Patti Arts, Commerce and Science College, Sant Tukaram Nagar, Pimpri, Pune, Maharashtra, India, 411018.

ABSTRACT

Breast and ovarian cancers are leading causes of cancer-related deaths among women worldwide. Both malignancies originate from epithelial tissue and share genetic mutations, primarily in BRCA1 and BRCA2. Understanding the common genetic interactions between these cancers is crucial for targeted therapeutic interventions. This study employs a bioinformatics approach to identify and analyze shared genes involved in breast and ovarian cancer. Using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database, 22 common genes were identified. Protein-protein interaction (PPI) networks were constructed using STRING and visualized in Cytoscape. Key topological parameters, including degree and betweenness centrality, were analyzed to determine hub genes. EGFR exhibited the highest connectivity, indicating its pivotal role in oncogenic pathways. Further clustering analysis using MCODE revealed a tightly interconnected module enriched in signaling and regulatory functions. Gene ontology (GO) and pathway enrichment analysis using ClueGO identified significant biological pathways, such as phosphatidylinositol 3-kinase signaling and Fc receptor signaling, associated with cancer progression. These findings enhance our understanding of common molecular mechanisms in breast and ovarian cancer and provide potential targets for therapeutic intervention.

Keywords: Breast Cancer, Ovarian Cancer, Protein-Protein Interaction Network

**Corresponding author*

ISSN: 0975-8585

INTRODUCTION

Epithelial ovarian cancer is the most common cause of gynecological cancer-associated death (Gordon et al.). It accounts for approximately 4% of all cancers in women and is the leading cause of death from gynecologic malignancies. One of the major challenges in diagnosing ovarian cancer is that early-stage disease is generally asymptomatic (Nicoletta Colombo et al.). Genes associated with a high risk of ovarian cancer include BRCA1 and BRCA2. Standard treatments for newly diagnosed ovarian cancer consist of cytoreductive surgery and platinum-based chemotherapy. In cases of recurrent cancer, treatment options include chemotherapy, anti-angiogenic agents, and poly (ADP-ribose) polymerase (PARP) inhibitors. Additionally, immunological therapies are currently being tested (Ursula A. Matulonis et al.). High-grade serous carcinoma (HGSC) is the most commonly diagnosed form of ovarian cancer, and at the time of diagnosis, it is typically very responsive to platinum-based chemotherapy.

In the United States, approximately 14% of all cancer diagnoses in women are breast cancer, making it the second leading cause of cancer-related death among women. Significant genes involved in breast cancer progression include BRCA1, RB1, TP53, PTEN, AKT1, CDH1, GATA3, and PIK3CA. Mutations in these genes can lead to disruptions in apoptosis, cell-cycle regulation, and transcriptional regulation. Standard treatments for breast cancer include chemotherapy, radiotherapy, immunotherapy, targeted therapies, and surgery (Mona Zamanian-Azodi et al.).

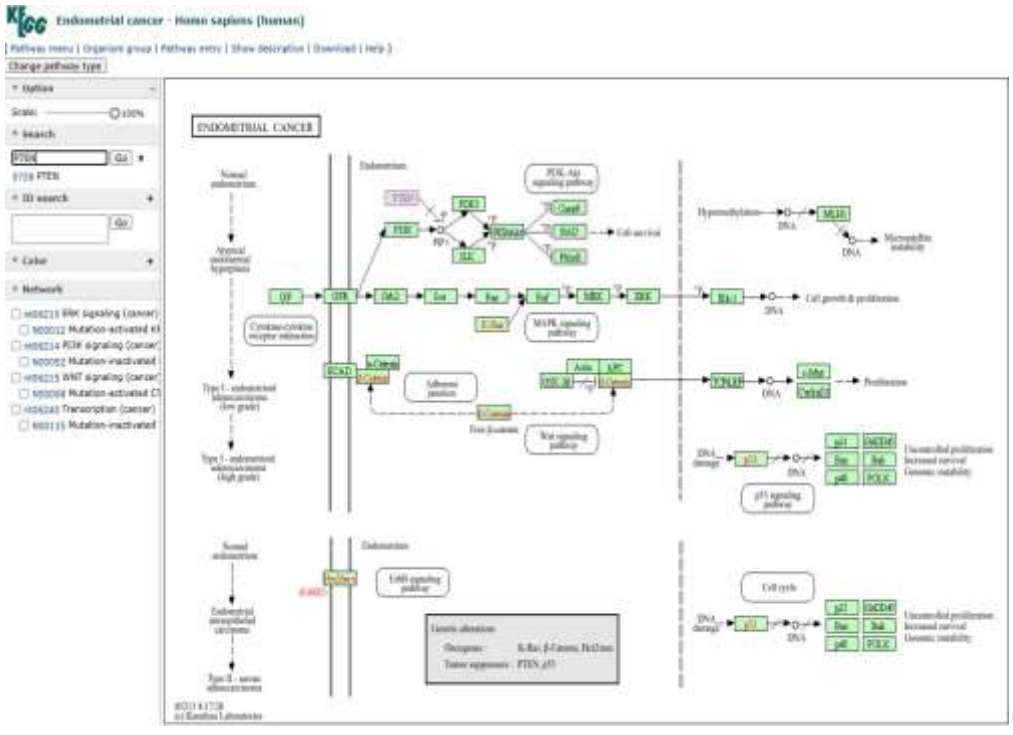
Both breast and ovarian cancers originate from epithelial tissue, which is highly susceptible to tumor development (Mona Zamanian-Azodi and Ayşe Çelik et al.). Several studies have aimed to identify common molecular mechanisms between breast and ovarian cancer, with many highlighting mutations in BRCA1 and BRCA2 as key genetic alterations in both diseases. Protein–protein interactions (PPIs) play a crucial role in coordinating various cellular processes (Karthik Raman et al.). Recently, PPI networks have gained significant attention due to their powerful ability to interpret disease-associated biological phenomena (Zamanian- Azodi et al.).

MATERIALS AND METHODS

Identification of Common Genes

Common gernes of breast and ovarian cancer for KEGG Pathway downloaded from websites (http://www.genome.jp /kegg/pathway.html) and compared manually. Twenty-two genes were identified.

Figure 1: Breast cancer Pathway



Protein Accession Number Retrieval

Uniprot accession number of selected genes was retrieved from (uniprot.org). The UniProt database was used to retrieve the accession numbers, protein names, and functional descriptions.

Protein-Protein Interaction (PPI) Network Construction

To explore the functional interaction among proteins and gain insights into their biological roles, PPl network was constructed using the STRING database.

Cytoscape is used for PPI network analysis.

The Cytoscape platform actively supports the development of plug-in tools that extend the core functionality. Topological centralities (degree and betweenness centrality) were evaluated to distinguish the biological value of genes, pathways and clusters. The number of edges that are connected to a designated node is the degree. The high degree indicates the significance of the gene in biological interactions, known as hub. In addition, the number of shortest paths that pass through each node implies betweenness centrality value.

Cluster Identification

Molecular complex detection (MCODE) is a useful method to identify clusters of highly connected nodes.Compute MCODE scores to assess cluster significance.

Functional Enrichment Analysis

Perform Gene Ontology (GO) and pathway enrichment analysis using ClueGO in Cytoscape. Interpret the biological significance of enriched pathways and gene functions (Mona Zamanian-Azodi et al.).

RESULTS AND DISCUSSIONS

Identification of common genes

A total of 22 common genes were identified between breast and ovarian cancer based on KEGG pathway analysis, as listed in the following table (Table 1).

Table 1: A number of twenty-two common genes between breast and ovarian cancer derived from KEGG pathway.

Gene Name	Protein Name	Uniport Accession ID
BRCA1	Breast cancer type 1 susceptibility protein	P38398
BRCA2	Breast cancer type 2 susceptibility protein	P51587
PTEN	Phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase and dual-specificity protein phosphatase PTEN	P60484
mTOR	Serine/threonine-protein kinase mTOR	P42345
EGF	Pro-epidermal growth factor	P01133
CASP9	Caspase-9	P55211
BAD	Bcl2-associated agonist of cell death	Q92934
MAPK1	Mitogen-activated protein kinase 1	P28482
AKT1	RAC-alpha serine/threonine-protein kinase	P31749
EGFR	Epidermal growth factor receptor	P00533
CCND1	G1/S-specific cyclin-D1	P24385
IGF1	Insulin-like growth factor 1 receptor	P08069
KIT	Mast/stem cell growth factor receptor Kit	P10721
GRB2	Growth factor receptor-bound protein 2	P62993
FCGR1A	High affinity immunoglobulin gamma Fc receptor I	P12314

ISSN: 0975-8585

XYLT2	Xylosyltransferase 2	Q9H1B5
MAP2K7	Dual specificity mitogen-activated protein kinase kinase 7	O14733
RSPO4	R-spondin-4	Q2I0M5
FRAT2	GSK-3-binding protein FRAT2	O75474
ZHX2	Zinc fingers and homeoboxes protein 2	Q9Y6X8
HTRA1	Serine protease HTRA1	Q92743
FGFR1	Fibroblast growth factor receptor 1	P1136

Protein Accession Number Retrieval

The UniProt database was used to retrieve the accession numbers, protein names, and functional descriptions of the 22 common genes identified between breast and ovarian cancer. These details provide a deeper understanding of the biological roles of each gene product.

Protein-Protein Interaction (PPI) Network Construction

RSPO4 is a Wnt signaling activator. Wnt signaling is a major oncogenic pathway: overactivation of Wnt can lead to breast cancer, ovarian cancer, colorectal cancer, and leukemia. FRAT2 is a **Wnt/β-catenin signaling regulator.** FRAT2 inhibits **GSK3β**, preventing β-catenin degradation and enhancing Wnt signaling. HTRA1 is a **tumor suppressor gene.** HTRA1 is a serine protease that regulates **TGF-β signaling**, cell migration, and apoptosis (Fig 2).

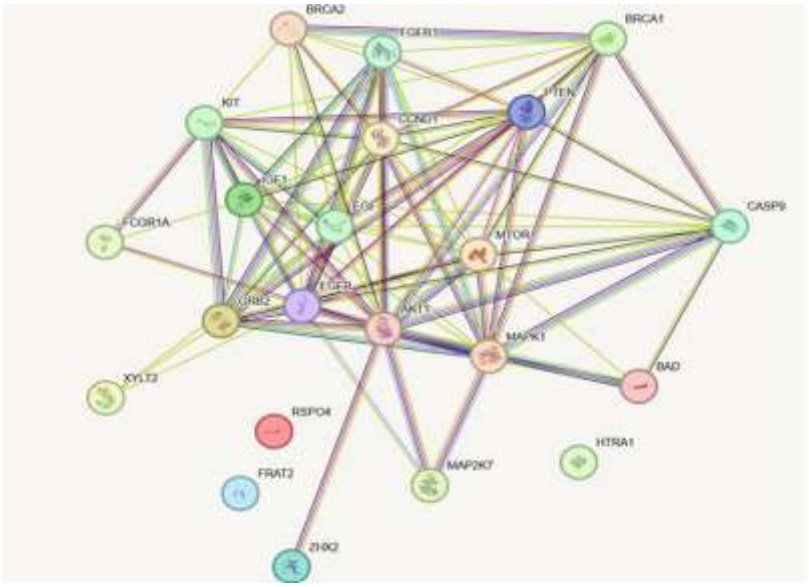


Figure 2: This PPI network consists of 22 nodes and 96 edges. A number of Twenty-two genes are connected to each other except, RSPO4, HTRA1, FRAT2.

Cytoscape is used for PPI network analysis.

EGFR (Epidermal Growth Factor Receptor) has the highest degree (16) and betweenness centrality (0.1004), suggesting it is a hub protein and plays a crucial role in signal transduction pathways, especially in cancer-related processes. EGF (Epidermal Growth Factor) is also highly connected (degree = 14), supporting its role in activating EGFR. GRB2 (Growth Factor Receptor-Bound Protein 2) is moderately connected and serves as an adapter protein, helping in EGFR-mediated signaling. XYLT2 (Xylosyltransferase 2) has the lowest connectivity, indicating it may not be as central in this network (Table2).

ISSN: 0975-8585

Table 2: Degree and betweeness centrality value derived from Cytoscape.

Gene Name	Degree	Betweeness centrality
EGFR	16	0.10047817106640637
EGF	14	0.03947599241716889
GRB2	11	0.02277018551528355
XYLT2	3	0.0

Cluster Identification

Genes in the clusters are as follows: MAPK1, KIT, AKT1, MTOR, PTEN, BRCA1, BRCA2, FGFR1, EGFR, CCND1, CASP9, EGF, IGF1, GRB2, IGF1. A higher score indicatesa tightly connected cluster, suggesting a biologically relevant protein module. It represents a critical signaling pathway, regulatory module, or protein complex (Fig 3).

Figure 3: Shows the MCODE algorithm analysis based on the number of interconnections in the large network of protein-protein interactions.

Score: - 12.462
Nodes: - 14
Edges: - 81

Functional Enrichment Analysis

Figure 4: ClueGO Functional Enrichment Results

GO terms are functionally clustered, meaning similar processes are grouped. These are the enriched pathways Fc receptor signaling pathway, Positive regulation of protein tyrosine kinase activity, Signal transduction in the absence of ligand, Phosphatidylinositol 3-kinase signaling, Response to UV-A Cell, and cycle regulation processes. These results provide insights into biological mechanisms that may be altered in the studied condition (Fig4).

The columns include:

Ontology Source: Indicates that GO terms are sourced from Gene Ontology Biological Process (GO_BP).
Term P-Value: Raw p-values representing the significance of GO term enrichment.
Term P-Value Corrected: Adjusted p-values (likely using multiple testing correction methods like Benjamini-Hochberg).
Group P-Value: Significance of a group of related GO terms.
Group P-Value Corrected: Adjusted significance for a group of terms.

Lowest p-values are the most significant and biologically relevant. Phosphatidylinositol 3-kinase signaling P value is lower 1.9E-5(Table 3).

Table 3: The table appears to be related to Gene Ontology (GO) enrichment analysis results, specifically generated using ClueGO

Ontology Source	Term PValue	Term PValue Corrected wit...	Group PValue	Group PValue Corrected wi...
GO_BiologicalProcess-EBI-U...	1.2 E-5	6.4 E-5	1.2 E-5	2.5 E-5
GO_BiologicalProcess-EBI-U...	2.0 E-5	2.0 E-5	2.0 E-5	2.0 E-5
GO_BiologicalProcess-EBI-U...	1.7 E-5	3.5 E-5	9.5 E-8	2.8 E-7
GO_BiologicalProcess-EBI-U...	4.8 E-6	6.2 E-5	9.5 E-8	2.8 E-7
GO_BiologicalProcess-EBI-U...	1.4 E-5	5.9 E-5	9.5 E-8	2.8 E-7
GO_BiologicalProcess-EBI-U...	1.7 E-5	3.5 E-5	9.5 E-8	2.8 E-7
GO_BiologicalProcess-EBI-U...	3.0 E-6	4.2 E-5	9.5 E-8	2.8 E-7
GO_BiologicalProcess-EBI-U...	8.4 E-6	7.6 E-5	9.5 E-8	2.8 E-7
GO_BiologicalProcess-EBI-U...	2.0 E-5	2.0 E-5	4.9 E-11	2.4 E-10
GO_BiologicalProcess-EBI-U...	1.4 E-5	5.9 E-5	4.9 E-11	2.4 E-10
GO_BiologicalProcess-EBI-U...	1.0 E-11	1.9 E-10	4.9 E-11	2.4 E-10
GO_BiologicalProcess-EBI-U...	2.0 E-10	3.7 E-9	4.9 E-11	2.4 E-10
GO_BiologicalProcess-EBI-U...	6.2 E-6	6.8 E-5	4.9 E-11	2.4 E-10
GO_BiologicalProcess-EBI-U...	5.2 E-7	7.9 E-6	4.9 E-11	2.4 E-10
GO_BiologicalProcess-EBI-U...	5.1 E-6	6.1 E-5	4.9 E-11	2.4 E-10
GO_BiologicalProcess-EBI-U...	1.2 E-8	2.1 E-7	4.9 E-11	2.4 E-10
GO_BiologicalProcess-EBI-U...	2.0 E-5	2.0 E-5	6.4 E-8	2.5 E-7
GO_BiologicalProcess-EBI-U...	1.7 E-5	5.1 E-5	6.4 E-8	2.5 E-7
GO_BiologicalProcess-EBI-U...	2.0 E-7	3.2 E-6	6.4 E-8	2.5 E-7
GO_BiologicalProcess-EBI-U...	1.1 E-5	8.1 E-5	6.4 E-8	2.5 E-7
GO_BiologicalProcess-EBI-U...	1.2 E-5	7.3 E-5	6.4 E-8	2.5 E-7
GO_BiologicalProcess-EBI-U...	1.4 E-5	5.9 E-5	6.4 E-8	2.5 E-7
GO_BiologicalProcess-EBI-U...	6.2 E-6	6.8 E-5	6.4 E-8	2.5 E-7
GO_BiologicalProcess-EBI-U...	5.1 E-6	6.1 E-5	6.4 E-8	2.5 E-7
GO_BiologicalProcess-EBI-U...	5.1 E-6	6.1 E-5	6.4 E-8	2.5 E-7
GO_BiologicalProcess-EBI-U...	7.9 E-6	7.9 E-5	6.4 E-8	2.5 E-7
GO_BiologicalProcess-EBI-U...	9.9 E-6	7.9 E-5	6.4 E-8	2.5 E-7

DISCUSSIONS

The identification of common genes between breast and ovarian cancer provides valuable insights into shared oncogenic pathways. EGFR emerged as a key hub protein, suggesting its critical role in tumorigenesis and signal transduction. The involvement of genes such as BRCA1, PTEN, and mTOR in the PPI network underscores their significance in maintaining cellular homeostasis. The MCODE clustering analysis revealed a highly connected module consisting of 14 genes, further supporting the hypothesis that certain molecular interactions play essential roles in cancer progression.

The ClueGO functional enrichment analysis confirmed that critical pathways, such as phosphatidylinositol 3-kinase signaling and cell cycle regulation, are enriched in the common gene set. The Wnt signaling activators RSPO4 and FRAT2, along with the tumor suppressor gene HTRA1, showed

weak connectivity, which may indicate their independent roles in tumorigenesis. The role of these less-connected genes warrants further investigation to determine their specific contributions.

Comparing centrality measures, EGFR had the highest degree and betweenness centrality, emphasizing its role in signal transduction pathways. This is consistent with previous research highlighting EGFR as a therapeutic target in multiple cancers. Additionally, the presence of key regulators such as MAPK1, AKT1, and CCND1 in the PPI network indicates their involvement in proliferative and survival pathways.

Overall, this study provides a bioinformatics-driven approach to understanding shared oncogenic mechanisms in breast and ovarian cancer. These findings align with existing literature and offer new insights into potential therapeutic targets.

CONCLUSION

This study successfully identified and analyzed twenty-two common genes involved in breast and ovarian cancer through a comprehensive bioinformatics approach. The PPI network analysis revealed key hub proteins, with EGFR emerging as the most significant. Functional enrichment analysis confirmed the involvement of crucial oncogenic pathways, including phosphatidylinositol 3-kinase signaling and Fc receptor signaling. The identification of highly connected modules within the network further supports the existence of common regulatory mechanisms between these cancers. These results contribute to a better understanding of molecular interactions in cancer and may aid in the development of targeted therapies.

REFERENCES

[1] Çelik, Ayşe, et al. "Relationship of breast cancer with ovarian cancer." Concise Rev. Mol. Pathol. Breast Cancer (2015): 87-202.
[2] Colombo, Nicoletta, et al. "Ovarian oncology/hematology 60.2 (2006): 159-179
[3] Dobhal, Shivashish, et al. "Insilico Identification of MHC displayed Tumor Associated Peptides in Ovarian Cancer for Multi-Epitope Vaccine Construct." (2023).
[4] Matulonis, Ursula A., et al. "Ovarian cancer." Nature reviews Disease primers 2.1 (2016): 1-22
[5] Raman, Karthik. "Construction and analysis of protein–protein interaction networks." Automated experimentation 2 (2010): 1-11.
[6] Sun, Jingchun, and Zhongming Zhao. "A comparative study of cancer proteins in the human protein-protein interaction network." BMC genomics 11 (2010): 1-10.
[7] Zamanian-Azodi, Mona, et al. "Protein-Protein Interaction Network could reveal the relationship between the breast and colon cancer." Gastroenterology and Hepatology from bed to bench 8.3 (2015): 215.