Frontiers in Life Sciences and Related Technologies 5(1) (2024) 74-82

Contents lists available at Dergipark

Frontiers in Life Sciences and Related Technologies

Journal homepage: http://dergipark.org.tr/en/pub/flsrt

Research article

Identification of potential hub genes as biomarkers for breast, ovarian, and endometrial cancers

Sema Atasever*[1]

[1] *Nevsehir Haci Bektas Veli University, Faculty of Engineering-Architecture, Department of Computer Engineering, 50300, Nevsehir, Türkiye*

Abstract

Breast cancer (BC) and gynecological cancers have emerged as significant threats to women's health and are known to be among the primary causes of cancer-related fatalities in women. Innovative treatments and early detection may significantly cut mortality rates for these diseases. In this study, potential hub genes were thoroughly evaluated in the contexts of BC, ovarian cancer (OC), and endometrial cancer (EC). Initially, a total of 374 overlapping differentially expressed genes (DEGs) were identified within the microarray datasets. The STRING database and Cytoscape software analyzed protein-protein interaction (PPI) network structure, whereas FunRich found hub genes. The five hub genes that were ultimately discovered are *PTEN*, *SMAD2*, *FASN*, *CYCS*, and *KRAS*. As a result, these genes may serve as potential biomarkers for the aforementioned diseases. Importantly, this study offers valuable insights into all three diseases based on recent molecular advancements. However, further investigation is required to precisely measure these biomarkers' effectiveness.

Keywords: *Breast cancer; DEGs; endometrial cancer; enriched pathway; hub gene; ovarian cancer; potential biomarker*

1. Introduction

Today, breast cancer (BC) and gynecological cancers have become a major threat to women's health (Zhang et al., 2019). In addition to the aging population, the incidence increases due to risk factors such as smoking, excess weight, and physical inactivity (Arakal et al., 2021). According to statistics in 2020, more than 2 million new cases of breast cancer were seen in women (BCS, 2023). Statistics show that BC is the most common cancer seen worldwide. However, gynecological cancers are among the deadliest cancers in the world (Yadav et al., 2020). Again, statistics in 2020 show that EC is the 6th most common cancer in women (ECS, 2023). With more than 313,000 new cases in 2020, OC is the 8th most common cancer in women and the 18th most common cancer overall (OCS, 2023). The incidence of these cancers, which seriously threaten women's health, continues to increase worldwide despite medical advances (Zhang et al., 2022).

Both OC and EC occur as part of Lynch syndrome or hereditary nonpolyposis colorectal cancer (Gayther et al., 2010). With the increase in obesity, mortality rates caused by EC are increasing in developed countries (Li et al., 2020). Studies have shown that *BRCA1*- and *BRCA2* gene mutations are important hereditary risk factors for the development of breast and ovarian cancer (Petrucelli et al., 2022). Microarray and sequencing technologies are commonly used today to identify biomarkers (Xue et al., 2021), but the hub genes shared by BC and EC have not yet been fully clarified (Rahman et al., 2019). Despite numerous researches into the molecular mechanisms of cancer, the processes of gynecological cancers are still not well understood. Therefore, the search for new biomarkers for early detection is very important (Zhang et al., 2019).

Next-generation sequencing technologies offer remarkable sequencing speed and lower costs (Lee et al., 2013; Toss et al.,

* Corresponding author.
E-mail address: sema@nevsehir.edu.tr (S. Atasever).
https://doi.org/10.51753/flsrt.1405816 Author contributions
Received 16 December 2023; Accepted 02 April 2024
Available online 30 April 2024

2015). Despite these developments, the survival rate of OC is low. Therefore, it is of great importance to detect OC at an early stage (Yadav et al., 2020).

This study aims to use bioinformatics tools to analyze microarray data, identify new biomarkers, and explore the molecular mechanisms of BC, OC, and EC. It is collected microarray datasets of BC, OC, and EC from Gene Expression Omnibus (GEO). Searching for a possible biological link between BC, OC, and EC disease, sequencing data from patients with specified diseases from GEO databases was used.

2. Materials and methods

The microarray datasets GEO42568, GEO27651, and GEO17025 were downloaded from GEO database. These three datasets were analysed and normalized using the GEO2R tool to find common DEGs. PPI network structure was performed using STRING (Szklarczyk et al., 2010) database and Cytoscape (Shannon et al., 2003) software, respectively. A brief workflow is indicated in Fig. 1.

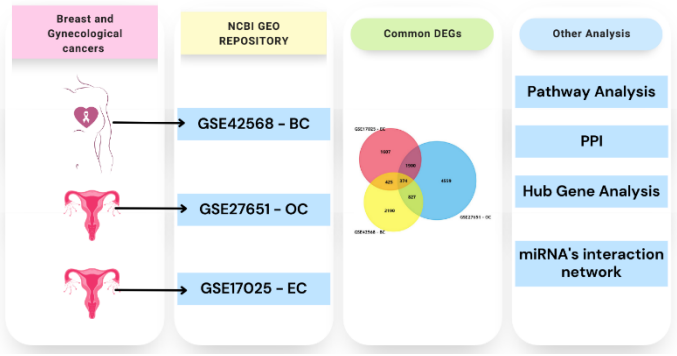

Fig. 1. Overall workflow of methodology. Microarray data was obtained from the GEO database. The number of samples in the study was 273, of which 238 were cancerous and 35 were normal tissue. The largest samples in the database belong to the breast tissues.

2.1. Data acquisition

In this study, three cancer-related gene expression profiles with GEO accession numbers GSE42568 (GSE42568, 2023), GSE27651 (GSE27651, 2023), and GSE17025 (GSE17025, 2023) were extracted from the GEO database, and DEGs were obtained using GEO2R. There is total of 121 patient samples in dataset GSE42568, 49 patient samples in dataset GSE27651, and 103 patient samples in dataset GSE17025 (see Table 1).

2.2. Data processing

All patient records were selected as test and control using the GEO2R tool, which was used to compare two or more sets of samples to identify DEGs under experimental conditions (see Fig. 2). Results are obtained as a table of genes ordered in order of importance. After processing each examined dataset with the GEO2R tool, thousands of DEGs were obtained.

2.3. Selection of DEGs

DEG lists downloaded using the GeoR tool, as shown in Fig. 3, were filtered according to p-value and log2 fold-change (Log2FC) values, and genes belonging to the related disease were obtained (Sarkar et al., 2021). The Benjamini–Hochberg approach was used to adjust the p-values (Benjamini et al., 1995).

2.4. Method for identification of DEGs

A systematic approach using the FunRich tool (version 3.1.3) (Fonseka et al., 2021) was used to identify DEGs in this study. Initially, gene expression data were extracted from three different datasets from GEO database. The main criteria for selecting these datasets were their relevance to the research topic and the quality of the data.

For each dataset, gene expression data were normalized for comparability. Then, statistical methods were used to find genes with significant expression differences between test and control groups, using criteria of a log2 fold change$| \geq 0.5$ and an adjusted p-value of < 0.05 to identify DEGs.

The Venn diagram function in FunRich was then used to visually represent and identify common genes across the three datasets. This step was crucial for locating hub genes that consistently showed different expressions across multiple datasets. The Venn diagram (see Fig. 4) shows the overlap of DEGs and helps isolate the most important candidates for further analysis.

In summary, the identification of DEGs involved several steps: normalizing the data, analyzing it statistically to find differences in expression, and using graphs to identify common genes in different datasets. This comprehensive strategy guaranteed that DEGs were statistically significant.

2.5. Network analysis of DEGs

Network analysis of DEGs was necessary for the study to better understand the complex gene interactions. All shared DEGs from the FunRich tool were brought into STRING during this step. This allowed for a comprehensive exploration of the dynamic interactions and functional relationships that shape the molecular landscape of BC, OC, and EC.

3. Results

3.1. Identification of DEGs

After the data was downloaded, a significance threshold was defined with a p-value<0.05 and a log2FC (fold change) $\geq$ 0.5. Adhering to these criteria, a total of 5676 genes out of 54675 in BC (GSE42568), 10219 genes out of 54675 in OC (GSE27651), and 5824 genes out of 54675 in EC (GSE17025)

Table 1
Datasets obtained from GEO.

Disease	Accession number of the dataset	#of patients	Diseased samples / Test species	Healthy samples / Control species
Breast Cancer	GSE42568	121	104	17
Ovarian Cancer	GSE27651	49	43	6
Endometrial Cancer	GSE17025	103	91	12
Total:		**273**	**238**	**35**

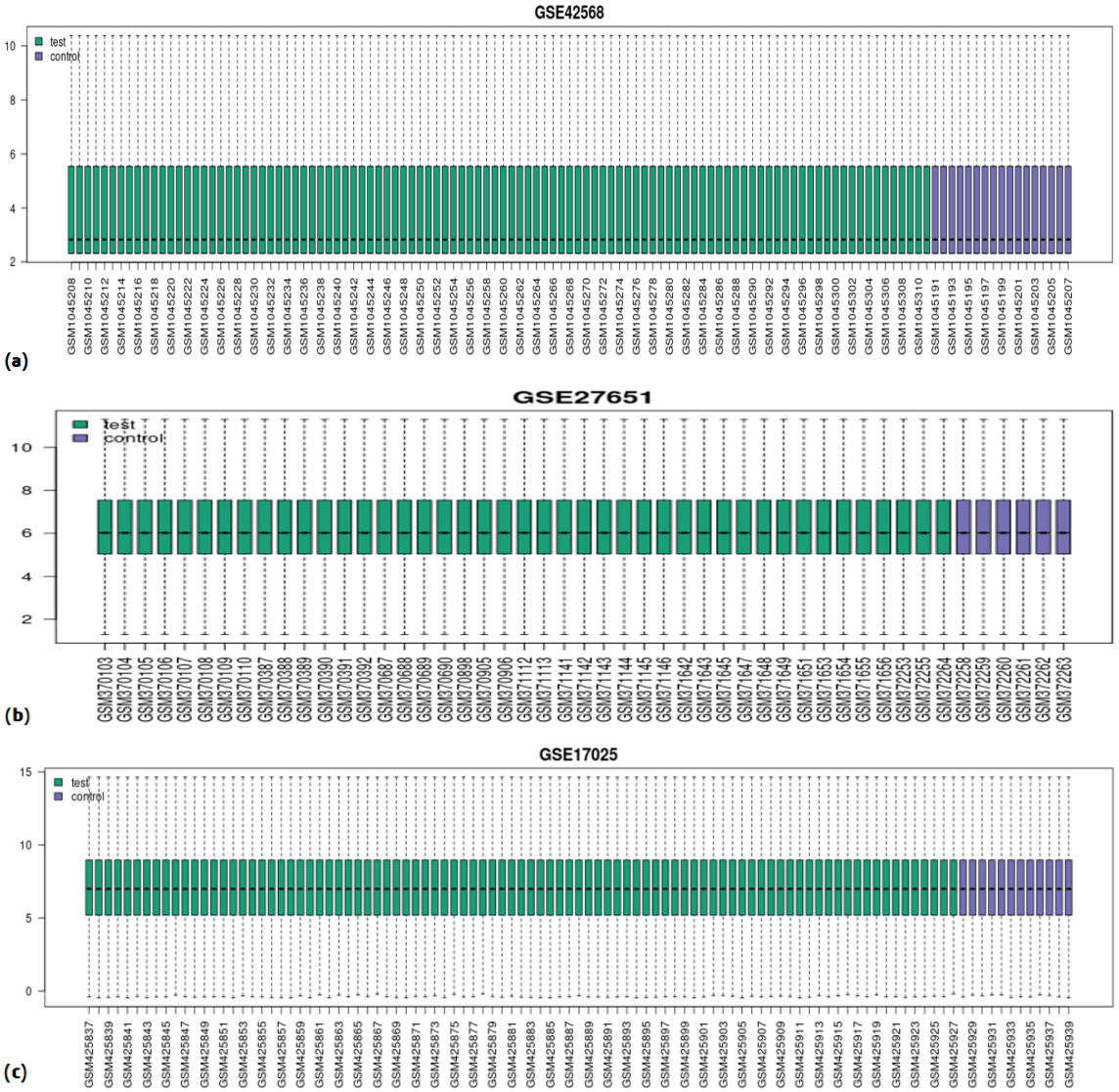

Fig. 2. Box plots of gene expression values (a) BC (GSE42568), (b) OC (GSE27651), (c) EC (GSE17025).

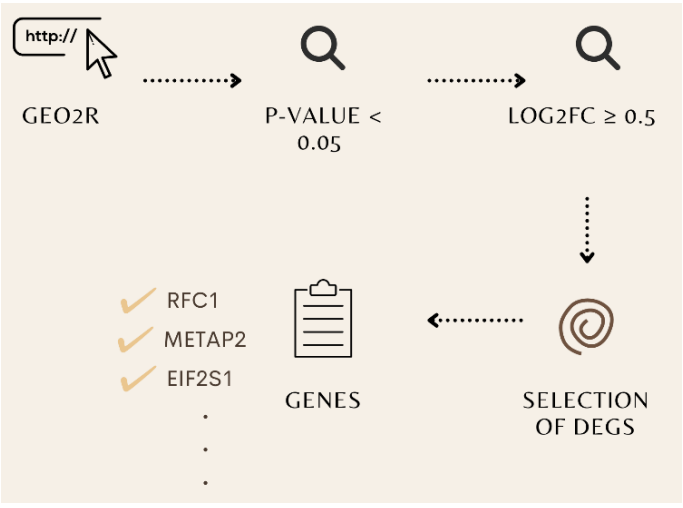

Fig. 3. Basic steps of selection of DEGs.

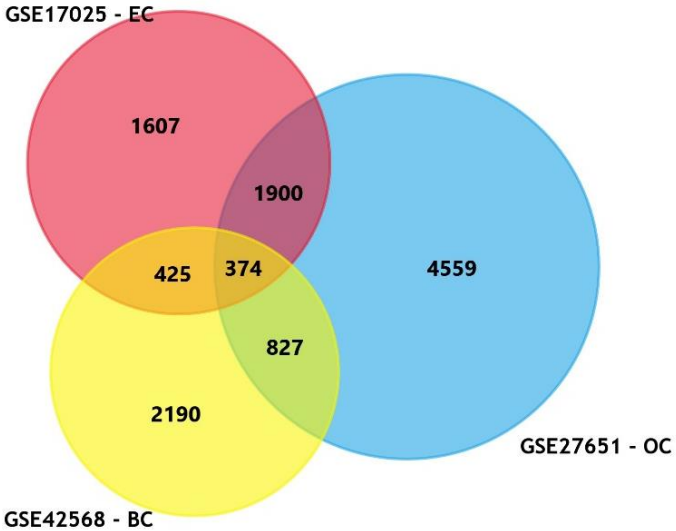

Fig. 4. The Venn diagram shows 374 common hub genes.

were selected for further analysis. This rigorous filtering process resulted in the identification of 3816 DEGs in BC, 7660 DEGs in OC, and 4306 DEGs in EC, providing a more refined subset for subsequent investigation.

3.2. Common DEGs analysis

Venn diagram analysis yielded 374 genes reflecting dataset

intersections.

3.3. Common DEGs network

3.3.1. Construction of PPIs network

Protein-Protein Interaction (PPI) networks helped uncover hub genes for various illnesses. All 374 common DEGs related to these disorders were imported into the STRING, constructing a comprehensive network with 355 nodes and 964 edges. The resulting PPI network exhibited an average node degree of 5.43 and a clustering coefficient of 0.328, visually represented in Fig. 5 and Fig. 6. This network analysis provides a holistic view of the interconnected relationships among the differentially expressed genes, offering valuable insights into the molecular dynamics underlying breast cancer, ovarian cancer, and endometrial cancer.

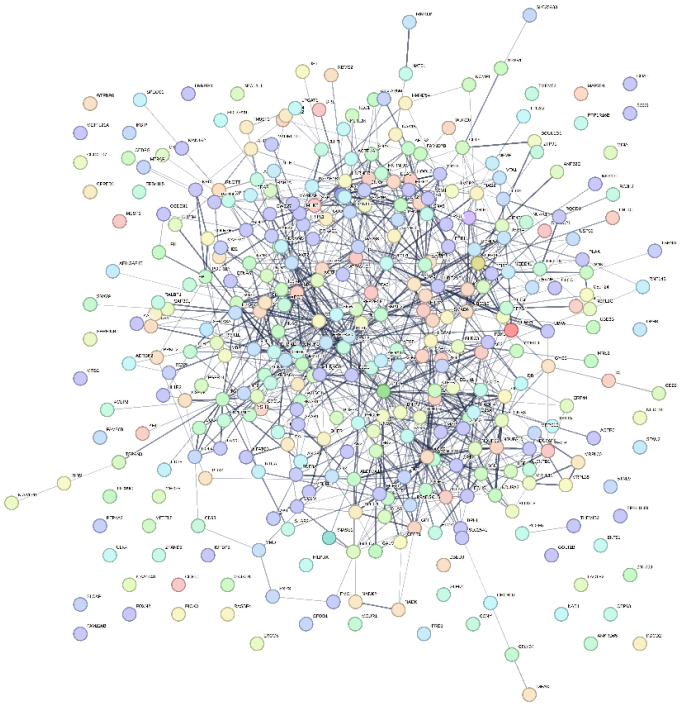

Fig. 5. PPIs network produced by STRING.

3.3.2. Hub gene analysis

In this study, the identification of hub genes was achieved by combining advanced bioinformatics tools and network analysis techniques. Initially, the STRING database (Szklarczyk et al., 2010) was used to construct a PPI network based on DEGs associated with BC, OC, and EC. This network forms the basis for identifying key genes that play central roles in these cancers.

Cytoscape software (Shannon et al., 2003), a powerful tool, was then used to develop the analysis and visualize molecular interaction networks. The Cyto-Hubba plugin was used in Cytoscape, a module specifically designed to identify hub genes in a network. Cyto-Hubba ranks genes within the network based on various topological algorithms such as degree centrality, closeness centrality, and betweenness centrality (Zhou et al., 2021). These algorithms calculate the importance of each gene in the network based on their connections and positions. In this study, degree filters were used. The top 20 genes with the highest scores in these degree centrality measurements were selected as centrality genes. These genes are shown in Fig. 7, which shows their positions and interactions within the network. By integrating these computational tools, the precision of hub gene selection has been increased and focused information has been provided on key players within the PPI network.

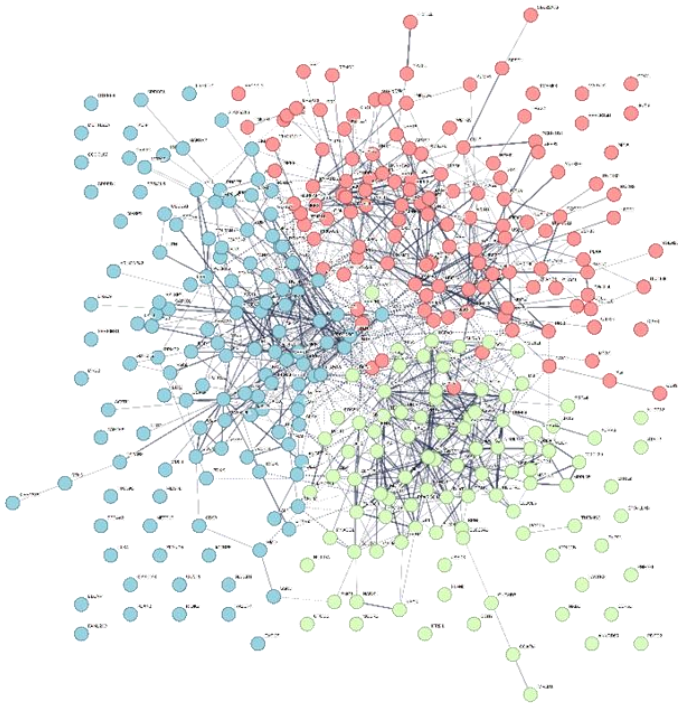

Fig. 6. K-means clustering applied in PPIs network.

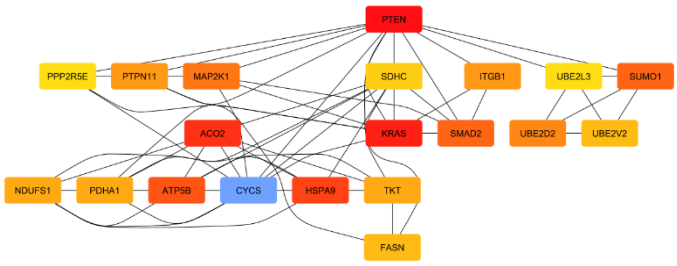

Fig. 7. Hub genes were identified using the plug-in cytoHubba in Cytoscape software.

Table 2
List of common Hub genes using degree centrality ranking methods.

No.	Gene Name	No.	Gene Name
1	*"CYCS"*	11	*"PTPN11"*
2	*"PTEN"*	12	*"ITGB1"*
3	*"KRAS"*	13	*"TKT"*
4	*"ACO2"*	14	*"PDHA1"*
5	*"HSPA9"*	15	*"NDUFS1"*
6	*"ATP5B"*	16	*"FASN"*
7	*"SMAD2"*	17	*"UBE2V2"*
8	*"SUMO1"*	18	*"SDHC"*
9	*"MAP2K1"*	19	*"PPP2R5E"*
10	*"UBE2D2"*	20	*"UBE2L3"*

Table 2 serves as a comprehensive overview, summarizing the list of common hub genes based on degree centrality ranking methods. Notably, *CYCS* emerges as the hub gene with the highest degree of interaction, underscoring its potential significance, while *UBE2L3* stands out as the least connected node among the top 20 hub genes.

3.3.3. miRNA's interaction network

Upon identification, the top 20 hub genes were subsequently integrated into the miRNet 2.0 software (Chang et al., 2022), allowing for a comprehensive exploration of their intricate interactions with miRNAs, as visually depicted in Fig. 8. This multi-layered analysis not only sheds light on the hub genes but also unravels potential regulatory relationships between these genes and miRNAs, offering a more holistic understanding of their roles in the molecular landscape of breast cancer, ovarian cancer, and endometrial cancer.

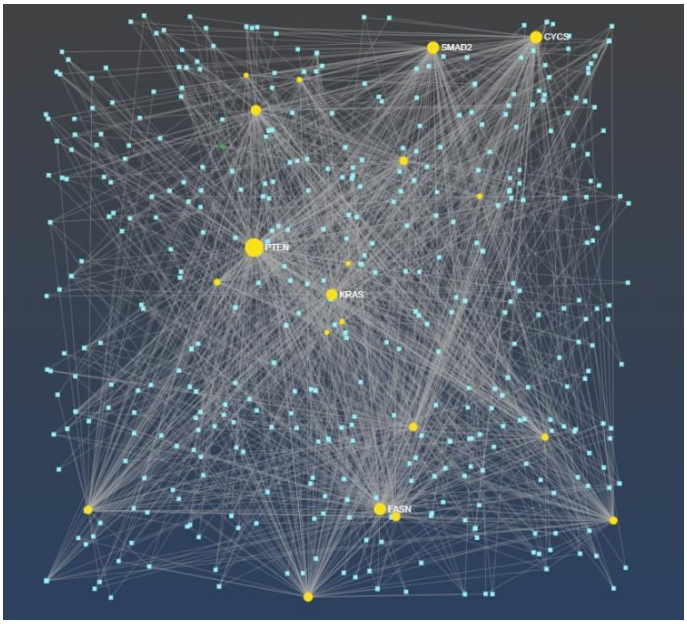

Fig. 8. Interaction between miRNA and 20 common genes of BC, OC and, EC obtained via miRNet 2.0. The five large nodes indicate the hub genes.

Table 3
miRNet 2.0 screening results.

Hub genes	Degree	Betweenness
PTEN	160	23099.09
SMAD2	107	12475.8
FASN	106	12336.65
CYCS	104	12544.66
KRAS	99	10806.33

Table 4
A list of common biomarker candidates.

Hub genes	Roles of the Biomolecules	Reference
PTEN	Phosphatase and tensin homolog	(Chou et al., 2014; Smith et al., 2016)
SMAD2	Cancer-related genes	(Martinez-Ledesma et al., 2015)
FASN	Fatty acid synthase	(Fernández et al., 2020)
CYCS	Cytochrome c, somatic	(Emmanuel et al., 2011)
KRAS	Regulation of cell proliferation	(Emmanuel et al., 2011)

Table 3 shows a list of central genes with their connections (degree) and importance (betweenness) in the network. Nodes with higher node degree act as hubs in a network (miRNet, 2024). *PTEN*, with the highest degree of 160, is the most connected gene in this network, followed by *SMAD2*, *FASN*, *CYCS*, and *KRAS*. A higher betweenness centrality score can indicate a gene's strategic role in the communication within the network. *PTEN* has the highest betweenness centrality at 23099.09, indicating it may play a significant role in the flow of information in the network.

3.3.4. Survival analysis

A method using both statistical significance and biological importance selected 5 hub genes for survival analysis (Table 4). These genes, central in the PPI network, are vital for cancer-related cellular processes. Their high interaction and central role in the network highlight their potential regulatory importance in cancer pathways. This study analyzed overall and disease-free survival (DFS) data for 1,668 patients with BC, OC, and EC using the GEPIA online service.

Fig. 9A shows a Kaplan-Meier curve for the OS of patients grouped by their high or low *PTEN*, *SMAD2*, *FASN*, *KRAS*, and *CYCS* gene expression levels. Fig. 9B does the same for DFS. In both figures, the blue line shows patients with low expression, and the red line shows those with high expression. Both groups start with a DFS probability of 1.0 (or 100%), which decreases over time as events (recurrence of disease) occur.

In summary, Kaplan-Meier survival curves showed that higher expression levels of the *PTEN*, *SMAD2*, *FASN*, and *KRAS* genes were associated with longer OS, whereas higher expression of *CYCS* was associated with lower OS. Therefore, the *PTEN*, *SMAD2*, *FASN*, and *KRAS* genes could be considered a positive prognostic biomarker for patients with BC and gynecological cancers.

In summary, Kaplan-Meier survival curves showed that higher expression levels of the *PTEN*, *SMAD2*, *FASN*, and *KRAS* genes were associated with a longer period of DFS. Therefore, the expression levels of these genes (*PTEN*, *SMAD2*, *FASN*, and *KRAS*) may serve as a prognostic biomarker for BC and gynecological cancers.

3.4. Pathways analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were conducted using the "Enrichr" online platform (Enrichr, 2024) to decipher crucial biological processes. This was accomplished for the top 20 hub genes, thereby illuminating significant insights.

The 374 overlapping DEGs were analyzed for GO enrichment, revealing the top 10 enriched terms in biological processes, cellular components, and molecular functions in Fig. 10A, including significant enrichment in "negative regulation of organ growth" and "cell-cell adhesion mediated by integrin". KEGG analysis showed significant enrichment in pathways like the "citrate cycle" and "central carbon metabolism in cancer" (Fig. 10B).

Fig. 10(a) provides a list of biological processes from GO, a major bioinformatics initiative that aims to unify the representation of gene and gene product features across all species. Each listed process is accompanied by a unique identifier (e.g., GO:0046621 for "negative regulation of organ growth"), allowing easy reference and cross-database comparisons.

Highlighted processes include various regulatory functions such as "negative regulation of organ growth", "cell-cell adhesion mediated by integrin", and "neuroinflammatory response". There are also specific pathways like "ribosome phosphate metabolic process", "regulation of inward rectifier

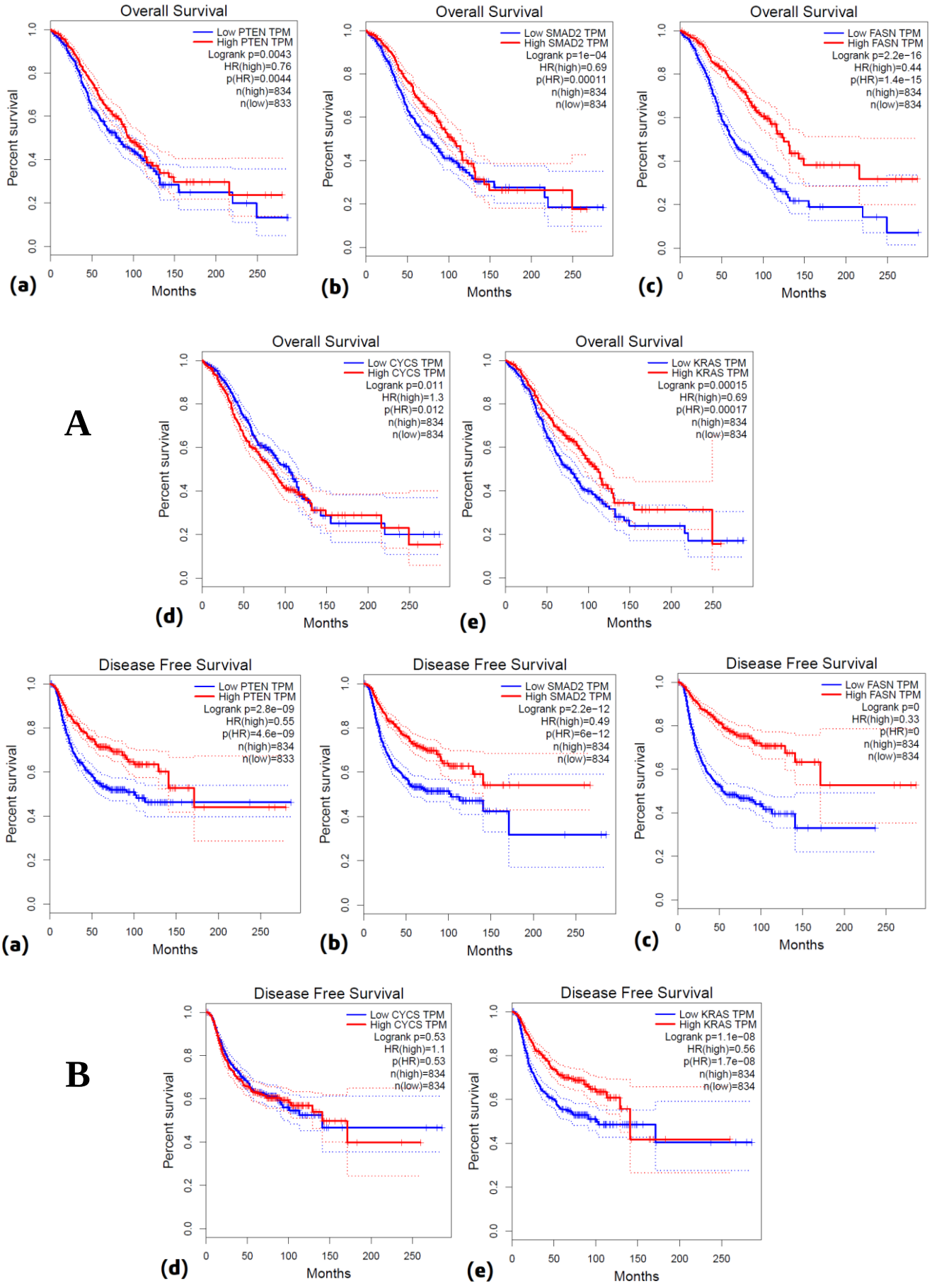

Fig. 9. Overall survival (OS) analysis of the candidate hub genes. (A) OS analysis. (a) *PTEN*, (b) *SMAD2*, (c) *FASN*, (d) *CYCS*, and (e) *KRAS*. (B) DFS analysis. (a) *PTEN*, (b) *SMAD2*, (c) *FASN*, (d) *CYCS*, and (e) *KRAS*. The threshold Log-rank p<0.05 was considered as statistically significant.

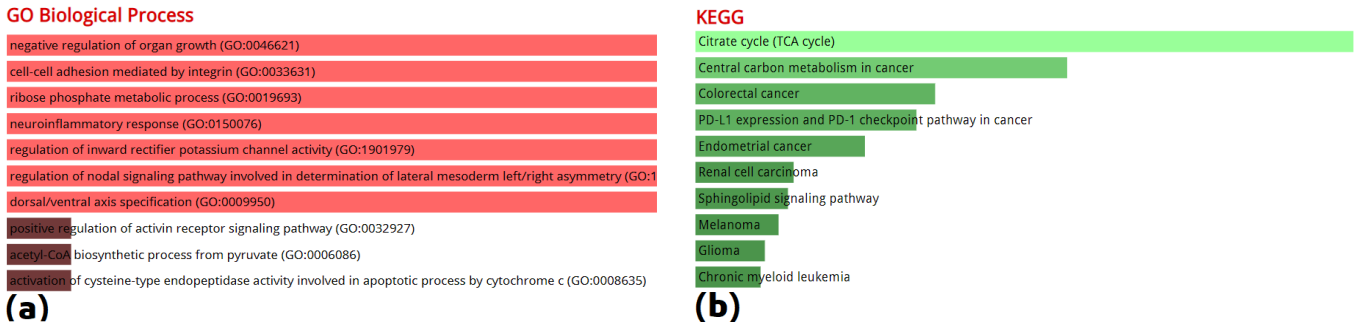

Fig. 10. Bioinformatics analysis of DEGs in the progression of BC, OC and EC cancers. Functional enrichment analysis of the overlapping DEGs. The GO enrichment analysis of DEGs in the categories of (a) biological process, (b) The KEGG pathway enrichment analysis of the overlapping DEGs. The top 10 enriched KEGG pathways were shown.

potassium channel activity". The functional enrichment analysis indicated that the overlapping DEGs were mainly associated with "negative regulation of organ growth" and "cell-cell adhesion mediated by integrin", "ribose phosphate metabolic process", and so on.

The analysis used in Fig. 10(b) is used to identify important biological pathways enriched in a set of DEGs. The bar chart shows the top 10 enriched KEGG pathways among the overlapping DEGs analyzed in this study. From the list, we can see that it includes metabolic pathways such as the "Citrate cycle (TCA cycle)" and "Central carbon metabolism in cancer", as well as pathways directly related to various cancers like "Colorectal cancer", "Endometrial cancer", "Renal cell carcinoma", "Melanoma", "Glioma", and "Chronic myeloid leukemia". The "PD-L1 expression and PD-1 checkpoint pathway in cancer" indicates a focus on immunological pathways that are targeted in cancer immunotherapy. In cancer treatment, tumor microenvironment is sensitive to treatment with immune checkpoint such as the PD-1/PD-L1 pathway because of radiotherapy (Du et al., 2020). The "Sphingolipid signaling pathway" is involved in signaling mechanisms that can affect cell growth, survival, differentiation, and apoptosis, which are processes relevant to cancer biology. Members of the sphingolipid family are widely involved in cancer cell growth, migration, invasion, and other biological processes (Sun et al., 2022).

4. Discussion

Numerous studies have been conducted in recent years to identify genetic markers for cancer (Banno et al., 2012; Toss et al., 2015; Walsh et al., 2016; Zhang et al., 2022). The US National Cancer Institute (NCI) defines a biomarker as "a biological molecule found in the blood, other body fluids, or tissues that is a sign of a normal or abnormal process, or a condition or disease" (Banno et al., 2012). New biomarkers are revealed by next-generation sequencing. This will allow clinicians to present in the most informative way recommendations for administering new treatments modifying existing treatments and adjusting dosage (Walsh et al., 2016). The generation of high-throughput technologies has improved our understanding of complex biological features such as tumors (Wang et al., 2022). Despite surgical and chemotherapy treatment applications in various cancer types such as EC, mortality rates are still increasing in recent years. Therefore, it is necessary to better understand the mechanisms that cause cancer (Li et al., 2020). In this study, by downloading three cancer datasets, 374 common DEGs were screened out. From their study of colorectal cancer patients treated with cetuximab, Li et al. (2010) concluded that *KRAS* mutation and *PTEN* protein expression were significantly associated with patient response rate and survival time.

The study by Zhang et al. (2024) identified 13 ferroptosis-related genes in Crohn's disease (CD) using bioinformatics analysis of two Gene Expression Omnibus datasets and further validated three of these genes (*IL-6*, *PTGS2*, and *DUOX2*) as key regulators of ferroptosis in CD through qPCR analysis of clinical samples. These findings suggest new biomarkers and therapeutic targets for CD, offering fresh insights into its pathogenesis and potential treatment strategies.

In recent years, efforts have been made to identify new biomarkers for gynecological cancer. For example, upregulation of *TRIM44* predicts poor prognosis in OC. Previous studies have found *SMYD2* to be an oncogene in several types of cancer. *MMP8* has been reported to be associated with BC. *HSDL2* acts as an oncogene in OC (Zhang et al., 2019). *PTEN* gene encodes a tumor suppressor phosphatase that has been found to be frequently mutated in patients with OC and EC (Smith et al., 2016).

In this study, a Venn diagram obtained using the FunRich tool was analyzed to identify hub genes. By determining 20 hub genes by the degree centrality method, in the analysis of common hub genes, "negative regulation of organ growth" and "cell-cell adhesion mediated by integrin", "ribose phosphate metabolic process", etc. It has been determined that there are genes associated with five unreported genes in BC, OC, and EC. The results may help us understand the development of BC, OC, and EC and guide further experiments.

In Fig. 8, the five large nodes (*PTEN*, *SMAD2*, *FASN*, *CYCS*, and *KRAS*) represent hub genes that are likely central in the network due to their high degree of connections, indicating they may play key roles in the regulatory processes across the three types of cancer. The study by Davies et al. (2014) showed that mutations in *PTEN* and *KRAS* alone predispose mice to a spectrum of serrated lesions reflective of the serrated pathway of colorectal cancer progression in humans. The study by Qian et al. (2022) stated that *SMAD2* was related to colorectal cancer, *KRAS* was related to ovarian cancer, and *KRAS* and *PTEN* were related to endometrial cancer. In another study (Stebbing et al., 2014), the alterations observed in phosphatases and the resulting malignancies were associated with the *PTEN* gene in cervical, ovarian, and breast cancers.

In the literature, studies have been carried out examining various gene expressions to evaluate their ability as prognostic markers. Yndestad et al. (2017) conducted a study indicating that high *PTEN* gene expression is a negative prognostic marker

in human primary breast cancers with preserved p53 function. Another study showed that loss of *PTEN* expression was associated with worse survival (Ferraldeschi et al., 2015). Although Liu et al. (2020) found that *SMAD2*, *p-SMAD2*, and *SMAD4* are not independent predictors by multivariate analysis, *SMAD4* positivity correlates with longer OS and progression-free survival. They also found that combined *p- SMAD2* and *SMAD4* expression can serve as an independent prognostic factor, suggesting that testing for these proteins in breast ductal carcinoma biopsies could offer extra prognostic insights. Ramanathan et al. (2017) showed that although *VEGFA* alone did not correlate with survival, high *ANG2* and high *VEGFA* co-expression correlated with decreased OS for breast cancer.

Inferences made with Kaplan-Meier curves are generally not sufficient to fully understand the prognostic value of a gene. Kaplan-Meier analysis evaluates OS or DFS by grouping patients based on a given gene expression level and produces survival curves for each group. This analysis is useful for visualizing the impact of a gene's high or low expression on patient survival, but this relationship is not definitive and does not account for other influencing factors. Multivariate analysis is necessary to support Kaplan-Meier results because it considers other important variables such as age, gender, type of treatment, stage, and other potential confounding factors. In their study, Scaglia et al. (2013) in addition to the Kaplan-Meier method, also used the Cox proportional hazards regression model to evaluate the effect of potential confounding factors and adjust their effects in the comparison between genders. Methods such as the Cox proportional hazards model (Liu et al., 2020) are used in multivariate analysis to determine whether the effect of a gene on survival is independent of all other variables. Therefore, complementing Kaplan-Meier analysis with multivariate analysis is crucial to confirm whether a gene is an independent prognostic marker, which is one of the limitations of this study.

5. Conclusion

In this study, it was aimed to analyze microarray samples by using bioinformatics tools to identify new biomarkers. Searching for a possible biological link between BC, OC, and EC disease, sequencing data from patients with specified diseases from GEO databases was used. 374 DEGs were common, and five of them came to the fore. These include the *PTEN*, *SMAD2*, *FASN*, *CYCS*, and *KRAS* genes. The functional enrichment analysis indicated that the overlapping DEGs were mainly associated with "negative regulation of organ growth" and "cell-cell adhesion mediated by integrin", "ribose phosphate metabolic process", and so on. Higher *PTEN*, *SMAD2*, *FASN*, and *KRAS* gene expressions correlate with increased overall and disease-free survival, unlike *CYCS*, which shows reduced overall survival with no significant impact on disease-free survival. However, further work is required to quantify the potency of these biomarkers. Understanding the role of the key genes identified in this study in signal transduction could help in creating targeted drugs for cancer treatment, either alone or combined with other therapies.

Conflict of interest: The author declares that she has no conflict of interests.

Informed consent: The author declares that this manuscript did not involve human or animal participants and informed consent was not collected.

References

Arakal, N. G., Sharma, V., Kumar, A., Kavya, B., Devadath, N. G., Kumar, S. B., ... & Murahari, M. (2021). Ligand-based design approach of potential Bcl-2 inhibitors for cancer chemotherapy. *Computer Methods and Programs in Biomedicine*, *209*, 106347.

Banno, K., Kisu, I., Yanokura, M., Tsuji, K., Masuda, K., Ueki, A., ... & Aoki, D. (2012). Biomarkers in endometrial cancer: Possible clinical applications. *Oncology letters*, *3*(6), 1175-1180.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289-300.

BCS, (2023). Breast Cancer Statistics, https://www.wcrf.org/cancer-trends/breast-cancer-statistics/, Last Accessed on December 16, 2023.

Chang, L., & Xia, J. (2022). MicroRNA regulatory network analysis using miRNet 2.0. In: Song Q., Tao Z. (eds) *Transcription Factor Regulatory Networks* (pp. 185-204). Springer, Humana, New York.

Chou, W. C., Cheng, A. L., Brotto, M., & Chuang, C. Y. (2014). Visual gene-network analysis reveals the cancer gene co-expression in human endometrial cancer. *BMC Genomics*, *15*(1), 1-12.

Davies, E. J., Marsh Durban, V., Meniel, V., Williams, G. T., & Clarke, A. R. (2014). PTEN loss and KRAS activation leads to the formation of serrated adenomas and metastatic carcinoma in the mouse intestine. *The Journal of Pathology*, *233*(1), 27-38.

Du, Z., Yan, D., Li, Z., Gu, J., Tian, Y., Cao, J., & Tang, Z. (2020). Genes involved in the PD-L1 pathway might associate with radiosensitivity of patients with gastric cancer. *Journal of Oncology*, *2020*.

Emmanuel, C., Gava, N., Kennedy, C., Balleine, R. L., Sharma, R., Wain, G., ... & deFazio, A. (2011). Comparison of expression profiles in ovarian epithelium in vivo and ovarian cancer identifies novel candidate genes involved in disease pathogenesis. *PloS One*, *6*(3), e17617.

ECS, (2023). Endometrial Cancer Statistics, https://www.wcrf.org/cancer-trends/endometrial-cancer-statistics/, Last Accessed on December 16, 2023.

Enrichr, (2024). Enrichr Database, https://maayanlab.cloud/Enrichr/, Last Accessed on December 16, 2023.

Fernández, L. P., de Cedron, M., & de Molina, A. (2020). Alterations of lipid metabolism in cancer: Implications in prognosis and treatment. *Frontiers in Oncology*, *10*, 577420.

Ferraldeschi, R., Rodrigues, D. N., Riisnaes, R., Miranda, S., Figueiredo, I., Rescigno, P., ... & de Bono, J. (2015). PTEN protein loss and clinical outcome from castration-resistant prostate cancer treated with abiraterone acetate. *European Urology*, *67*(4), 795-802.

Fonseka, P., Pathan, M., Chitti, S. V, Kang, T., & Mathivanan, S. (2021). FunRich enables enrichment analysis of OMICs datasets. *Journal of Molecular Biology*, *433*(11), 166747.

Gayther, S. A., & Pharoah, P. D. P. (2010). The inherited genetics of ovarian and endometrial cancer. *Current Opinion in Genetics & Development*, *20*(3), 231-238.

GSE17025, (2023). National Center for Biotechnology Information, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE17025, Last Accessed on December 16, 2023.

GSE27651, (2023). National Center for Biotechnology Information, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE27651, Last Accessed on December 16, 2023.

GSE42568, (2023). National Center for Biotechnology Information, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE42568, Last Accessed on December 16, 2023.

Li, F. H., Shen, L., Li, Z. H., Luo, H. Y., Qiu, M. Z., Zhang, H. Z., ... & Xu, R. H. (2010). Impact of KRAS mutation and PTEN expression on cetuximab-treated colorectal cancer. *World Journal of Gastroenterology: WJG*, *16*(46), 5881.

Li, Y., & Li, L. (2020). Bioinformatic screening for candidate biomarkers and their prognostic values in endometrial cancer. *BMC Genetics*, *21*(1), 1-13.

Liu, N., Qi, D., Jiang, J., Zhang, J., & Yu, C. (2020). Expression pattern of p-Smad2/Smad4 as a predictor of survival in invasive breast ductal carcinoma. *Oncology Letters*, *19*(3), 1789-1798.

Martinez-Ledesma, E., Verhaak, R. G. W., & Treviño, V. (2015). Identification of a multi-cancer gene expression biomarker for cancer clinical outcomes using a network-based algorithm. *Scientific Reports*, *5*(1), 11966.

miRNet, (2024). miRNet Tutorial Starting with a List, https://www.mirnet.ca/miRNet/resources/data/tutorials/Start_with_list.pdf, Last Accessed on January 24, 2024.

OCS, (2023). Ovarian Cancer Statistics, https://www.wcrf.org/cancer-trends/ovarian-cancer-statistics/, Last Accessed on December 16, 2023.

Petrucelli, N., Daly, M. B., & Pal, T. (2022). *BRCA1-and BRCA2-associated hereditary breast and ovarian cancer*. GeneReviews.

Qian, F., Kong, W., & Wang, S. (2022). Exploring autophagy-related prognostic genes of Alzheimer's disease based on pathway crosstalk analysis. *Bosnian Journal of Basic Medical Sciences*, *22*(5), 751.

Rahman, M. F., Rahman, M. R., Islam, T., Zaman, T., Shuvo, M. A. H., Hossain, M. T., ... & Moni, M. A. (2019). A bioinformatics approach to decode core genes and molecular pathways shared by breast cancer and endometrial cancer. *Informatics in Medicine Unlocked*, *17*, 100274.

Ramanathan, R., Olex, A. L., Dozmorov, M., Bear, H. D., Fernandez, L. J., & Takabe, K. (2017). Angiopoietin pathway gene expression associated with poor breast cancer survival. *Breast Cancer Research and Treatment*, *162*, 191-198.

Sarkar, D., Chakraborty, S., Bhowmick, S., & Maiti, T. (2021). *In-silico* analysis: common biomarkers of NDs. *BioRxiv*, 2021-2029.

Scaglia, N. C., Chatkin, J. M., Pinto, J. A., Tsukazan, M. T. R., Wagner, M. B., & Saldanha, A. F. (2013). Role of gender in the survival of surgical patients with nonsmall cell lung cancer. *Annals of Thoracic Medicine*, *8*(3), 142.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., ... & Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, *13*(11), 24982504.

Smith, I. N., & Briggs, J. M. (2016). Structural mutation analysis of PTEN and its genotype-phenotype correlations in endometriosis and cancer. *Proteins: Structure, Function, and Bioinformatics*, *84*(11), 1625-1643.

Stebbing, J., Lit, L. C., Zhang, H., Darrington, R. S., Melaiu, O., Rudraraju, B., & Giamas, G. (2014). The regulatory roles of phosphatases in cancer. *Oncogene*, *33*(8), 939-953.

Sun, Y., Xu, Y., Che, X., & Wu, G. (2022). Development of a novel sphingolipid signaling pathway-related risk assessment model to predict prognosis in kidney renal clear cell carcinoma. *Frontiers in Cell and Developmental Biology*, *10*, 881490.

Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., ... & Mering, C. V. (2010). The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Research*, *39*(suppl_1), D561-D568.

Toss, A., Tomasello, C., Razzaboni, E., Contu, G., Grandi, G., Cagnacci, A., ... & Cortesi, L. (2015). Hereditary ovarian cancer: not only BRCA 1 and 2 genes. *BioMed Research İnternational*, *2015*.

Walsh, M. F., Nathanson, K. L., Couch, F. J., & Offit, K. (2016). Genomic biomarkers for breast cancer risk. *Novel Biomarkers in the Continuum of Breast Cancer*, 1-32.

Wang, Y., Wang, J., Hu, Y., Shangguan, J., Song, Q., Xu, J., ... & Zhang, Y. (2022). Identification of key biomarkers for STAD using filter feature selection approaches. *Scientific Reports*, *12*(1), 19854.

Xue, H., Sun, Z., Wu, W., Du, D., & Liao, S. (2021). Identification of hub genes as potential prognostic biomarkers in cervical cancer using comprehensive bioinformatics analysis and validation studies. *Cancer Management and Research*, 117-131.

Yadav, G., Vashisht, M., Yadav, V., & Shyam, R. (2020). Molecular biomarkers for early detection and prevention of ovarian cancer—A gateway for good prognosis: A narrative review. *International Journal of Preventive Medicine*, *11*.

Yndestad, S., Austreid, E., Knappskog, S., Chrisanthar, R., Lilleng, P. K., Lønning, P. E., & Eikesdal, H. P. (2017). High PTEN gene expression is a negative prognostic marker in human primary breast cancers with preserved p53 function. *Breast Cancer Research and Treatment*, *163*, 177-190.

Zhang, S., Jiang, H., Gao, B., Yang, W., & Wang, G. (2022). Identification of diagnostic markers for breast cancer based on differential gene expression and pathway network. *Frontiers in Cell and Developmental Biology*, *9*, 811585.

Zhang, W., Li, Z., Li, H., & Zhang, D. (2024). Identification of differentially expressed genes associated with ferroptosis in Crohn's disease. *Experimental and Therapeutic Medicine*, *27*(2), 1-12.

Zhang, X., & Wang, Y. (2019). Identification of hub genes and key pathways associated with the progression of gynecological cancer. *Oncology Letters*, *18*(6), 6516-6524.

Zhou, C., Guo, H., & Cao, S. (2021). Gene Network Analysis of Alzheimer's Disease Based on Network and Statistical Methods. *Entropy*, *23*(10), 1365.

Cite as: Atasever, S. (2024). Identification of potential hub genes as biomarkers for breast, ovarian, and endometrial cancers. *Front Life Sci RT*, *5*(1), 74-82.