RESEARCH ARTICLE

# Bioinformatics and In silico approaches to identify novel biomarkers and key pathways for cancers that are linked to the progression of female infertility: A comprehensive approach for drug discovery

**Md. Arju Hossain[1], Md Sohel[2], Md Habibur Rahman[3]\*, Md Imran Hasan[3], Md. Sharif Khan[1], Md. Al Amin[1], Md. Zahidul Islam[4], Silong Peng[5]**

**1** Department of Biotechnology and Genetic Engineering, Mawlana Bhashani Science and Technology University, Tangail, Bangladesh, **2** Department of Biochemistry and Molecular Biology, Mawlana Bhashani Science and Technology University, Tangail, Bangladesh, **3** Department of Computer Science and Engineering, Islamic University, Kushtia, Bangladesh, **4** Department of Electronics, Graduate School of Engineering, Nagoya University, Nagoya, Japan, **5** Institute of Automation, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Beijing, China

\* habib@iu.ac.bd

## Abstract

Despite modern treatment, infertility remains one of the most common gynecologic diseases causing severe health effects worldwide. The clinical and epidemiological data have shown that several cancerous risk factors are strongly linked to Female Infertility (FI) development, but the exact causes remain unknown. Understanding how these risk factors affect FI-affected cell pathways might pave the door for the discovery of critical signaling pathways and hub proteins that may be targeted for therapeutic intervention. To deal with this, we have used a bioinformatics pipeline to build a transcriptome study of FI with four carcinogenic risk factors: Endometrial Cancer (EC), Ovarian Cancer (OC), Cervical Cancer (CC), and Thyroid Cancer (TC). We identified FI sharing 97, 211, 87 and 33 differentially expressed genes (DEGs) with EC, OC, CC, and TC, respectively. We have built gene-disease association networks from the identified genes based on the multilayer network and neighbour-based benchmarking. Identified TNF signalling pathways, ovarian infertility genes, cholesterol metabolic process, and cellular response to cytokine stimulus were significant molecular and GO pathways, both of which improved our understanding the fundamental molecular mechanisms of cancers associated with FI progression. For therapeutic intervention, we have targeted the two most significant hub proteins VEGFA and PIK3R1, out of ten proteins based on Maximal Clique Centrality (MCC) value of cytoscape and literature analysis for molecular docking with 27 phytoestrogenic compounds. Among them, sesamin, galangin and coumestrol showed the highest binding affinity for VEGFA and PIK3R1 proteins together with favourable ADMET properties. We recommended that our identified pathway, hub proteins and phytocompounds may be served as

new targets and therapeutic interventions for accurate diagnosis and treatment of multiple diseases.

## Introduction

In modern civilization, infertility is a significant threat globally, including in developing and developed countries. Infertility, according to Zegers-Hochschild et al. is the result of a malfunction in the reproductive system. There are two types of infertility in women: primary and secondary. Women with primary infertility have never been clinically diagnosed with pregnancy, but women with secondary infertility could not develop a pregnancy clinically even before being diagnosed [1]. Secondary infertility is more frequent than primary infertility for women [1–3] Inhorn (2014) reported that the prevalence of infertility up to 186 million people around the world with more incidence rate was confined to developing countries [4], including South Asia and Central Asia, Central and Eastern Europe, North Africa and the Middle East, and some regions of sub-Saharan Africa [1]. The women's infertility rate was calculated at around 8 percent, 13-14 percent, and 18 percent at the age of 19-26, 27-34, and 35-39 years, respectively [5]. Although the exact etiology of infertility in women was unknown, generic and cancerous risk factors might be associated with infertility. Ovarian cancer is one of the fatal of all gynecological diseases worldwide. Most commonly, two mutated genes, including breast cancer 1 (BRCA1) and breast cancer 2 (BRCA2) who were inherited, causing ovarian cancer and breast cancer in women [6]. Cirillo et al. found a linkage between ovarian cancer and irregular menstruation in their study [7]. Cervical cancer is the world's fourth most prevalent disease among females [8]. According to global report in 2018, approximately 570,000 patients were diagnosed with cervical cancer and 31,000 were died [9]. The study by J. Dor et al. mentions that cervical cancer has been linked to infertility-causing pelvic infections or adhesions in the past [10].

Endometrial cancer (EC) is the fifth most common cancer in women from developed countries, accounting for 4.8 percent of new cases and 2.1 percent of deaths [11]. Autosomal dominant mutations cause this disease in DNA mismatch repair (MMR) genes [12]. Patients who have a germline mutation in the MMR gene have a 20–70 percent lifetime risk of developing EC, depending on their individual circumstances [13]. Women who have endometriosis are more likely to have infertility, according to a study conducted by Bulun et al. [14]. The most frequent endocrine malignancy is thyroid cancer (TC) accounts for 3.4 percent of all malignancies diagnosed each year [15]. Thyroid cancer is caused by genetic and epigenetic changes, including mutations in genes of BRAF, RAS, PIK3CA, PTEN and so on [16]. Menstrual disorder and an increased risk of miscarriage in infertile women due to thyroid abnormalities are more frequent [17]. These mentioned risk factors can influence female infertility when they commonly share dysregulated gene expression (DEG) [18], and the molecular pathways that may trigger the influential factors to promote infertility and cancer can be exposed by analyzing protein-protein interaction (PPI), metabolic pathway and gene ontologies [19].

Additionally, several pathways may correlate important genes involved with the progression of infertility to various types of cancer. PIK3R1 is a p85 regulatory protein encoded by the Phosphatidylinositol-kinase regulatory subunit alpha gene that regulates the p110 catalytic subunit [20], where most frequent mutation occurs in ovarian cancer [21] and endometrial carcinomas [22], within the iSH2 domain. So, PIK3R1 maybe act as a therapeutic target for infertility and infertility-related cancer. On the other hand, VEGFA can also be a therapeutic

target in various infertility mediated gynecological cancer like ovarian [23] and cervical cancer [24].

Network pharmacology is a new approach that combines computer science and medicine by building and visualizing a "multi-gene, multi-target, multi-pathway interaction network" to assess the drug's molecular mechanism [25]. The molecular docking method means that a small molecule is spatially attached to a macromolecular system and can detect the additional value at bindings used in structural drug design [26]. Researchers have recently made many integrative network analysis approaches to classify biomolecules' potential functions in other diseases [27, 28].

However, genetic experiments were carried out to study the effect of risk factors on female infertility, but network-based approaches were not implicated for such type of study [29]. Integrative research is crucial for understanding and identifying the disease-causing molecular pathways. Therefore we aimed to use a network-based bioinformatics pipeline to elucidate the cancerous risk factors and genes of female infertility; those mediate disease progression through gene expression profiling, metabolic pathway analysis, gene ontologies, and PPI sub-network interaction analysis. Moreover, we targeted hub proteins from PPI analysis for molecular modelling and ADMET analyses with 27 phytoestrogenic compounds for therapeutic intervention. Finally, we have used gold benchmark databases OMIM and DisGeNET, as well as literature to validate the known FI associated genes and molecular pathways.

## Materials and methods

### Details information of GEO datasets

In our study, five different microarray datasets were analyzed, including Female Infertility (FI), Endometrial Cancer (EC), Ovarian Cancer (OC), Cervical Cancer (CC), and Thyroid Cancer (TC) with the accession numbers GSE92324, GSE63678, GSE124766, GSE29570, and GSE6004, respectively from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database. Table 1 provides a summary of the information contained in the datasets.

### Data preprocessing and identification of DEGs

Based on microarray data, global transcriptome analysis was implicated in investigating the FI's gene expression profiles with four cancerous risk factors to determine the molecular characteristics of human disorders. As different errors are typically accounted for in the preparation and analysis of microarray data in our study. In each sample, the disease group or control group must be standardized. One of the most common methods for standardizing gene expression matrices is the Z-score transform [30]. If $X_{ij}$ is the expression value of the i-th gene

**Table 1. An overview of datasets with their geo-features and quantitative measurements is provided.**

| GSE number | Disease name | Source name | Control sample | Case sample | DEGs | Up regulated genes | Down regulated genes | References |
|---|---|---|---|---|---|---|---|---|
| GSE92324 | Female Infertility | endometrial tissue | 8 | 10 | 1201 | 651 | 550 | [46] |
| GSE63678 | Endometrial Cancer | endometrial tissue | 5 | 7 | 723 | 456 | 267 | [47] |
| GSE124766 | Ovarian Cancer | ovarian tissue | 8 | 20 | 2101 | 1045 | 1057 | [48] |
| GSE29570 | Cervical Cancer | biopsy tissue | 17 | 45 | 880 | 535 | 345 | [49] |
| GSE6004 | Thyroid Cancer | frozen surgical tissue | 3 | 12 | 541 | 338 | 203 | [50] |

in sample j, then Z-score transform standardization is obtained as follows:

$$Z_{ij} = \frac{X_{ij} - X_i}{\sigma_i} \tag{1}$$

Where $\sigma_i$ and $X_i$ are considered the standard deviation and mean of the expression value of the i-th gene expected inclusive samples, respectively. This transformation enables a clear comparison of the values of gene expression in various models and diseases. We have conducted a linear regression technique on data from a time series and this Z-score transformation for achieving a joint t-testing statistic between two groups. Data were converted into log2, and the linear regression model was employed to compute expression levels of each gene through the following formula:

$$Y_i = \beta_0 + \beta_1 X_i \tag{2}$$

Here, $Y_i$ is the gene expression value, and $X_i$ is the disease state in this case (disease or control). The model parameters $\beta_0$ and $\beta_1$ were calculated applying least squares.

In this study, we first compared diseased tissue against controls to identify differentially expressed genes (DEGs) associated with their respective pathology. For clarification, the inclusion criteria of a study subject might be female groups between the ages of 21 to 45 who have been diagnosed with different stages of diseases. Exclusion criteria for this study might be cell line data and male groups for thyroid cancers. To make consistent expression data from different platforms and avoid the problems of experimental systems, we normalized the gene expression data comprising disease state and control data by using the quantile normalization and Z-score transformation techniques through the NCBIs GEO2R online tool. We performed the analysis of the microarray data using well established Linear Models for Microarray Data (LIMMA). Then we used an unpaired t-test in which essential genes were selected to see if any genes are differentially expressed in disease and control by setting P-value < 0.05 and an absolute value of $|logfold - change| > 1$. Moreover, a two-way analysis of variance (ANOVA) test was performed to determine the statistical significance between groups. P-values were adjusted by the well-established Benjamini-Hochberg method and Bonferroni correction method as indicated. Based on varying the False discovery rate (FDR) threshold and standard statistical criteria; we considered P-value< 0.05 and $|logfold - change| \geq 1$ for up-regulated genes, while P-value<0.05 and $|logfold - change| \leq -1$ for finding downregulated genes. Then we have provided the rationale and justification for the selection of common DEGs by hypergeometric tests. We performed hypergeometric tests for the DEGs to establish their role as predictive diagnostic biomarkers for female infertility and four selected cancers represented in Table 2. Further, to determine the shared DEGs, we compared the FI dataset with four other selected diseases using Venny v2.1 web tool [31]. Then the gene-disease network (GDN) was created and visualized with Cytoscape v3.8.2 [32].

**Table 2. Validation of DEGs by hypergeometric test.**

| Disease Pair | Female infertility DEGs | Cancer DEGs | Common DEGs | Hypergeometric p-value | Representation factor |
|---|---|---|---|---|---|
| FI+OC | 1201 | 2101 | 211 | 1.531e-53 | 3.3 |
| FI+EC | 1201 | 723 | 97 | 6.089e-36 | 4.5 |
| FI+CC | 1201 | 880 | 87 | 2.316e-22 | 3.3 |
| FI+TC | 1201 | 541 | 33 | 5.293e-05 | 2.1 |

https://doi.org/10.1371/journal.pone.0265746.t002

## Pathway and functional enrichment analysis

The signaling pathways and gene ontology of FI were assessed using the web-based gene set enrichment analysis tool EnrichR for all the genes that were differentially expressed in FI and cancerous risk factors [33]. The research included selecting gene ontology (GO) (biological processes, molecular function and cellular component) and signalling pathways (KEGG, WiKi, and Reactome) as an annotation source [34]. For statistical significance, the adjusted p-value was considered to achieve enhancement results.

## Protein interactome analysis

For the assembly and study of the network protein-protein interaction, the web-based visualization software STRING has been employed, and Cytoscape (v3.8.2) was used for further evaluation [32, 35, 36]. The PPI network used a graph with no direction for representation, where the nodes denoted the proteins, and the edges indicated the proteins' interactions. To classify strongly interconnected proteins (i.e., hub proteins), we conducted a topology study applying the Cyto-Hubba plugin, and the Maximal Clique Centrality (MCC) algorithm was implicated.

## Protein and ligand preparation

The two proteins, VEGFA (PDB ID: 1FLT) and PIK3R1 (PDB ID: 5M6U) were prepared by retrieving the three-dimension crystal structure from RCSB PDB [36–38]. The 3D structure of target proteins was prepared by removing water through Discovery Studio (Studio 2015), and Pymol [39] software package and minimized energy with GROMOS96 43B1 force field through SWISS PDB Viewer [40]. We prepared a phytoestrogen dataset from previous experimental studies searching related literature in PubMed, Scopus databases, the web of science, Google scholar; those were used as significant compounds in several cancer treatments. Then SDF format of all ligand molecules was retrieved from the PubChem database [41]. Pymol Software was used to convert SDF format compounds to PDB format. For optimization and preparation of ligands, we used PyRx integrated mmff94 (Merck molecular force field) [42, 43].

## Virtual screening

Molecular docking by AutoDock wizard has been done to understand the link between the ligand and the drug compounds with PyRx. Ligand and target protein was considered as flexible and rigid, respectively, during ducking. Ligands with the lowest RMSD values and the highest negative docking scores were considered for ADMET evaluation. Finally, Discovery studio and Pymol tools were used to examine the docked pose for molecular interactions between ligands and receptors.

## ADMET analysis

Compounds with higher binding affinity were selected for the ADMET study. SwissADME was used to measure the Absorption, Distribution, Metabolism, and Elimination (ADME) properties of potent drug candidates by PubChem provided canonical SMILES of selected ligands [44], while pkCSM toxicity prediction tools were used to investigate toxicity [45].

## Overview of a proposed integrated pipeline

This research work developed and applied a multi-step quantitative approach as an integrated pipeline of bioinformatics and molecular docking approaches methodologies, as shown in Fig 1.
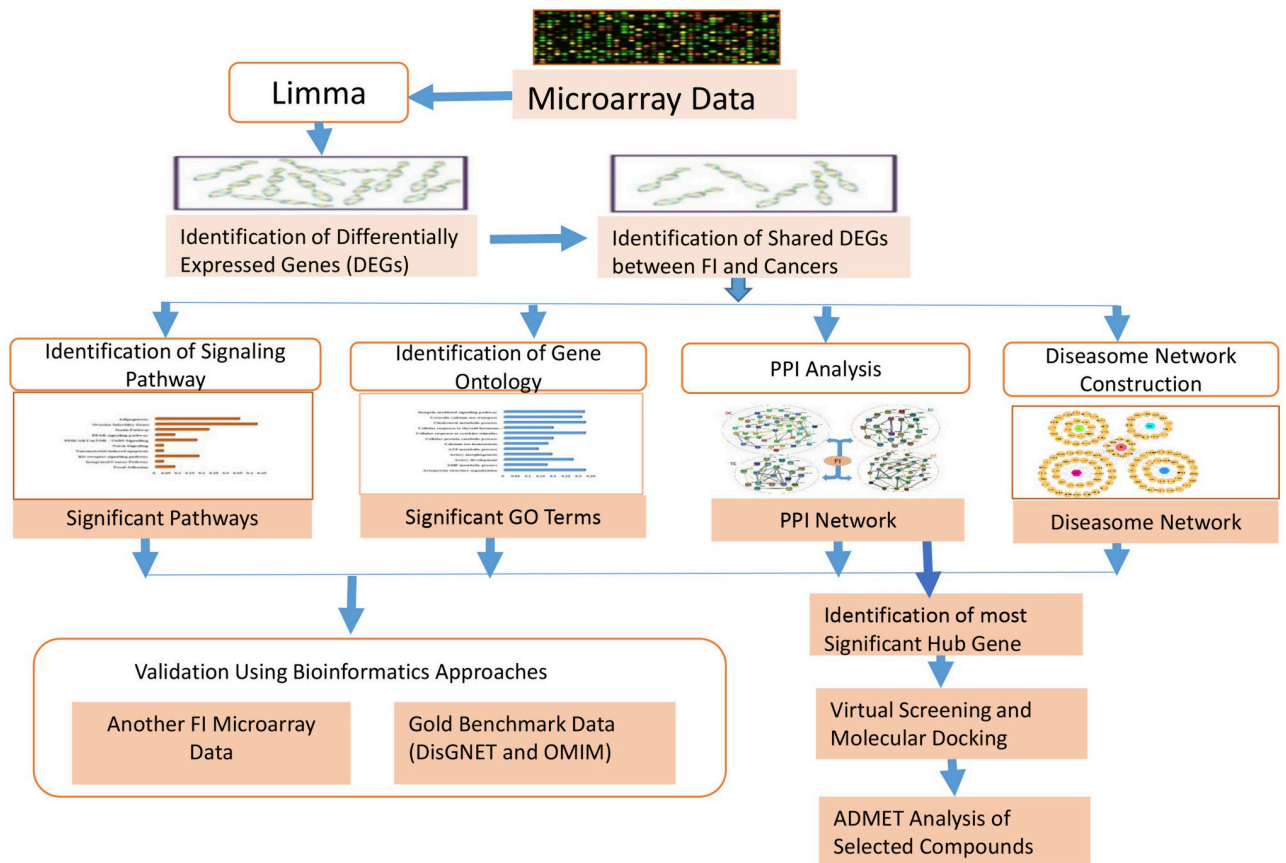
**Fig 1. Block diagram of the proposed bioinformatics and system biology pipeline.**

https://doi.org/10.1371/journal.pone.0265746.g001

## Results

### Gene expression analysis

The genetic variation patterns from endometrial tissues were used to identify dysregulated genes linked to FI during implantation failure cases of infertile patients were investigated and differentiated with the standard subject [46] by using R Bioconductor package Limma through NCBI GEO2R online tool. Compared to normal subjects, 1201 genes were differentially expressed (651 genes upregulated and 550 genes down-regulated). To investigate the relationship between the FI transcriptome and each risk factor, we implemented mRNA microarray data through a series of cross comparative analyses. The Venn diagram of Fig 2 shows that FI shared 97, 211, 87 and 33 genes with EC, OC, CC, and TC, respectively. Using Cytoscape, a gene-disease relationship network (GDN) based on FI data were created to find statistically meaningful associations among these risk factors. The relation between over and under-expressed genes is represented through the networks shown in Fig 3a and 3b. The most critical DEGs are defined using our proposed method, summarized in Table 1.

    Three fundamental genes were discovered in our research, including ABCC3, AEN, and ADAMTS1 are generally over-expressed among the FI, CC, and EC; ATP13A2 and AIFM1 are two upregulated genes found in the FI, CC, and OC; two crucial genes ALCAM and FAM13A are frequently upregulated in FI, TC, and OC. Three important genes PINLYP, GEMIN8, and ATP2A2 are associated with FI, EC, and OC.
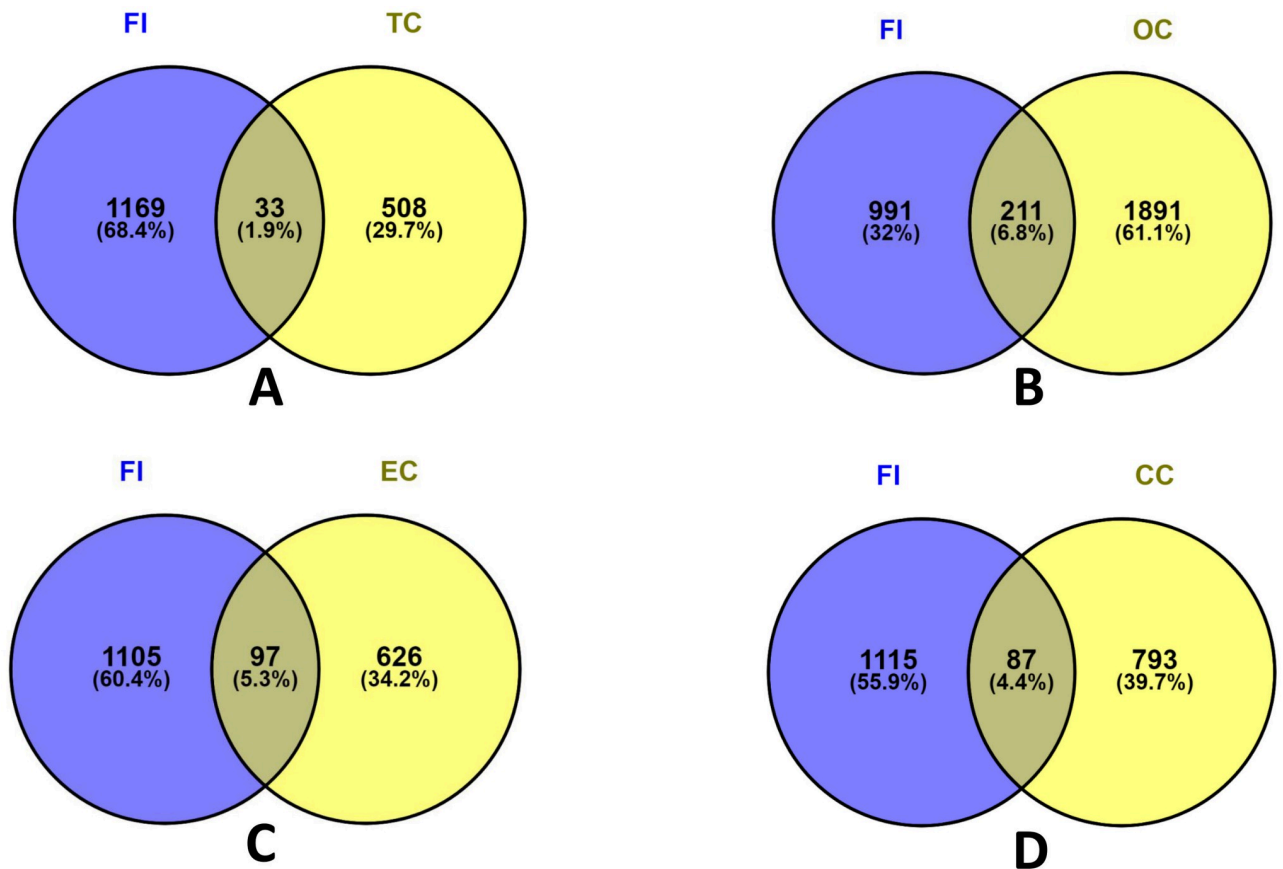
**Fig 2.** Venn diagram was applied to present all candidate targets of Female Infertility (FI) and four cancerous risk factors, where FI shared a common dysregulated gene between A. Thyroid cancer (TC), B. Ovarian cancer (OC), C. Endometrial cancer (EC), and D. Cervical cancer (CC).

On the other hand, the three down-regulated AK4, ACAA2, and GLIS3 genes are prevalent in the FI, CC, and OC; Two under-expressed genes BRE and ADGRL1 are shared in the FI, CC, and EC. The FI, TC, and CC are typical for one down-regulated gene ADAMTS9; one down-regulated FBXO2 gene is expected in the FI, OC, and TC. Besides, FI, OC, and EC are frequently shared one down-regulated gene AGPS.

### Pathway and functional association analysis

EnrichR online platform uses all differentially expressed common genes to find significant molecular pathways linked to FI and four risk factors through KEGG, WiKi, and the Reactome pathway database. The enrichment study identified 234, 253, and 102 pathways among KEGG, WiKi, and Reactome databases, respectively. Particularly, we considered only ten significant pathways of each pathway database after p-value adjustments associated with the cancer progression. The most considerable pathways have been found which are Proteoglycans in cancer (hsa05205), Notch Signaling (hsa04330), PPAR signalling pathway (hsa03320), Pathways in cancer (hsa05200), Diseases of signal transduction (R-HSA-5663202), CD28 dependent PI3K/Akt signalling (R-HSA-389357), ABC transporters in lipid homeostasis (R-HSA-1369062), VEGFA-VEGFR2 Signaling Pathway (hsa04370), and Thyroid hormone signalling pathway (hsa04919). These mentioned pathways and other significant mutual pathways are given in Fig 4a–4c.
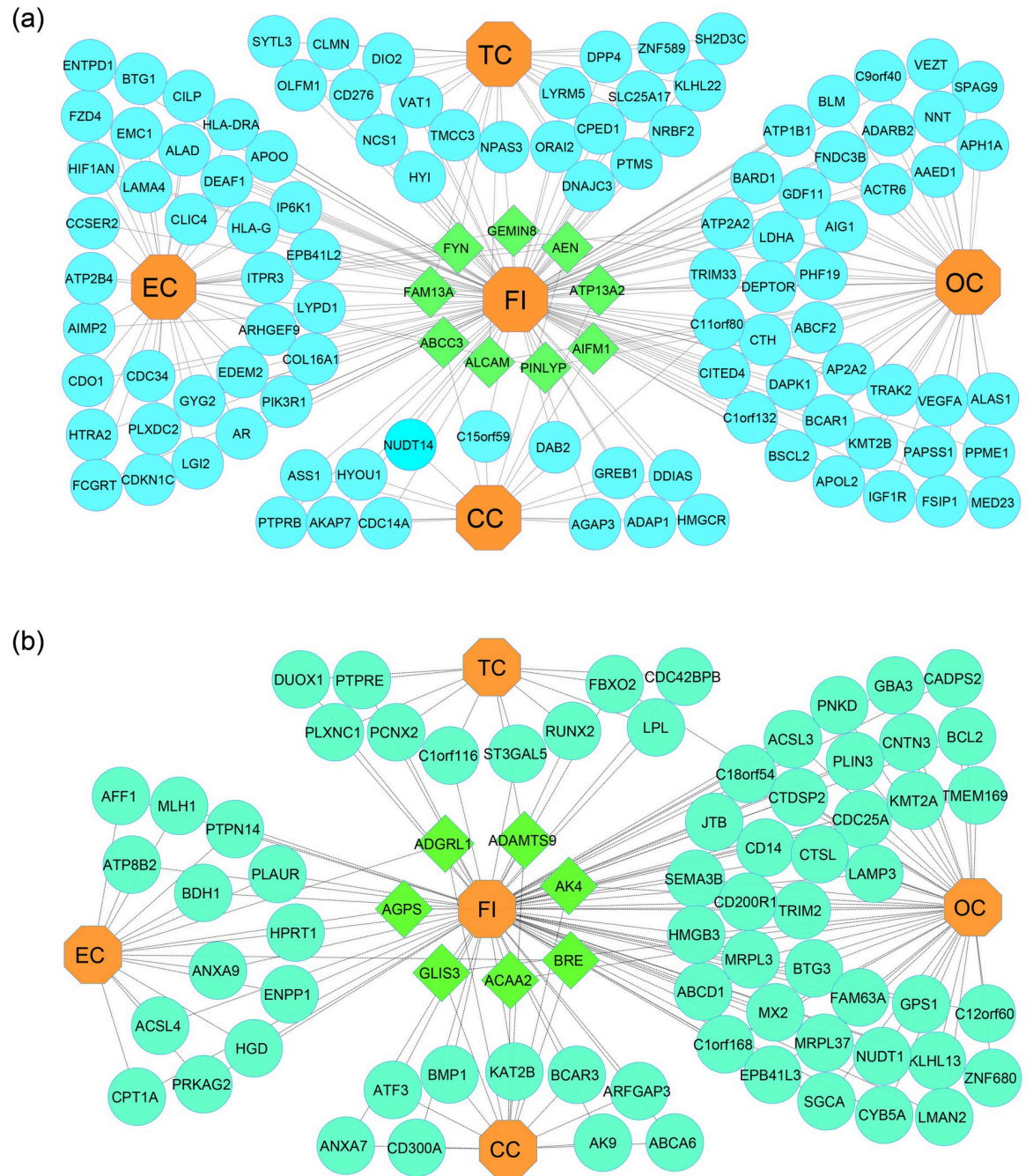
**Fig 3. a**: Gene-Disease network of common DEGs having upregulated genes between Female Infertility (FI) with Endometrial Cancer (EC), Ovarian Cancer (OC), Cervical Cancer (EC), and Thyroid Cancer (TC). Octagon-shaped and light red color nodes represent four risk factors, while round-shaped and sky color nodes represent DEGs. Square-shaped and green color nodes indicate DEGs are common among FI and four risk factors. **b**: Gene-Disease network of common DEGs having down-regulated genes between Female Infertility (FI) with Endometrial Cancer (EC), Ovarian Cancer (OC), Cervical Cancer (EC) and Thyroid Cancer (TC). Octagon-shaped and light red color nodes represent four risk factors, while round-shaped and sky color nodes represent DEGs. Square-shaped and green color nodes indicate DEGs are common among FI and four risk factors.

https://doi.org/10.1371/journal.pone.0265746.g003

**(a) KEGG Pathways associated with significantly common differentially expressed genes of the FI with risk factors**



**(b) WiKi Pathways associated with significantly common differentially expressed genes of the FI with risk factors**



**(c) Reactome Pathways associated with significantly common differentially expressed genes of the FI with risk factors**

**Fig 4.   a**: The top 20 signalling pathways from KEGG enrichment analysis were showed by the bar diagram with p-values.
**b**: The top 20 signalling pathways from WiKi enrichment analysis were showed by the bar diagram with p-values. **c**: The
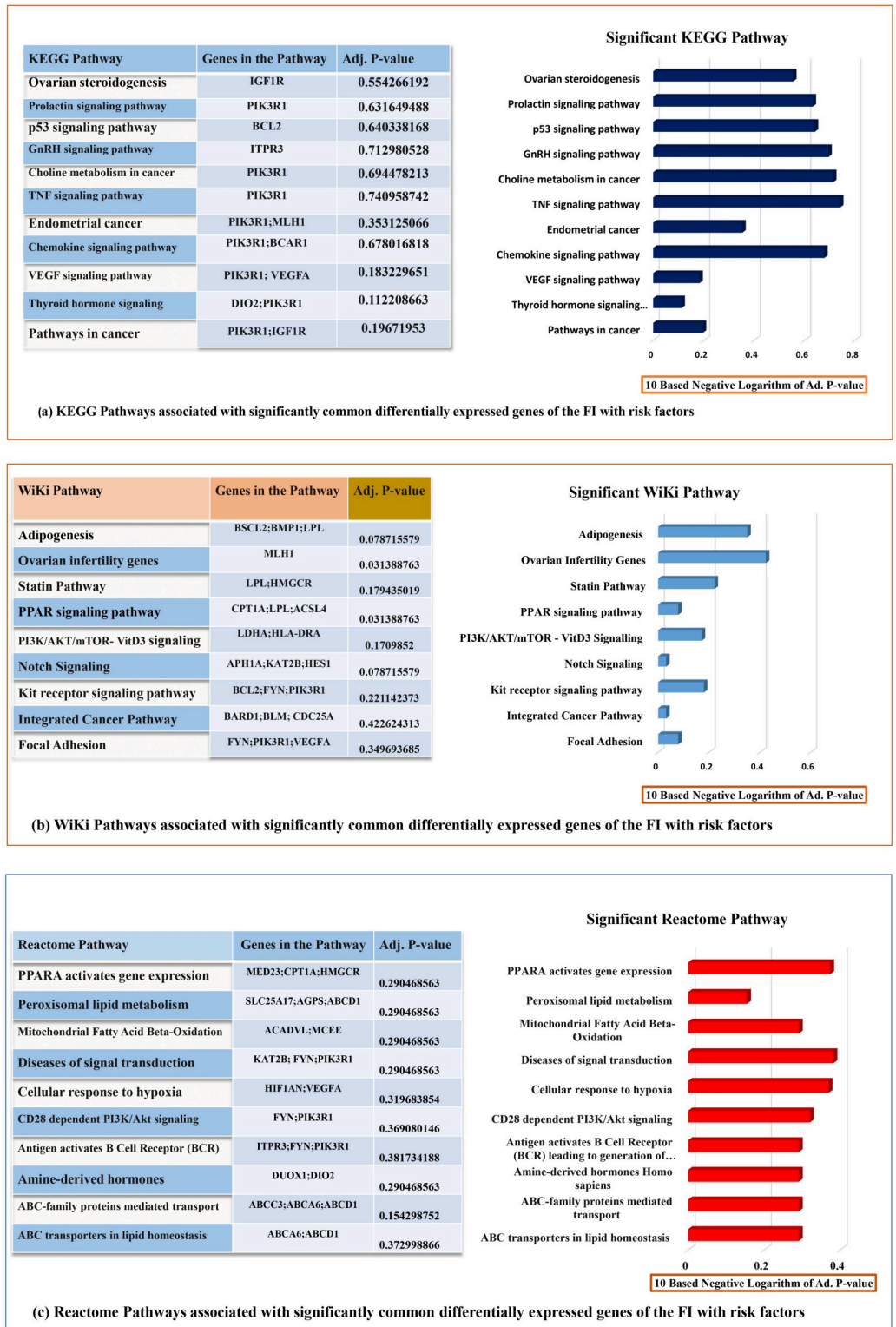top 20 signalling pathways from Reactome enrichment analysis were showed by the bar diagram with p-values.

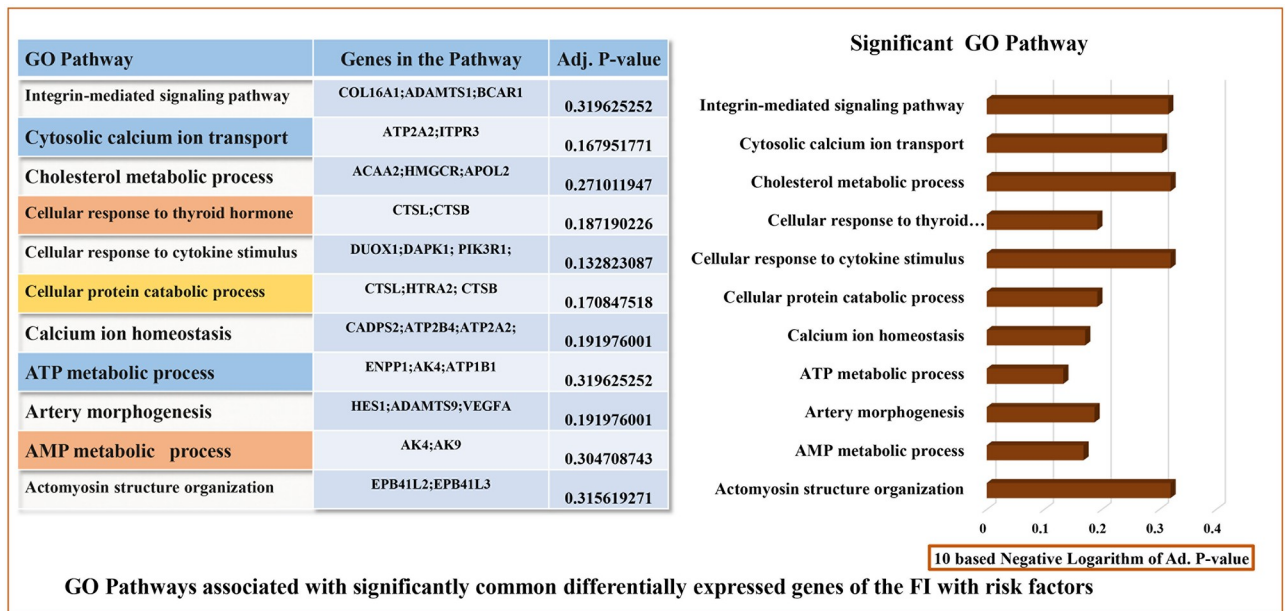https://doi.org/10.1371/journal.pone.0265746.g004

**Fig 5. The top 20 signalling pathways from GO biological process enrichment analysis were showed by the bar diagram with p-values.**

We used an ontology enrichment study to find 1586 GO terms (biological processes) for the commonly dysregulated genes among FI and cancerous risk factors. The primary significant GO classes are actomyosin structure organization, integrin-mediated signalling pathway, cholesterol metabolic process, ATP metabolic process, and cellular response to cytokine stimulus depicted in Fig 5.

## Establishment protein-protein interaction network

Both differentially expressed genes found in the FI and other risk factors to build the PPI network are shown in Fig 6. Each node in the network represents a protein and an edge represents the connection between two proteins. In addition, the network is split into four clusters, each of which represents a risk factor.

Using the Cyto-Hubba plugin, a generalized PPI network was created for topological analysis [61], displaying the ten most essential hub proteins in Fig 7: VEGFA, PIK3R1, BCAR1, AR, CPT1A, ACSL3, ACSL4, IGF1R, LPL, and HMGCR. Interestingly, each of the RUNX2, BCAR1, VEGFA, and PIK3R1 proteins belongs to one of the four clusters, suggesting that the FI shares them and the other three risk factors have interacted with in various clusters by other protein. On the other hand, BCL2 and ITPR3 belong to three groups and interrelate with other proteins in the network.

Four out of the ten hub proteins are dysregulated due to OC; EC dysregulates four, and two proteins are dysregulated by CC and TC. For docking purposes, these hub proteins may be the target proteins. The functionally significant ten hub proteins with molecular function were tabulated in Table 3.

## Virtual screening of retrieved compounds against VEGFA and PIK3R1 proteins

Virtual skimming with molecular docking is another technique to identify a lead compound in the drug discovery process. In our study, we took ten proteins with their rank of significance
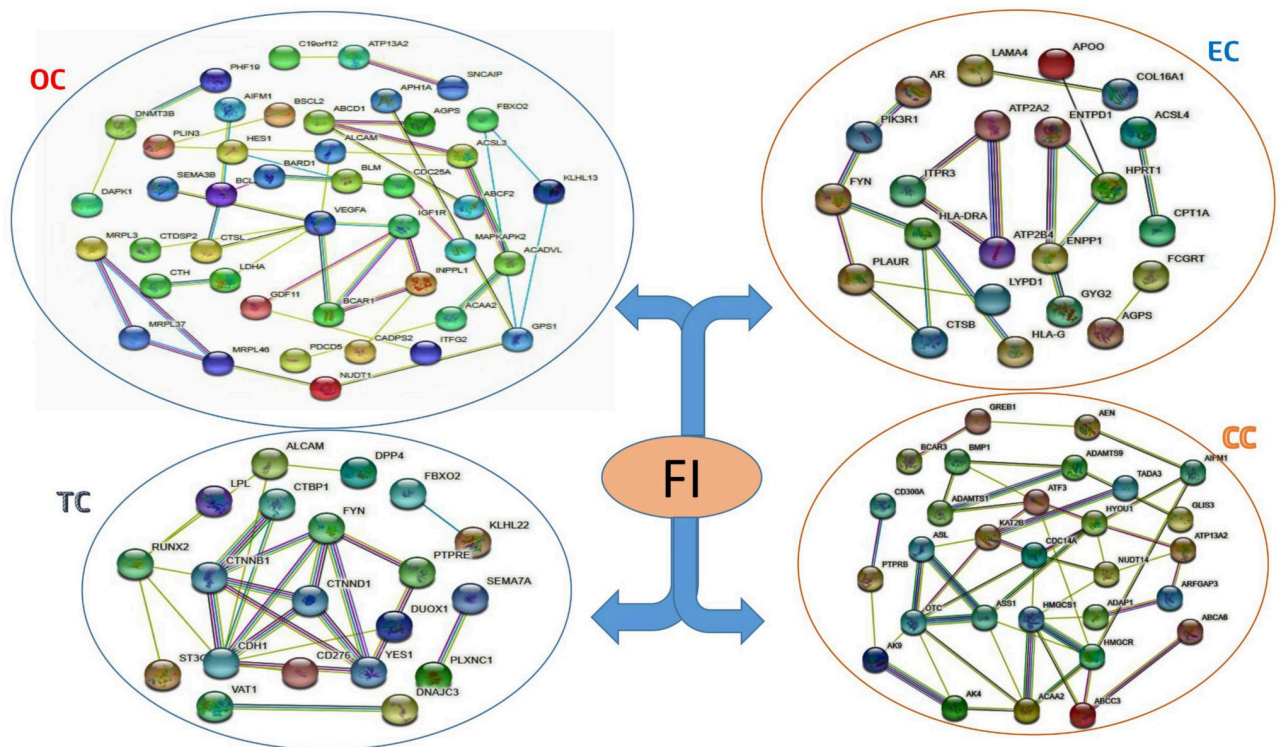
**Fig 6. Protein-Protein interaction network for the common paired DEGs between Female Infertility (FI) and Endometrial Cancer (EC), Ovarian Cancer (OC), Cervical Cancer (EC), and Thyroid Cancer (TC).** The network nodes depict target proteins, and the edges represent protein-protein relationships.

by cyto-Hubba plugin analysis. The significant protein, VEGFA and PIK3R1, were selected to conduct molecular docking purposes based on two criteria. First, maximal clique centrality (MCC) algorithm of Cytohubba plugin, VEGFA and PIK3R1 protein in a ranking of first and second position respectively which indicate most two significant proteins. Second, by searching the literature, we found VEGFA and PIK3R1 protein are the most interconnected among our selected cancer type risk factors, including Endometrial cancer [62, 63], Ovarian cancer [23, 64], Cervical cancer [24, 65] and Thyroid cancer [66, 67]. Then, a total of 27 phytoestrogenic compounds with control were chosen for molecular docking, showing an anti-cancer activity through literature analysis. Based on binding affinity, four compounds, including sesamin, alpha-mangostin, galangin, coumestrol, and quercetin, are considered for further analysis. This research used bevacizumab and wortmannin as a positive control ligand for VEGFA and PIK3R1 proteins, respectively. Wortmannin was a potent inhibitor (binding affinity -6.8 kcal/mol) against PIK3R1 protein that encodes P85 regulatory subunit, which regulates the P110 catalytic subunit in inter-Src homology-2 (iSH2) domain to the plasma membrane. Besides, bevacizumab (binding affinity -5.5 kcal/mol) can be used as a first approved angiogenesis inhibitor via VEGFA targeting protein [68]. The selected compounds with compound names and binding affinity are given in Fig 8a and 8b.

## Molecular interaction analysis of selected compounds

The active site of VEGFA (PDB ID: 1FLT) and PIK3R1 (PDB ID: 5M6U) proteins were predicted using CASTp server [69]. The domain part of PIK3R1 (chain B) protein is inter-Src homology-2 (iSH2) provided 400 to 600 amino acid sequence identified through Interpro
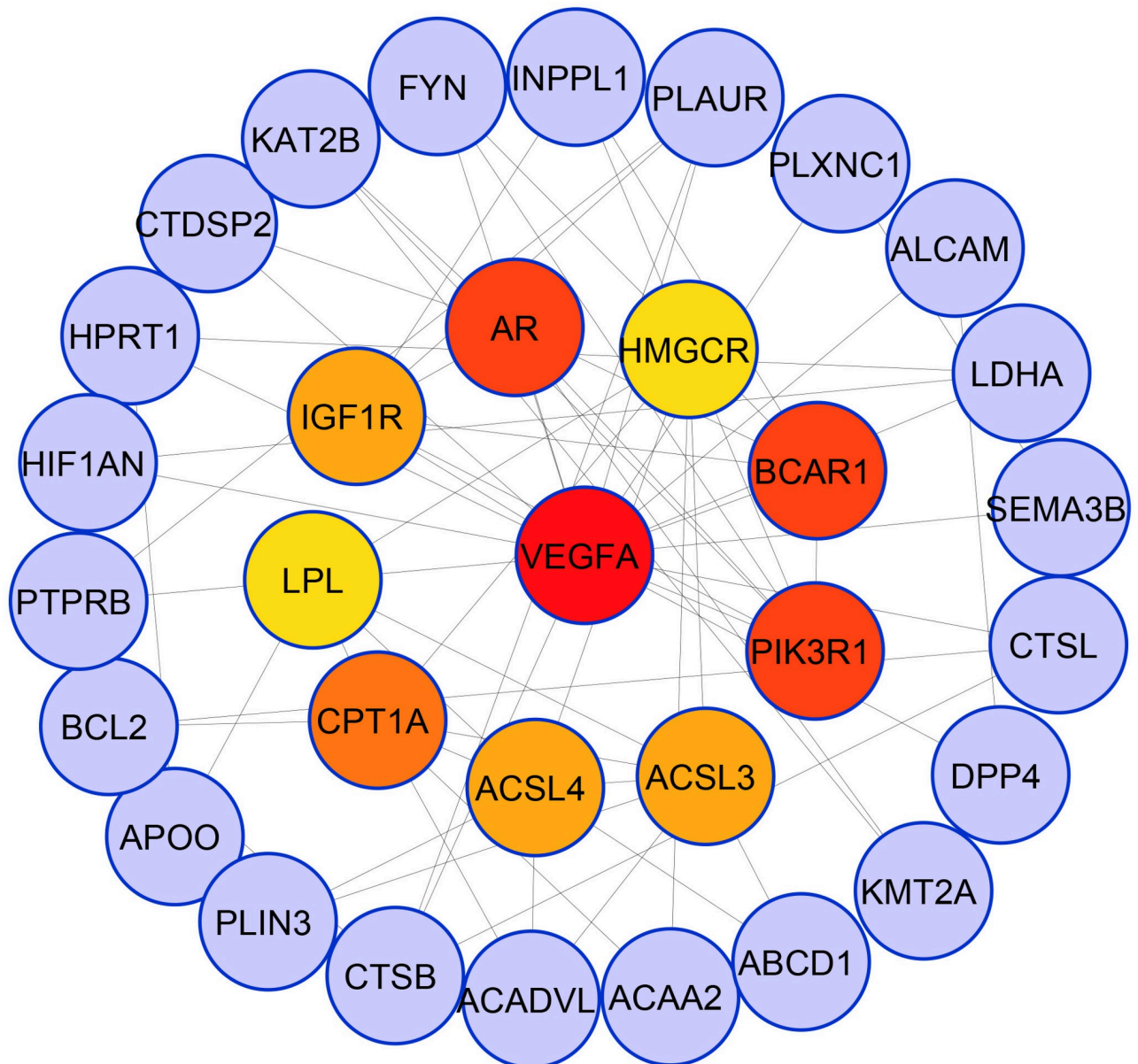
**Fig 7. The simplified PPI network depicting the hub proteins.** The ten most significant hub proteins are ranked with red to the yellow-colored gradient.

https://doi.org/10.1371/journal.pone.0265746.g007

server [70, 71]. Several frequent mutations in the inter-Src homology-2 (iSH2) domain, including Y504D, Q552K, I559N, D560Y, N564D, D569Y, R574T, T576del, W583del, N595K, and N600H mutants may disrupt the inhibitory interaction of the C2 domain with iSH2 and this residue act as a hotspot to occur mutation for oncogenesis like breast, endometrium and ovarian cancer particularly at amino acids 456-469 and 564-575 [70, 72–74]. The role of some functionally important mutated amino acids in the iSH2 domain were tabulated in Table 4 which are responsible for several tumor progression. In our investigation, we did not include the mutated form of PIK3R1 protein. In addition, all the selected phytochemicals bind to the only active site of this protein, and no one ligand can be bound to mutate amino acids. Our result demonstrated that the amino acid position 436-599 is expected to be conserved in the PIK3R1

**Table 3. Molecular functions of the ten hub proteins associated with the four risk factors.**

| Protein name | Description | MCC value | Molecular function | Ref. |
|---|---|---|---|---|
| VEGFA | Vascular endothelial growth factor A | 58 | Responsible for tumor growth and angiogenesis | [51] |
| PIK3R1 | Phosphoinositide-3-Kinase Regulatory Subunit 1 | 57 | Underexpression may cause PI3K pathway activation and contribute to growth and progression | [52] |
| BCAR1 | Breast cancer anti-estrogen resistance protein 1 | 38 | Overexpression performs a key role in carcinogenesis, most likely via activating the p38 MAPK pathway. | [53] |
| AR | Androgen receptor | 38 | Stimulates cell proliferation prostate cancer via regulating the G1/S transition in the cell cycle | [54] |
| CPT1A | Carnitine Palmitoyltransferase 1A | 35 | CPT1A expression is required for adipocytes to promote tumor growth and initiation | [55] |
| ACSL3 | Acyl-CoA Synthetase Long-Chain Family Member 3 | 34 | Overexpression of ACSL3 enhanced tumor cell proliferation, migration, and invasion, all of which favoured malignancy. | [56] |
| ACSL4 | Acyl-CoA Synthetase Long Chain Family Member 4 | 34 | Increase uncontrolled cell proliferation, facilitate tumor invasion and evade programmed cell death | [57] |
| IGF1R | Insulin Like Growth Factor 1 Receptor | 34 | Regulates cell proliferation and survival | [58] |
| LPL | Lipoprotein Lipase | 28 | Plays important roles in inflammation, obesity and cancer development | [59] |
| HMGCR | 3-Hydroxy-3-Methylglutaryl-CoA Reductase | 28 | Positively regulates the growth, migration, and tumorigenesis of several cancer cells | [60] |

protein active site. In a VEGFA protein, the domain site has resided in 39 to 135; a conserved site is found in 75 to 87 amino acid positions. Therefore, in our study, the docked compounds interaction showed that all compounds interact same binding pocket with the identical catalytic residues, including Gln 497, Ser 505, Tyr 508, Ile 524, Asn 527, and Tyr 528 for PIK3R1 protein (Table 5), while Glu 64, Ser 50, Cys 68, Ile 46 for VEGFA protein (Table 6). The ligand forms interaction with substrate-binding pocket residues were visualized using the BIOVIA discovery studio visualizer.

Sesamin formed two hydrogen bonds at Glu 502, Ser 505, and two hydrophobic bonds at Lys 506 and Tyr 528. On the contrary, alpha -mangostin showed one hydrogen bond at Ser 505 and one pi-pi-T shaped at Tyr 528, four alkyl bonds at Lys 506, Ile 509, Met525 and Lys 532. Besides, PIK3R1 protein and Galangin complex stabilized by one hydrogen bond at Gln497 and two hydrophobic bonds at Tyr 508 and Ile 524 positions. Coumestrol formed two hydrogen bonds at Gln 501, Ser 505, one pi-pi-T shaped at Tyr 508 and one Pi-alkyl bond at Ile 524 positions. The positive control Wortmannin formed two hydrogen bonds at Gln 497 and Asn 527 with two hydrophobic interactions, Pi-Pi Stacked and Pi-alkyl bond at Tyr 504 position (Fig 9). In the VEGFA protein, sesamin formed seven hydrogens and four hydrophobic bonds, while galangin formed five hydrogens and one hydrophobic interaction. Cumestrol and quercetin formed four and eight hydrogen bonds, respectively (Fig 10).

## Chemo-informatics properties analysis

The ADMET properties, including physicochemical, lipophilicity, water-solubility, pharmacokinetics, drug-likeness, medicinal chemistry, and toxicity of the selected potent four compounds was tabulated in Table 7. The physicochemical properties of our targeted compound remained in expected value. The pharmacokinetics parameters show the high gastrointestinal (GI) absorption rate of the selected compounds. In terms of drug-like activity, none of the four compounds violates Lipinski's rule of five. The water solubility reveals that three compounds are soluble in water, except alpha-mangostin are poorly soluble. The toxicity tests revealed that the human ether go-go-gene (hERG) is not inhibited by these compounds. The Ames test result data also express that only galangin compounds are not mutagenic compared with
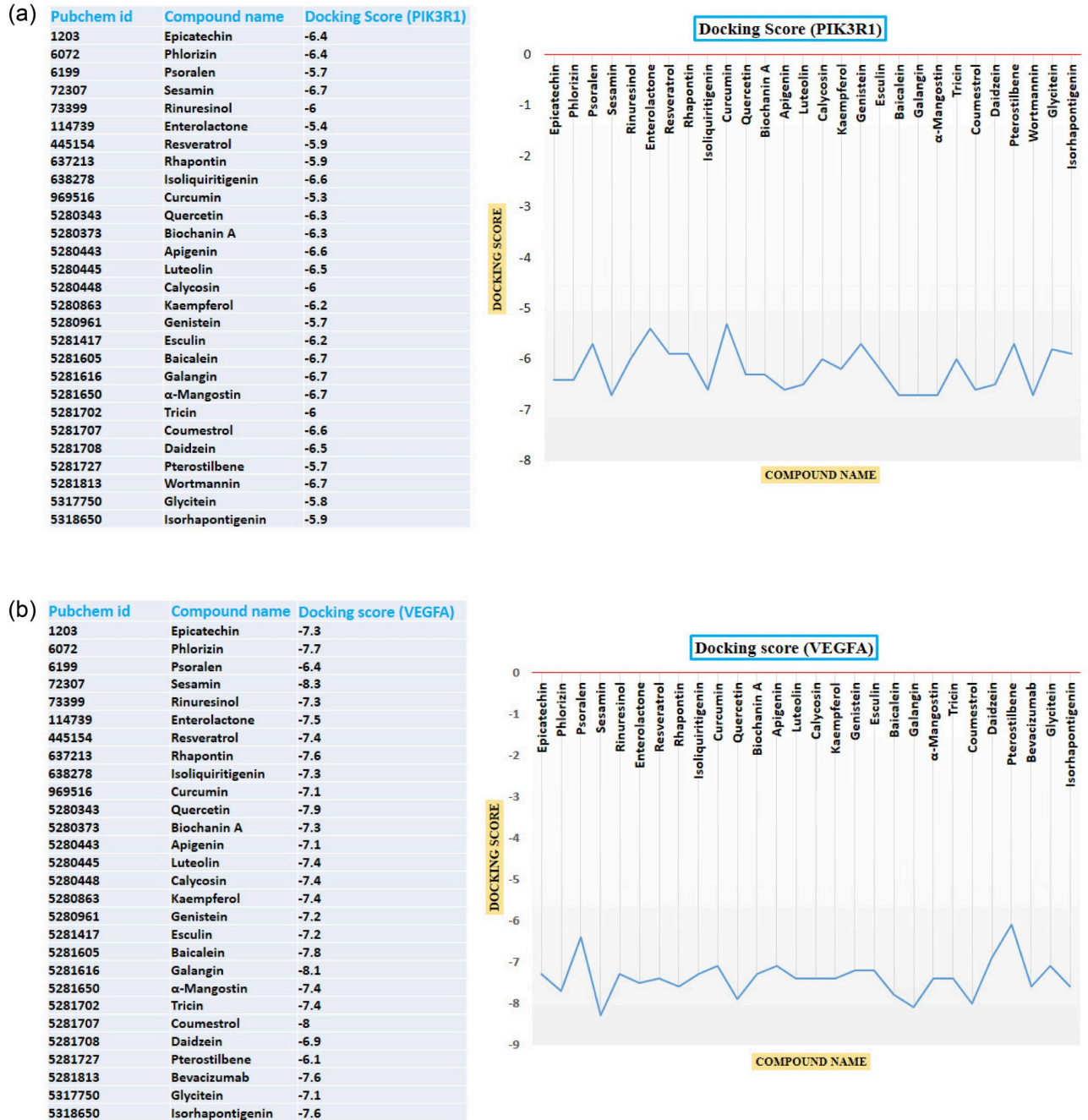
(a)

| Pubchem id | Compound name | Docking Score (PIK3R1) |
|---|---|---|
| 1203 | Epicatechin | -6.4 |
| 6072 | Phlorizin | -6.4 |
| 6199 | Psoralen | -5.7 |
| 72307 | Sesamin | -6.7 |
| 73399 | Rinuresinol | -6 |
| 114739 | Enterolactone | -5.4 |
| 445154 | Resveratrol | -5.9 |
| 637213 | Rhapontin | -5.9 |
| 638278 | Isoliquiritigenin | -6.6 |
| 969516 | Curcumin | -5.3 |
| 5280343 | Quercetin | -6.3 |
| 5280373 | Biochanin A | -6.3 |
| 5280443 | Apigenin | -6.6 |
| 5280445 | Luteolin | -6.5 |
| 5280448 | Calycosin | -6 |
| 5280863 | Kaempferol | -6.2 |
| 5280961 | Genistein | -5.7 |
| 5281417 | Esculin | -6.2 |
| 5281605 | Baicalein | -6.7 |
| 5281616 | Galangin | -6.7 |
| 5281650 | α-Mangostin | -6.7 |
| 5281702 | Tricin | -6 |
| 5281707 | Coumestrol | -6.6 |
| 5281708 | Daidzein | -6.5 |
| 5281727 | Pterostilbene | -5.7 |
| 5281813 | Wortmannin | -6.7 |
| 5317750 | Glycitein | -5.8 |
| 5318650 | Isorhapontigenin | -5.9 |



(b)

| Pubchem id | Compound name | Docking score (VEGFA) |
|---|---|---|
| 1203 | Epicatechin | -7.3 |
| 6072 | Phlorizin | -7.7 |
| 6199 | Psoralen | -6.4 |
| 72307 | Sesamin | -8.3 |
| 73399 | Rinuresinol | -7.3 |
| 114739 | Enterolactone | -7.5 |
| 445154 | Resveratrol | -7.4 |
| 637213 | Rhapontin | -7.6 |
| 638278 | Isoliquiritigenin | -7.3 |
| 969516 | Curcumin | -7.1 |
| 5280343 | Quercetin | -7.9 |
| 5280373 | Biochanin A | -7.3 |
| 5280443 | Apigenin | -7.1 |
| 5280445 | Luteolin | -7.4 |
| 5280448 | Calycosin | -7.4 |
| 5280863 | Kaempferol | -7.4 |
| 5280961 | Genistein | -7.2 |
| 5281417 | Esculin | -7.2 |
| 5281605 | Baicalein | -7.8 |
| 5281616 | Galangin | -8.1 |
| 5281650 | α-Mangostin | -7.4 |
| 5281702 | Tricin | -7.4 |
| 5281707 | Coumestrol | -8 |
| 5281708 | Daidzein | -6.9 |
| 5281727 | Pterostilbene | -6.1 |
| 5281813 | Bevacizumab | -7.6 |
| 5317750 | Glycitein | -7.1 |
| 5318650 | Isorhapontigenin | -7.6 |



**Fig 8. a**: Docking scores of all compounds. Most of the compounds have a range between -5 to -7 kcal/mol binding affinities. **b**: Docking scores of all compounds. Most of the compounds have a range between -5 to -8.3 kcal/mol binding affinities.

others. All the compounds show a negative response in skin sensitization and hepatotoxicity properties which is a good sign of a predictable drug molecule.

## Potential validation targets using FI dataset and gold benchmark databases

To prove the common DEGs linked with FI, we added another dataset (GSE16532) for FI expression. The validation of the FI dataset is an Expression profiling array of data based on

**Table 4. Role of some functionally essential mutated amino acids in the iSH2 domain.**

| Protein name | Mutant form | Function |
|---|---|---|
| PIK3R1 | N564D | to increase PI3K activity in endometrial tumors as a positive control |
| | R574T and T576del | to enhance roughly cell growth for oncogenic transformation |
| | G560Y | to disrupt the helix or change the length of the iSH2 domain as a single mutant |
| | E542K and E545K | contribute to various aspects of tumor development via regulating cellular pathways |

https://doi.org/10.1371/journal.pone.0265746.t004

endometrium biopsy tissues from 4 infertile patients and 5 fertile women during the mid-secretory phase (LH +7) [3]. The suggested method then uses two gold-standard databases, such as Online Man Mendelian Heritage (OMIM) and DisGeNET as well as literature to authenticate the genes found in our research that indicate potential disease risk. We examined the overlap between all DEGs and a gene validation expression dataset. We found 16 DEGs correlated with each risk factors, including FI shared 9 DEGs with OC and 4 DEGs with EC, while TC and CC shared 21 DEGs, respectively. To verify our established findings, we have also incorporated Online Mendelian Inheritance for Man (OMIM) and DisGeNET databases to validate identified drug-gene associations [75]. As seen in Fig 11 genes linked with CC, OC, TC, and EC are likely to positively co-related with FI. Overall, our results fill in significant deficiencies in our knowledge of FI pathobiology and could open up new directions for establishing mechanistic correlations between FI and various cancerous risk factors.

**Table 5. Docking simulation results of 4 compounds with docking score energy and interaction with amino acids against PIK3R1 protein.**

| Compound Name | Docking energy (kcal.mol-1) | Hydrogen Bond Ligand atom-Amino Acid Distance (Å) | Hydrophobic Bond Interaction-Amino Acid Distance (Å) |
|---|---|---|---|
| Sesamin | -6.7 | SER505:HG(2.99288) | Pi-Pi T- shaped-TYR528 (4.79069) |
| | | GLU502:HA(2.58332) | Pi-Alkyl—LYS506 (3.77723) |
| | | GLU502: O (2.34831) | |
| | | GLU502:OE2 (3.06726) | |
| α-mangostin | -6.7 | SER505:HG(2.51696) | Pi-Pi T-shaped -TYR528(4.7623) |
| | | SER505:HB2(2.66301) | Pi-Pi T-shaped -TYR528(5.12308) |
| | | | Alkyl—ILE509(5.44374) |
| | | | Alkyl—MET525(4.67183) |
| | | | Alkyl—LYS506(4.17836) |
| | | | Alkyl—ILE509(5.41536) |
| | | | Alkyl—LYS532 (4.48014) |
| Galanzin | -6.7 | GLN497:HE21(2.26174) | Pi-Pi T-shaped—TYR508(4.75895) |
| | | | Pi-Alkyl—ILE524(4.34365) |
| Coumestrol | -6.6 | GLN501:HA (2.89717) | Pi-Pi T-shaped—TYR508(4.75166) |
| | | SER505:HB1(2.96075) | Pi-Alkyl—ILE524(4.51341) |
| | | | Pi-Alkyl—ILE524(5.34046) |
| | | | Pi-Alkyl—ILE524 (4.21717) |
| Wortmannin | -6.8 | GLN497:HE21(2.82878) | Pi-Pi Stacked—TYR504(4.64417) |
| | | ASN527:OD1(3.04664) | Pi-Alkyl—TYR504(5.17453) |
| | | GLN497:HE21(2.82878) | |
| | | ASN527:OD1(3.04664) | |

https://doi.org/10.1371/journal.pone.0265746.t005

**Table 6. Docking simulation results of 4 compounds with docking score energy and interaction with amino acids against VEGFA protein.**

| Compound Name | Docking energy (kcal.mol-1) | Hydrogen Bond Ligand atom-Amino Acid Distance (A) | Hydrophobic Bond Interaction-Amino Acid Distance (A) |
|---|---|---|---|
| Bevacizumab (Control) | -5.5 | Y:ARG133:HH11 (2.31056) | Y:TYR139 (3.80127) |
| | | Y: LYS217:HN (2.29829) | V: LYS16 (4.51013) |
| | | Y:TYR139:OH (2.3838) | V:MET18 (5.29144) |
| | | Y:LEU215:O (2.6249) | V:MET18 (4.84131) |
| | | | Y: TYR139 (4.87068) |
| | | | Y: LYS217 (4.78722) |
| Coumestrol | -8 | W: GLU64:HN (2.10073) | W: GLU64:OE2 (4.70255) |
| | | W: ASP63:HA (2.67885) | W: GLU64:OE2 (4.81957) |
| | | X:THR226:HA (2.96673) | V:ILE46 (3.85383) |
| | | V:PHE47:HN (2.86934) | V:ILE46 (3.62054) |
| | | | V: ILE46 (5.07809) |
| | | | V: ILE46 (4.35264) |
| Galangin | -8.1 | V: CYS61:HN (1.98732) | V: GLU64:OE1 (4.51529) |
| | | V: CYS68:HN (2.11427) | |
| | | V: CYS60:HA (2.50449) | |
| | | V: GLU67:HA (2.53521) | |
| | | W: SER50:HG (2.86959) | |
| Quercitin | -7.9 | W:ASN62:O (2.9349) | W:ASP63:OD1 (3.75656) |
| | | V:ASP34:OD2 (1.99649) | V:ILE46 (1.99649) |
| | | V:TYR45:O (4.98128) | V:ILE46 (4.98128) |
| | | V:PHE36:HN (3.19653) | |
| | | V: SER50:HG (2.56624) | |
| | | V: SER50:HB1 (4.77306) | |
| | | V: PHE47:HN (2.9349) | |
| | | W: GLU64:HN (2.73851) | |
| Sesamin | -8.3 | V: CYS61:HN (2.47451) | V: GLU64:OE1 (4.20541) |
| | | V:GLY59:O (2.54908) | C- W:ILE46 (4.22532) |
| | | V:CYS57:O (2.9038) | C- V:CYS68 (4.53856) |
| | | V:CYS68:O (2.9161) | V:CYS68 (4.92696) |
| | | V: CYS57:O (3.08045) | |
| | | V: CYS68:HN (2.9202) | |
| | | W: SER50:HG (2.63393) | |

https://doi.org/10.1371/journal.pone.0265746.t006

## Discussion

Network-based bioinformatics and in- silico approaches offer more outstanding guidance for solid interaction and give a systemic view of mechanistic role in disease development. These studies can uncover new connections between the FI and other likelihood factors. Therefore, we analyzed dysregulation in the FI patients' gene expression, molecular pathways, ontologies, and PPIs. Our results found multiple essential genes to determine the FI's treatment objectives. Aside from that, our research identified and explained numerous metabolic roles associated with these genes. Our study of gene expression has shown that FI is closely linked to endometrial cancer (97 genes), ovarian cancer (211 genes), cervical cancer (87 genes), and thyroid cancer (33 genes) with sharing a maximum gene list. We created and examined the PPI linkage to recognize the central mechanism behind the FI better.

Additionally, diseases-associated genes play a critical role in human interaction via the pathways. In this study, we identified several significant pathways linked to cancer
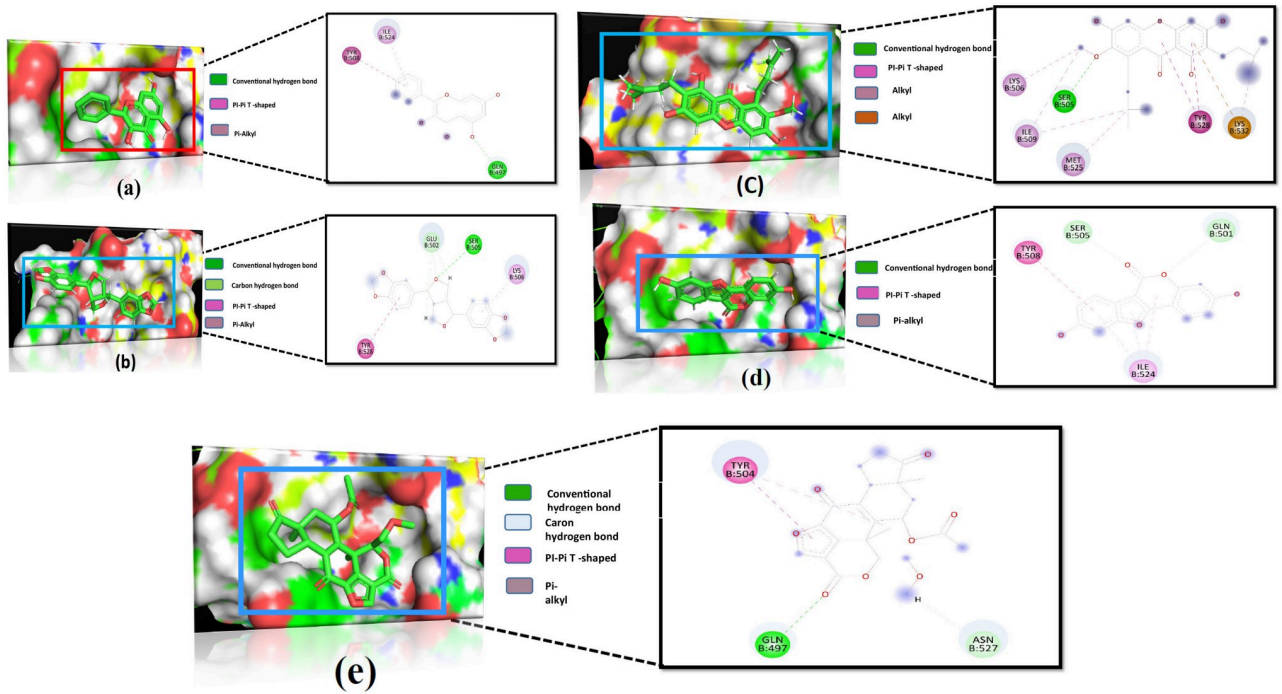
**Fig 9. Molecular interactions analysis of selected compounds against PIK3R1 (PDB ID: 5M6U) of (a) galangin, (b) sesamin, (c) alpha-mangostin, (d) coumestrol and (e) control.**
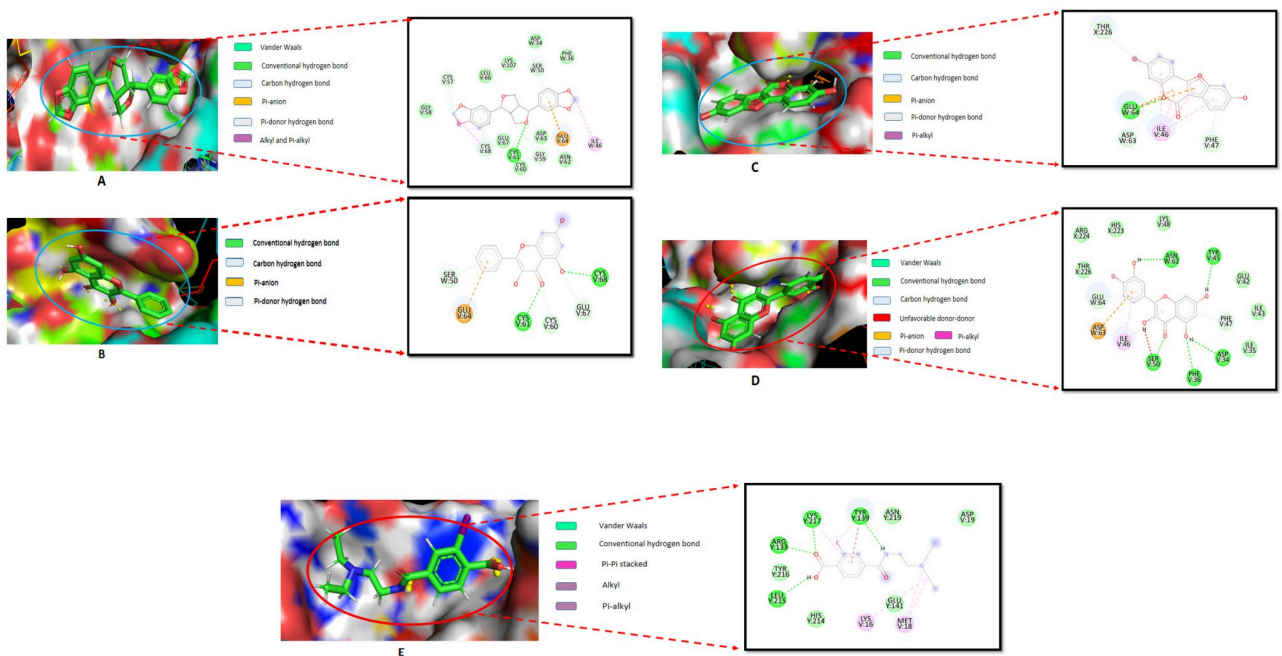
**Fig 10. Molecular interactions of selected compounds against VEGFA (PDB ID: 1FLT) of (A) sesamin, (B) galangin, (C) coumestrol, (D) quercetin and (E) control. All selected compounds interact with the vital substrate management catalytic site.**

**Table 7. Chemo-informatics analysis of selected compounds.**

| ADMET Properties | | Sesamin | α-Mangostin | Galanzin | Coumestrol | Wortmannin |
|---|---|---|---|---|---|---|
| Physicochemical properties | MW (g/mol) | 354.35 | 410.46 | 270.24 | 268.22 | 428.43 |
| | NHA | 26 | 30 | 20 | 20 | 31 |
| | NAHA | 12 | 14 | 16 | 17 | 5 |
| | HBA | 06 | 06 | 5 | 5 | 8 |
| | HBD | 00 | 03 | 3 | 2 | 0 |
| | MR | 90.00 | 119.99 | 73.99 | 73.81 | 105.71 |
| | TPSA (A2) | 55.38 | 100.13 | 90.90 | 83.81 | 109.11 |
| Lipophilicity | iLOGP | 3.46 | 4.14 | 2.08 | 1.80 | 2.70 |
| | XLOGP3 | 2.68 | 6.27 | 2.25 | 2.76 | 1.18 |
| | WLOGP | 2.57 | 5.09 | 2.58 | 3.10 | 2.54 |
| | MLOGP | 1.98 | 2.19 | 0.52 | 1.76 | 0.94 |
| | Silicos-IT Log P | 3.25 | 5.52 | 2.52 | 2.88 | 3.64 |
| | Consensus Log | 2.79 | 4.64 | 1.99 | 2.46 | 2.20 |
| Water Solubility | ESOL Log S | -3.93 | -6.35 | -3.46 | -3.87 | -3.10 |
| | ESOL Class | Soluble | Poorly Souble | Soluble | Soluble | Soluble |
| | ALI Log S | -3.50 | -8.16 | -3.79 | -4.18 | -3.07 |
| | ALI Class | Soluble | Poorly Souble | Soluble | Moderately soluble | Soluble |
| | Silicos-IT Log S | -4.60 | -6.14 | -4.40 | -5.03 | -5.15 |
| | Silicos-IT Class | Moderately soluble | Poorly Souble | Moderately soluble | Moderately soluble | Moderately soluble |
| Pharmacokinetics Properties | GI absorption | High | High | High | High | High |
| | CYP1A2 inhibitor | No | No | Yes | Yes | No |
| | CYP2C19 inhibitor | Yes | No | No | No | No |
| | CYP2C9 inhibitor | No | yes | No | No | No |
| | CYP2D6 inhibitor | Yes | No | Yes | Yes | No |
| | CYP3A4 inhibitor | Yes | No | Yes | No | No |
| | Log Kp (skin permeation) | -6.56cm/s | -4.35 cm/s | -6.35 cm/s | -5.98 cm/s | -8.08 cm/s |
| Druglikeness | Lipinski | Yes; 0 violation | Yes; 0 violation | Yes; 0 violation | Yes; 0 violation | Yes; 0 violation |
| | Ghose | Yes | Yes | Yes | Yes | Yes |
| | Egan | Yes | Yes | Yes | Yes | Yes |
| | Bioavailability Score | 0.55 | 0.55 | 0.55 | 0.55 | 0.55 |
| Medicinal Chemistry | Pains | 0 alert | 0 alert | 0 alert | 0 alert | 0 alert |
| | Leadlikeness | No;1 violation: MW>350 | No;2 violation: MW>350, XLOGP3>3.5 | Yes | Yes | No;1 violation: MW>350 |
| | Synthetic accessibility | 4.12 | 3.91 | 3.12 | 3.16 | 5.66 |
| Toxicity | Ames test | Yes | Yes | No | Yes | Yes |
| | hERG I inhibitor | No | No | No | No | No |
| | hERG II inhibitor | No | Yes | No | No | No |
| | Skin Sensitisation | No | No | No | No | No |
| | Hepatotoxicity | No | No | No | No | No |

https://doi.org/10.1371/journal.pone.0265746.t007

progressions such as cancer proteoglycans, TNF signalling pathway, ovarian infertility genes, diseases of signal transduction, focal adhesion, and notch signalling pathway. Several gene ontologies classes were also discovered in our research, including actomyosin structure organization, integrin-mediated signalling pathway, cholesterol metabolic process, ATP metabolic process, and cellular response to cytokine stimulus which are closely related to the cancer
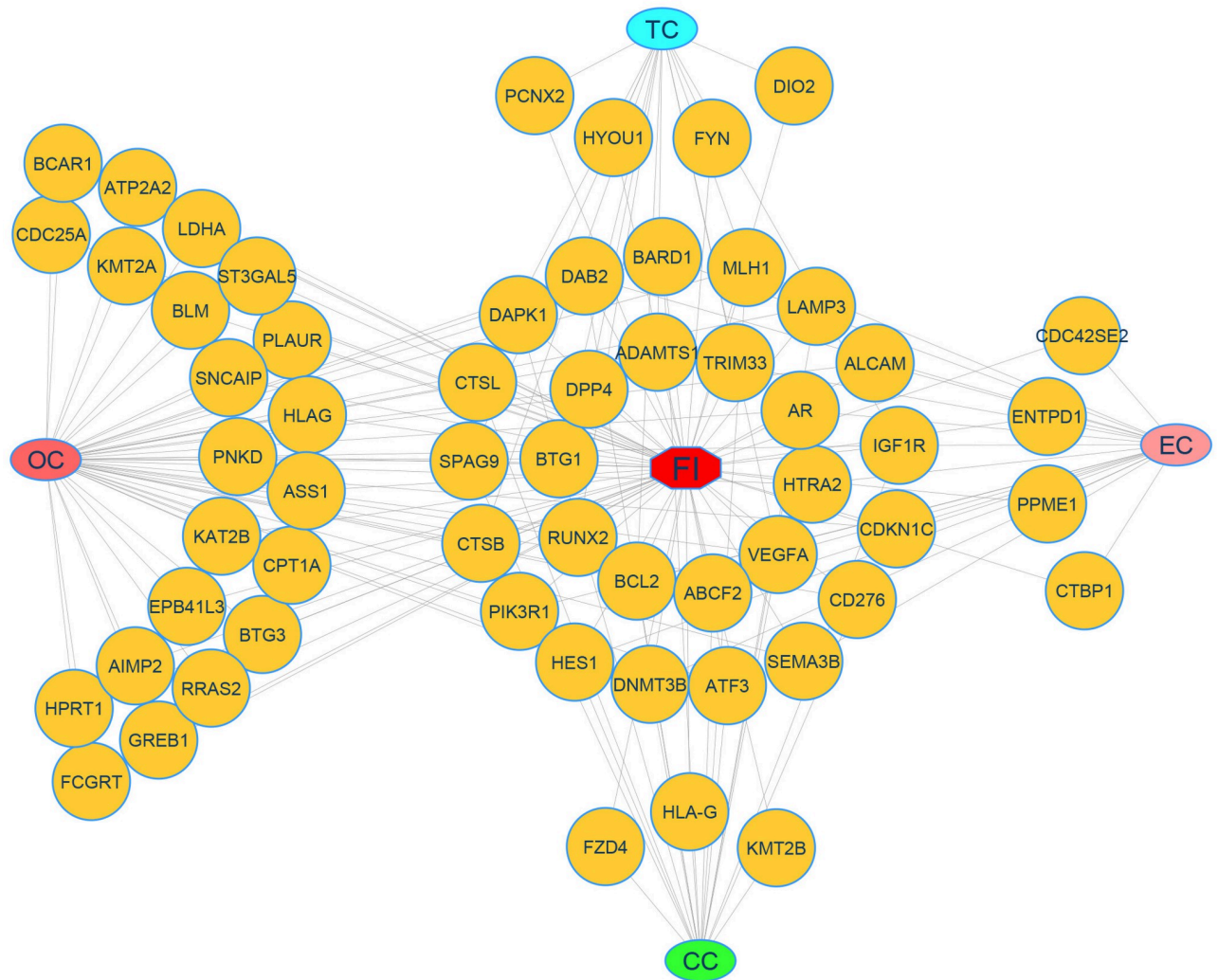
**Fig 11. Gene-disease association analysis of DEGs of 4 cancerous risk factors with FI using OMIM and DisGNET databases.** Ellipse-shaped nodes represent risk factors, and round-shaped nodes represent DEGs. Deep-green color and Round-shaped indicates common genes between FI and four cancerous risk factors.

progression. Transplantable tumors in mice undergo fast hemorrhagic necrosis through the TNF pathway which also plays a role in cell transformation, survival, proliferation and metastasis stage of tumor growth [76, 77]. Changes influence cancer pathogenesis in mechanisms involved in signal transduction such as cell differentiation, proliferation and cell death [78]. The cytokine pattern generated by tumor cells indicated that immune stimulation coincides and might provide new therapeutic options for cancer patients [79]. Moreover, dysfunctional cholesterol metabolism in women may play a role in certain types of infertility [80] and gynecological cancer [81].

In this study, statistical analysis findings were used to develop a PPI network based on the differentially expressed genes. Topological research techniques were used to find core proteins (i.e., hubs). These identified hubs proteins may be used as possible drug targets or biomarkers. Ten hub proteins were seen in the PPI network study (VEGFA, PIK3R1, BCAR1, AR, CPT1A, ACSL3, ACSL4, IGF1R, LPL, and HMGCR) are involved in the FI. Endothelial cells (EC) mature in response to vascular endothelial cell growth factor (VEGF2) via the VEGF2 receptor,

which is activated by cyclooxygenase (COX), while COX2 has been attributed to the onset of a variety of cancers, predominantly colon cancer [82]. The insulin-like growth factor receptor (IGF-1R) is a tyrosine kinase that regulates cell proliferation and survival. It is thought that this receptor inhibition might be a potential target in cancer therapy [58]. PI3K, a phosphoinositide 3-kinase, is upregulated in several human cancers, with oncogenic mutations discovered in the p110-alpha catalytic and p85-alpha regulatory subunits that serve as a docking site during the drug development phase [70, 83]. The serum anti-estrogen resistance protein 1 (BCAR1 also known as p130-cas) is a molecular marker for pulmonary disease diagnosis and prognosis [84]. The variant of the androgen receptor (AR), especially AR-V7 that facilitates AR targeted drug resistance.

In castration-resistant prostate cancer, AR-V9 could be a critical component of therapeutic resistance (CRPC) [85]. In vivo, statins target HMGCR (HMG-CoA reductase) in breast cancer cells and they may also have an anti-proliferative property in HMGCR progressive tumors [86]. The presence of an active FAO is a significant indicator of radiation resistance in nasopharyngeal carcinoma (NPC). So, CPT1A targeting may be a good way to improve radiotherapy's beneficial efficacy in patients with nasopharyngeal carcinoma (NPC) [87]. A family of enzymes catalyses the activation of fatty acids by adding coenzyme A named the acyl CoA synthetase ligases (ACSL) showing more aggressive cancerous features [88].

The hub proteins we have identified in this study play a significant role in infertility and cancer biology. However, the molecular mechanism of FI and its link to the cancerous risk factor is controversial. Besides, several proteins are being explored as therapeutic approaches. This helps our network system to define genes of interest in our disease progression with a specific function. Then we selected VEGFA and PIK3R1 proteins for the docking approach based on the Maximal Clique Centrality (MCC) value of the cytohubba plugin and literature analysis indicates that VEGFA and PIK3R1 most common genes responsible for the progression of several cancer, including Endometrial cancer [62, 63], Ovarian cancer [23, 64], Cervical cancer [24, 65] and Thyroid cancer [66, 67]. Several frequent mutations, including Y504D, Q552K, I559N, D560Y, N564D, D569Y, R574T, T576del, W583del, N595K, and N600H mutants may regulate the iSH2 domain for the creation of multiple cancer like breast, endometrium and ovarian cancer particularly at amino acids 456-469 and 564-575 [70, 72–74] in PIK3R1 protein. In our analysis, we did not include the mutated form of PIK3R1 protein.

Further, selected phytoestrogenic compounds were used before as a therapeutic approach in several cancer treatments. Additionally, ADMET properties show a positive sign as a drug compound for treating the mentioned disease in our study.

Molecular modelling has become an integral part of modern drug development networks especially docking study [89]. Furthermore, time and laboratory costs can be reduced dramatically by reducing the choices of drug molecules in the silico based drug development phase [90]. Besides, the investigator can draw the binding insights into the target protein of the drug molecules and rationally define treatment objectives. However, of the 27 phytochemical compounds evaluated, the peak four are chosen based on their highest docking score compared to positive control inhibitors bevacizumab and wortmannin and found that sesamin, galangin, and coumestrol could be used as promising inhibitors against VEGFA and PIK3R1 proteins.

The three compounds have chosen exhibit strong non-covalent interactions with other predicted binding site residues by the CASTp server, including Lys 506, Ile 509, Tyr 504, Tyr 508, Ile 524, Tyr 528, Met 525, Lys 532 residues at the iSH2 domain of the B chain of PIK3R1 protein. All compounds that bind to the catalytic site of PIK3R1 protein and the same binding pockets ignoring the mutation region were found in our investigation. On the other hand, Glu 64, Ser 50, Cys 68, and Ile 46 were the active site of the VEGFA protein, where the top three compounds were bound. Sesamin can be used as an important therapeutic agent to prevent or

treat different of types cancers [91]. This result indicates that sesamin inhibits cervical cancer cell proliferation via the induction of PTEN-mediated apoptosis [92]. A recent study concluded that, sesamin destroy thyroid cell expansion and initiates apoptosis by hindering STAT-3 translocation [93]. A recent study demonstrated that natural antioxidants, including sesamin, along with galangin, serve as a drug compound to prevent Covid-19 severity [94]. Shu Lio et al. showed that Coumestrol, not only has an inhibitory impact in terms of growth of three cancerous cell line but also serve as a cell penetrable CK2 inhibitor with submicromolar IC50 and a new ATP competitive [95]. Coumestrol suppresses proliferation of ES2 human epithelial ovarian cancer cells via PI3K and ERK1/2 MAPK pathways. These findings indicate that coumestrol showing an anti-cancer effect through the direct targeting of haspin kinase [96, 97]. H. Huang et al. established that galangin showed anti-cancerous activity via p53-dependent pathway; as a result, it can be used for the treatment purposes of platinum-resistant ovarian cancers in humans [98]. This result demonstrated that galangin could be used as an effective adjuvant to increase the efficacy of chemotherapeutic agents in several cancer treatments [99].

In the biopharmaceutical sector, computational biology techniques enhanced the drug discovery process not only to find but also to develop lead compounds against disease [100]. The physicochemical and drug-likeness properties of any compounds were determined by ADMET analysis in a productive and cost-saving approach way [101]. ADMET testing gives a good perception of potential drug applicants. A potential drug should have an optimum molecular weight between 150 and 500 g/mol (Dalton), the higher number of heavy atoms (NHA) and higher number of heavy aromatic atoms (NAHA), hydrogen bonding acceptor (HBA) $\leq$ 10, hydrogen bonding donor (HBD) $\leq$ 5, TPSA (topological polar surface area) around 20 to 130 Å2, and molar refractory (MR) range of 40 to 130 is a good indication of drug performance. Analysis of water solubility indicated that three compounds are soluble in water, except alpha-mangostin is poorly soluble. The medicinal chemistry properties of four compounds indicate that they have no PAINS (pan assay interference compounds) alert that refer to provide a strong propensity to attach to their targets with no false-positive results. Moreover, sesamin and alpha-mangostin infringe lead likeness properties, whereas coumestrol and galangin do not break these properties. The compounds studied in this research are moderately synthetic. (Assume value 1 for low and 10 for high synthetic accessibility). We examined the possible treatment mechanisms of the 27 phytochemicals in this network-based bioinformatics study against VEGFA and PIK3R1 hub proteins. By analyzing DEGs, molecular docking and ADMET properties, we suggest that identified hub proteins, especially VEGFA and PIK3R1 can be used as a target to treat infertility mediated cancer progression. Besides, the top 3 compounds, including sesamin, galangin and coumestrol, could contribute to new treatment techniques. Because previously, these compounds were used in several cancer treatments, but it was not treated in infertility mediated cancer progression. So, we investigate the transcriptomic data for identifying DEGs as the molecular target with therapeutic agents. However, further clarification is still needed for the mechanism of these active compounds through in vitro study which can facilitate the design of new drugs with a wide range of anti-cancerous activity.

## Conclusion

In this research, the microarray transcriptomic data were used to determine the genetic linkage of FI with multiple cancer. According to our findings, these network-based bioinformatics techniques can mention positive interaction of FI mediated cancer risk factors. We identified several mechanistic pathways and gene ontology biological processes like TNF signalling

pathway, signal transduction, cholesterol metabolic process and cellular response to cytokine stimulus which provide a clear view of the genotypic link among multiple cancers and FI. The study of gene expression has also established ten hub proteins which significantly contribute to FI and cancer progression; among them VEGFA and PIK3R1 are the most significant for disease progressions. Molecular docking and ADMET properties against VEGFA (PDB ID: 1FLT) and PIK3R1 (PDB ID: 5M6U) showed sesamin, galangin, and coumestrol were the top three compounds that might play a role on therapeutic target in the treatment of infertility and infertility mediated cancer progression. This approach permits us to determine how their interactions can better understand FI processes from a mechanical perspective with cancer. We recommended that our identified pathway, hub proteins, and phytocompounds serve as new targets and therapeutic interventions for accurate diagnosis and treatment of multiple diseases. Further molecular dynamic simulation and clinical/in vitro study must be needed to confirm the activity of compounds, along with the co-relation of DEGs between FI and the formation of several cancers.

## Acknowledgments

## Author Contributions

**Conceptualization:** Md Habibur Rahman.

**Data curation:** Md. Sharif Khan.

**Formal analysis:** Md. Arju Hossain.

**Methodology:** Md. Arju Hossain.

**Project administration:** Md Habibur Rahman.

**Resources:** Md. Al Amin.

**Supervision:** Md Habibur Rahman.

**Visualization:** Md Imran Hasan.

**Writing – original draft:** Md. Arju Hossain, Md Sohel.

**Writing – review & editing:** Md Habibur Rahman, Md. Zahidul Islam, Silong Peng.

## References

1. Mascarenhas MN, Flaxman SR, Boerma T, Vanderpoel S, Stevens GA. National, Regional, and Global Trends in Infertility Prevalence Since 1990: A Systematic Analysis of 277 Health Surveys. PLoS Medicine. 2012; 9. https://doi.org/10.1371/journal.pmed.1001356 PMID: 23271957

2. Schmidt L, Muenster K. Infertility, involuntary infecundity, and the seeking of medical advice in industrialized countries 1970–1992: a review of concepts, measurements and results. Human Reproduction. 1995 Jun 1; 10(6):1407–18. https://doi.org/10.1093/HUMREP/10.6.1407 PMID: 7593506

3. Altmae S, Martinez-Conejero J, Salumets A, Simon C, Horcajadas J, Stavreus-Evers A. Endometrial gene expression analysis at the time of embryo implantation in women with unexplained infertility. Molecular Human Reproduction. 2009; 16: 178–187. https://doi.org/10.1093/molehr/gap102 PMID: 19933690

4. Inhorn MC, Patrizio P. Infertility around the globe: new thinking on gender, reproductive technologies and global movements in the 21st century. Human Reproduction Update. 2015; 21: 411–426. https://doi.org/10.1093/humupd/dmv016 PMID: 25801630

5. Dunson DB, Baird DD, Colombo B. Increased Infertility With Age in Men and Women. Obstetrics Gynecology. 2004; 103: 51–56. https://doi.org/10.1097/01.AOG.0000100153.24061.45 PMID: 14704244

6. Worzfeld T, Strandmann EPV, Huber M, Adhikary T, Wagner U, Reinartz S, et al. The Unique Molecular and Cellular Microenvironment of Ovarian Cancer. Frontiers in Oncology. 2017; 7. https://doi.org/10.3389/fonc.2017.00024 PMID: 28275576

7. Cirillo PM, Wang ET, Cedars MI, Chen L-M, Cohn BA. Irregular menses predicts ovarian cancer: Prospective evidence from the Child Health and Development Studies. International Journal of Cancer. 2016; 139: 1009–1017. https://doi.org/10.1002/ijc.30144 PMID: 27082375

8. Yang S, Wu Y, Deng Y, Zhou L, Yang P, Zheng Y, et al. Identification of a prognostic immune signature for cervical cancer to predict survival and response to immune checkpoint inhibitors. OncoImmunology. 2019; 8. https://doi.org/10.1080/2162402X.2019.1659094 PMID: 31741756

9. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians. 2018; 68: 394–424. https://doi.org/10.3322/caac.21492 PMID: 30207593

10. Dor J, Lerner-Geva L, Rabinovici J, Chetrit A, Levran D, Lunenfeld B, et al. Cancer incidence in a cohort of infertile women who underwent in vitro fertilization. Fertility and Sterility. 2002; 77: 324–327. https://doi.org/10.1016/S0015-0282(01)02986-7 PMID: 11821091

11. Johnatty SE, Tan YY, Buchanan DD, Bowman M, Walters RJ, Obermair A, et al. Family history of cancer predicts endometrial cancer risk independently of Lynch Syndrome: Implications for genetic counselling. Gynecologic Oncology. 2017; 147: 381–387. https://doi.org/10.1016/j.ygyno.2017.08.011 PMID: 28822557

12. Ring KL, Bruegl AS, Allen BA, Elkin EP, Singh N, Hartman A-R, et al. Germline multi-gene hereditary cancer panel testing in an unselected endometrial cancer cohort. Modern Pathology. 2016; 29: 1381–1389. https://doi.org/10.1038/modpathol.2016.135 PMID: 27443514

13. Stoffel E, Mukherjee B, Raymond VM, Tayob N, Kastrinos F, Sparr J, et al. Calculation of Risk of Colorectal and Endometrial Cancer Among Patients With Lynch Syndrome. Gastroenterology. 2009; 137: 1621–1627. https://doi.org/10.1053/j.gastro.2009.07.039 PMID: 19622357

14. Ballweg ML. News from the Endometriosis Movement. Journal of Endometriosis. 2010; 2: 164–167. https://doi.org/10.1177/228402651000200308

15. Seib CD, Sosa JA. Evolving Understanding of the Epidemiology of Thyroid Cancer. Endocrinology and Metabolism Clinics of North America. 2019; 48: 23–35. https://doi.org/10.1016/j.ecl.2018.10.002 PMID: 30717905

16. Xing M. Molecular pathogenesis and mechanisms of thyroid cancer. Nature Reviews Cancer. 2013; 13: 184–199. https://doi.org/10.1038/nrc3431 PMID: 23429735

17. Unuane D, Poppe K. Female infertility: do we forget the thyroid? Journal of Endocrinological Investigation. 2015; 38: 571–574. https://doi.org/10.1007/s40618-015-0280-0 PMID: 25847323

18. Xu H, Moni MA, Liò P. Network regularised Cox regression and multiplex network models to predict disease comorbidities and survival of cancer. Computational Biology and Chemistry. 2015; 59: 15–31. https://doi.org/10.1016/j.compbiolchem.2015.08.010 PMID: 26611766

19. Podder NK, Rana HK, Azam MS, Rana MS, Akhtar MR, Rahman MR, et al. A system biological approach to investigate the genetic profiling and comorbidities of type 2 diabetes. Gene Reports. 2020; 21: 100830. https://doi.org/10.1016/j.genrep.2020.100830

20. Miled N, Yan Y, Hon W-C, Perisic O, Zvelebil M, Inbar Y, et al. Mechanism of Two Classes of Cancer Mutations in the Phosphoinositide 3-Kinase Catalytic Subunit. Science. 2007; 317: 239–242. https://doi.org/10.1126/science.1135394 PMID: 17626883

21. Li X, Mak VCY, Zhou Y, Wang C, Wong ESY, Sharma R, et al. Deregulated Gab2 phosphorylation mediates aberrant AKT and STAT3 signaling upon PIK3R1 loss in ovarian cancer. Nature Communications. 2019; 10. https://doi.org/10.1038/s41467-019-08574-7 PMID: 30755611

22. Cheung LW, Hennessy BT, Li J, Yu S, Myers AP, Djordjevic B, et al. High Frequency of PIK3R1 and PIK3R2 Mutations in Endometrial Cancer Elucidates a Novel Mechanism for Regulation of PTEN Protein Stability. Cancer Discovery. 2011; 1: 170–185. https://doi.org/10.1158/2159-8290.CD-11-0039 PMID: 21984976

23. Jang K, Kim M, Gilbert CA, Simpkins F, Ince TA, Slingerland JM. VEGFA activates an epigenetic pathway upregulating ovarian cancer-initiating cells. EMBO molecular medicine. 2017 Mar; 9(3):304–18. https://doi.org/10.15252/emmm.201606840 PMID: 28179359

24. Chen B, Zhang C, Dong P, Guo Y, Mu N. Molecular regulation of cervical cancer growth and invasion by VEGFa. Tumor Biology. 2014 Nov; 35(11):11587–93. https://doi.org/10.1007/s13277-014-2463-2 PMID: 25135429

**25.** Yuan H, Ma Q, Cui H, Liu G, Zhao X, Li W, et al. How Can Synergism of Traditional Medicines Benefit from Network Pharmacology? Molecules. 2017; 22: 1135. https://doi.org/10.3390/molecules22071135 PMID: 28686181

**26.** Saikia S, Bordoloi M. Molecular Docking: Challenges, Advances and its Use in Drug Discovery Perspective. Current Drug Targets. 2019; 20: 501–521. https://doi.org/10.2174/1389450119666181022153016 PMID: 30360733

**27.** Rahman MH, Peng S, Hu X, Chen C, Uddin S, Quinn JMW, et al. Bioinformatics Methodologies to Identify Interactions Between Type 2 Diabetes and Neurological Comorbidities. IEEE Access. 2019; 7: 183948–183970. https://doi.org/10.1109/ACCESS.2019.2960037

**28.** Torkamani A, Topol EJ, Schork NJ. Pathway analysis of seven common diseases assessed by genome-wide association. Genomics. 2008; 92: 265–272. https://doi.org/10.1016/j.ygeno.2008.07.011 PMID: 18722519

**29.** Mallepaly R, Butler PR, Herati AS, Lamb DJ. Genetic Basis of Male and Female Infertility. Genetics of Human Infertility Monographs in Human Genetics. 2017;: 1–16. https://doi.org/10.1159/000477275

**30.** Rahman MH, Peng S, Hu X, Chen C, Rahman MR, Uddin S, et al. A Network-Based Bioinformatics Approach to Identify Molecular Biomarkers for Type 2 Diabetes that Are Linked to the Progression of Neurological Diseases. International Journal of Environmental Research and Public Health. 2020; 17: 1035. https://doi.org/10.3390/ijerph17031035

**31.** Oliveros JC. VENNY. An interactive tool for comparing lists with Venn Diagrams. http://bioinfogp.cnb.csic.es/tools/venny/index.html. 2007.

**32.** Smoot ME, Ono K, Ruscheinski J, Wang P-L, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics. 2010; 27: 431–432. https://doi.org/10.1093/bioinformatics/btq675 PMID: 21149340

**33.** Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Research. 2016; 44. https://doi.org/10.1093/nar/gkw377 PMID: 27141961

**34.** Rahman MH, Rana HK, Peng S, Hu X, Chen C, Quinn JMW, et al. Bioinformatics and machine learning methodologies to identify the effects of central nervous system disorders on glioblastoma progression. Briefings in Bioinformatics. 2021 https://doi.org/10.1093/bib/bbaa365 PMID: 33406529

**35.** Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. Nucleic Acids Research. 2016; 45. https://doi.org/10.1093/nar/gkw937 PMID: 27924014

**36.** Mahmud SMH, Al-Mustanjid M, Akter F, Rahman MS, Ahmed K, Rahman MH, et al. Bioinformatics and system biology approach to identify the influences of SARS-CoV-2 infections to idiopathic pulmonary fibrosis and chronic obstructive pulmonary disease patients. Briefings in Bioinformatics. 2021 https://doi.org/10.1093/bib/bbab115 PMID: 33847347

**37.** Burley SK, Bhikadiya C, Bi C, Bittrich S, Chen L, Crichlow GV, et al. RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. Nucleic acids research. 2021 Jan 8; 49(D1):D437–51. https://doi.org/10.1093/nar/gkaa1038 PMID: 33211854

**38.** Wiesmann C, Fuh G, Christinger HW, Eigenbrot C, Wells JA, de Vos AM. Crystal structure at 1.7 Å resolution of VEGF in complex with domain 2 of the Flt-1 receptor. Cell. 1997 Nov 28; 91(5):695–704. https://doi.org/10.1016/S0092-8674(00)80456-0 PMID: 9393862

**39.** Alexander N, Woetzel N, Meiler J. Bcl::Cluster: A method for clustering biological molecules coupled with visualization in the Pymol Molecular Graphics System. 2011 IEEE 1st International Conference on Computational Advances in Bio and Medical Sciences (ICCABS). 2011.

**40.** Kaplan W. Swiss-PDB Viewer (Deep View). Briefings in Bioinformatics. 2001; 2: 195–197. https://doi.org/10.1093/bib/2.2.195 PMID: 11465736

**41.** Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, et al. PubChem 2019 update: improved access to chemical data. Nucleic Acids Research. 2018; 47. https://doi.org/10.1093/nar/gky1033

**42.** Dallakyan S, Olson AJ. Small-Molecule Library Screening by Docking with PyRx. Methods in Molecular Biology Chemical Biology. 2014;: 243–250. https://doi.org/10.1007/978-1-4939-2269-7-19

**43.** Halgren TA. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. Journal of Computational Chemistry. 1996; 17: 490–519.

**44.** Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific Reports. 2017; 7. https://doi.org/10.1038/srep42717 PMID: 28256516

45. Pires DEV, Blundell TL, Ascher DB. pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures. Journal of Medicinal Chemistry. 2015; 58: 4066–4072. https://doi.org/10.1021/acs.jmedchem.5b00104 PMID: 25860834

46. Pathare ADS, Zaveri K, Hinduja I. Downregulation of genes related to immune and inflammatory response in IVF implantation failure cases under controlled ovarian stimulation. American Journal of Reproductive Immunology. 2017; 78. https://doi.org/10.1111/aji.12679 PMID: 28370781

47. Pappa KI, Polyzos A, Jacob-Hirsch J, Amariglio N, Vlachos GD, Loutradis D, et al. Profiling of Discrete Gynecological Cancers Reveals Novel Transcriptional Modules and Common Features Shared by Other Cancer Types and Embryonic Stem Cells. Plos One. 2015; 10. https://doi.org/10.1371/journal.pone.0142229 PMID: 26559525

48. Hoffmann K, Berger H, Kulbe H, Thillainadarasan S, Mollenkopf H-J, Zemojtel T, et al. Preservation of stemness in high-grade serous ovarian cancer organoids requires low Wnt environment. 2019. https://doi.org/10.1101/741397

49. Guardado-Estrada M, Medina-Martínez I, Juárez-Torres E, Roman-Bassaure E, Macías L, Alfaro A, et al. The Amerindian mtDNA haplogroup B2 enhances the risk of HPV for cervical cancer: de-regulation of mitochondrial genes may be involved. Journal of Human Genetics. 2012; 57: 269–276. https://doi.org/10.1038/jhg.2012.17 PMID: 22357541

50. Vasko V, Espinosa AV, Scouten W, He H, Auer H, Liyanarachchi S, et al. Gene expression and functional evidence of epithelial-to-mesenchymal transition in papillary thyroid carcinoma invasion. Proceedings of the National Academy of Sciences. 2007; 104: 2803–2808. https://doi.org/10.1073/pnas.0610733104 PMID: 17296934

51. Li X, Hu Z, Shi H, Wang C, Lei J, Cheng Y. Inhibition of VEGFA Increases the Sensitivity of Ovarian Cancer Cells to Chemotherapy by Suppressing VEGFA-Mediated Autophagy. OncoTargets and therapy. 2020; 13:8161. https://doi.org/10.2147/OTT.S250392 PMID: 32884298

52. Taniguchi CM, Winnay J, Kondo T, Bronson RT, Guimaraes AR, Alemán JO, et al. The phosphoinositide 3-kinase regulatory subunit p85 can exert tumor suppressor properties through negative regulation of growth factor signaling. Cancer research. 2010 Jul 1; 70(13):5305–15. https://doi.org/10.1158/0008-5472.CAN-09-3399 PMID: 20530665

53. Huang W, Deng B, Wang RW, Tan QY, He Y, Jiang YG, et al. BCAR1 protein plays important roles in carcinogenesis and predicts poor prognosis in non-small-cell lung cancer. PloS one. 2012 Apr 27; 7 (4):e36124. https://doi.org/10.1371/journal.pone.0036124 PMID: 22558353

54. Wang Q, Li W, Zhang Y, Yuan X, Xu K, Yu J, et al. Androgen receptor regulates a distinct transcription program in androgen-independent prostate cancer. Cell. 2009 Jul 23; 138(2):245–56. https://doi.org/10.1016/j.cell.2009.04.056 PMID: 19632176

55. Xiong X, Wen YA, Fairchild R, Zaytseva YY, Weiss HL, Evers BM, et al. Upregulation of CPT1A is essential for the tumor-promoting effect of adipocytes in colon cancer. Cell death disease. 2020 Sep 10; 11(9):1–6. https://doi.org/10.1038/s41419-020-02936-6

56. Fernández LP, Merino M, Colmenarejo G, Moreno-Rubio J, Sánchez-Martínez R, Quijada-Freire A, et al. Metabolic enzyme ACSL3 is a prognostic biomarker and correlates with anticancer effectiveness of statins in non-small cell lung cancer. Molecular oncology. 2020 Dec; 14(12):3135–52. https://doi.org/10.1002/1878-0261.12816 PMID: 33030783

57. Tang Y, Zhou J, Hooi SC, Jiang YM, Lu GD. Fatty acid activation in carcinogenesis and cancer development: Essential roles of long-chain acyl-CoA synthetases. Oncology letters. 2018 Aug 1; 16 (2):1390–6. https://doi.org/10.3892/ol.2018.8843 PMID: 30008815

58. Miller LM, Mayer SC, Berger DM, Boschelli DH, Boschelli F, Di L, et al. Lead identification to generate 3-cyanoquinoline inhibitors of insulin-like growth factor receptor (IGF-1R) for potential use in cancer treatment. Bioorganic Medicinal Chemistry Letters. 2009; 19: 62–66. https://doi.org/10.1016/j.bmcl.2008.11.037 PMID: 19041240

59. Takasu S, Mutoh M, Takahashi M, Nakagama H. Lipoprotein lipase as a candidate target for cancer prevention/therapy. Biochemistry research international. 2012 Jan 1; 2012. https://doi.org/10.1155/2012/398697 PMID: 22028972

60. Yeganeh B, Wiechec E, Ande SR, Sharma P, Moghadam AR, Post M, et al. Targeting the mevalonate cascade as a new therapeutic approach in heart disease, cancer and pulmonary disease. Pharmacology therapeutics. 2014 Jul 1; 143(1):87–110. https://doi.org/10.1016/j.pharmthera.2014.02.007 PMID: 24582968

61. Chin C-H, Chen S-H, Wu H-H, Ho C-W, Ko M-T, Lin C-Y. cytoHubba: identifying hub objects and subnetworks from complex interactome. BMC Systems Biology. 2014; 8. https://doi.org/10.1186/1752-0509-8-S4-S11 PMID: 25521941

**62.** Holland CM, Day K, Evans A, Smith SK. Expression of the VEGF and angiopoietin genes in endometrial atypical hyperplasia and endometrial cancer. British journal of cancer. 2003 Sep; 89(5):891–8. https://doi.org/10.1038/sj.bjc.6601194 PMID: 12942123

**63.** Rick ME, Rudd ML, Godwin AK, Sgroi D, Merino M, Bell DW. PIK3R1 (p85) is somatically mutated at high frequency in primary endometrial cancer. Cancer research. 2011 Jun 15; 71(12):4061–7. https://doi.org/10.1158/0008-5472.CAN-11-0549

**64.** Xinran LI, Mak VC, Zhou Y, Lu Y, Cheung AN, Mills GB, et al. PIK3R1 loss activates AKT and STAT3 signaling in ovarian cancer.

**65.** Roszik J, Ring KL, Wani KM, Lazar AJ, Yemelyanova AV, Soliman PT, et al. Gene expression analysis identifies novel targets for cervical cancer therapy. Frontiers in immunology. 2018 Sep 19; 9:2102. https://doi.org/10.3389/fimmu.2018.02102 PMID: 30283446

**66.** Song YS, Kim MJ, Sun HJ, Kim HH, Shin HS, Kim YA, et al. Aberrant thyroid-stimulating hormone receptor signaling increases VEGF-A and CXCL8 secretion of thyroid cancer cells, contributing to angiogenesis and tumor growth. Clinical Cancer Research. 2019 Jan 1; 25(1):414–25. https://doi.org/10.1158/1078-0432.CCR-18-0663 PMID: 30314969

**67.** Murugan AK, Dong J, Xie J, Xing M. Uncommon GNAQ, MMP8, AKT3, EGFR, and PIK3R1 mutations in thyroid cancers. Endocrine pathology. 2011 Jun; 22(2):97–102. https://doi.org/10.1007/s12022-011-9155-x PMID: 21487925

**68.** Garcia J, Hurwitz HI, Sandler AB, Miles D, Coleman RL, Deurloo R, et al. Bevacizumab (Avastin®) in cancer treatment: A review of 15 years of clinical experience and future outlook. Cancer treatment reviews. 2020 Jun 1; 86:102017. https://doi.org/10.1016/j.ctrv.2020.102017 PMID: 32335505

**69.** Binkowski TA. CASTp: Computed Atlas of Surface Topography of proteins. Nucleic Acids Research. 2003; 31: 3352–3355. https://doi.org/10.1093/nar/gkg512 PMID: 12824325

**70.** Fruman D. Faculty Opinions recommendation of Somatic mutations in p85alpha promote tumorigenesis through class IA PI3K activation. Faculty Opinions—Post-Publication Peer Review of the Biomedical Literature. 2009. https://doi.org/10.3410/f.1284956.752054

**71.** Apweiler R. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. Nucleic Acids Research. 2001; 29: 37–40. https://doi.org/10.1093/nar/29.1.37 PMID: 11125043

**72.** Sun M, Hillmann P, Hofmann BT, Hart JR, Vogt PK. Cancer-derived mutations in the regulatory subunit p85 of phosphoinositide 3-kinase function through the catalytic subunit p110. Proceedings of the National Academy of Sciences. 2010 Aug 31; 107(35):15547–52. https://doi.org/10.1073/pnas.1009652107

**73.** Huang CH, Mandelker D, Schmidt-Kittler O, Samuels Y, Velculescu VE, Kinzler KW, et al. The structure of a human p110/p85 complex elucidates the effects of oncogenic PI3K mutations. Science. 2007 Dec 14; 318(5857):1744–8. https://doi.org/10.1126/science.1150799 PMID: 18079394

**74.** Koboldt DC, Fulton R, McLellan M, Schmidt H, Kalicki-Veizer J, McMichael J, et al. Comprehensive molecular portraits of human breast tumours. Nature. 2012; 490(7418):61–70. https://doi.org/10.1038/nature11412

**75.** Rahman Md Habibur, Rana Humayan Kabir, Peng Silong, Kibria Md Golam, Islam Md Zahidul, et al. "Bioinformatics and System Biology Approaches to Identify Pathophysiological Impact of COVID-19 to the Progression and Severity of Neurological Diseases". Computers in Biology and Medicine (2021): 104859. https://doi.org/10.1016/j.compbiomed.2021.104859 PMID: 34601390

**76.** Carswell EA, Old LJ, Kassel R, Green S, Fiore N, Williamson B. An endotoxin-induced serum factor that causes necrosis of tumors. Proceedings of the National Academy of Sciences. 1975 Sep 1; 72(9):3666–70. https://doi.org/10.1073/pnas.72.9.3666 PMID: 1103152

**77.** Sethi G, Sung B, Aggarwal BB. TNF: a master switch for inflammation to cancer. Front biosci. 2008 May 1; 13(2):5094–107. https://doi.org/10.2741/3066 PMID: 18508572

**78.** Liem AA, Chamberlain MP, Wolf CR, Thompson AM. The role of signal transduction in cancer treatment and drug resistance. European Journal of Surgical Oncology (EJSO). 2002 Sep 1; 28(6):679–84. https://doi.org/10.1053/ejso.2002.1303 PMID: 12359208

**79.** Lippitz BE. Cytokine patterns in patients with cancer: a systematic review. The lancet oncology. 2013 May 1; 14(6):e218–28. https://doi.org/10.1016/S1470-2045(12)70582-X PMID: 23639322

**80.** Yesilaltay A, Dokshin GA, Busso D, Wang L, Galiani D, Chavarria T, et al. Excess cholesterol induces mouse egg activation and may cause female infertility. Proceedings of the National Academy of Sciences. 2014 Nov 18; 111(46):E4972–80. https://doi.org/10.1073/pnas.1418954111 PMID: 25368174

**81.** Xu H., Zhou S., Tang Q., Xia H. Bi F. Cholesterol metabolism: New functions and therapeutic approaches in cancer. Biochim. Biophys. Acta—Rev. Cancer 1874, 188394 (2020).

82.  Murphy JF, Fitzgerald DJ. Vascular endothelial cell growth factor (VEGF) induces cyclooxygenase (COX)-dependent proliferation of endothelial cells (EC) via the VEGF-2 receptor. The FASEB Journal. 2001 Jul; 15(9):1667–9. https://doi.org/10.1096/fj.00-0757fje PMID: 11427521

83.  Mandelker D, Gabelli SB, Schmidt-Kittler O, Zhu J, Cheong I, Huang C-H, et al. A frequent kinase domain mutation that changes the interaction between PI3K and the membrane. Proceedings of the National Academy of Sciences. 2009; 106: 16996–17001. https://doi.org/10.1073/pnas.0908444106

84.  Deng B, Huang W, Tan Q-Y, Fan X-Q, Jiang Y-G, Liu L, et al. Breast Cancer Anti-Estrogen Resistance Protein 1 (BCAR1/p130cas) in Pulmonary Disease Tissue and Serum. Molecular Diagnosis Therapy. 2011; 15: 31–40. https://doi.org/10.1007/BF03257191 PMID: 21469768

85.  Kohli M, Ho Y, Hillman DW, Etten JLV, Henzler C, Yang R, et al. Androgen Receptor Variant AR-V9 Is Coexpressed with AR-V7 in Prostate Cancer Metastases and Predicts Abiraterone Resistance. Clinical Cancer Research. 2017; 23: 4704–4715. https://doi.org/10.1158/1078-0432.CCR-17-0017 PMID: 28473535

86.  Bjarnadottir O, Romero Q, Bendahl P-O, Jirström K, Rydén L, Loman N, et al. Targeting HMG-CoA reductase with statins in a window-of-opportunity breast cancer trial. Breast Cancer Research and Treatment. 2013; 138: 499–508. https://doi.org/10.1007/s10549-013-2473-6 PMID: 23471651

87.  Tan Z, Xiao L, Tang M, Bai F, Li J, Li L, et al. Targeting CPT1A-mediated fatty acid oxidation sensitizes nasopharyngeal carcinoma to radiation therapy. Theranostics. 2018; 8: 2329–2347. https://doi.org/10.7150/thno.21451 PMID: 29721083

88.  Radif Y, Ndiaye H, Kalantzi V, Jacobs R, Hall A, Minogue S, et al. The endogenous subcellular localisations of the long chain fatty acid-activating enzymes ACSL3 and ACSL4 in sarcoma and breast cancer cells. Molecular and Cellular Biochemistry. 2018; 448: 275–286. https://doi.org/10.1007/s11010-018-3332-x PMID: 29450800

89.  Oany A, Pervin T, Emran A. Design of an epitope-based peptide vaccine against spike protein of human coronavirus: an in silico approach. Drug Design, Development and Therapy. 2014;: 1139. https://doi.org/10.2147/DDDT.S67861 PMID: 25187696

90.  Talele T, Khedkar S, Rigby A. Successful Applications of Computer Aided Drug Discovery: Moving Drugs from Concept to the Clinic. Current Topics in Medicinal Chemistry. 2010; 10: 127–141. https://doi.org/10.2174/156802610790232251 PMID: 19929824

91.  Majdalawieh AF, Massri M, Nasrallah GK. A comprehensive review on the anti-cancer properties and mechanisms of action of sesamin, a lignan in sesame seeds (Sesamum indicum). European Journal of Pharmacology. 2017; 815: 512–521. https://doi.org/10.1016/j.ejphar.2017.10.020 PMID: 29032105

92.  Kuo TN, Lin CS, Guan-De Li CY, Kao SH. Sesamin inhibits cervical cancer cell proliferation by promoting p53/PTEN-mediated apoptosis. International Journal of Medical Sciences. 2020; 17(15):2292. https://doi.org/10.7150/ijms.48955 PMID: 32922194

93.  Ma Y, Karunakaran T, Veeraraghavan VP, Mohan SK, Li S. Sesame inhibits cell proliferation and induces apoptosis through inhibition of STAT-3 translocation in thyroid cancer cell lines (FTC-133). Biotechnology and Bioprocess Engineering. 2019 Aug; 24(4):646–52. https://doi.org/10.1007/s12257-019-0151-1

94.  Pandey AK, Verma S. An in-silico evaluation of dietary components for structural inhibition of SARS-Cov-2 main protease. Journal of Biomolecular Structure and Dynamics. 2020;: 1–7. https://doi.org/10.1080/07391102.2020.1809522 PMID: 32811367

95.  Liu S, Hsieh D, Yang Y-L, Xu Z, Peto C, Jablons DM, et al. Coumestrol from the national cancer Institute's natural product library is a novel inhibitor of protein kinase CK2. BMC Pharmacology and Toxicology. 2013; 14. https://doi.org/10.1186/2050-6511-14-36 PMID: 23845105

96.  Castro CC, Pagnussat AS, Orlandi L, Worm P, Moura N, Etgen AM, et al. Coumestrol has neuroprotective effects before and after global cerebral ischemia in female rats. Brain Research. 2012; 1474: 82–90. https://doi.org/10.1016/j.brainres.2012.07.025

97.  Kim JE, Lee SY, Jang M, Choi HK, Kim JH, Chen H, et al. Coumestrol epigenetically suppresses cancer cell proliferation: Coumestrol is a natural haspin kinase inhibitor. International journal of molecular sciences. 2017 Oct; 18(10):2228. https://doi.org/10.3390/ijms18102228 PMID: 29064398

98.  Huang H, Chen AY, Ye X, Guan R, Rankin GO, Chen YC. Galangin, a Flavonoid from Lesser Galangal, Induced Apoptosis via p53-Dependent Pathway in Ovarian Cancer Cells. Molecules. 2020; 25: 1579. https://doi.org/10.3390/molecules25071579 PMID: 32235536

99.  Hajipour H, Nouri M, Ghorbani M, Bahramifar A, Emameh RZ, Taheri RA. Targeted Nanostructured Lipid Carrier Containing Galangin as a Promising Adjuvant for Improving Anticancer Effects of Chemotherapeutic Agents.

**100.** Bansal M, Gatta GD, Bernardo DD. Inference of gene regulatory networks and compound mode of action from time course gene expression profiles. Bioinformatics. 2006; 22: 815–822. https://doi.org/10.1093/bioinformatics/btl003 PMID: 16418235

**101.** Rahman A, Naheed NH, Raka SC, Qais N, Momen AZMR. Ligand-based virtual screening, consensus molecular docking, multi-target analysis and comprehensive ADMET profiling and MD stimulation to find out noteworthy tyrosine kinase inhibitor with better efficacy and accuracy. Advances in Traditional Medicine. 2019; 20: 645–661. https://doi.org/10.1007/s13596-019-00406-9