Article

Not peer-reviewed version

Integrating Network Pharmacology and Experimental Evaluation of *Ocimum tenuiflorum* (Tulsi) compounds targeting Breast Cancer Markers

AMRITA MAHAJAN *, Dr. Pravin Badhe , Prashant Mali

Posted Date: 23 May 2024

doi: 10.20944/preprints202405.1511.v1

Keywords: Breast cancer; network pharmacology; text mining; drug therapy; genes; pathway analysis; biological process; protein-protein interaction; degree and betweenness; molecular docking; meristematic cells; anti-mitotic; hepatotoxicity

 doi:10.20944/preprints202405.1511.v1

Article

Integrating Network Pharmacology and Experimental Evaluation of *Ocimum tenuiflorum* (Tulsi) compounds Targeting Breast Cancer Markers

Amrita Mahajan [1,*], Pravin D Badhe [2] and Prashant D Mali [3]

Department of pharmacology, Sinhgad College of Pharmacy, Pune, INDIA.
* Correspondence: amritamahajan02@gmail.com

Abstract: One of the most prevalent and deadly types of cancer is breast cancer. The value and effectiveness of the current pharmacological therapy are, however, constrained. Using computational methods and publicly available data, the current study set out to identify the genes and molecular pathways linked to breast cancer. It also looked into potential treatments for the disease that would target these molecular processes. In this study, we mined genes that were strongly associated with breast cancer using text mining and GeneCodis. Using STRING and Cytoscape, protein-protein interaction (PPI) study was carried out. Candidate medications were then derived based on the drug-gene interaction study of the final genes. 2,658 genes linked to breast cancer were found by our investigation using text mining searches. Out of which 166 genes have been taken which are relevant to breast cancer. Ten genes representing ten pathways—which a total of ten proteins could target—were found by gene enrichment analysis. We have taken Holy Basil as our lead as it contains various bioactive compounds such as flavonoids, terpenoids, and phenolics, which have shown antioxidant, anti-inflammatory, and anticancer properties. For this analysis we choose the molecular docking technique to check the effects of different chemical constituents of Holy Tulsi on breast cancer targeted protein and compare their results. We've performed the molecular docking in between chemical constituents of Holy Basil and the targeted proteins and determined the binding affinity with the help of PyRx and BIOVIA Discovery studio software. Anti-mitotic activity of *Pisum sativum* using tulsi extract to check the cytotoxicity effect was done. To determine the toxicity level of the extract on rat liver was performed in the end. In conclusion, investigating candidate medications that target the genes/pathways relevant to breast cancer in order to uncover possible treatments may be accomplished through drug discovery employing in silico text mining and pathway analysis technologies along with the wet lab experimentation with plants and preclinicals.

Keywords: Breast cancer; network pharmacology; text mining; drug therapy; genes; pathway analysis; biological process; protein-protein interaction; degree and betweenness; molecular docking; meristematic cells; anti-mitotic; hepatotoxicity

Introduction

One of the most prevalent malignancies in women globally, breast cancer claimed around 570,000 lives in 2015. Worldwide, more than 1.5 million women (or 25% of all women with cancer) receive a breast cancer diagnosis each year. According to estimates, breast cancer accounted for 30% of all new cancer diagnoses (252,710) among women in the United States in 2017. Breast cancer is incurable mostly because it is a metastatic cancer that frequently spreads to distant organs such the liver, brain, lung, and bone. A favourable prognosis and a high chance of survival can result from early detection of the illness. As the early identification of breast cancer, the 5-year relative survival rate for patients in North America is above 80%. In order to detect breast cancer, mammography is a commonly used screening method that has been shown to effectively lower mortality [1]. In the last ten years, additional screening techniques have also been used and researched, such as Magnetic

 doi:10.20944/preprints202405.1511.v1

Resonance Imaging (MRI), which is more sensitive than mammography. Many variables can raise the risk of breast cancer, including sex, aging, estrogen, family history, gene mutations, and an unhealthy lifestyle. The majority of incidences of breast cancer occur in women, who also account for 100 times more cases than males do. Despite the fact that breast cancer is becoming more common in America, fewer people die from the disease as a result of broad early detection programs and cutting-edge medical treatments. Recent developments in biological therapy have shown promise in treating breast cancer [2].

The choice of a suitable treatment to optimize function preservation and reduce the risk of recurrence and metastasis remains difficult in light of the rising incidence of breast cancer and rising patient expectations in recent years. Surgical excision is often regarded as the gold standard in the clinic for treating breast cancer [3]. When treating late-stage breast cancer, non-surgical methods such as photodynamic therapy (PDT), radiation treatment, cryotherapy, and chemotherapy are frequently employed. Nonetheless, there is still a dearth of study on medication therapy, and more investigation is required to support the creation of new therapeutic approaches [4].

Drug repositioning may expedite the process of discovering additional conditions that existing drugs could treat more effectively and potentially at a lower cost, even though the efficacy of the currently available drug therapies is limited and the discovery of new drug therapies using traditional methods is likely to take a long time [5]. Using computational methods such as text mining, biological process and pathway analysis, protein-protein interaction (PPI) analysis to mine public databases, and bioinformatics tools to systematically identify interaction networks between drugs and gene targets, this study aimed to investigate new drug therapies for breast cancer. For the purpose of choosing a medicine, we were able to examine the features of potential genes utilizing data analytical methods [6]. Based on the final genes' drug-gene interaction analysis, candidate drugs were then derived.

The field of network pharmacology encompasses systems biology, network analysis, connection, redundancy, and pleiotropy in its drug design methodology. A new way of approaching drug discovery is provided by network pharmacology, which simultaneously incorporates efforts to increase clinical efficacy and comprehend toxicity and side effects, two of the main causes of failure. Network research have been seen in a very effective discovery analysis in discovering the biological science. Numerous research have demonstrated the effectiveness of network [7,8].

Furthermore, synthetic behaviours, combinations, and molecular biology probes have been used to identify emergent phenotypes beyond those reported in single-gene deletion investigations [9]. Although there is a strong biological case for multitarget tactics to be preferred over single-target approaches, the pharmaceutical industry currently employs few multitarget strategies [10].

Fast, iterative structure-based drug discovery required advancements in virtual screening, computational processing capacity, high-power radiation sources, refinement methods, compute graphics, and cryocrystallography [11]. It will be necessary to improve a separate set of tools, which deal with combinatorial and network search algorithms and techniques for biological profile prediction, before network pharmacology becomes widely used. The notion that comprehending the drug's biological and kinetic profile is more significant than validating specific targets or combinations of targets is one that network pharmacology brings back [12].

Large datasets are regularly gathered, saved, and analyzed in the big data era in order to support scientific discoveries and verify theories in the field of biomedicine. Without a question, the introduction of new technologies and open data efforts has resulted in a significant rise in data volume and diversity [13]. Big data are employed in every step of the drug development process, from finding new leads and therapeutic candidates to identifying targets and mechanisms of action. The purpose of illustrating and discussing these approaches is to give an overview of the many databases and computational tools that are available [14]. We believe that personalized care and cost-effectiveness are the two main goals of big data leveraging. To address this, we suggest utilizing information technologies in conjunction with (chemo)informatic tools, leveraging their complementary abilities. Identification of a potential target for the treatment is the first step in the drug discovery process [15]. This could be a specific protein or molecule that is involved in the disease

 doi:10.20944/preprints202405.1511.v1

or condition that the drug is intended to treat. Identifying a potential target for a drug is an important step in the drug discovery and development process. Chemoinformatic technologies have the potential to significantly progress in silico drug design and discovery by facilitating multi-level information integration that improves the accuracy of data results [16].

To name a few, chemical structure similarity searching, data mining/machine learning, gene ontology and enrichment analysis, STRING database for protein-protein interaction, Cytoscape have been routinely and successfully implemented [17,18].

In silico study using Pyrx and open Babel software was used to observe targets and binding affinity. It is helpful in virtual screening of libraries of compounds against potential drug targets. [19,20]

To understand and study the genetic pattern of an individual DNA molecule or chromosomes. To determine the mitotic activity of *Pisum sativum* and the effect of tulsi extract on it, the cytotoxicity test has been performed [21,22].

At the end, the liver toxicity of rats using the extract has been determined to evaluate any liver damage due to the extract [23].

Materials and Methods

Text mining: Text mining enables the automatic collection of disease-gene correlations from extensive biological literature. The UniProt Knowledgebase (UniProtKB) integrates the unreviewed UniProtKB/TrEMBL entries, annotated by automatic methods, such as our rule-based systems, with the reviewed UniProtKB/Switzerland-Prot entries, to which data have been contributed by researchers [19,20]. Of the over 120 million entries in UniProtKB/TrEMBL, the majority are the results of large-scale sequencing operations. For data exploration, UniProt creates connections between genes and the literature. This implies that when queries are run, all the genes from the accessible biological literature that are relevant to the search terms are extracted. We used the idea of "breast cancer" in our inquiry in this investigation [20–22]. After selecting "Ensembl Gene ID" and "Associated Gene Name" under GENE, we decided to use "Homo sapiens" as the species dataset. We chose to "search for UniProt genomics" after typing "breast cancer" into the search box, and then we clicked "download" from the top choices. The query then yielded all of the gene hits that were utilized for the following action [23–25].

***Examination of biological pathways and processes*:** An essential tool for the biological interpretation of high-throughput experiments, GeneCodis (http://genecodis.cnb.csic.es/) is a web-based tool that combines multiple information sources to search for annotations that frequently co-occur in a group of genes and rank them by statistical significance [26].

We conducted an enrichment analysis of the genes linked to cSCC using GeneCodis. The genes from the text mining step were added to the input set, and the GO biological process categories were used to analyze this gene set. The biological processes that showed the greatest enrichment were chosen [27,28]. The second phase involved using the genes with the chosen annotations for an additional GeneCodis analysis using the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways' annotations. The pathways that were most pertinent to the pathophysiology of cSCC were chosen among those that were highly enriched above the P-value limit. Genes associated with the chosen pathways were employed in additional analysis [29–31].

Network for protein-protein interaction: The protein-protein interactions of certain genes are integrated in the STRING database (http://string-db.org). We entered the genes we had chosen in the previous stage, chose "Homo sapiens" as the organism, and chose "Multiple proteins" from the left menu bar on the STRING database's first page. In terms of the confidence score, larger confidence scores are seen when there is more evidence that two proteins interact with one another [32–34]. Nonetheless, the study's confidence level was set at medium (score 0.400), even though a lower score would make the network less confident given that it might widen the inclusion criteria. Next, the target genes' protein-protein interaction network was discovered [35,36].

Identifying targets by acknowledging the degree & betweenness: Next, we visualized and analyzed the interaction network using the Cytoscape software platform. A software program called

 doi:10.20944/preprints202405.1511.v1

Cytoscape allows users to visually explore biological networks made up of genes, proteins, and other sorts of interactions [37,38]. The tool is backed by a variety of annotations and experimental data. We imported data from the STRING EXPORT channel in the ".tsv" format. After that, the topological features of every node were examined and the key nodes were chosen using CentiScaPe, an application that computes a greater variety of network parameters [39–41]. To choose the important genes, we used "Degree and "Betweenness" as a criterion. Greater gene products that interact with one another indicate a higher node degree, which increases the node's contribution to breast cancer. The node's Betweenness value shows how likely a node is to link to the core of other nodes [42,43]. The nodes for which Degree and Betweenness were both larger than or equal to the mean were designated as important nodes in this study according to the criteria for selecting key genes [44,45].

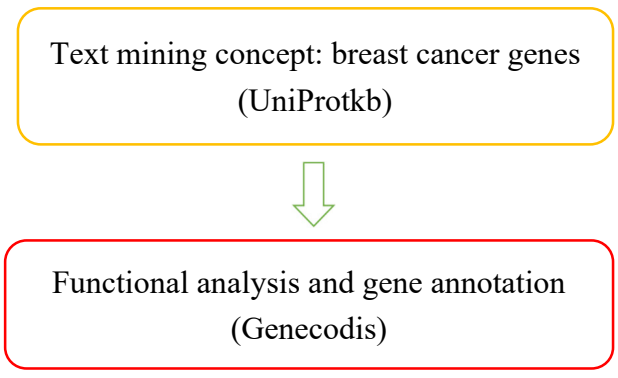
Molecular docking of the targets and lead compounds: Then, we have started with structure-based virtual screening, structures of the target molecules and small molecules [46]. For downloading the 3D structures we detailed from pubchem, protein data bank. Open Babel is used for converting sdf.Format to pdbqt.Format. Then the ligand and macromolecules are added in Pyrx [47]. Started the docking of each ligand one by one. After the completion of docking, process 'File saved' in Csv.Format.xl. Here we have checked the binding affinity between the lead and macromolecules [48,49].

Cytotoxic potential of extract on P. sativum: Seeds were kept in distilled water for 48h to obtain 1-2 cm root length and grown in sprouts maker box. Roots were exposed to the extract for 1 hour. The root tips of length between 1-2 cm were cut from 4-5 peas, and then the tips were taken out from the extract and exposed to carnoy's fixing solution (3 ethanol. 1 glacial acetic acid, v/v) for 1-2 hours. Root tips were taken out and these were incubated in a hot air oven for 15-20 mins with 6N HCL at 30-40 degree. Then the roots were exposed to carmine dye by discarding the HCL, again incubated for 10 mins. The excess dye was washed with glacial acetic acid. The effect of extract on the root tips were observed under microscope [50–52].

Precision-cut liver slice cultures in toxicity testing: Rat was dissected open after cervical dislocation, the liver lobes were removed and transferred to prewarmed kreb's Ringer HEPES buffer (KRH) (2.5 m**µ** HEPES, PH 7.4, 118 m**µ** Nacl, 2.25 m**µ** Kcl, 2.5 m**µ** $Cacl_2$, 1.5 m**µ** KH_2PO_4, 1.18 m**µ** $MgSO_4$, 5 m**µ** beta-hydroxybutyrate, 4.0 m**µ** glucose). Incubated for 10 mins in a shaker [53,54]. Washed it with KRH for 3 times, again dipped into KRH with 2ml methanol for 30 mins in the shaker. Transferred it to the eppendorf tube and homogenize it with the homogenizer stick [55]. The semi-solid was poured on the slide and smeared on it. It was left to dry in the oven for 10 mins. The slide was again dipped in methanol and freezed for 10 mins [56]. Then taken out from methanol and dipped it in the acridine solution (1.2 ml acridine + 50 ml water + 0.5 ml acetic acid). The last part was to wash off the excessive stain with distilled water and the damage was observed under the microscope [57].

Results

Result of text mining: 166 genes were discovered to be connected to breast cancer by text mining searches during the investigation of possible breast cancer treatment options (Figure 1). A total of 166 genes were taken because no duplication were discovered in the genes.



 doi:10.20944/preprints202405.1511.v1

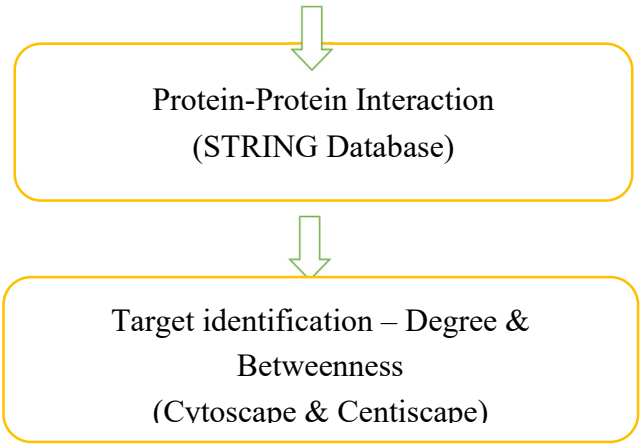
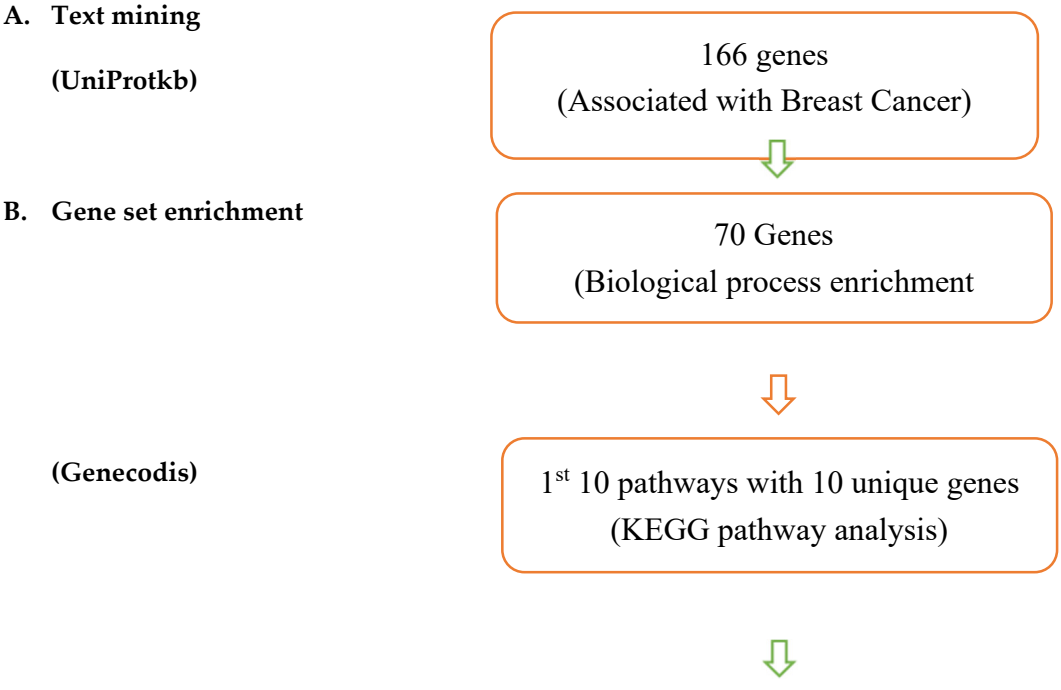


Figure 1. Overall data mining procedure.

In the (Figure 1), Overall process of data mining has been described. Text mining was utilized in conjunction with UniProtkb to uncover genes linked to breast cancer. Following extraction, GeneCodis was used to examine each gene's function. Using Cytoscape, the target identification was observed by determining the degree and betweenness of the target proteins, and further enrichment was achieved using protein interaction analysis with STRING. Thus, one can obtain potential genes.

The result from the text mining helped us to analyse the large number of proteomics and genomics from the data, articles, documents and genes associated with disease. With the help of extraction/mining of genes through text mining we found the genes that are linked with breast cancer and were sent for enrichment analysis.

Result of biological process and pathway: All the 166 genes from text mining (Table 1) are pasted in genecodis (Figure 2). In order to identify the most enriched phrases associated with breast cancer pathology, biological processes analysis was initially used in GeneCodis gene enrichment analysis. To ensure that only the most enriched annotations were chosen during this procedure. The analysis of enriched pathway annotations resulted in 1st 10 pathways containing a total of 10 unique genes (Table 2).



 doi:10.20944/preprints202405.1511.v1

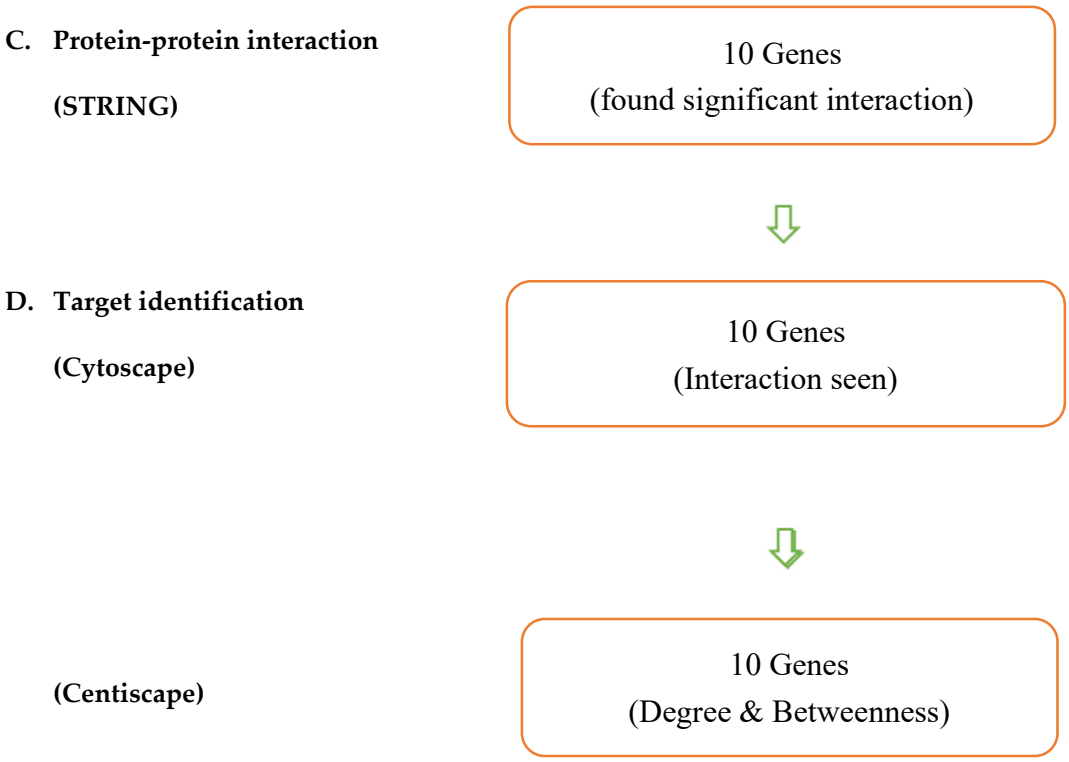


Figure 2. (A) Text mining: It was performed using the search term 'breast cancer' and 166 genes were found using UniProtkb**.** (B) Genecodis: pathway analysis was performed with 1st 10 unique genes. (C) STRING: protein-protein interaction of 10 genes were found. (D) Cytoscape: Candidate gene identification was done along with the degree and betweenness of the target genes using CentiScape.

Table 1. 166 entries with genes of Breast Cancer extracted from UniProt database.

Entry	Gene Names	Organism	Length
Q8IV36	HID1 C17orf28 DMC1	Homo sapiens (Human)	788
Q8IV76	PASD1	Homo sapiens (Human)	773
Q8IVG5	SAMD9L C7orf6 DRIF2 KIAA2005 UEF	Homo sapiens (Human)	1584
Q8IVH2	FOXP4 FKHLA	Homo sapiens (Human)	680
Q8IVL5	P3H2 LEPREL1 MLAT4	Homo sapiens (Human)	708
Q8IVL8	CPO	Homo sapiens (Human)	374
Q8IVM8	SLC22A9 hOAT4 OAT7 UST3	Homo sapiens (Human)	553
Q8IVT5	KSR1 KSR	Homo sapiens (Human)	923
Q8IW00	VSTM4 C10orf72	Homo sapiens (Human)	320
Q8IWA4	MFN1	Homo sapiens (Human)	741
Q8IWU5	SULF2 KIAA1247 UNQ559/PRO1120	Homo sapiens (Human)	870
Q8IWU6	SULF1 KIAA1077	Homo sapiens (Human)	871
Q8IWV2	CNTN4	Homo sapiens (Human)	1026
Q8IWW8	ADHFE1 HMFT2263	Homo sapiens (Human)	467
Q8IWX7	UNC45B CMYA4 UNC45	Homo sapiens (Human)	931
Q8IX03	WWC1 KIAA0869	Homo sapiens (Human)	1113
Q8IX12	CCAR1 CARP1 DIS	Homo sapiens (Human)	1150
Q8IXB3	TRARG1 IFITMD3 LOST1 TUSC5	Homo sapiens (Human)	177
Q8IXJ6	SIRT2 SIR2L SIR2L2	Homo sapiens (Human)	389
Q8IY92	SLX4 BTBD12 KIAA1784 KIAA1987	Homo sapiens (Human)	1834
Q8IYB4	PEX5L PEX5R PXR2	Homo sapiens (Human)	626
Q8IYF1	ELOA2 TCEB3B TCEB3L	Homo sapiens (Human)	753
Q8IYH5	ZZZ3	Homo sapiens (Human)	903

 doi:10.20944/preprints202405.1511.v1

Q8IYK4	COLGALT2 C1orf17 GLT25D2 KIAA0584	Homo sapiens (Human)	626
Q8IYT3	CCDC170 C6orf97	Homo sapiens (Human)	715
Q8IZ41	RASEF RAB45	Homo sapiens (Human)	740
Q8IZ69	TRMT2A	Homo sapiens (Human)	625
Q8IZF3	ADGRF4 GPR115 PGR18	Homo sapiens (Human)	695
Q8IZJ1	UNC5B P53RDL1 UNC5H2 UNQ1883/PRO4326	Homo sapiens (Human)	945
Q8IZL8	PELP1 HMX3 MNAR	Homo sapiens (Human)	1130
Q8IZW8	TNS4 CTEN PP14434	Homo sapiens (Human)	715
Q8N0W4	NLGN4X KIAA1260 NLGN4 UNQ365/PRO701	Homo sapiens (Human)	816
Q8N104	DEFB106A BD6 DEFB106 DEFB6; DEFB106B	Homo sapiens (Human)	65
Q8N108	MIER1 KIAA1610	Homo sapiens (Human)	512
Q8N136	DAW1 ODA16 WDR69	Homo sapiens (Human)	415
Q8N158	GPC2	Homo sapiens (Human)	579
Q8N163	CCAR2 DBC1 KIAA1967	Homo sapiens (Human)	923
Q8N1B3	CCNQ FAM58A	Homo sapiens (Human)	248
Q8N1L9	BATF2	Homo sapiens (Human)	274
Q8N2A8	PLD6	Homo sapiens (Human)	252
Q8N2M8	CLASRP SFRS16 SWAP2 UNQ2428/PRO4988	Homo sapiens (Human)	674
Q8N2U9	SLC66A2 PQLC1	Homo sapiens (Human)	271
Q8N371	KDM8 JMJD5	Homo sapiens (Human)	416
Q8N3A8	PARP8	Homo sapiens (Human)	854
Q8N3F8	MICALL1 KIAA1668 MIRAB13	Homo sapiens (Human)	863
Q8N427	NME8 SPTRX2 TXNDC3	Homo sapiens (Human)	588
Q8N474	SFRP1 FRP FRP1 SARP2	Homo sapiens (Human)	314
Q8N488	RYBP DEDAF YEAF1	Homo sapiens (Human)	228
Q8N4F0	BPIFB2 BPIL1 C20orf184 LPLUNC2 UNQ2489/PRO5776	Homo sapiens (Human)	458
Q8N554	ZNF276 CENP-Z ZFP276 ZNF477	Homo sapiens (Human)	614
Q8N556	AFAP1 AFAP	Homo sapiens (Human)	730
Q8N5H7	SH2D3C NSP3 UNQ272/PRO309/PRO34088	Homo sapiens (Human)	860
Q8N695	SLC5A8 AIT SMCT SMCT1	Homo sapiens (Human)	610
Q8N6D2	RNF182	Homo sapiens (Human)	247
Q8N752	CSNK1A1L	Homo sapiens (Human)	337
Q8N7J2	AMER2 FAM123A	Homo sapiens (Human)	671
Q8N7W2	BEND7 C10orf30	Homo sapiens (Human)	519
Q8N807	PDILT	Homo sapiens (Human)	584
Q8N8S7	ENAH MENA	Homo sapiens (Human)	591
Q8N9N5	BANP BEND1 SMAR1	Homo sapiens (Human)	519
Q8N9N8	EIF1AD	Homo sapiens (Human)	165
Q8NA54	IQUB	Homo sapiens (Human)	791
Q8NAP8	ZBTB8B	Homo sapiens (Human)	495
Q8NAX2	KDF1 C1orf172	Homo sapiens (Human)	398
Q8NB49	ATP11C ATPIG ATPIQ	Homo sapiens (Human)	1132
Q8NBU5	ATAD1 FNP001	Homo sapiens (Human)	361

Table 2. Summary of KEGG process gene set enrichment analysis.

 doi:10.20944/preprints202405.1511.v1

	Description	Pval adj	Relative enrichment	Genes
2	mitochondrial membrane fusion	4.14E-02	562.59	MFN1
3	positive regulation of mitochondrial fusion	4.14E-02	375.06	PLD6
4	regulation of cell division	5.90E-02	51.14	ZZZ3
5	cell differentiation	1.12E-01	3.33	BATF2,PLD6,PDILT
6	regulation of protein localization to membrane	5.61E-02	187.53	GPC2
7	smoothened signaling pathway	4.14E-02	31.7	GPC2,IQUB
8	protein auto-ADP-ribosylation	5.61E-02	93.77	PARP8
9	neuron cell-cell adhesion	5.61E-02	66.19	CNTN4
10	RNA methylation	6.38E-02	41.67	TRMT2A
11	translational initiation	1.13E-01	15.85	EIF1AD

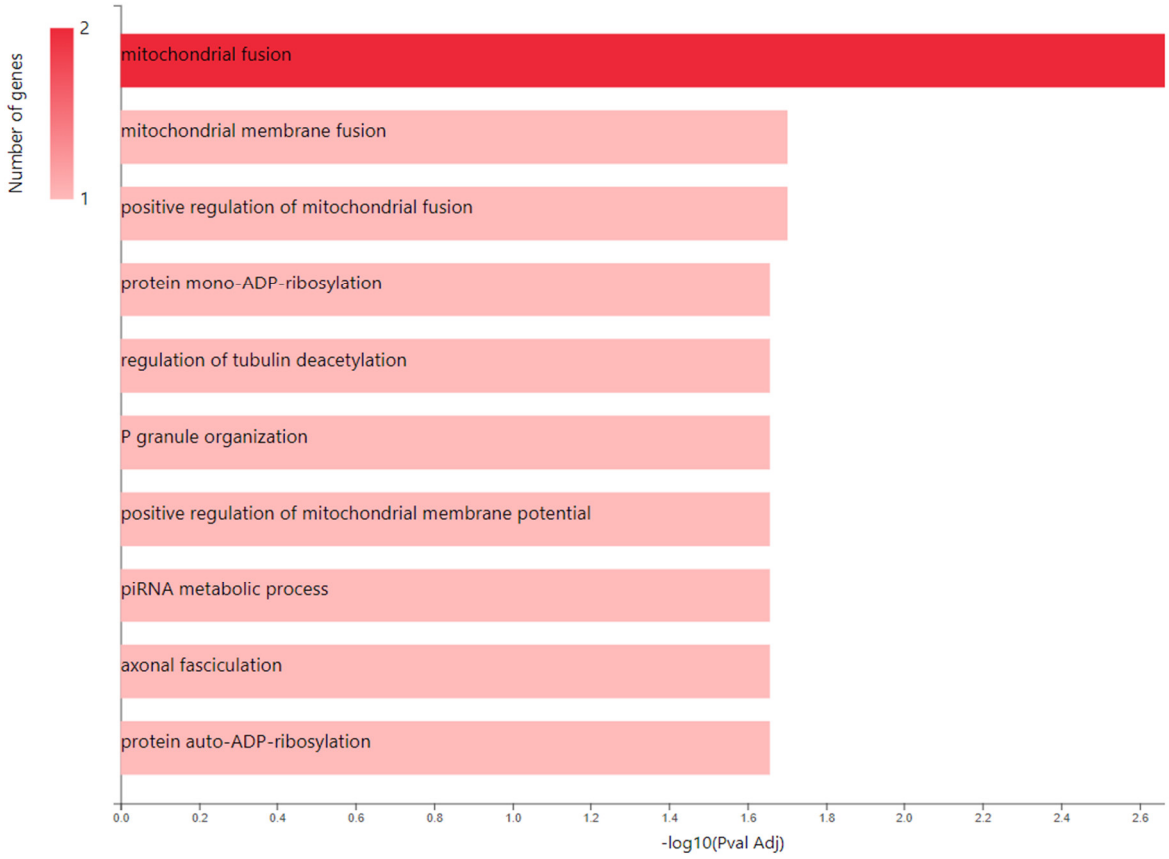


Figure 3. Bar-graph depicting the relationship between the genes and Pvalue.

Result of protein-protein interaction: The protein-protein interaction of 10 genes that were extracted from the KEGG enrichment analysis were autodetected by STRING database and the nodes and their length between the targets are studied and shown accordingly in (Figure 4).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 May 2024 doi:10.20944/preprints202405.1511.v1

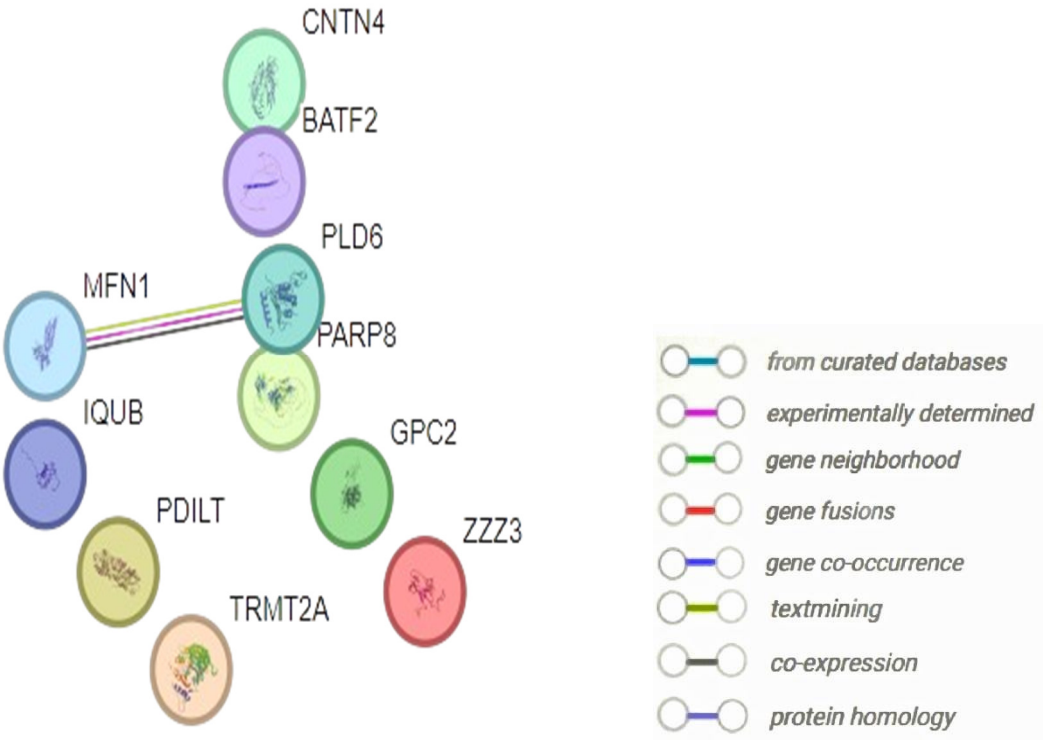


Figure 4. The protein-protein medium (confidence score 0.511) interaction network of the 10 targeted genes, produced using STRING. Network nodes represent proteins and different colored edges represent protein-protein interaction.

Result of target identification with degree and betweenness: Here we have identified the targets and candidate genes (Figure 5). We loaded data into Cytoscape in the ".tsv" format from the STRING EXPORT channel. Then, to evaluate each node's topological characteristics and identify the important nodes, we employed CentiScaPe, a software that computes a greater number of network characteristics, with this the degree and betweenness values of all the proteins were identified (Table 3). The total number of edges occurring to the node is represented by the node degree (Figure 6).

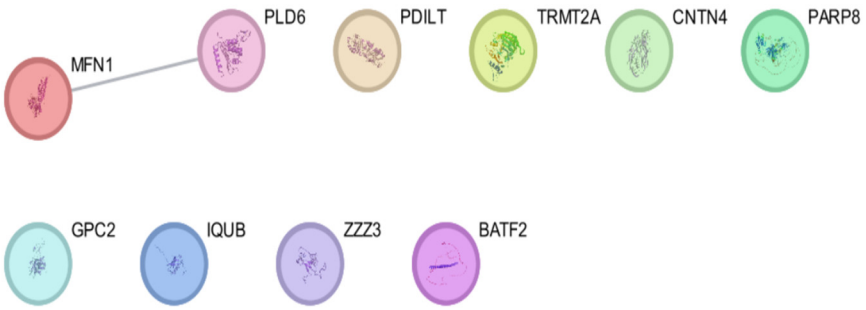


Figure 5. The protein-protein interaction network of 10 genes and identification of candidate genes, produced using Cytoscape. Network nodes represent proteins and edges represent protein-protein associations.

 doi:10.20944/preprints202405.1511.v1

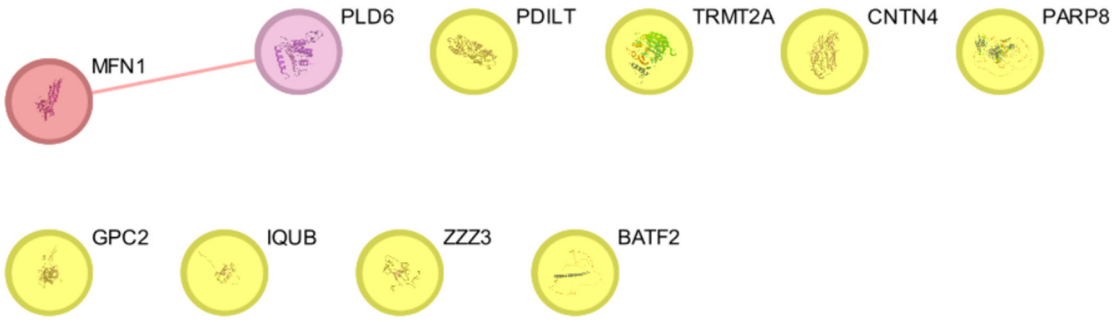


Figure 6. The degree and betweenness of the proteins are produced by using CentiScaPe.

Table 3. Degree and betweenness values of the candidate genes.

Genes/Proteins	Betweenness	Degree
MFN1	0	1
PDILT	0	0
TRMT2A	0	0
CNTN4	0	0
PARP8	0	0
GPC2	0	0
IQUB	0	0
ZZZ3	0	0
BAFT2	0	0
PLD6	0	1

Result of Molecular Docking with the targets and the compounds: 10 Genes that were storted out using various softwares and centiscape were then studied with virtual screening and from which 4 genes were able to dock with 28 chemical constituents of Tulsi and have shown binding affinity. The genes were: MFN1, PDILT, GPC2, ZZZ3. All these 4 candidate genes had shown greater binding affinity with Ursolic acid (compound of Tulsi extract).

Table 4. Binding affinity of the genes with the lead compound.

Ligand	Binding Affinity
GPC2 + Ursolic acid	8.7
PDILT + Ursolic acid	8.6
MFN1 + Ursolic acid	8.2
ZZZ3 + Ursolic acid	7.1

 doi:10.20944/preprints202405.1511.v1

Result of cytotoxic activity with P.sativum: The concentration of the extract was 100**µ**g/ml and from that 3 ml of the extract solution was taken and diluted upto 10 ml with distilled water. Upon observing the slide under microscope, it was seen that the extract has no cytotoxic effect in the plant hence the result was found to be negative (Figure 7).

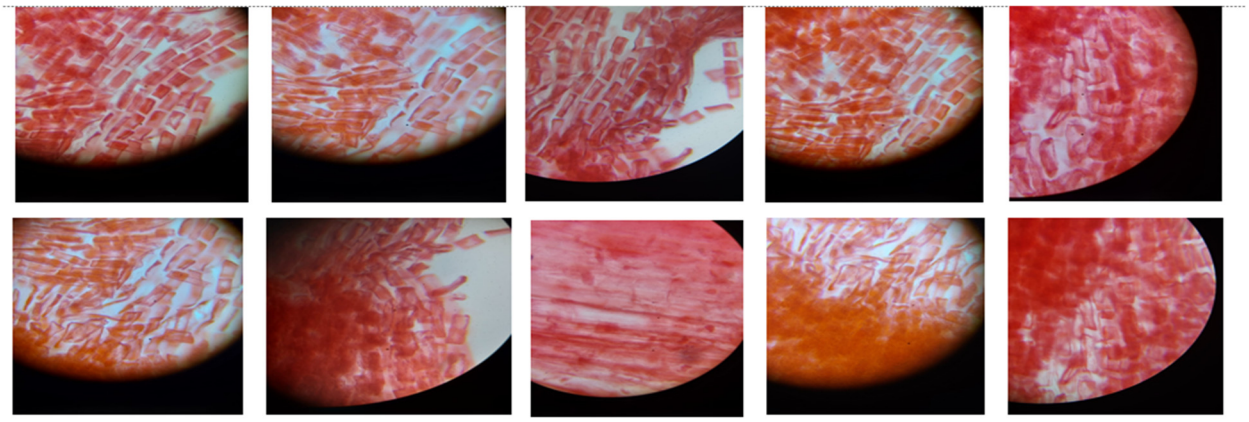


Figure 7. No cytotoxic effect of Tulsi extract has been observed in the mitotic activity of *P.sativum*.

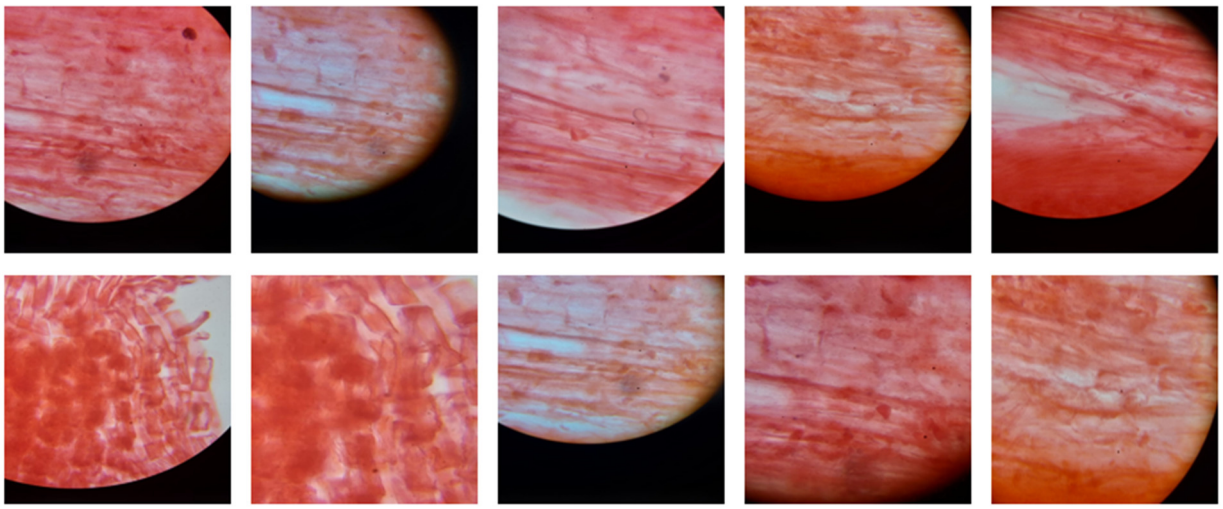


Figure 8. Mitotic activity of *P.sativum* (Control).

Result of liver slice culture in toxicity testing: When the slices were installed into methanol certain damage to the liver has been seen. The green colour in the pictures show the 'Debris', the yellow colour in the pictures show 'RNA Damage', The deep yellow colour depicts 'DNA Damage'.

 doi:10.20944/preprints202405.1511.v1

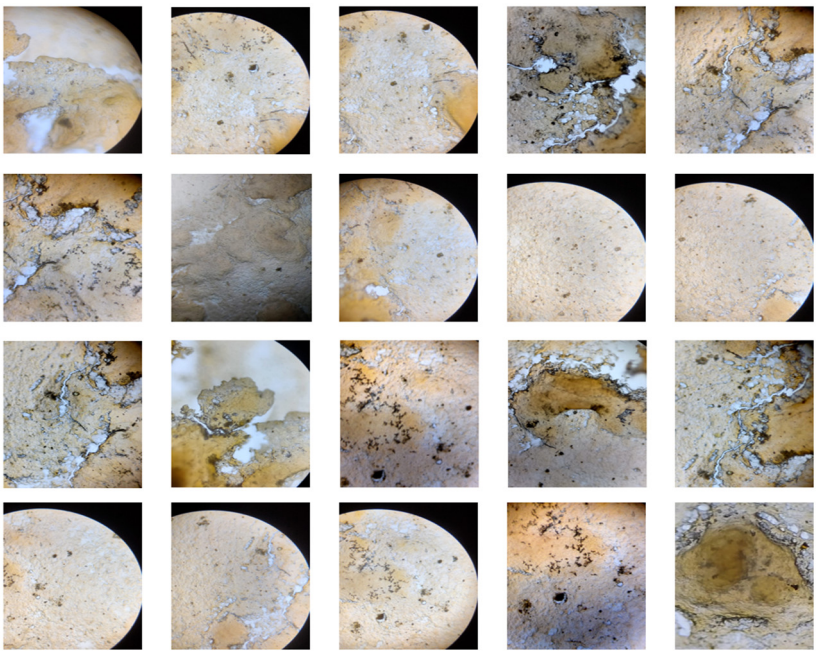


Figure 8. Liver damage due to methanol has been seen under the microscope.

Discussion

The most prevalent cancer of the mammary gland, breast cancer has a high rate of metastasis and a fatality rate of over 70% when it spreads. It is well acknowledged that surgical excision is the primary treatment for most cases of breast cancer. There hasn't been much study done on medications as adjuvant therapy, though [58,59]. After doing a gene set enrichment analysis, we were able to identify 166 target genes in this study, of which 70 genes could potentially be used to treat breast cancer. The carcinogenic process's genetic modifications affect how cells function, enabling self-sufficiency in growth signals, insensitivity to antigrowth signals, escape from apoptosis, infinite replicative potential, invasion, angiogenesis, and metastasis. These changes are consistent with the enriched biological processes—such as "cell differentiation," "cell division," "cell adhesion," and "signal transduction"—that were found using GeneCodis analysis (Table I) [60–62]. The appropriate genes are then exported to the STRING database to check and analyse the protein-protein interaction along with the study of different types of nodes present. The same genes are as well transferred to the Cytoscape to visualize the interaction of targets more flexibly [63,64]. With the targeted gene we came across the degree and betweenness of the genes, which helped us to imply the candidate genes for the further analysis with drugs [65–67]. After getting the genes, molecular docking has been performed of those targets with the compounds of Tulsi extract and the compound showing highest binding affinity with the protein was found as lead compound [68,69]. After the network pharmacology, In-vitro cell culture was performed to check the cytotoxicity of the extract, and it was observed that there was no toxic effect of the extract on the plant. Then we moved towards ex-vivo with rat liver. We have checked the toxicity of the extract on the liver slices, and no such damage has been observed with the liver slice cultures [70,71]. With the new approaches of personalised medicine, or mitigated medication, these techniques can show high-yielding measures with which we can help curing breast cancer and prevent it from further spreading [72,73].

Conclusion

1. The proteins that have shown the molecular docking are: **MFN1, PLD6, GPC2, ZZZ3.**
2. The compound found as the lead was: **Ursolic Acid.**
3. Cytotoxicity of the extract was found to be: **Negative.**
4. Liver damage of rat with the extract was found to be: **Negative.**

 doi:10.20944/preprints202405.1511.v1

From the above network pharmacology and experimental pharmacology with various analytical tools and wet lab synthesis, we came to the conclusion that, with the incidence of breast cancer on the rise and patient expectations rising in recent years, selecting an appropriate treatment to maximize preservation of function and minimize the risk of metastasis and recurrence remains challenging. In the clinic, surgical excision is frequently considered the best course of action for treating breast cancer. Non-surgical techniques like photodynamic therapy (PDT), radiation therapy, cryotherapy, and chemotherapy are commonly used to treat late-stage breast cancer. However, research on pharmaceutical therapy is still lacking, and more studies are needed to help develop novel therapeutic strategies.

Even though the efficacy of the currently available drug therapies is limited, and the discovery of new drug therapies using traditional methods is likely to take a long time, drug repositioning may speed up the process of discovering additional conditions that existing drugs could treat more effectively and potentially at a lower cost. This study aimed to investigate new novel plant based drug therapies for breast cancer by means of computational methods including text mining, biological process and pathway analysis, protein-protein interaction (PPI) analysis to mine public databases, bioinformatics tools and experimental methods with in-vitro and ex-vivo to systematically identify interaction networks between drug-gene targets and finding out the lead, including working with the compounds in wet lab. We were able to use data analytical techniques to look at the characteristics of possible genes and work through plant based models in order to select a medication.

With the help of such experimentations, we can further take the analysis towards the novel drug therapy and treat the genes that can be used as a biomarker in correspondence with the drugs compounded for breast cancer.

References

1. Sun, Y. S., Zhao, Z., Yang, Z. N., Xu, F., Lu, H. J., Zhu, Z. Y., ... & Zhu, H. P. (2017). Risk factors and preventions of breast cancer. *International journal of biological sciences*, *13*(11), 1387.
2. Stewart BW, and Wild CP. World Cancer Report 2014. Geneva, Switzerland: WHO Press; 2014.
3. Siegel RL, Miller KD, and Jemal A. Cancer Statistics, 2017. CA Cancer J Clin. 2017; 67: 7-30.
4. Polyak K. Breast cancer: origins and evolution. J Clin Invest. 2007; 117: 3155-3163.
5. . Valenti G, Quinn HM, Heynen G, et al. Cancer Stem Cells Regulate Cancer-Associated Fibroblasts via Activation of Hedgehog Signalling in Mammary Gland Tumors. Cancer Res. 2017; 77: 2134-2147.
6. Eisemann N, Waldmann A, Geller AC, Weinstock MA, Volkmer B, Greinert R, Breitbart EW and Katalinic A: Non-melanoma skin cancer incidence and impact of skin cancer screening on incidence. J Invest Dermatol 134: 43-50, 2014.
7. Muzic JG, Schmitt AR, Wright AC, Alniemi DT, Zubair AS, Olazagasti Lourido JM, Sosa Seda IM, Weaver AL and Baum CL: Breast cancer: A population-based study in olmsted county, minnesota, 2000 to 2010. Mayo Clin Proc 92: 890-898, 2017.
8. Burton KA, Ashack KA and Khachemoune breast cancer: A review of high-risk and metastatic disease. Am J Clin Dermatol 17: 491-508, 2016.
9. Diepgen T, Fartasch M, Drexler H and Schmitt J: Occupational breast cancer induced by ultraviolet radiation and its prevention. Br J Dermatol 167 (Suppl 2): S76-S84, 2012.
10. Boeckx C, Baay M, Wouters A, Specenier P, Vermorken JB, Peeters M and Lardon F: Anti-epidermal growth factor receptor therapy in head and neck squamous cell carcinoma: Focus on potential molecular mechanisms of drug resistance. Oncologist 18: 850-864, 2013.
11. Moosavinasab S, Patterson J, Strouse R, Rastegar-Mojarad M, Regan K, Payne PR, Huang Y and Lin SM: 'RE:fine drugs': An interactive dashboard to access drug repurposing opportunities. Database (Oxford) 2016: pii: baw083, 2016.
12. Andronis C, Sharma A, Virvilis V, Deftereos S and Persidis A: Literature mining, ontologies and information visualization for drug repurposing. Brief Bioinform 12: 357-368, 2011.
13. Liu H, Beck TN, Golemis EA and Serebriiskii IG: Integrating in silico resources to map a signaling network. Methods Mol Biol 1101: 197-245, 2014.
14. Baran J, Gerner M, Haeussler M, Nenadic G and Bergman CM: pubmed2ensembl: A resource for mining the biological literature on genes. PLoS One 6: e24716, 2011.
15. Carmona-Saez P, Chagoyen M, Tirado F, Carazo JM and Pascual-Montano A: GENECODIS: A web-based tool for finding significant concurrent annotations in gene lists. Genome Biol 8: R3, 2007.

 doi:10.20944/preprints202405.1511.v1

16. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-CepasJ, Simonovic M, Roth A, Santos A, Tsafou KP, et al.: STRING v10: Protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 43: D447-D452, 2015.
17. Harbeck N, Beckmann MW, Rody A, Schneeweiss A, Müller V, Fehm T, Marschner N, Gluz O, Schrader I, Heinrich G, et al.: HER2 dimerization inhibitor pertuzumab-mode of action and clinical data in breast cancer. Breast Care (Basel) 8: 49-55, 2013.
18. Balduzzi S, Mantarro S, Guarneri V, Tagliabue L, Pistotti V, Moja L and D'Amico R: Trastuzumab-containing regimens for metastatic breast cancer. Cochrane Database Syst Rev: CD006242, 2014.
19. Jacob, R. B., Andersen, T., & McDougal, O. M. (2012). Accessible high-throughput virtual screening molecular docking software for students and educators. *PLoS computational biology*, *8*(5), e1002499.
20. Kondapuram, S. K., Sarvagalla, S., & Coumar, M. S. (2021). Docking-based virtual screening using PyRx Tool: autophagy target Vps34 as a case study. In *Molecular Docking for Computer-Aided Drug Design* (pp. 463-477). Academic Press.
21.
1. Majewska, A., Wolska, E., Śliwińska, E., Furmanowa, M., Urbańska, N., Pietrosiuk, A., ... & Kuraś, M. (2003). Antimitotic effect, G2/M accumulation, chromosomal and ultrastructure changes in meristematic cells of Allium cepa L. root tips treated with the extract from Rhodiola rosea roots. *Caryologia*, *56*(3), 337-351.
22. Van de Bovenkamp, M., Groothuis, G. M. M., Meijer, D. K. F., & Olinga, P. (2007). Liver fibrosis in vitro: cell culture models and precision-cut liver slices. *Toxicology in vitro*, *21*(4), 545-557.
23. Naik, R. S., Mujumdar, A. M., & Ghaskadbi, S. (2004). Protection of liver cells from ethanol cytotoxicity by curcumin in liver slice culture in vitro. *Journal of ethnopharmacology*, *95*(1), 31-37.
24. Wood JP, Smith AJ, Bowman KJ, Thomas AL and Jones GD: Comet assay measures of DNA damage as biomarkers of irino- tecan response in colorectal cancer in vitro and in vivo. Cancer Med 4: 1309-1321, 2015.
25. Cogle CR, Scott BL, Boyd T and Garcia-Manero G: Oral azaciti- dine (CC-486) for the treatment of myelodysplastic syndromes and acute myeloid leukemia. Oncologist 20: 1404-1412, 2015.
26. Su G, Morris JH, Demchak B and Bader GD: Biological network exploration with Cytoscape 3. Curr Protoc Bioinformatics 47: 8.13.1-24, 2014.
27. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, et al.: Integrative analysis of complex cancer genomics and clinical profiles using the cBio- Portal. Sci Signal 6: pl1, 2013.
28. Wagner AH, Coffman AC, Ainscough BJ, Spies NC, Skidmore ZL, Campbell KM, Krysiak K, Pan D, McMichael JF, Eldred JM, et al.: DGIdb 2.0: Mining clinically relevant drug-gene interactions. Nucleic Acids Res 44: D1036-D1044, 2016.
29. Ribas A and Flaherty KT: BRAF targeted therapy changes the treatment paradigm in melanoma. Nat Rev Clin Oncol 8: 426-433, 2011.
30. UniProt Consortium. (2019). UniProt: a worldwide hub of protein knowledge. *Nucleic acids research*, *47*(D1), D506-D515.
31. Karsch-Mizrachi I., Takagi T., Cochrane G.International Nucleotide Sequence Database Collaboration The international nucleotide sequence database collaboration. Nucleic Acids Res. 2018; 46:D48–D51.
32. Giraldo-Calderón G.I., Emrich S.J., MacCallum R.M., Maslen G., Dialynas E., Topalis P., Ho N., Gesing S.VectorBase ConsortiumVectorBase ConsortiumMadey G.et al. VectorBase: an updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. Nucleic Acids Res. 2015; 43:D707–D713.
33. Zerbino D.R., Achuthan P., Akanni W., Amode M.R., Barrell D., Bhai J., Billis K., Cummins C., Gall A., Girón C.G.et al. Ensembl 2018. Nucleic Acids Res. 2018; 46:D754–D761.
34. Chen C., Natale D.A., Finn R.D., Huang H., Zhang J., Wu C.H., Mazumder R. Representative proteomes: a stable, scalable and unbiased proteome set for sequence analysis and functional annotation. PLoS One. 2011; 6:e18910.
35. Chen C., Huang H., Mazumder R., Natale D.A., McGarvey P.B., Zhang J., Polson S.W., Wang Y., Wu C.H., Consortium UniProt Computational clustering for viral reference proteomes. Bioinformatics. 2016; 32:2041–2043.
36. Mitchell A.L., Scheremetjew M., Denise H., Potter S., Tarkowska A., Qureshi M., Salazar G.A., Pesseat S., Boland M.A., Hunter F.M.I.et al. EBI Metagenomics in 2017: enriching the analysis of microbial communities, from sequence reads to assemblies. Nucleic Acids Res. 2018; 46:D726–D735.
37. Gene Ontology Consortium The Gene Ontology project in 2008. Nucleic Acids Res. 2008; 36:D440–D444.
38. Hastings J., Owen G., Dekker A., Ennis M., Kale N., Muthukrishnan V., Turner S., Swainston N., Mendes P., Steinbeck C. ChEBI in 2016: Improved services and an expanding collection of metabolites. Nucleic Acids Res. 2016; 44:D1214–D1219.

 doi:10.20944/preprints202405.1511.v1

39. Poux S., Arighi C.N., Magrane M., Bateman A., Wei C.-H., Lu Z., Boutet E., Bye-A-Jee H., Famiglietti M.L., Roechert B.et al. On expert curation and scalability: UniProtKB/Swiss-Prot as a case study. Bioinformatics. 2017; 33:3454–3460.
40. Orchard S., Kerrien S., Abbani S., Aranda B., Bhate J., Bidwell S., Bridge A., Briganti L., Brinkman F.S.L., Cesareni G.et al. Protein interaction data curation: the International Molecular Exchange (IMEx) consortium. Nat. Methods. 2012; 9:345–350.
41. Rao, V. S., Srinivas, K., Sujini, G. N., & Kumar, G. N. (2014). Protein-protein interaction detection: methods and analysis. *International journal of proteomics*, *2014*.
42. Pillutla, R. C., Fisher, P. B., Blume, A. J., & Goldstein, N. I. (2002). Target validation and drug discovery using genomic and protein–protein interaction technologies. *Expert Opinion on Therapeutic Targets*, *6*(4), 517-531.
43. Feng, Y., Wang, Q., & Wang, T. (2017). Drug target protein-protein interaction networks: a systematic perspective. *BioMed research international*, *2017*.
44. Archakov, A. I., Govorun, V. M., Dubanov, A. V., Ivanov, Y. D., Veselovsky, A. V., Lewi, P., & Janssen, P. (2003). Protein-protein interactions as a target for drugs in proteomics. *Proteomics*, *3*(4), 380-391.
45. Scott, D. E., Bayly, A. R., Abell, C., & Skidmore, J. (2016). Small molecules, big targets: drug discovery faces the protein–protein interaction challenge. *Nature Reviews Drug Discovery*, *15*(8), 533-550.
46. Wang R, Lai L, Wang S (2002) Further development and validation of empirical scoring functions for structure-based binding affinity prediction. J Comput-Aided Mol Des 16: 11–26.
47. Goodsell DS, Olson AJ (1990) Automated docking of substrates to proteins by simulated annealing. Proteins 8: 195–202.
48. Ounthaisong, U., & Tangyuenyongwatana, P. (2017). Cross-docking study of flavonoids against tyrosinase enzymes using PyRx 0.8 virtual screening tool. *TJPS*, *41*(2017).
49. Morris G, Goodsell D, Halliday R, Huey R, Hart W (1998) Automated docking using a Lamarckian genetic algorithm and an empirical free energy function. J Comput Chem 19: 1639–1662.
50. Kavitha, M., Srinivasan, P. T., Renuga, G., & Jayakumar, L. V. (2014). Evaluation of Antimitotic Activity of Mukia maderaspatana L. Leaf Extract in Allium cepa Root Model. *International Journal of Pharmacy Research & Technology (IJPRT)*, *4*(2), 1-4.
51. Aşkin Çelik, T., & Aslantürk, Ö. S. (2006). Anti-mitotic and anti-genotoxic effects of Plantago lanceolata aqueous extract on Allium cepa root tip meristem cells. *Biologia*, *61*, 693-697.
52. Thenmozhi, A., Nagalakshmi, A., & Rao, U. M. (2011). Study of cytotoxic and antimitotic activities of Solanum nigrum by using Allium cepa root tip assay and cancer chemopreventive activity using MCF-7-human mammary gland breast adenocarcinoma cell lines. *Int J Sci Technol*, *1*(2), 26-48.
53. Nordmann R, Ribièrea C, Rouach H. Implication of free radical mechanisms in ethanol-induced cellular injury. Free Radic Biol Med 1992;12:3219-40
54. Artee GE. Oxidants and antioxidants in alcohol-induced liver disease. Gastroenterol 2003;124:778-90.
55. Wu D, Cederbaum AI. Alcohol, oxidative stress, and free radical damage. Alcohol Res Health 2003;27:274-84
56. Kurose I, Higuchi H, Kato S, Miura S, Ishii H. Ethanol-induced oxidative stress in the liver. Alcohol Clin Exp Res 1996;20:77-85.
57. Kumaran A, Karunakaran RJ. In vitro antioxidant activities of methanol extracts of five Phyllanthus species from India. Food Sci Technol-LEB 2007;40:344-52.
58. Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., ... & Von Mering, C. (2015). STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic acids research*, *43*(D1), D447-D452.
59. Safari-Alighiarloo, N., Taghizadeh, M., Rezaei-Tavirani, M., Goliaei, B., & Peyvandi, A. A. (2014). Protein-protein interaction networks (PPI) and complex diseases. *Gastroenterology and Hepatology from bed to bench*, *7*(1), 17.
60. Saito, R., Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L., Lotia, S., ... & Ideker, T. (2012). A travel guide to Cytoscape plugins. *Nature methods*, *9*(11), 1069-1076.
61. 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), 2498-2504.
62. Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L., & Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*, *27*(3), 431-432.
63. Noor, F., Tahir ul Qamar, M., Ashfaq, U. A., Albutti, A., Alwashmi, A. S., & Aljasir, M. A. (2022). Network pharmacology approach for medicinal plants: review and assessment. *Pharmaceuticals*, *15*(5), 572.
64. S Azmi, A. (2013). Adopting network pharmacology for cancer drug discovery. *Current drug discovery technologies*, *10*(2), 95-105.
65.

 doi:10.20944/preprints202405.1511.v1

2. Pei, C., Yang, K., Chen, Y., Dong, Y., Meng, X., & Song, N. (2024). Mechanisms of action of Qinghao in treating breast cancer and doxorubicin-induced cardiotoxicity based on network pharmacology and molecular docking.
66. Hopkins, A. L. (2008). Network pharmacology: the next paradigm in drug discovery. *Nature chemical biology*, *4*(11), 682-690.
67. Li, Z., Han, P., You, Z. H., Li, X., Zhang, Y., Yu, H., ... & Chen, X. (2017). In silico prediction of drug-target interaction networks based on drug chemical structure and protein sequences. *Scientific reports*, *7*(1), 11174.
68. Irwin JJ, Shoichet BK (2004) ZINC – A free database of commercially available compounds for virtual screening. J Chem Inf Model 45: 177–182.
69. Collignon B, Schulz R, Smith JC, Baudry J (2011) Task-parallel message passing interface implementation of Autodock4 for docking of very large databases of compounds using high-performance super-computers. J Comp Chem 32: 1202–1209.
70. Aşkin Çelik, T., & Aslantürk, Ö. S. (2006). Anti-mitotic and anti-genotoxic effects of Plantago lanceolata aqueous extract on Allium cepa root tip meristem cells. *Biologia*, *61*, 693-697.
71. Palamanda JR, Kehre JP. Inhibition of protein carbonyl formation and lipid peroxidation by glutathione in rat liver microsomes. Arch Biochem Biophys 1992;293:103-9.
72. Baraya, Y. U. S. A. B., Wong, K. K., & Yaacob, N. S. (2017). The immunomodulatory potential of selected bioactive plant-based compounds in breast cancer: a review. *Anti-Cancer Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry-Anti-Cancer Agents)*, *17*(6), 770-783.
73. Kapinova, A., Stefanicka, P., Kubatka, P., Zubor, P., Uramova, S., Kello, M., ... & Kruzliak, P. (2017). Are plant-based functional foods better choice against cancer than single phytochemicals? A critical review of current breast cancer research. *Biomedicine & Pharmacotherapy*, *96*, 1465-1477.