Indian Journal of Biochemistry & Biophysics
Vol. 60, May 2023, pp. 393-405
DOI: 10.56042/ijbb.v60i5.1356

Screening of phytochemicals for potential breast cancer targets BRCA1 and BARD1: A network pharmacology approach

Kikrusenuo Kiewhuo[1,2], Esther Jamir[1,2], Lipsa Priyadarsinee[1,2], Selvaraman Nagamani[1,2] & G Narahari Sastry[1,2]*

[1]Advanced Computation and Data Sciences Division,
CSIR – North East Institute of Science and Technology, Jorhat – 785 006, Assam, India

[2]Academy of Scientific and Innovative Research (AcSIR), Ghaziabad-201002, Uttar Pradesh, India

Received 10 May 2023; *revised 12 July 2023*

Breast cancer disease onset and progression has a complex pathophysiology. The recent findings reveal that the genetic variations and mutations in the ring domains of BRCA1 and BRCA2 appear to be of critical significance. These genes are essential for DNA repair and adverse variations or mutations in the BRCA1 and BRCA2 genes may increase the likelihood of developing breast cancer. Virtual screening of phytochemicals that are interacting with BRCA1 and BRCA2 mutations can be an effective strategy in identifying potential leads against the genes associated with breast cancer. Thus, the current study focuses on the efficacy of the phytochemicals from north east region of India for their potential against breast cancer. The Phytochemicals reported in North East India Medicinal Plant Database (NEIMPDB) were chosen against the BRCA1 and BARD1 proteins which are implicated in human breast cancer. Maclurin, a phytophenol phytochemical known for its antioxidant activity was identified as a top scoring phytochemical in the screening of both BRCA1 and BARD1 proteins. The traditional practices of these plants revealed the use of *Ricinis communis* in the treatment of glandular cancer. Further, the network pharmacology and gene enrichment analysis were also performed for the target genes of top phytochemicals res with Cytoscape to highlight the constituent target gene in diseases pathways. Eight (8) phytochemicals namely Acteoside (*Barleria cristata*), Anthraquinones Emodin (*Stellaria media*), L-(+)-Quercitol (*Aloe barbadensis*), Maclurin (*Ricinus communis*), Myoinositol (*Dioscorea sp.*), Neoisoastilbin (*Goniothalamus sesquipedalis*), Okanin (*Urtica magellanica*), and Plantamajoside (*Coccinia grandis*) were identified as the compounds with the most number of breast cancer target genes.

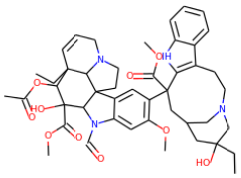
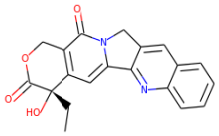
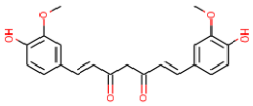
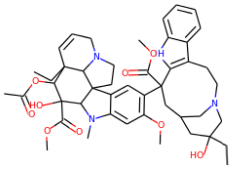
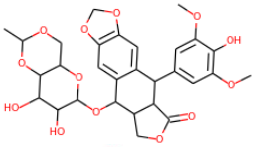
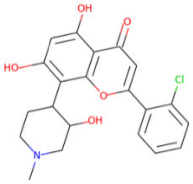
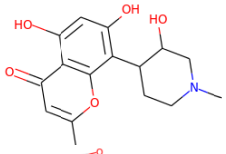
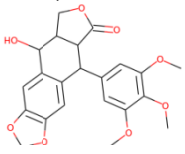
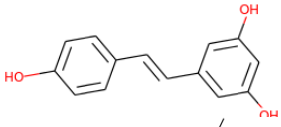
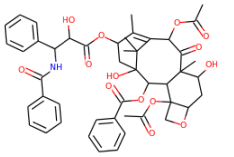
Keywords: Drug targets, Gene enrichment analysis, Medicinal plants, Network pharmacology, Physicochemical properties, Virtual Screening

Medicinal plants have been used for centuries by millions of communities around the world to treat various ailments. About 80% of the citizens of emergent nations rely on medicinal plants as the primary source of health management[1] and these practices have greatly influenced the culture, religion and ethnicity of the communities. These plants release an arsenal chemicals as defensive mechanisms against infections and predators[2]. These chemicals are the natural products, and they play significant role in the chemical biology for the treatment of diseases and infections[3]. The identification and isolation of natural products has led to the development of a variety of commonly used drugs. Use of products as anti-cancer agents have harnessed interest among pharmaceutical industries in the past decades. There are currently over a 100 natural product derived compounds under clinical trials for therapeutics for cancer[4]. Some important natural products used as anti-cancer agents are betulinic acid, camptothecin, curcumin, etoposide, indirubins, flavopiridol, rohitukine, podophyllotoxin, resveratrol, taxol, vinblastine, and Vincristine (Table 1)[5].

A wide spectrum of medicinal plants is known to treat cancer, with approximately 3000 plant species[6]. The *in silico* analysis of the natural products of medicinal plants can serve as a connecting bridge in prioritizing the plants for cancer treatment. Cancer is responsible for nearly 10 million deaths worldwide [WHO report, 2020], among which breast cancer is the most invasive and developing type of cancer (2.26 million cases, WHO report)[7]. Mutations in the genes BRCA1 and BRCA2 are accountable for about 30-50% of hereditary breast cancer. Breast cancer type 1 susceptibility protein (BRCA1) interacts with BRCA1 Associated RING Domain 1 (BARD1) at the N and C terminal, where the N-terminal alpha helix of BRCA1 is aligned anti-parallel to the C-terminal alpha helix of the BARD1 and vice versa[8]. The four-helix bundle

*Correspondence:
E-mail: gnsastry@gmail.com; gnsastry@neist.res.in
Suppl. Data available on respective page of NOPR

Table 1 — List of some commonly used anti-cancer drugs derived from plants

Sl No	Drug	Plant name	Medicinal use	Structure
1	Betulinic acid	*Dillenia indica*	Cuts and wounds	
2	Camptothecin	*Natsiatum herpeticum*	Carminative	
3	Curcumin	*Cucurma angustifolia*	Tonic	
4	Vinblastine	*Cataranthus roseus*	Tuberculosis	
5	Etoposide	*Sinopodophyllum hexandrum*	Cardiovascular Agents	
6	Flavopiridol	*Dysoxylum binectariferum*	Bladder cancer	
7	Rohitukine	*Aphanamixis polystachya*	Tumours	
8	Podophyllotoxin	*podophyllum peltatum*	Tuberculosis	
9	Resveratrol	*Michelia champaca*	Cancer, heart disease	
10	Taxol	*Taxus brevifolia*	Breast cancer, ovarian cancer	

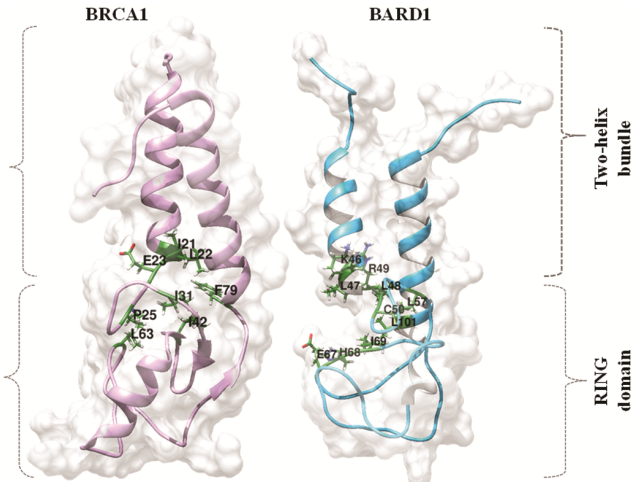


Fig. 1 — Crystal structure of BRCA1 and BARD1 proteins with the active sites (PDB ID-IJM7)

stabilizes the heterodimer formed and the interaction formed increases the UBE2 ligase activity of BRCA1 which is responsible for the tumour suppressive function of the protein (Fig. 1). This heterodimeric complex is important in catalyzing DNA double strand break (DSB) repair by homologous recombination and ubiquitin transfer by E3-ubiquitin protein ligase activity[9]. The E3 ubiquitin ligase binds to the BRCA1-BARD1 heterodimer and activates the interaction with DNA, the DNA damage response components help target DSBs, particularly those on damaged replication forks, into the HR pathway for repair. These activities contribute to the antitumorigenic properties of this complex[10]. The mutations terminate the ubiquitin ligase activity and the domain loses its ability to restore the G2-M checkpoint. Thus, mutation in this complex is directly associated with breast and ovarian cancers. The BRCA1 and BARD1 proteins interact at the RING domain which further complexes with the ubiquitin-conjugating enzyme E2D3. This complex associates with the histone proteins in the nucleosomal core complex, particularly with the H2A and H2B histone proteins[11]. The interaction of the BRCA1-BARD1 and nucleosomes induces high-affinity chromatin binding and increases the ubiquitination ability of the histone proteins. Enhancing the interaction between the enzyme and substrate complex lead to the structural stability, which appear to contribute to the enabling of ubiquitination-mediated DNA damage repair mechanism[9]. The BRCA1-BARD1 complex can regulate transcriptional control and DNA repair mechanisms by reading and writing nucleosomes. This complex can affect the structure of chromatin and the expression of genes by identifying particular histone modifications and changing H2A. The interactions between BRCA1-BARD1, nucleosomes, and other protein complexes indicate its multifunctional role in preserving genome integrity and avoiding the onset of breast cancer[12].

Network pharmacology is a novel approach based on interdisciplinary theories implemented for the construction and visualization of networks that help evaluate the molecular mechanisms of drugs from multiple perspectives[13]. Network pharmacology can be used identify the potential target genes of medicinal plants[14] and help in identifying the important pathways and genes of breast cancer.

The current study focuses on screening the phytochemicals present in medicinal plants reported for the treatment of various ailments and documented in North East India Medicinal Plant Database (NEI-MPDB)[15] against breast cancer. A total of 4242 phytochemicals from NEI-MPDB were screened against BRCA1 and BARD1 protein of human breast cancer. The target genes were identified for the top scoring phytochemicals and network was generated. The gene enrichment analysis was also performed for the dataset. Maclurin, a phytophenol phytochemical known for its antioxidant activity was identified as a top scoring phytochemical in the screening of both BRCA1 and BARD1 proteins of human breast cancer. The medicinal plants that report the presence of maclurin were identified from the database and the traditional practices of these plants revealed the traditional use of *Ricinis communis* in the treatment of glandular cancer. Further, the network pharmacology aspect of the phytochemicals and gene analysis was also performed for the dataset to highlight the plant constituent target gene diseases pathways.

Materials and Methods

Screening of phytochemicals

Data of phytochemicals were obtained from NEI-MPDB[15], a collective database of medicinal plants found in North East India. The database reported a total of 5410 active phytochemicals found in medicinal plants of the North East region. This dataset was screened and curated where only those phytochemicals whose compound summary is listed in PubChem were identified. The curation led to the identification of 4242 unique phytochemicals.

Target protein for breast cancer

The structure of BRCA1 and BARD1 protein (PDB ID: 1JM7) was retrieved from Protein Data Bank (PDB)[16] (Table S1). The retrieved structure is a heterodimeric RING-RING complex responsible for ubiquitination and DNA double stranded break repair. The complex was prepared using UCSF Chimera[17] for the generation of separate chains and removal of heteroatoms, ions and water molecules.

Active site prediction and grid generation

Active site prediction of the protein targets was done using CASTp[18]. CASTp is an open access server which computes the surface pockets of proteins. The grid box was generated by selecting the active sites as the centroid of the box in the receptor-grid generation module of Maestro[19] (Table S2).

Virtual screening

Virtual screening is a technique used in the identification of novel bioactive molecules[20-23]. This approach is more efficient than other methods at obtaining lead compounds inexpensively and quickly while consuming less resources[24-26]. Virtual screening was conducted using Glide of Maestro v11.5 available in Schrödinger Release 2018-1 (Schrödinger LLC, New York, USA; https://www.schrodinger.com)[19].

Physicochemical and ADMET properties calculation

The physicochemical and ADMET properties of the compounds with the highest docking scores were calculated using ADMETlab 2.0[27]. ADMETlab 2.0 is an open access server for the predictions of ADMET properties of compounds[28-30].

Prediction of target genes of top 20 molecules

The targets of the top 20 best molecules were retrieved and predicted for constructing protein network. The potential targets were retrieved from target databases such as DrugBank[31], Drug repurposing hub[32]. The unavailable targets were predicted through online server DIGEB-Prèd[33] by using SMILES as input file of the molecules. The retrieval of breast cancer target genes was done using two human gene database- GeneCards (https://www.genecards.org/)[34] and MalaCards (https://www.malacards.org/)[35] using keywords "breast cancer" "mammary carcinoma". The targets obtained from these datasets was compared with the original target genes of the top 20 phytochemicals.

Network construction and analysis

The network interaction analysis for biological entities has made the identification of new drugs a reality[36,37]. The complex organization of gene-gene, protein-protein interactions can be visualized and studies using network pharmacology approaches[14]. The list of selected genes was taken to construct protein-protein network using Cytoscape 3.9.1[38]. For gene biological interpretation, a Cytoscape plugin ClueGo[39] was used to integrate KEGG[40] pathways and create gene ontology pathway term network. Further, gene enrichment analysis was carried out using online tools such as ShinyGO 0.76[41] and Metascape[42]. The workflow adapted for the study is given in (Fig. 2).

Results and discussion

Virtual screening

The phytochemicals dataset was screened against two breast cancer targets – BRCA1 and BARD1 respectively. The top docking scores for both screening was identified (Table 2). The highest docking score for BRCA1 was –8.033 kcal/mol and –8.695 kcal/mol for BARD1 protein depicting the high affinity of the compounds. The highest scores belong to phytochemicals Maclurin and Plantamajoside, respectively. Maclurin also had a high docking score for BARD1 protein and it is the only phytochemical with

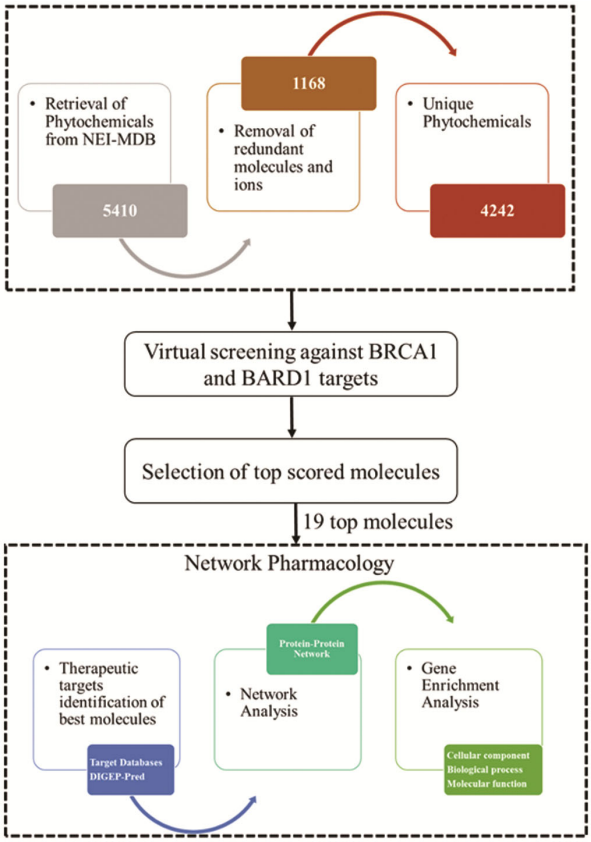


Fig. 2 — Schematic representation of workflow adopted in this study to identify potential phytochemicals against breast cancer

high docking scores in both BRCA1 and BARD1 protein. Maclurin is a phytochemical belonging to phytophenol family, known for its antioxidant activity and in the treatment of hyperuricemia or gout. It is presented in plants like *Ricinus communis, Mangifera indica,* and in *Garcinia* of Clusiaceae family. The comparison of the docking scores between the two breast cancer proteins revealed that BARD1 had a higher average docking score (–7.74 kcal/mol) than BRCA1 (–6.44 kcal/mol). The 2D representations of the docking interaction with BRCA1 and BARD1 targets is shown in (Figs 3 & 4). The 2D interaction study was carried out to analyse the interaction pattern of protein and ligand complexes using Discovery studio[43]. It can be observed that most of the interacting residues are among the active site residues with various non-covalent interactions between the top 10 phytochemicals and the interacting protein BRCA1 and BARD1 respectively[8,9]. The residues in the interacting surface of BRCA1 and BARD1 is reported to play important role in maintaining the stability and function of the complex[9]. Hence, we have also identified few of the residues from the 2D interaction studies which are found to be present among the interface residues. Residues namely Arg78, Asp40, Phe79, Asp40 and Ile21 were identified among the BRAC1 and its interacting phytochemicals and residues such as Ile69, Glu67, Leu47, and Cys66 were identified among the BRAD1 and its interacting residues respectively. Among these residues, Phe79, Ile21 and Leu47 are reported to be hotspot residues present in the BRAC1-BARD1 complexes[59]. This gives an overview of the identified residues which can be targeted for designing potential inhibiters against the breast cancer associated proteins.

In addition, the electrostatic potential map of the top 10 phytochemicals along with the interacting amino acids of BRCA1 and BARD1 was carried out to map the interacting region which are involved protein-ligand interaction binding affinity (Figs S1 & S2).

Physicochemical and ADMET properties

The physicochemical properties of the top scoring phytochemicals were calculated, and these phytochemicals was also mapped with the medicinal plants listed in NEI-MPDB (Table 2). It was observed that the plants *Camellia sinensis* and *Ricinus communis* were among the most occurring plants. The optimal molecular weight of a potential drug candidates ranges between 100-600, and it was observed that fifteen phytochemicals among the top scorers had the optimal range of molecular weight.

Table 2 — The docking score and physicochemical properties of the identified top 20 phytochemicals interacting with breast cancer associated targets BRCA1 and BARD1

Proteins	Compound name	Medicinal Plants	Docking score (kcal/mol)	LogS	LogD	LogP	MW	TPSA
BRCA1	Maclurin	*Ricinus communis*	–8.033	–3.111	1.700	1.875	262.05	118.22
	6-thioguanosine5'-diphosphate	*Allium ascalonicum*	–7.374	–1.642	–1.270	–2.924	459.00	229.91
	6-Biopterin	*Foeniculum vulgare*	–7.363	–2.896	–0.890	–2.040	237.09	138.74
	Gallocatechin	*Camellia sinensis*	–6.449	–2.88	0.905	0.736	306.07	130.61
	Paeonidin-3-5,-Diglucoside	*Begonia palmata*	–6.232	–2.746	0.824	–0.843	625.18	259.75
	Anthraquinones Emodin	*Stellaria media*	–5.902	–5.594	2.845	3.856	270.05	94.83
	2′-Deoxycytidine	*Gelsemium elegans*	–5.755	–0.556	–1.847	–3.032	307.06	154.70
	Adenine	*Cajanus cajan*	–5.735	–2.458	–0.633	–1.686	135.05	81.21
	(+)-Scytalone	*Streblus asper*	–5.729	–1.226	1.012	1.006	194.06	77.76
	Clavatol	Begonia inflata	–5.689	–1.821	1.302	2.244	180.08	57.53
BARD1	Plantamajoside	*Coccinia grandis*	–8.695	–1.821	–0.256	–0.346	640.20	265.52
	3,4-dihydroxyphenyl	*Camellia sinensis*	–8.608	–4.065	2.017	2.964	716.14	281.95
	Myoinositol	*Dioscorea sp.*	–8.083	–0.353	–2.328	–3.057	180.06	121.38
	L-(+)-Quercitol	*Aloe barbadensis*	–7.848	–0.08	–1.863	–2.364	164.07	101.15
	D-Glucosamine	*Costus speciosus*	–7.534	–0.005	–2.208	–2.539	179.08	116.17
	Maclurin	*Ricinus communis*	–7.338	–3.111	1.700	1.875	262.05	118.22
	Okanin	*Urtica magellanica*	–7.248	–2.389	1.026	2.149	288.06	118.22
	Neoisoastilbin	*Goniothalamus sesquipedalis*	–7.231	–3.871	1.113	–0.101	450.12	186.37
	Acteoside	*Barleria cristata*	–7.055	–2.051	1.055	–0.328	624.21	245.29
	Leucoanthocyanidin	*Aegle marmelos*	–6.864	–3.026	0.198	–0.272	322.07	150.84

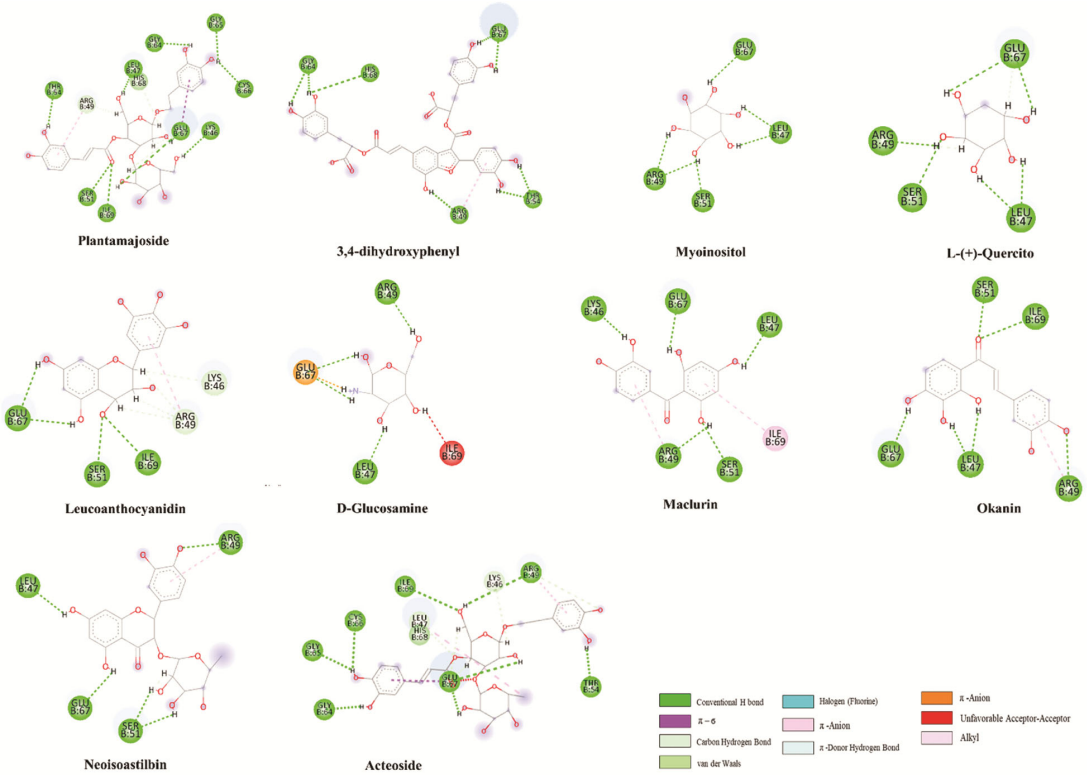


Fig. 3 — 2D representation depicting docking interactions of top 10 phytochemicals with breasts cancer associated target BRCA1

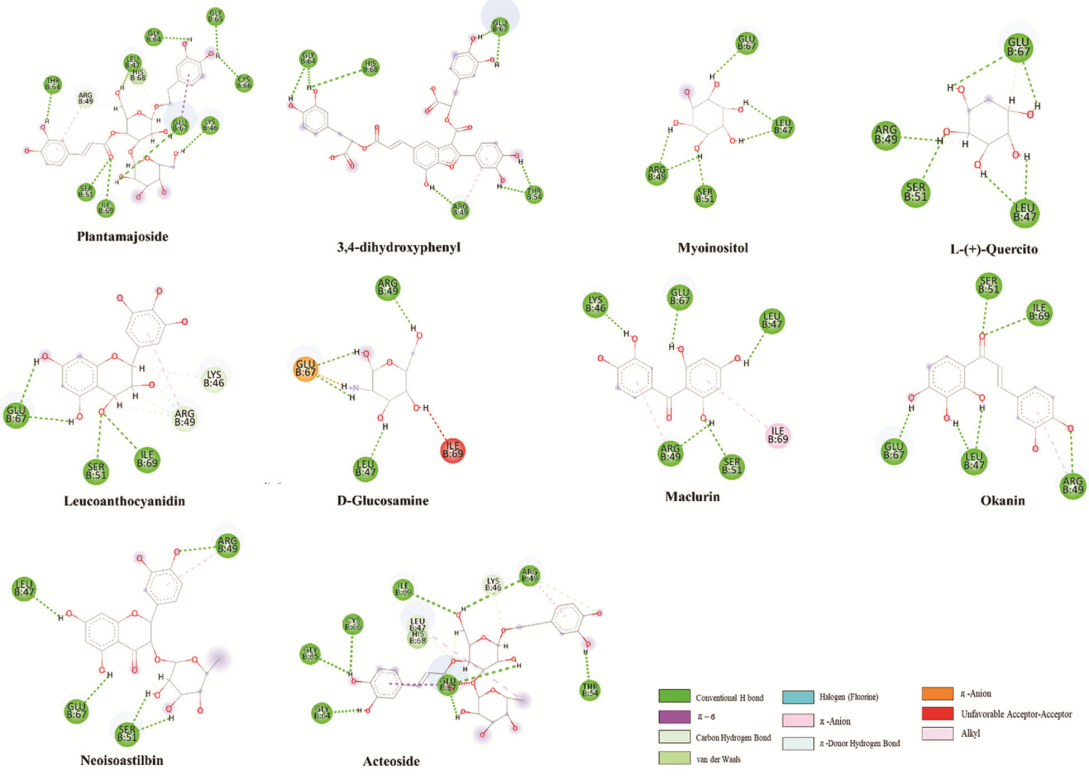


Fig. 4 — 2D representation depicting docking interactions of top 10 phytochemicals with breasts cancer associated target BARD1

Table 3 — The ADMET properties of the top 20 phytochemicals

Compound name	Medicinal Plants	ADME						Toxicity	
		BBB	Caco-2	HIA	MDCK	PPB	SkinSen	Ames	Carcinogenicity
Maclurin	*Ricinus communis*	0.021	–5.463	0.036	6.63E-06	96.54%	0.942	0.778	0.105
6-thioguanosine 5'-diphosphate	*Allium ascalonicum*	0.809	–6.516	0.743	3.54E-05	15.47%	0.301	0.384	0.941
6-Biopterin	*Foeniculum vulgare*	0.885	–6.093	0.018	3.49E-05	42.87%	0.049	0.063	0.947
Gallocatechin	*Camellia sinensis*	0.019	–6.306	0.274	3.89E-06	91.16%	0.958	0.437	0.036
Paeonidin-3-5,-Diglucoside	*Begonia palmata*	0.454	–6.453	0.953	0.00012	69.37%	0.009	0.461	0.162
Anthraquinones Emodin	*Stellaria media*	0.056	–5.149	0.016	1.10E-05	99.29%	0.678	0.823	0.301
Clavatol	Begonia inflata	0.141	–4.709	0.007	1.53E-05	94.48%	0.904	0.029	0.385
2′-Deoxycytidine	*Gelsemium elegans*	0.82	–6.194	0.966	0.00127	7.79%	0.244	0.017	0.081
Adenine	*Cajanus cajan*	0.288	–5.072	0.024	4.29E-06	11.43%	0.928	0.346	0.032
(+)-Scytalone	Streblus asper	0.309	–4.917	0.107	6.31E-06	48.79%	0.9	0.146	0.095
Plantamajoside	*Coccinia grandis*	0.12	–6.656	0.995	2.45E-05	93.19%	0.962	0.339	0.031
3,4-dihydroxyphenyl	*Camellia sinensis*	0.012	–6.652	0.914	8.20E-06	97.28%	0.872	0.036	0.196
Myoinositol	*Dioscorea* sp.	0.693	–5.334	0.974	0.01187	12.30%	0.02	0.038	0.003
L-(+)-Quercitol	*Aloe barbadensis*	0.773	–5.224	0.919	0.00151	10.68%	0.06	0.071	0.013
Leucoanthocyanidin	*Aegle marmelos*	0.033	–6.462	0.473	3.81E-06	89.74%	0.94	0.418	0.024
D-Glucosamine	*Costus speciosus*	0.373	–5.274	0.885	0.00201	10.65%	0.083	0.342	0.016
Okanin	*Urtica magellanica*	0.02	–5.375	0.051	9.01E-06	99.08%	0.967	0.829	0.607
Neoisoastilbin	*Goniothalamus sesquipedalis*	0.066	–6.495	0.761	7.98E-06	91.50%	0.276	0.732	0.246
Acteoside	*Barleria cristata*	0.336	–6.449	0.993	6.37E-05	93.34%	0.687	0.389	0.077

*BBB: blood–brain barrier, Caco-2: cancer coli-2, HIA: human intestinal absorption, MDCK: Madin-Darby canine kidney, PPB: Plasma Protein Binding, SkinSen: skin sensitisation, Ames: Ames

The topological polar surface area (TPSA) was also calculated based on Veber's rule and twelve phytochemicals had the optimal range of scores (0-140). The log S, log D and log P values were also calculated and it was observed that sixteen, eight and seven phytochemicals scored values in the optimal value range for potential drug candidate.

The ADMET properties was calculated using the following parameters – blood brain barrier (BBB), colon adenocarcinoma cell lines (Caco-2) permeability, Madin−Darby Canine Kidney cells (MDCK) permeability, human intestinal absorption, plasma protein binding (PPB), skin sensitization, Ames test for mutagenicity, and carcinogencity. It has been observed that 47% of phytochemicals had the probability of being BBB-, while 42% of phytochemicals were predicted to be HIA+ and low probability of causing skin sensitization. 52% of phytochemicals were predicted to have a proper Caco-2 and PPB value, 36% of compounds showed high probability for MDCK permeability and low Ames toxicity. About 73% of phytochemicals were predicted as having low probability of being carcinogens (Table 3). In addition, drug likeliness properties of the top scoring phytochemicals were also calculated **(**Table S3). It was observed that all the top 19 phytochemicals followed two or more drug likeliness rules among which phytochemicals such as maclurin, 6-Biopterin, Gallocatechin, Anthraquinones Emodin, Leucoan-thocyanidin, 2′-Deoxycytidine and Okanin were found to be following all the four rules which also indicates their drug likeness properties to be potential leads for treating cancer diseases.

Therapeutic indications

The therapeutic indications of the available phytochemicals were obtained from DrugCentral and DrugBank. The therapeutic indications of the phytochemicals include a combination of Berger's disease, growth factor, Immune boosters, Polycystic Kidney and pseudovitamin. The traditional practices of these plants were also obtained from NEI-MPDB. It was observed that the screened plants were used in the treatment of thirty different ailments which includes - abdominal pain, abortifacient, anaemia, analgesic, arthritis, body ache, boils, bronchitis, burns, chest pain, cough/cold, cuts and wounds, diabetes, diarrhoea, dysentery, epilepsy, fever,

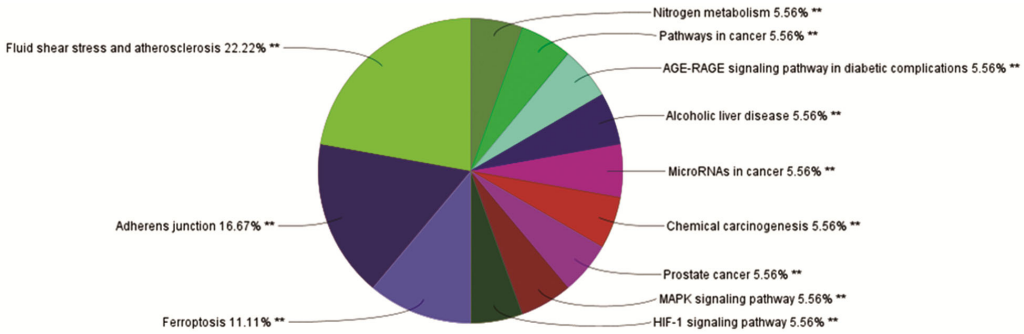


Fig. 5 — Percentage of genes involved in various KEGG pathways where the maximum percentage of 22.22% of genes are observed to be involved in fluid shear stress and atherosclerosis while the minimum percentage of 5.56% of gene is observed in multiple pathways

gastritis, glandular tumors, helminthiasis, inflammation, jaundice, kidney problems, menstrual disorder, muscle cramp, piles, pneumonia, rheumatism, skin diseases, and toothache. An interesting observation was seen where the plant *Ricinus communis* have been reported to be used in the treatment of glandular cancer. Therefore, this corresponds to the high docking score obtained from the phytochemical maclurin.

Screening of potential and breast cancer targets

A total of 326 potential targets were obtained from the top 20 phytochemicals Drug Bank, Drug repurposing hub and DIGEB-Prèd. The dataset was curated to identify unique potential targets, where 194 genes were obtained for the network construction. The screening of breast cancer targets from the human gene database identified 204 targets. This dataset was compared with the potential targets and 20 genes out of the 194 were identified as breast cancer genes.

Network construction and analysis

The complex drug-gene-disease interaction mechanism is addressed using network pharmacology. The network was constructed using 194 target genes from twenty phytochemicals and eighteen medicinal plants through cytoscape network visualization where FDR value was set to <0.05. The target genes interacted strongly with the fluid shear stress and atherosclerosis pathway. The top ten targets obtained from the network analysis are - fluid shear stress and atherosclerosis (22.22%), adherens junction (15.67%), ferroptosis (11.11%), nitrogen meta-bolism (5.56%), pathways in cancer (5.56%), AGE-RAGE signalling pathway in diabetic complications (5.56%), alcoholic liver diseases (5.56%), microRNAs in cancer (5.56%), chemical carcinogensis (5.56%), and prostate cancer (5.56%) (Fig. 5). The network diagram of the top 20 phytochemicals with 194 genes is shown in (Fig. 6).

It can be observed from the network generated that about 22.24% of genes are involved in cancer related pathways such as microRNAs in cancer, chemical carcinogensis, proteoglycans in cancer, and prostate cancer. Table 4 represents the pathways associated with the target genes identified from different databases. P-value < 0.05 indicates that the listed pathways and genes have significant correlation with the phytochemicals and cancer related target genes. Further, the set of genes were also analysed to identify the breast cancer genes among the genes involved in the four cancer pathways. It was observed that 30 different genes from 13 unique phytochemicals belonging to 13 medicinal plants were involved in the cancer related pathways. Among them, the phytochemical okarin reported in *Urtica magellanica* was the most occurring phytochemicals with 24 predicted target genes occurring in the network, followed by 11 genes of athraquinones of *Stellaria media* and 10 genes of myoinositol of *Dioscorea* sp. The analysis with the breast cancer genes revealed the presence of 15 genes involved in the cancer related pathways of the network constructed. The phytochemicals where the 15 breast cancer target genes were identified are Acteoside (*Barleria cristata*), Anthraquinones Emodin (*Stellaria media*), L-(+)-Quercitol (*Aloe barbadensis*), Maclurin (*Ricinus communis*), Myoinositol (*Dioscorea sp.*), Neoisoastilbin (*Goniothalamus sesquipedalis*), Okanin (*Urtica magellanica*), and Plantamajoside (*Coccinia grandis*). The structure of the eight identified phytochemicals are shown in (Fig. 7).

GO and pathway enrichment analysis

The crucial target genes were analysed by GO biological process enrichment and KEGG pathway

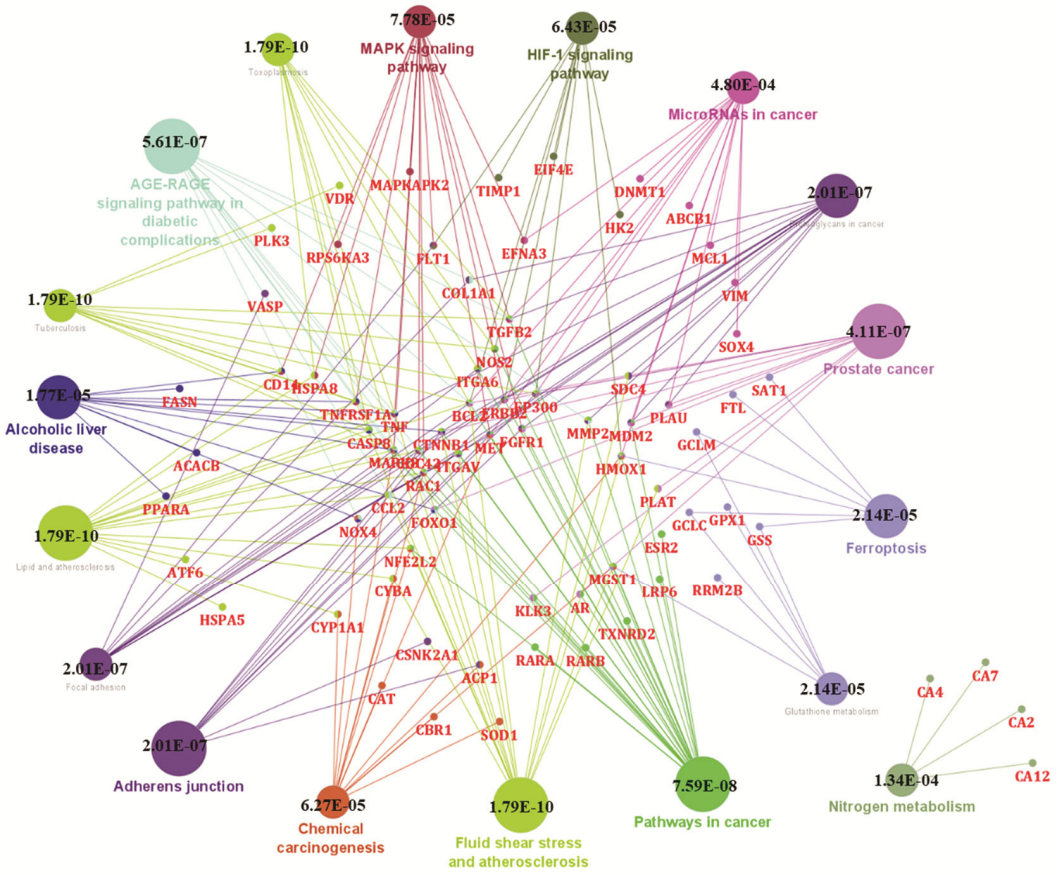


Fig. 6 — Gene enrichment analysis for genes associated with top screened phytochemicals. The network represents the KEGG pathways of the selected genes where FDA value was set to <0.05

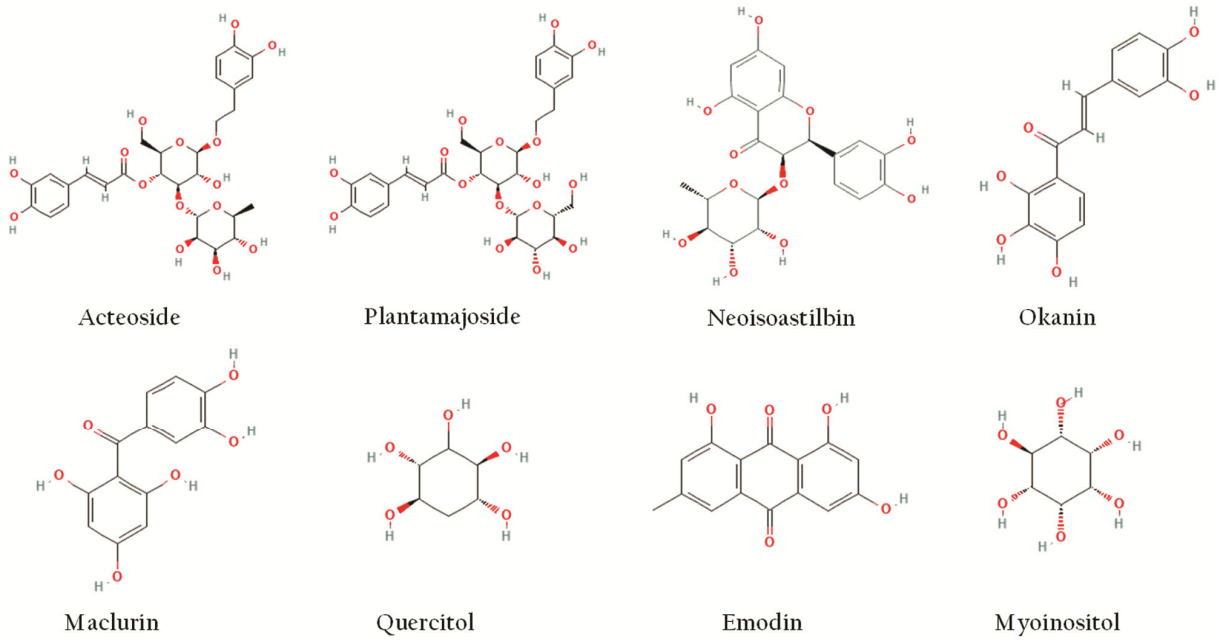


Fig. 7 — The 2D structures of the eight selected phytochemicals as potential breast cancer targets on network pharmacology

Table 4 — P-value of the pathways associated with the target genes of the top 20 phytochemicals along with the no. and type of genes

Term	P-Value	No. of Genes	Associated Genes Found
Toxoplasmosis	1.79E-10	9	BCL2, CASP8, HSPA8, ITGA6, MAPK8, NOS2, TGFB2, TNF, TNFRSF1A
Tuberculosis	1.79E-10	11	BCL2, CASP8, CD14, EP300, MAPK8, NOS2, PLK3, TGFB2, TNF, TNFRSF1A, VDR
Lipid and atherosclerosis	1.79E-10	15	ATF6, BCL2, CASP8, CCL2, CD14, CDC42, CYBA, CYP1A1, HSPA5, HSPA8, MAPK8, NFE2L2, RAC1, TNF, TNFRSF1A
Fluid shear stress and atherosclerosis	1.79E-10	15	BCL2, CCL2, CTNNB1, CYBA, HMOX1, ITGAV, MAPK8, MGST1, MMP2, NFE2L2, PLAT, RAC1, SDC4, TNF, TNFRSF1A
Pathways in cancer	7.59E-08	27	AR, BCL2, CASP8, CDC42, CTNNB1, EP300, ERBB2, ESR2, FGFR1, FOXO1, HMOX1, ITGA6, ITGAV, KLK3, LRP6, MAPK8, MDM2, MET, MGST1, MMP2, NFE2L2, NOS2, RAC1, RARA, RARB, TGFB2, TXNRD2
Focal adhesion	2.01E-07	12	BCL2, CDC42, COL1A1, CTNNB1, ERBB2, FLT1, ITGA6, ITGAV, MAPK8, MET, RAC1, VASP
Adherens junction	2.01E-07	9	ACP1, CDC42, CSNK2A1, CTNNB1, EP300, ERBB2, FGFR1, MET, RAC1
Proteoglycans in cancer	2.01E-07	14	CDC42, COL1A1, CTNNB1, ERBB2, FGFR1, ITGAV, MDM2, MET, MMP2, PLAU, RAC1, SDC4, TGFB2, TNF
Prostate cancer	4.11E-07	11	AR, BCL2, CTNNB1, EP300, ERBB2, FGFR1, FOXO1, KLK3, MDM2, PLAT, PLAU
AGE-RAGE signaling pathway in diabetic complications	5.61E-07	11	BCL2, CCL2, CDC42, COL1A1, FOXO1, MAPK8, MMP2, NOX4, RAC1, TGFB2, TNF
Alcoholic liver disease	1.77E-05	11	ACACB, CASP8, CD14, CTNNB1, FASN, FOXO1, MAPK8, NOX4, PPARA, TNF, TNFRSF1A
Glutathione metabolism	2.14E-05	6	GCLC, GCLM, GPX1, GSS, MGST1, RRM2B
Ferroptosis	2.14E-05	6	FTL, GCLC, GCLM, GSS, HMOX1, SAT1
Chemical carcinogenesis	6.27E-05	13	ACP1, CAT, CBR1, CYBA, CYP1A1, HMOX1, MAPK8, MET, MGST1, NFE2L2, NOX4, RAC1, SOD1
HIF-1 signaling pathway	6.43E-05	9	BCL2, EIF4E, EP300, ERBB2, FLT1, HK2, HMOX1, NOS2, TIMP1
MAPK signaling pathway	7.78E-05	15	CD14, CDC42, EFNA3, ERBB2, FGFR1, FLT1, HSPA8, MAPK8, MAPKAPK2, MET, RAC1, RPS6KA3, TGFB2, TNF, TNFRSF1A
Nitrogen metabolism	1.34E-04	4	CA12, CA2, CA4, CA7
MicroRNAs in cancer	4.80E-04	14	ABCB1, BCL2, DNMT1, EFNA3, EP300, ERBB2, HMOX1, MCL1, MDM2, MET, PLAU, SOX4, TGFB2, VIM

Blue colour highlighted are the breast cancer target genes retrieved from human gene database GeneCards and MalaCards

enrichment analysis. The GO enrichment analysis comprised three parts: biological process (BP), cellular components (CC), and molecular function (MF). The GO and BP analysis predicted that the key targets are mainly involved in response to oxidative stress, nutrient levels and extracellular stimulus where 29 to 32 genes were involved. The analysis also showed that the target involved in the response to oxygen containing component had a low fold enrichment score but involved the highest number of genes (73 genes). The key targets involved in the cellular component analysis are perinulcear endoplasmic reticulum, ficolin-1-rich granule lumen and peroxisome. The component involving the greatest number of genes was the mitochondrion where 36 genes were involved in the activity. The molecular function analysis was also performed and the following activities were observed to have the highest fold enrichment scores – carbonate dehydratase, nuclear receptor, and ligand activated transcription factor activity. Signalling receptor binding activity had the highest number of gene involvement (34 genes). The top ten enriched biological process, cellular components and molecular function are shown in (Fig. 8).

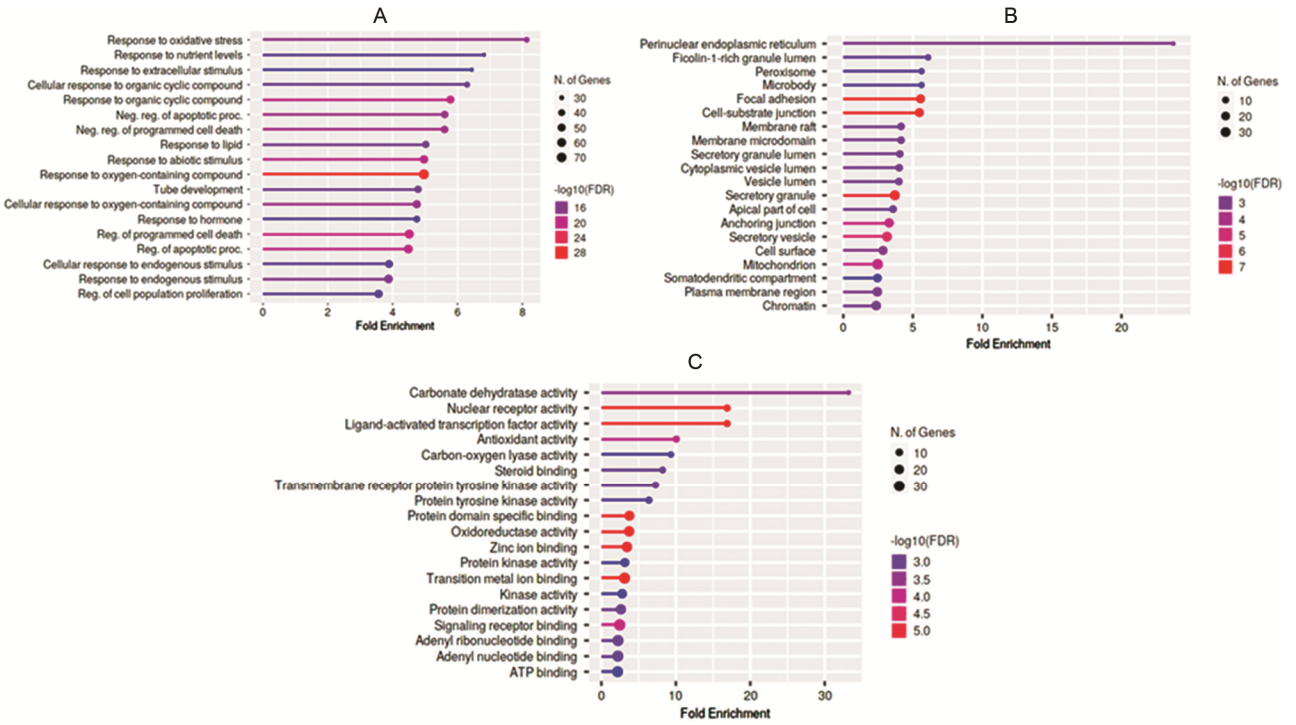


Fig. 8 — The gene enrichment analysis of the identified target proteins associated with top screened phytochemicals where (A) biological process; (B) cellular component; and (C) molecular function

Conclusion

Natural products are important sources of new therapeutics in modern day. The search for novel drugs with potential anti-cancer agents continues to be a challenge mainly due to the complications in the synthesis of natural products. Identifying potential medicinal plants with anti-cancer properties can be an alternative therapeutic use against the life-threatening disease[44,45]. In this study, phytochemicals documented in NEI-MPDB were screened against BRCA1 and BARD1 breast cancer targets. The synergistic effect of the top scoring phytochemicals was performed by constructing a network using Cytoscape. The phytochemicals with the top docking scores were identified and the target genes were predicted. Among the top scoring phytochemicals for BRCA1 and BARD1, a phytophenol maclurin known for its antioxidant activity was identified as a top scoring phytochemical for both targets. The back-mapping of the top 20 phytochemicals was done from NEI-MPDB to identify the medicinal plants and the therapeutic uses and it was observed that eighteen different plants were used in the treatment of thirty ailments including glandular tumors. An interaction network was constructed with the set of predicted target genes of the top 20 phytochemicals and breast cancer target genes were identified from the network. Eight phytochemicals namely Acteoside (*Barleria cristata*), Anthraquinones Emodin (*Stellaria media*), L-(+)-Quercitol (*Aloe barbadensis*), Maclurin (*Ricinus communis*), Myoinositol (*Dioscorea* sp.), Neoisoastilbin (*Goniothalamus sesquipedalis*), Okanin (*Urtica magellanica*), and Plantamajoside (*Coccinia grandis*) were identified as the compounds with the most number of breast cancer target genes. Therefore, based on the topology of the docking weighted phytochemical cancer target network, these compounds and their associated plants may be explored as potential therapeutic indications against breast cancer associated genes.

Conflicts-of-interest

The authors declare no conflict of interest.

Acknowledgement

KK thanks UGC for JRF Fellowship, EJ thanks DST for Inspire Fellowship. The Department of Biotechnology, New Delhi is thanked for the funding in the form of Centre of Excellence in Advanced Computation and Data Sciences (No. BT/PR40188/BTIS/137/27/2021).

References

1 Khan MSA & Ahmad I, Herbal medicine: current trends and future prospects. *In New look to phytomedicine*, Academic Press, (2019) 3.

2 Kade IJ, Johnson DO, Akpambang VO & Rocha JB, Polymerization of gallic acid enhances its antioxidant capacity: Implications for plant defence mechanisms. *Biokemistri*, 24 (2012)15.

3 Ma L, Zhang M, Zhao R, Wang D, Ma Y & Li A, Plant Natural Products: Promising Resources for Cancer Chemoprevention. *Mol. (Basel, Switzerland)*, 26 (2021) 933.

4 Atanasov AG, Zotchev SB & Dirsch VM, International Natural Product Sciences Taskforce, & Supuran, C. T. Natural products in drug discovery: advances and opportunities. *Nat. Rev. Drug Discov*, 20 (2021) 200.

5 Liu EH, Qi LW, Wu Q, Peng YB & Li P, Anticancer agents derived from natural products. *Mini-Rev Med Chem*, 9 (2009) 1547.

6 Cragg GM, Grothaus PG & Newman DJ, Impact of natural products on developing new anti-cancer agents. *Chem Rev*, 109 (2009) 3012.

7 Key TJ, Verkasalo PK & Banks E, Epidemiology of breast cancer. *Lancet Oncol*, 2(2001):133.

8 Kiewhuo K, Priyadarsinee L, Sarma H & Sastry GN, Molecular Dynamics Simulations on the Effect of Mutations in the RING Domains of BRCA1-BARD1 Complex and its Relevance to the Prognosis of Breast Cancer. *J Biomol Struct Dyn*, (2023) 1.

9 Sarma H, Kiewhuo K, Jamir E & Sastry GN, In silico investigation on the mutational analysis of BRCA1-BARD1 RING domains and its effect on nucleosome recognition and ubiquitination. *Biophys Chem*, 300 (2023) 107070.

10 Ruffner H, Joazeiro CA, Hemmati D, Hunter T & Verma IM, Cancer-predisposing mutations within the RING domain of BRCA1: loss of ubiquitin protein ligase activity and protection from radiation hypersensitivity. *Proc Natl Acad Sci India*, 98 (2001) 5134.

11 Li W, Witus SR, Wang M, Brzovic PS, Klevit RE & Zhao W, BRCA1-BARD1 ubiquitylates histones for genome maintenance. *Cancer Res*, 83 (2023) 6096.

12 Witus SR, Zhao W, Brzovic PS & Klevit RE, BRCA1/BARD1 is a nucleosome reader and writer. *Trends Biochem. Sci*, 47 (2022) 582.

13 Zhang Q, Liu J, Li R, Zhao R, Zhang M, Wei S, Ran D, Jin W & Wu C, A network pharmacology approach to investigate the anticancer mechanism and potential active ingredients of Rheum palmatum L. against lung cancer via induction of apoptosis. *Front Pharmacol*, 11 (2020) 528308.

14 Santh Rani T, Premitha Rajya Lakshmi P & Manga Devi C, Network pharmacology and molecular docking study of the active ingredients in Saptasaram kashayam for the treatment of Polycystic ovary syndrome. *Indian J Biochem Biophys*, 60 (2023) 108.

15 Kiewhuo K, Gogoi D, Mahanta HJ, Rawal RK, Das D & Sastry GN, North East India medicinal plants database (NEI-MPDB). *Comput Biol Chem*, 100 (2022) 107728.

16 Sussman JL, Lin D, Jiang J, Manning NO, Prilusky J, Ritter O & Abola EE, Protein Data Bank (PDB): database of three-dimensional structural information of biological macromolecules. *Acta Crystallogr D Biol Crystallogr*, 54 (1998) 1078.

17 Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE, UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem*, 25 (2004) 1605.

18 Tian W, Chen C, Lei X, Zhao J & Liang J, CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Res*, 46 (2018) W363.

19 Maestro S, Maestro. Schrödinger, LLC, New York, NY. 2020;2020.

20 Jamir E, Sarma H, Priyadarsinee L, Nagamani S, Kiewhuo K, Gaur AS, Rawal RK, Murugan NA, Subramanian V & Sastry GN, Applying polypharmacology approach for drug repurposing for SARS-CoV-2. *J Chem Sci*, 134 (2022) 1.

21 Jamir E, Kiewhuo K, Priyadarsinee L, Sarma H, Nagamani S & Sastry GN, Structure-function relationships among selected human coronaviruses. *Indian J Biochem Biophys*, 59 (2022) 536.

22 Choudhury C, Priyakumar UD & Sastry GN. Dynamic ligand-based pharmacophore modeling and virtual screening to identify mycobacterial cyclopropane synthase inhibitors. *J Chem Sci*, 128 (2016) 719.

23 Madugula SS, Nagamani S, Jamir E, Priyadarsinee L & Sastry GN, Drug repositioning for anti-tuberculosis drugs: an in silico polypharmacology approach. *Mol Divers*, 26 (2021)1.

24 Badrinarayan P & Sastry GN, Virtual screening filters for the design of type II p38 MAP kinase inhibitors: a fragment based library generation approach. *J Mol Graph Model*, 34 (2012) 89.

25 Srivastava HK, Choudhury C & Sastry GN, The efficacy of conceptual DFT descriptors and docking scores on the QSAR models of HIV protease inhibitors. *Med Chem*, 8 (2012) 811.

26 Bohari MH & Sastry GN, FDA approved drugs complexed to their targets: evaluating pose prediction accuracy of docking protocols. *J Mol Model*, 18 (2012) 4263.

27 Kumar N, Sarma H and Sastry GN, Repurposing of approved drug molecules for viral infectious diseases: a molecular modelling approach. *J Biomol Struct Dyn*, 22 (2021)1.

28 Belkadi A, Kenouche S, Melkemi N, Daoud I & Djebaili R, K-means clustering analysis, ADME/pharmacokinetic prediction, MEP, and molecular docking studies of potential cytotoxic agents. *Struct Chem*, 32 (2021) 2235.

29 Nagamani S, Gaur AS, Tanneeru K, Muneeswaran G, Madugula SS, Consortium MP, Druzhilovskiy D, Poroikov VV & Sastry GN, Molecular property diagnostic suite (MPDS): Development of disease-specific open source web portals for drug discovery. *SAR QSAR Environ Res*, 28 (2017) 913.

30 Shah U, Patel S, Patel M, Gandhi K & Patel A, Identification of chalcone derivatives as putative non-steroidal aromatase inhibitors potentially useful against breast cancer by molecular docking and ADME prediction. *Indian J Chem - B Org Med*, 59 (2020) 283.

31 Wishart DS, DrugBank and its relevance to pharmacogenomics. *Pharmacogenomics*, 9 (2008) 1155.

32 Corsello SM, Bittker JA, Liu Z, Gould J, McCarren P, Hirschman JE, Johnston SE, Vrcic A, Wong B, Khan M & Asiedu J, The Drug Repurposing Hub: a next-generation drug library and information resource. *Nat Med*, 23 (2017) 405.

33 Lagunin A, Ivanov S, Rudik A, Filimonov D, Poroikov V & DIGEP-Pred: web service for in silico prediction of drug-induced gene expression profiles based on structural formula. *Bioinform*, 29 (2013) 2062.

34 Safran M, Dalah I, Alexander J, Rosen N, Iny Stein T, Shmoish M, Nativ N, Bahir I, Doniger T, Krug H & Sirota-Madi A, GeneCards Version 3: the human gene integrator. *Database (Oxford)*, (2010) baq020.
35 Rappaport N, Twik M, Plaschkes I, Nudel R, Iny Stein T, Levitt J, Gershoni M, Morrey CP, Safran M & Lancet D, MalaCards: an amalgamated human disease compendium with diverse clinical and genetic annotation and structured search. *Nucleic Acids Res*, 45 (2017) D877.
36 Badrinarayan P, Choudhury C & Sastry GN. Molecular modeling. *Syst Synth Biol*, (Springer Publishing) 2015, 93.
37 Choudhury C, Priyakumar UD & Sastry GN, Dynamics based pharmacophore models for screening potential inhibitors of mycobacterial cyclopropane synthase. *J Chem Inf Model,* 55 (2015) 848.
38 Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B & Ideker T, Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, 13 (2003) 2498.
39 Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pagès F, Trajanoski Z & Galon J, ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinform*, 25 (2009)1091.
40 Ogata H, Goto S, Sato K, Fujibuchi W, Bono H & Kanehisa M, KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*, 27 (1999) 29.
41 Ge SX, Jung D & Yao R, ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinform*, 36 (2020) 2628.
42 Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C & Chanda SK, Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*, 10 (2019) 1.
43 Studio D, Discovery studio. *Accelrys* [2.1]. 2008.
44 Sathelly K, Kumar KN, Kiranmayi MU, Obul RC & Poda S, Anticancer potential of *Solanum lycopersicum* L. extract in human lung epithelial cancer cells A549. *Indian J Biochem Biophys*, 60 (2022) 76.
45 Selvam P, Vijayakumar T, Wadhwani A & Muthulakshmi L, Bioreduction of silver nanoparticles from aerial parts of *Euphorbia hirta* L. (EH-ET) and its potent anticancer activities against neuroblastoma cell lines. *Indian J Biochem Biophys*, 56 (2019) 132.