Research Article

Structure Based Virtual Screening of Druggable Phytochemicals from Indian Medicinal Plants against the Protein Targets Involved in Breast Cancer Metastatsis, and Cytotoxicity and Antioxidant Studies of Gmelina arborea Extracts

Jaswanth Jenny P[1*] and Dhamotharan R[2]

[1]*Department of Plant Biology and Plant Biotechnology, Presidency College, Chennai (TN), India;* [2]*Department of Botany, Government Arts College of Men (Autonomous), Chennai (TN), India*

Corresponding author:
Jaswanth Jenny P, Department of Plant Biology and Plant Biotechnology, Presidency College, Chennai (TN), India;
E-mail: Jaswanth.p@gmail.com

Abstract

In this study, druggable phytochemicals from IMPPAT database were subjected to docking-based virtual screening against identified targets (S100Z, NAALADL2, MYH15 and OR2AK2) in human breast metastatic cancer from our previous exome data analysis. The present study clearly showed that this approach positively screened out 4-Hydroxysesamin, isoarboreol, *Gummadiol* and 4,8-Dihydroxysesamin fromIMPPAT database depending on various parameters such Lipinski's rule of five (Ro5), oral PhysChem score, Glaxo Smith Kline's 4/100, Pfizers 3/75, Veber rule and Egan rule. Phytochemicals from *Gmelina arborea*, 4-Hydroxysesamin against S100Z, NAALADL2 and OR2AK2 showed high binding affinity with Pyrx binding energy of -6.8 Kcal/mol, -9.9 Kcal/mol and -9.1 Kcal/mol, respectively. 4-Hydroxysesamin and isoarboreol against MYH15 demonstrated high binding affinity with binding energy values of -7.8 Kcal/mol and -9.7 Kcal/mol, respectively. All the phytochemicals that showed high binding energy values were from the plant *Gmelina arborea*. Methanolic extracts of *Gmelina arborea* showed remarkable cytotoxicity and antioxidant activity in human metastatic breast cancer cell line, MDA-MB-231. Further studies are needed to explore the mechanism involved.

Introduction

Breast cancer is the most common malignancy across globe. WHO reports 2.2 million breast cancer cases in 2020, affecting around 1 in 12 women in their life-span. Metastasis stage of breast cancer is the main cause of mortality because of breast cancer. Unlike initial stage of breast cancer, metastatic breast cancer is widely considered irremediable. Chemotherapy is the most effective treatment form in practice that kills the cancerous cells in multiple locations by travelling through the bloodstream. A systemic treatment like chemotherapy often leads to short term or long-term adverse events. [1,2] Plant derived anti-cancer agents are being recognized as an alternative treatment to chemotherapy with prolific benefits.

Traditional herbal medicines have evolved over many years. Phytochemicals, the major bioactive constituents obtained from plant sources has proved its efficiency in inhibiting carcinogenic process at specific stages. [3,4] Various mechanisms are studied on the anti-carcinogenic properties of dietary phytochemicals. Synergistic effects of phytochemicals from fruits and vegetables were found to be associated with potential antioxidant and anti-cancer activities. [5] Studies have also shown that no single phytochemical is a substitute for the additive effects of phytochemicals in fruits and vegetables. [6] Drug discovery from medicinal plants have been fruitful in treating various diseases such as cancer, HIV/AIDS, alzheimer's, malaria and pain. [7] Various epidemiological studies have proven the beneficial effects of vegetables and fruits in reducing the risk of cancers.

India having a rich heritage in the use of plant-derived molecules as traditional medicines, scientists at many instances has tried to explore the science or molecular impact involved. Similarly, traditional medicine across the globe got to evolve, and gained huge interests in discovering therapeutic phytochemicals. Consolidations of data as phytochemical databases were much needed by the scientific community to gain more insights. This in turn resulted in the availability wide range of herbal or phytochemical databases as a public resource. Indian Medicinal Plants, Phytochemistry and Therapeutics (IMPPAT), Duke's phytochemistry and ethnobotanical database, NAPRALERT, Phytochemica and MAPS database are some of the databases that has information on 3D structure of the phytochemicals, ADMET properties, pharmacological activity, chemistry etc.

Computational approaches in identifying small molecules

How to Cite this Article: Jenny JP, et al. Structure Based Virtual Screening of Druggable Phytochemicals from Indian Medicinal Plants against the Protein Targets Involved in Breast Cancer Metastatsis, and Cytotoxicity and Antioxidant Studies of Gmelina arborea Extracts. Ann Med Health Sci Res. 2021;11:S3:126-137.

that effectively interacts with protein targets are now a days unavoidable in the discovery of new drugs. Structure based design and virtual screening of small molecules are extensively used in the identification of lead identification and optimization. [8] Application of virtual screening is slowly replacing the empirical screening in the discovery of new ligands. [9] This involves a computational methodology which helps in reducing huge chemical libraries to handful of potentially interacting small molecules. [10] Various tools like autodock, dock, gold, pyrx are widely used for proteinbased virtual screening of chemical compounds. [11] The process of selection of small molecules involves identification of a chemical library, prediction of active sites in the protein target of interests, docking of each small molecule with the target and analysing the interaction based on the binding affinity scores. Scoring algorithms that reveals the strength of the interactions are vital in the docking process. [12,13]

Constituents from plants play a crucial role in cancer therapy today with myriad benefits. Drug discovery from medicinal plants have been fruitful in treating various diseases such as cancer, HIV/AIDS, alzheimer's, malaria and pain. [7] Various epidemiological studies have proven the beneficial effects of vegetables and fruits in reducing the risk of cancers. Globally the mortality rate caused by cancer has the increased the search for effective anticancer agents to fight cancer. Drug discovery from plants have resulted in identification of several lead molecules against pharmacological targets. There are various drugs from plant origin currently in clinical use such as paclitaxel, camptothec in derived analogues, galanthamine, triotropium and more drugs in Phase II and Phase III in clinical trials. [14] Natural products play a relevant role in cancer therapy today with substantial numbers of anticancer agents used in the clinic being either natural or derived from natural products from various sources such as plants, animals and microorganisms (also of marine origin). Pharmacological and phytochemical characterization contributes to both the safe use of herbal medicines and the identification of leads for drug development. Worldwide herbal medicine market is expected to reach around US$ 60 billion. [15] Pharmaceutical industries have started investing millions of dollars on research and development of traditional herbal medicines. The surge in the use of herbal medicines is mainly for its non-toxic properties.

Cytotoxic agents as targeted therapies are in the mainstay of treatment of patients affected with metastatic breast cancer. Herceptin (trastuzumab) targeting HER2 is one of the successful immune targeted therapies in practice. Several concepts have evolved in the identification of effective targets for the treatment of cancer. Exome sequencing is one of the promising techniques in the recent years that have proven records in the identification of candidate genes for specific diseases or phenotypes. It is a genomic technique coupled with multidisciplinary interpretation that efficiently detects secondary highly clinically significant mutations in cancer susceptibility genes without a family history of disease. [16] Mutations and its impact on the protein structure has revealed various insights in the identification protein targets for diseases. In our previous analysis of exome data from paired primary and metastatic breast cases led to identification of causative mutations in S100Z, NAALADL2, MYH15 and OR2AK2 genes that also has an impact on the protein structure. [1]

S100Z is dimeric, predominantly alpha helical proteins that have demonstrated its ability to bind to calcium ions. [17,18] S100Z was found to be involved in different stages of breast cancer development. [19] Abnormal expression of S100Z was previously found to be associated with gastric cancer and oral squamous cell carcinoma in two different studies, and S100Z rs7712957 significantly associated with ulcerative colitis in a Dutch study. [20-22]

N-acetyl-L-aspartyl-L-glutamate-peptidase-like 2 (NAALADL2) is a member of the glutamate carboxy peptidase II family that was first identified in a patient with mild mental retardation. [23,24] Pathogenic variants in NAALADL2 loci was found to be associated with breast cancer risk. [25] Significant association of NAALADL2 rs1870740 with Kawasaki disease was detected in a genome wide association study, whereas no NAALADL2 mutation was detected in Cornelia de Lange syndrome. [26,27] Two other genome-wide association studies demonstrated significant association of rs78943174 at 3q26.31 (NAALADL2) with prostate cancer aggressiveness, and NAALADL2 (rs3914502 and rs2222447) with autism spectrum disorder. [28,29]

Myosin heavy chain 15 (MYH15) protein is detected widely in skeletal muscle, more specifically in fibres of the orbital layer of extraocular muscles and in the extracapsular region of bag fibres. [30,31] A study in chicken breeds demonstrates involvement of MYH15 in the regulation of muscle growth. [32] MYH15 rs3900940 associated with coronary heart disease and coronary microvascular dysfunction, whereas MYH15 rs9288876, rs7635009 and rs1454197 were associated with asthma. [33-36] MYH15 was also identified as a causative gene for carotid paragangliomas. [37]

OR2AK2 belongs to the olfactory receptor gene family, which constitute the largest gene family of the mammalian genomes. [38] Nominal association of OR2AK2 mutations with obesity and diabetic and non-diabetic nephropathy has been observed previously. [39,40] No direct association of OR2AK2 with oncolytic diseases was published till date.

In this study, phytochemicals that has the properties of therapeutic drugs were selected from IMPPAT database and is explored for interaction against the protein targets, S100Z, NAALADL2, MYH15 and OR2AK2. Interestingly, phytochemicals from *Gmelinaarborea* demonstrated best interaction with the studied protein targets. Crude methanolic extract of *Gmelinaarborea* exhibited significant inhibition of tumour cell growth and radical scavenging activity in human breast cancer metastatic cell line.

Materials and Methods

IMPPAT

Indian Medicinal Plants Phytochemistry and Therapeutics (IMPPAT) is a manually curated phytochemicals library that comprises 9596 phytochemicals. Physiochemical properties of the phytochemicals along with the Absorption Distribution Metabolism Excretion Toxicity (ADMET) and drug likeliness properties calculated using cheminformatics approach is also presented in the database. ADMET properties were predicted using admetSAR open source tool. Drug likeliness properties of the phytochemicals were based on the various scoring systems

such as Lipinski's rule of five (Ro5), oral PhysChem score, Glaxo Smith Kline's: 4/100, Pfizers: 3/75, Veber rule and Egan rule.

CASTp

Computed Atlas of Surface Topography of proteins (CASTp) developed by University of Minnesota is one of the widely used tool to predict the active site of target molecules. CASTp finds the probable active sites of proteins and pockets based on the alpha shape and a specific algorithm to find the pockets on the surface of the protein. It also identifies the void region that is buried vacuum inside the proteins that cannot interact with water molecules since the hetero atoms are removed. CASTp provides a detailed delineation of the atoms that participates in the interactions. PDB coordinate files of the proteins should be submitted. CASTp is also mapped to the annotations from Protein Data Bank (PDB), Swiss-Prot, Online Mendelian Inheritance in Man (OMIM), there by the function of the active sites can also be explored. [41,42]

Prankweb

Prankweb is a web based server that incorporates a machine learning based method to identify the binding sites of the protein structure. It uses a P2 rank, a template free machine learning method based on the identification of nearest potential ligandabilitycentre placed in a solvent accessible protein surface. Ligand bind sites were predicted based on the combination of amino acids with high ligandability scores. [43]

Metapocket

Metapocket is also a web based server that is used to identify the appropriate binding sites of target molecules. This algorithm finds the potential probable binding pockets on a protein surface based on the consensus of results from LIGESITE(cs), PASS, Q-SiteFinder and SURFnet. Metapocket calculates z-score by comparing the ranking scores predicted by the four algorithms. Twelve sites *i.e,* the top three pockets from each algorithm are clustered using a simple hierarchical clustering algorithm based on their spatial similarity. This is followed by ranking each cluster by metaz-score, which is the sum of the z-scores of the identified pocket sites in a cluster (http://metapocket.eml.org). [44]

Pfam

Domain analysis of the target proteins were performed by comparing against the Pfam data warehouse of protein families, which is classified into manually curated Pfam-Aand automatically generated Pfam-B families. [45] Each family consists of two multiple sequence alignments and two profile hidden Markov models. [46] Pfam 32.0 (September 2018 release) holding 17929 entries was used in this study. As the number of families in Pfam started growing, protein families were further grouped based on the common evolutionary origin termed as Clans. [47]

Scanprosite

Scanprosite is a well-established web based tool developed using scripting language perl running on a unix operating system that presents in the results in a HTML page. Scanprosite identify the matching amino acid sequences against PROSITE database, which holds a humongous collection of biologically meaningful signatures. amino acid signatures in PROSITE database are described as patterns, short motifs and generalized profiles of larger domains. All the signatures available are based on manually curated annotation template called prorule, an internal rule followed by Swissprot group. [48]

Pyrx

Structure based virtual screening of ligands against protein targets are effectively used to drug discovery to identify the most potential binding ligand from a library. Pyrx is an open source tool developed using Python language that is compatible with windows, linux and mac operating systems. This docking algorithm keeps the target molecule static and the ligand molecule flexible. Steps involve loading of macromolecule and the library of the molecules, preparation of the molecules to docking through conversion of PDB coordinate files to PDBQT format, pyrx uses autodockvina is used to dock the molecules, active sites are marked in the macromolecule, make a grid around the active sites to do effective interaction followed by docking. Pyrx provides 9 conformations of the ligand with the top binding affinity scores. [49]

PyMol

PyMol is a molecular visualization tool being maintained by Schrodinger. It is known by the scientists worldwide for its high quality, speed and versatility. PyMol allows the user to upload more than one coordinate files and visualize 3D model, and each object can be controlled separately. This visualization tool is also supported by a command line window, which helps the user to control the display using scripts. It also helps in visualizing the different conformations of single structure mostly derived from docking or interaction studies, professional graphics in both windows and unix and produces quality images. [50]

Ligplot

Ligplot is a bio-software that automatically generates 2D images of proteins interaction with ligand. PDB format coordinate files are compatible with Ligplot. It also provides the strengths, hydrogen bonds, hydrophobic interactions and atom accessibilities. This tool is widely used in plotting multiple conformations of ligands binding to the protein target. [51]

Collection and identification of plant material

The leaves of Gmelina arborea is collected from south India, Kanya Kumari district during the month of January and February. The plants were identified and authenticated by a plant taxonomist.

Extract preparation

The freshly collected leaves were cleaned with water and dried in the shade, then coarsely powdered using a mortar and pestle. The powdered leaf was measured for 250 grams and taken separately. Soxhlet extraction used in a successive manner using the solvent water, methanol and petroleum ether. Soxhlet extraction is a hot continuous process of extraction that requires minimum amount of solvent for extraction. The extract was then filtered through Whatman no 1 filter paper. The filtrates obtained were dried at a temperature of 40±2 to haven gummy

concentrate of the crude extracts. The yield of extract using every solvent was calculated using the below formula.

Percentage of Yield=Weight of the dry extract ×100

Weight of the dried leaf powder and the obtained extract was protected against light and humidity.

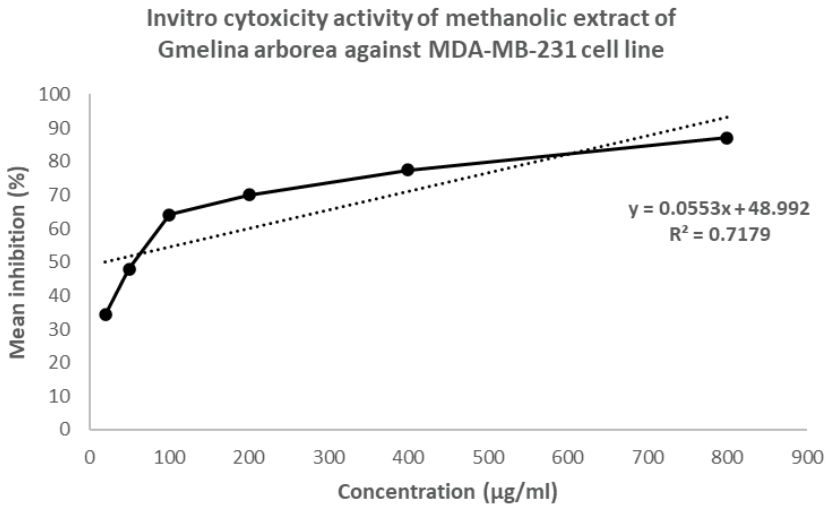
Cell lines

MDA-MB-231 (human breast carcinoma, ER, tumorigenic, and invasive) cell lines were obtained from the National Centre for Cell Science (NCCS), Pune, India, and as such, were cultured in Dulbecco's Modified Eagle's Medium (DMEM), supplemented with 10% FCS and pencillin (100 units/ml), streptomycin (100 mg/ml) in a humified incubator at 37°C. The cells were grown until 80%-90% confluence is reached. Confluency is measured based on 80% surface of the culture vessel is covered with cells. A single scratch was made in the cell monolayer using a 100 µl pipette tip and the image was captured using a light microscope.

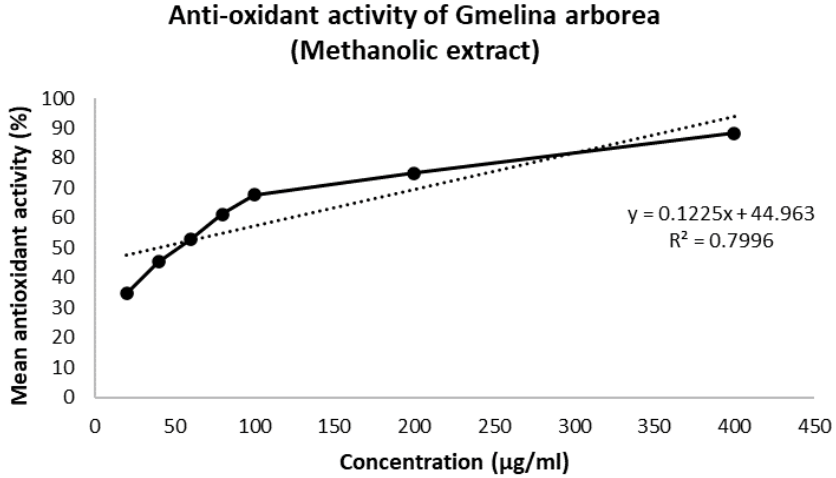
Cytotoxicity assay

Cytotoxicity assay is the most widely utilized method to measure the potential of a test material against a tumor cell population. [52] The process involves exposure of the tumor cells to test material, incubated for a period of time and the viable cells are analyzed by measuring a marker. In this study, MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) assay developed by Mossman was used to measure the efficacy of different concentrations of Gmelina arborea in human breast carcinoma MDA-MB-231 cell lines. [53] Water soluble dye MTT is converted into an insoluble purple formazan by mitochondrial reductase followed by the solubilization of formazan and analysis of the concentration by 540 nm optical density. Cultured cells were plated in 96 well plates at a concentration of 100 µL/well and subjected to incubation overnight thereby allowing the cells to attach. The diluted extracts of Gmelinaarborea were added to each wells with different serial concentrations (20 µg/ml, 50 µg/ml, 100 µg/ml, 200 µg/ml, 400 µg/ml, and 800 µg/ml). After 72 hours of incubation at 37°C in a 5% CO_2 incubator, MTT solution was added to each well and further incubated for 4 hours followed by the addition of 100 µL of DMSO to all the wells. DMSO solubilizes formazan crystals, and the optical density was measured at a wavelength of 540 nm. Cytotoxicity was measured by the concentration of *Gmelina arborea* extract causing 50% growth inhibition of the tumor cells and a dose dependent. Graphs 1 and 2 was plotted to determine the IC50 value. The image of the cells before and after treatment were recorded by inverted microscope attached to a camera system. IC50 value was calculated from the equation y=mx+C.

$$\%\ of\ reduction = \frac{Absorbance\ control - Absorbance\ test\ sample}{Absorbance\ Control} \times 100$$



Graph 1: *In vitro* cytoxicity activity of menthanolic extract of *Gmelina arborea* against MDA-MB-231 cell line.



Graph 2: Antioxidant activity of *Gmelina arborea*.

DPPH (2,2-Diphenyl-1-Picryl-Hydrazyl-Hydrate) assay

DPPH assay, the free radical scavenging is a widely accepted technique to evaluate the antioxidant activity of plant extracts. It is considered as an effective, accurate, easy and cost effective method to estimate the free radical scavenging activity of antioxidants. The mechanism involves reduction of odd electron of nitrogen atom in DPPH by accepting a hydrogen atom from the antioxidants and convert DPPH into 1-1,Diphenyl-2-Picrylhydrazyl. This reaction causes discoloration of purple coloured DPPH to yellow pale colour demonstrating the scavenging activity. The absorbance was measured at 517 nm, and lower absorbance specifies a higher radical scavenging activity. Antioxidant activity of methanolic extract of *Gmelina arborea* was estimated. Antioxidant L-ascorbic acid was used as a positive control.

Statistical analysis

Results were collected, analysed, and mentioned as averages and standard deviation from three independent experiments performed. The data are expressed as mean ± standard deviation. IC50 value was calculated using linear regression analysis.

Results

Selection of therapeutically eligible phytochemicals

Virtual screening plays a key role in the evolution of drug discovery. IMPPAT database was found to contain 960 potential druggable phytochemicals from Indian medicinal plants, out of which 29 phytochemicals were selected in this study [Table 1]. These 29 phytochemicals has 'zero violation' of Lipinski's role of five, '0' Oral PhysChem score, Quantitative Estimate of Drug likeness (QED) between 0.75 to 1, along with best categorization of Veber rule, Egan rule, GSK 4/400 and Pfizer 3/75. [54]

Active site prediction

Active sites in protein structures S100Z, NAALADL2, MYH15 and OR2AK2 were predicted using CASTp, Prankweb and Metapocket servers. CASTp and Metapocket detected 9 common amino acids active sites in S100Z. PDB sum and ScanProsite shows the interaction between S100Z protein and calcium ions, but these calcium binding amino acids were not matching with the list of selected amino acids. Twelve amino acids (338, 363, 364, 367, 373, 375, 408, 409, 410, 425, 427 and 428) in the M28 domain of NAALAD2 were selected as active site residues. Active site and functional analysis of MYH15 suggested 'GESGAGKT' (Pos. 196 to 203) as a nucleotide binding site, and the fragment was found to be present in one of the cavity. OR2AK2 active site analysis showed 49 common residues by the three servers used. Functional analysis using Pfam and ScanProsite did not reveal any active sites for OR2AK2, and literature search also confirmed that active sites for OR2AK2 has not been explored before.

Interaction studies

In molecular docking process, validation of a docking protocol

Table 1: Phytochemicals selected for interaction studies against breast cancer metastatic candidates.

Physiochemical Name	Plant	IMPPAT Identifier
(+)-arborone	*Gmelina arborea*	CASID:108069-03-0
1,5-dihydroxy-3,8-dimethoxyxanthone	*Swertiachirayita*	CASID:134779-25-2
O-desmethyl-N-formyl-desacetyl-beta-lumicolchicine	*Gloriosasuperba*	CASID:18172-25-3
Eupatin	*Mucunapruriens/ Pterocarpusmarsupium*	CASID:19587-65-6
Gummadiol	*Gmelina arborea*	CHEMSPIDER:10308017
Cornigerine	*Gloriosasuperba*	CID:100188
3-O-Demethyl-N-formyldesacetyl-beta-lumicolchicine	*Gloriosasuperba*	CID:101609250
Syringicacid	*Catharanthusroseus*	CID:10742
Thiamine	*Crocus sativus*	CID:1130
Carpusin	*Pterocarpusmarsupium*	CID:134369
3-Acetyl-7-O-methylaromadendrin	*Eucalyptus citriodora*	CID:15139424
Tinosponone	*Tinosporacordifolia*	CID:15215479
4-Hydroxysesamin	*Gmelina arborea*	CID:16745513
Pinoresinol	*Valerianajatamansi*	CID:17750970
Colchiceinamid	*Gloriosasuperba*	CID:18397
Isoarboreol	*Gmelina arborea*	CID:21722929
2-Demethylcolchicine	*Gloriosasuperba*	CID:23757
Gloriosine	*Gloriosasuperba*	CID:23890
Paulownin	*Gmelina arborea*	CID:3084131
Medioresinol	*Solanumnigrum*	CID:332425
3,5-Dihydroxy-7,3',4',5'-tetramethoxyflavone	*Artemisia annua*	CID:44259639
biochanin A	*Oroxylumindicum*	CID:5280373
1,8-Dihydroxy-2,6-dimethoxy-9H-xanthen-9-one	*Swertiachirayita*	CID:5281653
Swerchirin	*Swertiachirayita*	CID:5281660
Folerogenin	*Eucalyptus citriodora*	CID:5319509
Skullcapflavone I	*Andrographispaniculata*	CID:5320399
Colchicine	*Gloriosasuperba*	CID:6167
Sinapic acid	*Catharanthusroseus*	CID:637775
4,8-Dihydroxysesamin	*Gmelina arborea*	CID:73067510

is a critical step in order to make sure the potential ligands bind within the targets cavity in the correct conformation. Docking in this study was performed using Pyrx, interactions were visualized using PyMol and LIGPLOT was used to plot the 2D representations of the protein ligand complexes. S100Z structure was retrieved from Protein Data Bank (PDB), whereas modelled structures of NAALADL2, MYH15 and OR2AK2 were used from our previous study. [1] Active site analysis of MYH15 and OR2AK2 suggested 42 aminoacids and 49 aminoacids, respectively. Since, the numbers predicted were high, blind docking was performed for MYH15 and OR2AK2, and interestingly the amino acids predicted by the servers exhibited H-bond interaction with the phytochemicals. The docking process generated ten conformers for each phytochemical with targets studied, and interactions exhibiting high binding affinities were tabulated. 4-hydroxysesamin, isoarboreal, gummadiol and 4,8-dihydroxysesamin from *Gmelina arborea* exhibited highest binding affinity towards the candidate targets studied [Tables 2-5].

The compound 4-hydroxysesamin has-6.8 Kcal/mol binding energy against S100Z, where ARG 13, SER 19 and ARG 25 amino acid residues exhibited H-bond interaction. Similarly, 4-hydroxysesamin exhibited binding energy values of -9.9 Kcal/mol and -9.1 Kcal/mol with NAALADL2 and OR2AK2 through the formation of H-bond interaction with SER 364, THR 107, and THR 276, respectively. MYH15 demonstrated high binding affinity towards 4-Hydroxysesamin and isoarboreol in focused and blind docking approach with binding energy values of -7.8 Kcal/mol and -9.7 Kcal/mol, respectively. Interestingly, all the phytochemicals that showed high binding energy values were from the plant *Gmelina arborea.*

Collection of plant and extraction

The leaves of *Gmelina arborea*were collected in a moderate

Table 2: Results of virtual screening performed from IMPPAT database against breast cancer metastatic candidates.

Target	Ligand	Binding affinity	H Bonds	Phytochemical name	Plant source
S100Z	CID: 16745513	-6.8	ARG 13, SER 19, ARG 25	4-Hydroxysesamin	*Gmelina arborea*
NAALADL2	CID: 16745513	-9.9	SER 364	4-Hydroxysesamin	*Gmelina arborea*
	CID: 21722929	-9.8	GLN 425	Isoarboreol	*Gmelina arborea*
MYH15	CID: 10308017	-7.8	GLY 96, LYS 97	Gummadiol	*Gmelina arborea*
	CID: 73067510	-7.8	GLU 92, SER 93, GLY 96	4,8-Dihydroxysesamin	*Gmelina arborea*
	CID: 10308017	-7.7	GLY 96	Gummadiol	*Gmelina arborea*
	CID: 21722929	-7.7	GLY 96	Isoarboreol	*Gmelina arborea*
MYH15*	CID: 21722929	-9.7	LYS 178, SER 179, HIS 558	Isoarboreol	*Gmelina arborea*
	CID: 21722929	-9.7	SER 149, GLU 177, THR 370	Isoarboreol	*Gmelina arborea*
	CID: 73067510	-9.6	LEU 176, GLU 381, HIS 558	4,8-Dihydroxysesamin	*Gmelina arborea*
	CID: 73067510	-9.6	SER149, ARG 151, LEU 176, THR 370	4,8-Dihydroxysesamin	*Gmelina arborea*
	CID: 73067510	-9.6	SER149, ARG 151, GLU 177, ARG 180	Isoarboreol	*Gmelina arborea*
	CID: 10308017	-9.5	ARG 151, LEU 176, LYS 178	Gummadiol	*Gmelina arborea*
	CID: 73067510	-9.5	LYS 178, HIS 558	4,8-Dihydroxysesamin	*Gmelina arborea*
OR2AK2*	CID: 16745513	-9.1	THR 107, THR 276	4-Hydroxysesamin	*Gmelina arborea*
	CID: 10308017	-8.9	SER 148	Gummadiol	*Gmelina arborea*
	CID: 10308017	-8.8	ARG 169, TYR 192	Gummadiol	*Gmelina arborea*
	CID: 134369	-8.7	ARG 169	carpusin	*Pterocarpusmarsupium*
	CID: 17750970	-8.7	VAL 185, THR 276	Pinoresinol	*Valerianajatamansi*
	CID: 23890	-8.7	SER 148	Gloriosine	*Gloriosasuperba*

*Blind docking / Bold indicates predicted active sites

Table 3: Total yield obtained from different extracts of *Gmelina arborea.*

S. No	Extracts	Yield (gms)	% of yield (w/w)
1	Aqueous	25.2	10.08
2	Methanol	18.4	7.36
3	Petroleum ether	37.5	9.38

Table 4: Percentage inhibition of MDA-MB-231 cell line treated with different concentrations of aqueous, methanol and petroleum ether extracts of *Gmelina arborea.*

Concentration µg/ml	Aqueous extract ± SEM	Methanol extract ± SEM	Petroleum ether extract ± SEM
20	30.52 ± 0.84	34.33 ± 0.52	29.88 ± 0.55
50	38.62 ± 0.57	47.94 ± 0.43	36.57 ± 0.31
100	65.1 ± 0.65	64.07 ± 0.19	55.06 ± 0.73
200	68.98 ± 0.78	69.96 ± 0.61	64.02 ± 0.21
400	77.19 ± 0.22	77.43 ± 1.05	68.91 ± 0.46
800	87.56 ± 0.51	87.02 ± 1.28	89.27 ± 0.59
IC50	79.4 ± 3.72	18.25 ± 1.06	153.51 ± 1.75

Table 5: Percentage inhibition of cell line.

Concentration µg/ml	*Gmelina arborea* (Methanolic extract)	Ascorbic acid (Standard)
20	34.68 ± 0.44	42.15 ± 0.53
40	45.26 ± 0.61	51.08 ± 0.20
60	52.71 ± 0.55	55.59 ± 0.49
80	61.32 ± 0.87	66.42 ± 0.92
100	67.72 ± 0.57	71.53 ± 0.55
200	75.02 ± 0.41	78.73 ± 0.50
400	88.29 ± 1.02	95.84 ± 0.88
IC50	41.14 ± 2.2	2.92 ± 1.19

weather condition and subjected to soxhlet extraction using aqueous, methanol and petroleum ether solvents. The yield and yield percentage of lyophilized aqueous, methanol and petroleum ether *Gmelina arborea* leaves extracts were recorded. The yield obtained from *Gmelina arborea* leaves using water, methanol and petroleum ether was 25.2 g, 18.4 g and 37.5 g respectively.

Effects of *Gmelina arborea* extracts on cytotoxicity and proliferation

The cytotoxicity of aqueous, methanolic and petroleum ether extracts from *Gmelinaarborea*were initially screened on human breast carcinoma MDA-MB-231 cell line. Cells were placed in a 96 well plates. After 24 hrs, cells were introduced with different concentrations (20 µg/ml, 50 µg/ml, 100 µg/ml, 200 µg/ml, 400 µg/ml and 800 µg/ml) of *Gmelina arborea* extracts and then viable cells were measured using MTT assay followed by calculation of inhibitory concentration (IC50) value. Among the extracts evaluated, methanol extract of *Gmelina arborea*exhibited maximum inhibition with an IC50 value of 18.25 ± 1.06 in human breast carcinoma MDA-MB-231 cell line, followed by water and petroleum ether extracts at 79.4 ± 3.72 and 153.51 ± 1.75, respectively. All the extracts studied inhibited cell viability in a dose dependent manner. *Gmelina arborea* extract at 20 µg/ml to 100 µg/ml concentration did not demonstrate significant inhibitory effects, whereas extracts at 200 µg/ml to 800 µg/ml exhibited notable decreased cell viability. In all the three times tested, MTT assay of different concentrations of the methanolic extract exhibited highly cytotoxic activity against MDA-MB-231 cells, and aqueous and petroleum ether extracts showed moderate cytotoxic activity with reference to the NCI criteria and Geran protocol (extracts with IC50 ≤ 20 µg/mL=highly cytotoxic, IC50 ranged between 21 and 200 µg/mL=moderately cytotoxic, IC50 ranged between 201 µg/mL and 500 µg/mL=weakly cytotoxic, and IC50>501 µg/mL=no cytotoxicity). [55]

In vitro antioxidant assay

The stable organic free radical DPPH exhibits absorbance band between 515 nm to 528 nm. The DPPH radical scavenging activity (IC50) of the methanolic extract of *Gmelina arborea* at different concentrations were evaluated. Absorbance value was tabulated for the sample and the standard. Methanolic extract of DPPH was capable to neutralize the DPPH free radicals through hydrogen donating activity. The IC50 values for scavenging of free radicals were 41.14 ± 2.2 and 2.92 ± 1.19 for methanolic extract of *Gmelina arborea* and ascorbic acid, respectively. The results indicate DPPH scavenging efficiency of *Gmelina arborea.*

Discussion

Exploration of phytochemicals from plants with anti-cancer activities have recently received huge attention because of their probability of usage as functional foods and for being an important source of therapeutic medicines. A plethora of research is happening on the medicinal properties of herbs across the globe. This study focussed on the identification of potential phytochemicals that could be used to inhibit the expression of mutated targets, S100Z, NAALADL2, MYH15 and OR2AK2 identified through exome data analysis previously. Active site prediction of the targets followed by virtual screening of phytochemicals targeting the targets of interest, and cytotoxicity and antioxidant studies of the plant extract with multiple solvents were performed in this study.

Computational approach in the field of drug discovery, has not only expedited the process but also have created huge impact in the funds being spend. In silico active site prediction tools such as CASTp, PrankWeb and Metapocket servers were used in this study to identify most potential active sites. Active sites for interaction studies were selected based on the consensus of results from all three servers. Visualization of the active sites by PyMol confirmed the involvement of selected active sites forming a groove or pocket to accommodate the ligand molecule. IMPPAT database is utilized for the selection of druggable compounds because of the scientifically curated content on phytochemicals from Indian herbs. The most frequently used virtual screening software for computational drug discovery, PyRx is used in this study.

Collectively, 7 phytochemicals were found to have effective interaction with the targets studied. Out of these phytochemicals, Gummadiol (5), Isoarboreal (5), 4,8-dihydroxysesamin (4) and 4-hydroxysesamin (3) were the most common. Interestingly, all these phytochemicals were from the same source, *Gmelina arborea*. 4-hydroxysesamin against S100Z, NAALADL2 and OR2AK2 showed high binding affinity with Pyrxbinding energy of -6.8 Kcal/mol, -9.9 Kcal/mol and -9.1 Kcal/mol, respectively. Literature studies confirms the potential medicinal uses of *Gmelina arborea* which possess lots of therapeutic components. [56] *Gmelina arborea* commonly known as gamhar is a deciduous large sized bush which grows about 4 m to 8 m. [57] Arborea is native of north-eastern India, comprising Assam, Arunachal Pradesh and Mizoram, and is widely distributed across South Asia, Pakistan, Bangladesh, China, Japan, Myanmar, Nepal, Pakistan, Sri Lanka, Thailand and India. [58,59] A study exploring the biophysical performance of Arborea in Jharkhand, Indiastates that the plant canopy characteristics of Arborea was found to be 10.2 m ± 0.33 m tree height, 70.3 cm ± 4.17 cm trunk circumference at breast height, 2.5 m ± 0.07 m, and also the foliar characteristics include the leaf area of 170.80 cm^2 ± 4.03 cm^2 and leaf area index of 0.53 ± 0.03. [60] *arborea* G has its presence in the ayurvedic pharmacopoeia of India, and states it has medicinal properties that can be used against fever, breathing disorders and swelling. [61]

4-hydoxysesamin is one of the lignans that are isolated from the

heartwoods of Arborea [62] Pelter et al. solved the proposed the structure of 4-hydroxysesamin through HNMR spectra. [63] QEDw score of 4,8-dihydoxysesamin is 0.868 with good solubility and '0' violations of Lipinski's rule. This phytochemical is also found to be a biological constituent of *Antrodiacinnamomea*, which is also referred as Taiwan mushroom and *Cinnamomumcamphora* commonly called as Karpura [Figures 1-3]. [64,65]

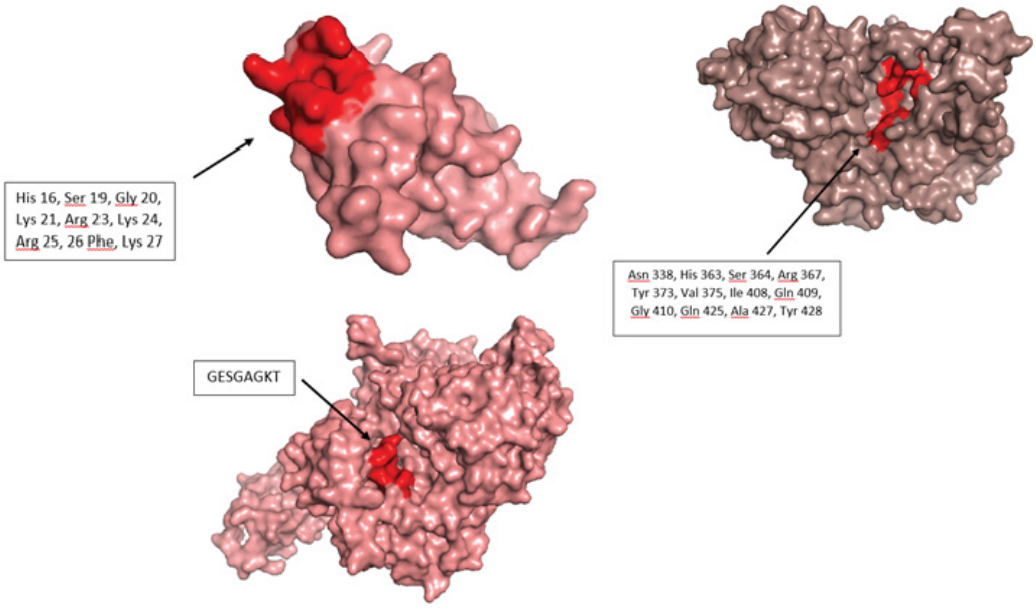


Figure 1: Visualization of active sites using PyMol.

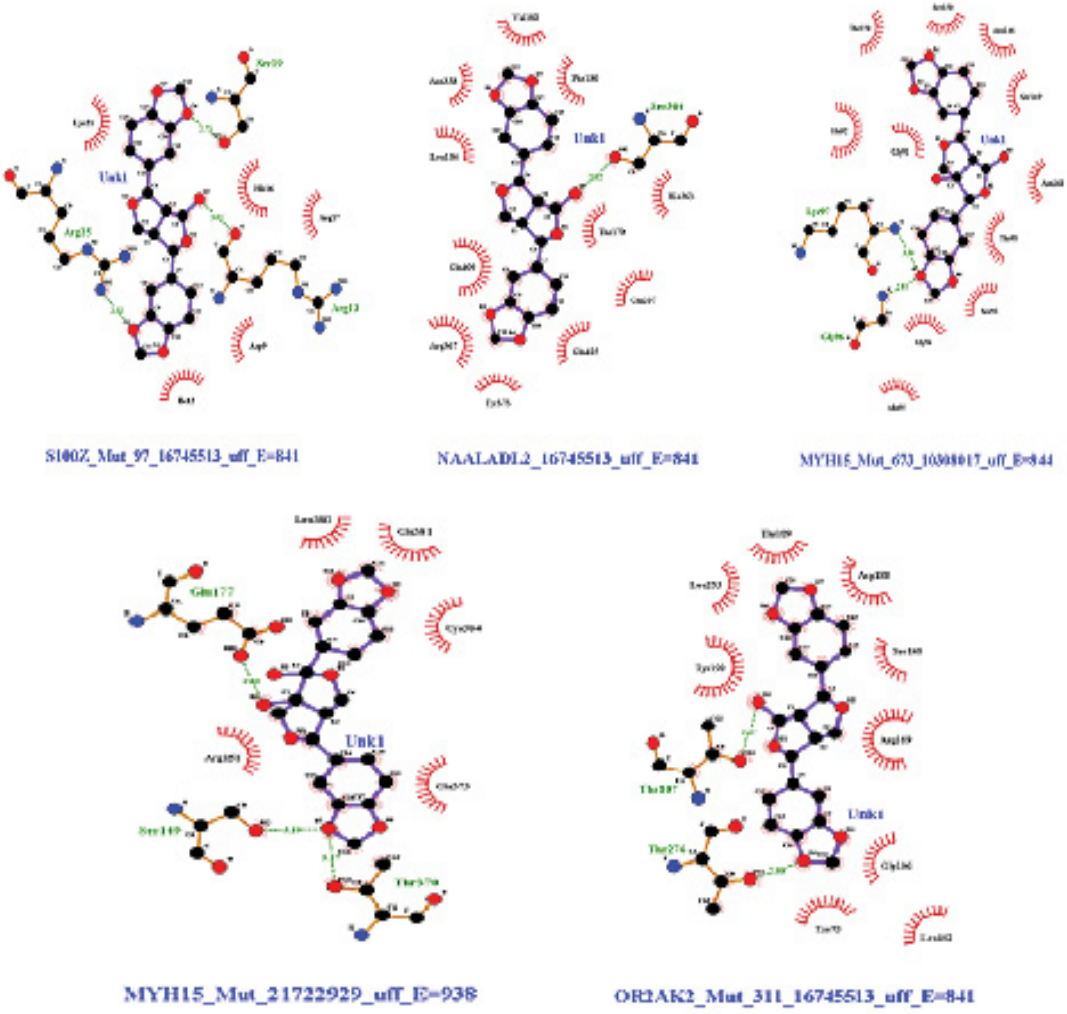


Figure 2: Top five interactions with phytochemicals.

US20170189417A1 patent mentions that 4-hydroxysesamin could be used as an effective treatment for cancer. [66] Cytotoxicity studies on 4-hydroxysesamin in human medullablastoma, MCF-7 and HeLa cell lines demonstrated to be inactive [Figure 4]. [67]

Gummadiol (1,4-dihyroxy-2,6-diperonyl-3,7-dioxabicyclo-(3,3,0)-octane) was the first member identified from *Gmelina arborea*. [68] IMPPAT states the QEDw score of gummadiol as 0.792 with good solubility and '0' violations of Lipinski's rule. ADMET properties predicted using admetSAR shows that Gummadial is a non-carcinogen [Figure 5].

Isoarboreol is a stereoisomer of arboreol based on ^{1}H NMR and ^{13}C NMR spectra that is found rich in the roots of *Arborea* . [69,70] This phytochemical belongs to the Furnoidlignans class. QEDw score of isoarboreol is 0.785 with good solubility and 0 violations of Lipinski's rule. Isoarboreol is also isolated from heartwoods of *Arborea* [Figure 6]. [71]

4,8-dihyroxysesamin was isolated in lesser quantities in the purification process of gummadiol from the heartwood of *Gmelina arborea*. [62] Vitexnegundo seeds also possess 4,8-dihyroxysesamin. [72] Pelter et al. have explored the synthesis of 4,8-dihyroxysesamin. [73] QEDw score of 4,8-dihydoxysesamin is 0.804 with good solubility and '0' violations of Lipinski's rule.

Chemical constituents of *Gmelina arborea* includes lignans,

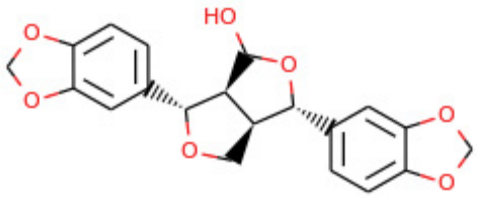


Figure 3: Structure of 4-Hydroxysesamin.

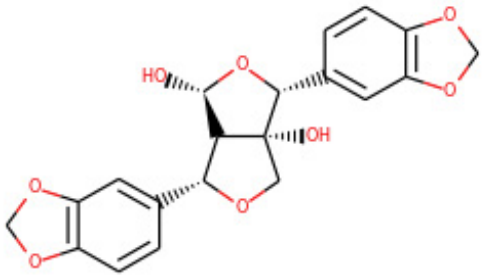


Figure 4: Structure of Gummadiol.

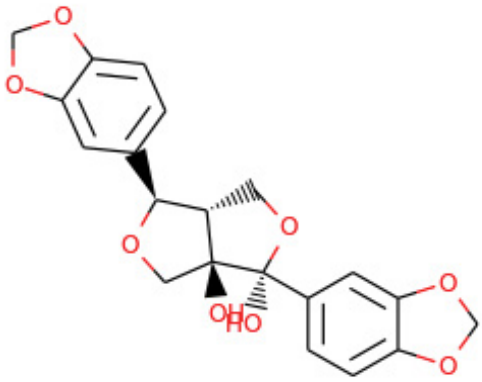


Figure 5: Structure of isoarboreol.

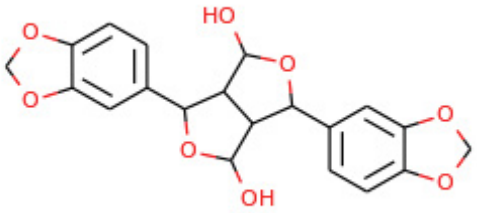


Figure 6: Structure of 4, 8-dihydoxysesamin.

iridoid glycoside, flavonoids, flavons, flavone glycoside and sterols. [74] Anjaneyulu et al. reported 6-hydroxy lignans from the heartwood of *Gmelina arborea*, 6-bromo-isoarboreol, 4-hydroxysesamin, 4,8-diydroxysesamin, 1,4-dihydroxysesamin (gummadiol), 2-piperonyl-3-hydroxymethyl-4-(α-hydroxy-3,4-methylenedioxybenzyl)-4-hydroxytetrahydrofuran, and 4-0-glucoside of 4-epigummadiol. [62,63] Hentriacontanol-1, a sesquiterpene, ceryl alcohol, β-sitosterol and n-octacosanol were isolated from the light petroleum extract, and gmelinol in aqueous extract. [75] Arborone and 7-oxo-dihydrogmelinol are the two keto-lignans identified. [76] 5,7-dihydroxy-4-methoxy flavone was isolated from fruit extract and found to exhibit diuretic activity. [77] Four lignans (+)-7-O-ethyl arboreol, (+)-paulownin, (+)-gmelinol, (+)-epieudesmin and (−)-β-sitosterol isolated from *Gmelina arborea* exhibited anti-fungal activity against *Trametes versicolor* and *Fomitopsis palustris.* [78] Two known compounds, 2,6-dimethoxy-p-benzoquinone (5) and 3,4,5-trimethoxyphenol were identified in methanolic extract. [79]

In the present study, the effect of aqueous, methanolic and petroleum ether extracts of *Gmelina arborea* was determined by using the MTT assay. The MTT assay helps in the quantification of metabolically viable tumour cells by their capability to decrease MTT. The results from the MTT assay in 72 hours incubation of MDA-MBA-231 cells with methanolic extract of *Gmelina arborea* leaves demonstrated best inhibition with an IC50 value of 18.25 ± 1.06. Aqueous and petroleum extracts showed low and moderate inhibition with an IC50 value of 79.4 ± 3.72 and 153.51 ± 1.75, respectively. Difference in inhibitory activity of the extracts may be due to the different capability of solvents in the extraction of chemical constituents. A wide variety of studies have proved the cytotoxic effects of plant extracts.

Medicinal effects specifically anti-cancer properties of *Gmelina arborea* has been explored and proved in various studies. Treatment of carrageenan induced inflammation albino rats with aqueous and methanolic stem bark extracts of *Gmelina arborea* exhibited maximum inhibition. [80] Methanolic extracts of *Gmelina arborea* leaves demonstrated high anticancer activity in hepatocellular carcinoma cell line. [81] Hydroacetonic crude leaf extract of *Gmelina arborea* with 75.50% of flavonoids and 24.49% of cinnamic acid derivatives exhibited anti-cancer potential against different cancerous cell lines. [82] Methanolic, n-butanol and ethyl acetate extracts exhibited cytotoxic activity in shrimp larvae with the IC50 values within the standards of American Cancer Institute criteria. [83] Methanolic bark extract exhibited significant anti-tumor activity in an ascites tumor mouse model. [84] Various compounds isolated from

different parts of *Gmelina arborea* have proved its efficiency in demonstrating significant radical scavenging activity. [79,85] Previous studies have also proved the antioxidant activities of *Gmelina arborea* leaves extract. [86] Seed, stem bark and fruit extracts have demonstrated efficient scavenging activity. [59,86,87]

Conclusion

In conclusion, the results of the current study indicate that methanolic extract of *Gmelina arborea* possess potential cytotoxicity and antioxidant activity against metastatic breast cancer cells. There are very few herbal drugs explored for their anticancer activities against metastatic breast cancer. Albeit the exact inhibitory mechanism of *Gmelina arborea* was not explored, results show optimistic evidence of involvement of active components of the extracts in the regulation of apoptosis. It may also due to interference of active components with cell signaling. Although further research is needed, a result of this study demonstrates that leaves of *Gmelina arborea* may have chemo-preventive or anti-metastatic properties.

References

1. Jaswanth JP, Dhamotharan R. Exome data analysis in the discovery of variants associated with breast cancer Metastasis and their implications on protein structure. Ann Rom Soc Cell Biol. 2021;25:1663-1682.
2. Partridge AH, Burstein HJ, Winer EP. Side effects of chemotherapy and combined chemohormonal therapy in women with early-stage breast cancer. J Natl Cancer Inst Monogr. 2001;30:135-142.
3. Surh YJ. Cancer chemoprevention with dietary phytochemicals. Nat Rev Cancer. 2003;3:768-780.
4. Sun J, Chu YF, Wu XZ, Liu RH. Antioxidant and anti-proliferative activities of common fruits. J Agric Food Chem. 2002;50:7449-7454.
5. Liu RH. Health benefits of fruit and vegetables are from additive and synergistic combinations of phytochemicals. Am J Clin Nutr. 2003;78:517-520
6. Liu RH. Potential synergy of phytochemicals in cancer prevention: Mechanism of action. J Nutr. 2004;134:3479-3485.
7. Balunas MJ, Kinghorn AD. Drug discovery from medicinal plants. Life Sci. 2005;78:431-441.
8. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: Methods and applications. Nat Rev Drug Discov. 2004;3:935-949
9. Shoichet BK. Virtual screening of chemical libraries. Nature. 2004;432:862-865.
10. Walters WP, Stahl MT, Murcko MA. Virtual screening: An overview. Drug Discov Today. 1998;3:160-178.
11. Bissantz C, Folkers G, Rognan D. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. J Med Chem. 2000;43:4759-67.
12. Kontoyianni M, McClellan LM, Sokol GS. Evaluation of docking performance: Comparative data on docking algorithms. J Med Chem. 2004;47:558-565.
13. Perola E, Walters WP, Charifson PS. A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. Proteins. 2004;56:235-249.
14. Jachak SM, Saklani A. Challenges and opportunities in drug discovery from plants. JSTOR. Current Science. 2007;92:1251-1257.
15. Tilburt JC, Kaptchuk TJ. Herbal medicine research and global health: An ethical analysis. Bull World Health Organ. 2008;86:594-599.
16. Johnston JJ, Rubinstein WS, Facio FM, Ng D, Singh LN, Teer JK, et al. Secondary variants in individuals undergoing exome sequencing: screening of 572 individuals identifies high-penetrance mutations in cancer-susceptibility genes. Am J Hum Genet. 2012;91:97-108.
17. Gribenko AV, Hopper JE, Makhatadze GI. Molecular characterization and tissue distribution of a novel member of the S100 family of EF-hand proteins. Biochemistry. 2001;40:15538-15548.
18. Streicher WW, Lopez MM, Makhatadze GI. Modulation of quaternary structure of S100 proteins by calcium ions. Biophys Chem. 2010;151:181-186.
19. Carlsson H, Petersson S, Enerback C. Cluster analysis of S100 gene expression and genes correlating to psoriasin (S100A7) expression at different stages of breast cancer development. Int J Oncol. 2005;27:1473-1481.
20. Wang C, Luo J, Rong JL, He SY, Zhang L, Zheng F. Distinct prognostic roles of S100 mRNA expression in gastric cancer. Pathology Research and Practice. 2019;215:127-136.
21. Raffat MA, Hadi NI, Hosein M, Mirza S, Ikram S, Akram Z. S100 proteins in oral squamous cell carcinoma. Clinica Chimica Acta. 2018;480:143-149.
22. Festen EAM, Stokkers PCF, van Diemen CC, van Bodegraven AA, Boezen HM, Crusius BJA, et al. Genetic analysis in a Dutch study sample identifies more ulcerative colitis susceptibility loci and shows their additive role in disease risk. Am J Gastroenterol. 2010;105:395-402.
23. Whitaker HC, Shiong LL, Kay JD, Gronberg H, Warren AY, Seipel A, et al. N-acetyl-L-aspartyl-L-glutamate peptidase-like 2 is overexpressed in cancer and promotes a pro-migratory and pro-metastatic phenotype. Oncogene. 2014;33:5274-5287.
24. Borg K, Stankiewicz P, Bocian E, Kruczek A, Obersztyn E, Lupski JR, et al. Molecular analysis of a constitutional complex genome rearrangement with 11 breakpoints involving chromosomes 3, 11, 12, and 21 and a similar to 0.5-Mb submicroscopic deletion in a patient with mild mental retardation. Hum Genet. 2005;118:267-275.
25. Walker LC, Marquart L, Pearson JF, Wiggins GAR, O'Mara TA, Parsons MT, et al. Evaluation of copy-number variants as modifiers of breast and ovarian cancer risk for BRCA1 pathogenic variant carriers. Eur J Hum Genet. 2017;25:432-438.
26. Burgner D, Davila S, Breunis WB, Ng SB, Li Y, Bonnard C, et al. A genome-wide association study identifies novel and functionally related susceptibility loci for kawasaki disease. Plos Genetics. 2009;5:1-15.
27. Tonkin ET, Smith M, Eichhorn P, Jones S, Imamwerdi B, Lindsay S, et al. A giant novel gene undergoing extensive alternative splicing is severed by a Cornelia de Lange-associated translocation breakpoint at 3q26.3. Hum Genet. 2004;115:139-148.
28. Berndt SI, Wang ZM, Yeager M, Alavanja MC, Albanes D, Amundadottir L, et al. Two susceptibility loci identified for prostate cancer aggressiveness. Nature Communications. 2015;6:1-7.
29. Kuo PH, Chuang LC, Su MH, Chen CH, Wu JY, Yen CJ, et al. Genome-wide association study for autism spectrum disorder in Taiwanese han population. Plos One. 2015;10:10-15.
30. Rossi AC, Mammucari C, Argentini C, Reggiani C, Schiaffino S. Two novel/ancient myosins in mammalian skeletal muscles: MYH14/7b and MYH15 are expressed in extraocular muscles and muscle spindles. J Physiol (London). 2010;588:353-364.

31. Mascarello F, Toniolo L, Cancellara P, Reggiani C, Maccatrozzo L. Expression and identification of 10 sarcomeric MyHC isoforms in human skeletal muscles of different embryological origin. Diversity and similarity in mammalian species. Ann Anat. 2016;207:9-20.
32. Zhang ZR, Du HR, Yang CW, Li QY, Qiu MH, Song XY, et al. Comparative transcriptome analysis reveals regulators mediating breast muscle growth and development in three chicken breeds. Anim Biotechnol. 2019;30:233-241.
33. Bare LA, Morrison AC, Rowland CM, Shiffman D, Luke MM, Iakoubova OA, et al. Five common gene variants identify elevated genetic risk for coronary heart disease. Genet Med. 2007;9:682-689.
34. Luke MM, Lalouschek W, Rowland CM, Catanese JJ, Bolonick JI, Bui ND, et al. Polymorphisms associated with both non cardioembolic stroke and coronary heart disease: Vienna Stroke Registry. Cerebrovascular Diseases. 2009;28:499-504.
35. Yoshino S, Cilluffo R, Best PJM, Atkinson EJ, Aoki T, Cunningham JM, et al. Single nucleotide polymorphisms associated with abnormal coronary microvascular function. Coron Artery Dis. 2014;25:281-289.
36. Chen G, Luo L, Zhang MM, Wu SQ, Wang Y, Sandford AJ, et al. Association study of myosin heavy chain 15 polymorphisms with asthma susceptibility in Chinese Han. Biomed Res Int. 2019:11.
37. Snezhkina AV, Lukyanova EN, Zaretsky AR, Kalinin DV, Pokrovsky AV, Golovyuk AL, et al. Novel potential causative genes in carotid paragangliomas. BMC Med Genet. 2019;20:1-9.
38. Gilad Y, Bustamante CD, Lancet D, Paabo S. Natural selection on the olfactory receptor gene family in humans and chimpanzees. Am J Hum Genet. 2003;73:489-501.
39. Wu YL, Duan HP, Tian XC, Xu CS, Wang WJ, Jiang WJ, et al. Genetics of Obesity Traits: A bivariate genome-wide association analysis. Front Genet.2018;9:1-9.
40. Bailey JNC, Palmer ND, Ng MCY, Bonomo JA, Hicks PJ, Hester JM, et al. Analysis of coding variants identified from exome sequencing resources for association with diabetic and non-diabetic nephropathy in African Americans. Hum Genet. 2014;133:769-79.
41. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J. CASTp: Computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic Acids Res. 2006;34:116-118.
42. Binkowski TA, Naghibzadeh S, Liang J. CASTp: Computed atlas of surface topography of proteins. Nucleic Acids Res. 2003;31:3352-3355.
43. Jendele L, Krivak R, Skoda P, Novotny M, Hoksza D. PrankWeb: A web server for ligand binding site prediction and visualization. Nucleic Acids Res. 2019;47:345-349.
44. Huang BD. MetaPocket: A meta approach to improve protein ligand binding site prediction. OMICS. 2009;13:325-330.
45. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, et al. The Pfam protein families database. Nucleic Acids Res. 2012;40:290-301.
46. Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, et al. The Pfam protein families database. Nucleic Acids Res. 2004;32:138-141.
47. Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, Lassmann T, et al. Pfam: clans, web tools and services. Nucleic Acids Res. 2006;34:247-251.
48. De-Castro E, Sigrist CJA, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, et al. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res. 2006;34:362-365.
49. Dallakyan S, Olson AJ. Small-molecule library screening by docking with PyRx. Methods Mol Biol. 2015;1263:243-250.
50. http://pymol.sourceforge.net/newman/userman.pdf
51. Wallace AC, Laskowski RA, Thornton JM. Ligplot-A program to generate schematic diagrams of protein ligand interactions. Protein Eng. 1995;8:127-134.
52. vanMeerloo J, Kaspers GJL, Cloos J. Cell sensitivity assays: The MTT assay. Methods Mol Biol. 2011;731:237-245.
53. Mosmann T. Rapid colorimetric assay for cellular growth and survival-application to proliferation and cyto-toxicity assays. J Immunol Methods. 1983;65:55-63.
54. Mohanraj K, Karthikeyan BS, Vivek-Ananth RP, Chand RPB, Aparna SR, Mangalapandi P, et al. IMPPAT: A curated database of Indian medicinal plants, phytochemistry and therapeutics. Sci Rep. 2018;8:4329-4332.
55. Geran RI, Greenberg NH, MacDonald MM, Schumacher A, Abbott BJ. Protocols for screening chemical agents and natural products against animal tumors and other biological systems. 1972.
56. Powar P, Rathod S, Ambikar R, Sharma P. Formulation and evaluation of a herbal anti-bacterial gel containing *gmelina arborea* leaf extract. Indian Drugs. 2015;52:10-14.
57. https://www.researchgate.net/publication/329340739_Anti-nociceptive_and_Anti-inflammatory_Effects_of_Gmelina_arborea_Roxb_e_x_Sm
58. Kumar A, Chawhaan PH, Matharoo AK. Improvement through selection of plus trees in *Gmelina arborea*. J Trop For Sci. 2003;15:441-449.
59. Kumar D, Sanghi A, Chandra R, Arora S, Sharma AK. Membrane stability and antioxidant activity of *Gmelina arborea* seed extracts and their fatty acid composition. Br J Pharm Res. 2015;6:261-268.
60. Das B, Sarkar PK, Kumari N, Dey P, Singh AK, Bhatt BP. Biophysical performance of different multipurpose trees species in Jharkhand, India. Current Science. 2019;116:82-88.
61. http://www.ayurveda.hu/api/API-Vol-2.pdf
62. Anjaneyulu A, Rao VK, Row LR, Pelter A, Ward RS. Novel hydroxylignans from the heartwood of *Gmelina arborea*. Tetrahedron Lett. 1977;33:133-143.
63. Pelter A, Ward RS, Watson DJ, Jack IR. Synthesis and nmr spectra of 2, 6-and 2, 4-diaryl-3, 7-dioxabicyclo-[3.3. 0] octanes. J Chem Soc (Perkin 1). 1982;14:183-190.
64. Wu DP, Chiang HC. Constituents of antrodia-cinnamomea. JCCS. 1995;42:797-800.
65. Jin L, Luo YD, Zhang L, Zhang SS, Luo YM. 2015. Studies on the chemical constituents of *cinnamomumcamphora*. proceedings of the 3rd international conference on material, mechanical and manufacturing engineering. Guangzhou. China.
66. Chen PL. Pharmaceutical composition for treating colorectal cancer. Google Patents; 2017.
67. Wu TS. Triterpenoid derivatives, benzenoid derivatives and pharmaceutical compositions containing the same. Google Patents; 2011.
68. Anjaneyulu ASR, Madhusudhana AR, Kameswara VR, Ramachandra LR, Andrew P, Robert SW. The structure of Gummadiol. Tetrahedron Lett. 1975;16:1803-1806.
69. Anjaneyulu A, Rao KJ, Rao VK, Row LR, Subrahmanyam C, Pelter A, et al. The structures of lignans from *Gmelina arborea* Linn. Tetrahedron Lett. 1975;31:1277-1285.

70. Shukla S, Saluja A, Pandya S. A tioxidat activity ad free radical scavegig potetial of *Gmelia arboreali*. Pharmacology online. 2009;1:1035-1043.

71. Kulkarni YA, Chavan D, Shah K. Toxicity study and analgesic activity of *Gmelina arborea* extract and its fractions. J. Nat. Pharm. 2013;4:71-74.

72. Zheng CJ, Lan XP, Cheng RB, Huang BK, Han T, Zhang QY, et al. Furanofuran lignans from *vitexnegundo* seeds. Phyto chem. Lett. 2011;4:298-300.

73. Pelter A, Ward RS, Watson DJ, Collins P, Kay IT. Unambiguous synthesis of lignans of the 2,6-diaryl-3,7-dioxabicyclo- 3.3.0 octane series-synthesis of eudesmin and 4,8-dihydroxysesamin. Tetrahedron Lett. 1979;24:2275-2278.

74. Arora C, Tamrakar V. *Gmelina arborea*: Chemical constituents, pharmacological activities and applications. Int J Phytomed. 2017;9:528-542.

75. Joshi KC, Singh P, Pardasani R. Chemical components of the roots of *tectonagrandis* and *Gmelina arborea*. Plantamedica. 1977;32:71-75.

76. Satyanarayana P, Rao PK, Ward R, Pelter A. Arborone and 7-oxo-dihydrogmelinol: Two new keto-lignans from *Gmelina arborea*. J Nat Prod. 1986;49:1061-1064.

77. Nayak BS, Ellaiah P, Dinda SC, Moharana BP, Khadanga M, Nayak S. Diuretic activity of flavonoid compound isolated from *Gmelina arborea* fruits extract. Eur J Pharm Med Res. 2017;4:616-622.

78. Kawamura F, Ohara S, Nishida A. Antifungal activity of constituents from the heartwood of *Gmelina arborea*: Part 1. Sensitive antifungal assay against Basidiomycetes. Holzforschung. 2004;58:189-192.

79. Falah S, Katayama T, Suzuki T. Chemical constituents from *Gmelina arborea* bark and their antioxidant activity. J Wood Sci. 2008;54:483-489.

80. Kulkarni YA, Panjabi R, Patel V, Tawade A, Gokhale A. Effect of *Gmelina arborea* Roxb in experimentally induced inflammation and nociception. J Ayurveda Integr Med. 2013;4:152-157.

81. Ghareeb MA, Shoeb HA, Madkour HM, Refahy LA, Mohamed MA, Saad AM. Antioxidant and cytotoxic activities of flavonoidal compounds from *Gmelina arborea* (Roxb). Glob J Pharmacol. 2014;8:87-97.

82. Arora C, Tamrakar V. *Gmelina arborea*: Chemical constituents, pharmacological activities and applications. Int J Phytomedicine. 2017;9:528-542.

83. Shoeb HA, Madkour HMF, Refahy LA, Mohamed MA, Saad AM, Ghareeb MA. Antioxidant and cytotoxic activities of *Gmelina arborea* ROXB. Leaves. Br J Pharm Res. 2014;4:125-144.

84. Nify F. Study of *in vitro* Cytotoxicity and *in vivo* Anti-tumour and anti-inflammatory activities of *Gmelina arborea* roxb. Stem Bark. 2014.

85. Yadava RN, Raghuvansi J. Antioxidant activity of a new flavone glycoside from the stems of *Gmelina arborea* roxb. Int J Pharm Sci Res. 2020;11:198-203.

86. Lawrence L, Menon S, Vincent S, Sivaram VP, Padikkala J. Radical scavenging and gastroprotective activity of methanolic extract of *Gmelina arborea* stem bark. J Ayurveda Integr Med. 2016;7:78-82.

87. Sinha S, Dixit P, Bhargava S, Devasagayam TPA, Ghaskadbi S. Bark and fruit extracts of *Gmelina arborea* protect liver cells from oxidative stress. Pharm. Biol. 2006;44:237-243.