Integrating network pharmacology and pharmacological evaluation to investigate the anticancer effects of Duranta erecta Linn. Verbenaceae in breast cancer

Yannick Luther Agbana
Thomas Abu
Toluwanimi Emmanuel Akinleye
Adekunle Johnson Adeniji
Omonike O. Ogbole (nikeoa@yahoo.com)

Research Article

Keywords: Breast cancer, Duranta erecta, Natural product, Repenin A, Network pharmacology

Posted Date: January 11th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2462558/v1

Additional Declarations:
Competing interests: The authors declare that there is no competing interests.

Integrating network pharmacology and pharmacological evaluation to investigate the anticancer effects of *Duranta erecta* Linn. Verbenaceae in breast cancer

ABSTRACT

Objective

Breast cancer is the most prevalent type of cancer among women in sub-Saharan Africa. Efforts are being made to tackle the disease. However, numerous challenges are still reported. *Duranta erecta* showed medicinal relevance in different ailments but its molecular mechanism of action in breast cancer is not unraveled. The objective of this study is to evaluate the anticancer effect of *Duranta erecta* on breast cancer cells and determine the molecular mechanism of action *in silico*.

Materials and Methods

The Phytochemical Interaction Database, published literature, and the Swiss TargetPrediction database, respectively, were used to identify the active ingredients and targets of *Duranta erecta*. GEO datasets and TCGA databases were searched for breast cancer-related targets. A protein-protein interaction (PPI) network was constructed to screen the primary targets. For GO and KEGG pathway enrichment analyses, ShinyGO was used. By using molecular docking, interactions between potential targets and active substances were evaluated. MTT assay was conducted to evaluate the cytotoxicity effect of *Duranta erecta*.

Results

Duranta erecta demonstrated a cytotoxic effect on breast cancer cells. The IC50 values are 9.99 µg/mL and 15.07 µg/mL for the fruit extract and the leaves extract respectively. A total of 102 common targets and 77 active plant compounds were discovered, of which 37 are potential drug candidates. There were 10 hub targets identified by the PPI network. The hub targets are linked to pathways in cell proliferation and cancer. The best overall binding affinity was demonstrated by repenin A in binding with AURKA, CDK1, and EGFR.

Conclusion

This study was able to accurately predict the active ingredients and potential targets used in *Duranta erecta*'s treatment of breast cancer. This study offers a fresh approach to future deeper studies on the molecular mechanisms of the plant and its compounds in breast cancer.

Keywords: Breast cancer; *Duranta erecta*; Natural product; Repenin A; Network pharmacology

List of Abbreviations

CADD: computer-aided drug discovery

DEGs: differentially expressed genes

FBS: fetal bovine serum

GEO: Gene Expression Omnibus

GEPIA: Gene Expression Profiling Interactive Analysis

GO: Gene Ontology

KEGG: Kyoto Encyclopedia of Genes and Genomes

LogP: Lipophilicity

MEM: minimum essential medium

NCDs: noncommunicable diseases

NHBA: number of Hydrogen bond acceptors

NHBD: number of Hydrogen bond donors

PPI: protein-protein interaction

QSAR: quantitative structure-activity relationship

SDF: structure data file

SMILES: simplified molecular-input line-entry system

TCGA: The Cancer Genome Atlas

1. Introduction

Cancer is a significant public health issue in Africa. It is becoming more prevalent in Sub-Saharan Africa, affecting a large portion of the one billion inhabitants [1]. The disease is one of the top three causes of premature death (aged 30-69 years) in almost all constituent countries, accounting for one in every seven premature deaths overall and one in every four deaths from noncommunicable diseases (NCDs) [1]. The burden of cancer is anticipated to nearly double in sub-Saharan Africa over the next 20 years as a result of population growth and aging, reaching 1.5 million new cases and 1 million deaths by the year 2040, even if cancer incidence rates remained constant [1].

Breast cancer is the most prevalent type of cancer among women in sub-Saharan Africa. A recent epidemiological comparative study of breast cancer burden globally over 28 years shows that central sub-Saharan Africa had the highest case fatality rate [2]. In east sub-Saharan Africa, by 2030, it is expected that Tanzania's incidence rates for breast cancer will have increased by more than 80% [3]. In 2018 there were 45,157 new cases of breast cancer in Western Africa, and there were 20,983 breast cancer-related deaths [4]. In Ghana, new cases of breast cancer in 2018 accounted for 33.6% of all female cancer cases and 20.4% of all cancer cases overall [4]. In Nigeria, over the past forty years, there has been an 80% increase in the incidence of breast cancer. A recent retrospective study in Nigeria among adolescents and young adult females aged 15 to 39 years showed that breast cancer was the most common cancer, accounting for 50% of all cancers. The adolescents and young adult females breast cancer patients constituted 51% of all breast cancer cases [5,6]. Similar findings are made in Togo, where breast cancer ranks first in female cancers, accounting for 21.2%, and strucks 38.3% of women under the age of 40 [7].

These observations indicates that the incidence and mortality rate of breast cancer is increasing in sub-Saharan Africa and breast cancer is claiming the life of numerous young female adults. Subsequently, breast cancer can impose drawbacks on the socioeconomic status of sub-Saharan Africa population. Currently, the treatment regimen of breast cancer in sub-Saharan Africa includes surgery, radiotherapy, and chemotherapy [8]. Unfortunately, these treatment routes present a lot of challenges. Many countries do not have access to radiotherapy; in the regions where it is available the access is low because user fees are extremely expensive. The cancer drugs used for chemotherapy are highly cost. Cancer medicines are more expensive in African

countries than in other regions with comparable income levels [9]. In addition, surgery and radiotherapy can be invasive for the patients, and the current drugs demonstrated some level of toxicity. Therefore, in addition to improving diagnostic services and breaking sociocultural and economic barriers, drug discovery in the African context should be encouraged and advanced for cancer treatment in general, and breast cancer treatment in particular.

Drug discovery refers to the process of discovering medications to cure and treat diseases. The procedure begins when there is an unmet clinical need, such as a disease or clinical condition for which no relevant medical products are available. Target discovery is the first step in the drug discovery and development process, followed by hit generation and lead identification, lead optimization, when necessary, preclinical studies, clinical trials, and pharmacovigilance [10].

The advancement of machine learning and artificial intelligence has revolutionized the field of drug discovery. Many tools have been created using their algorithms to ease the process of drug discovery and development, and to improve the success rate of drug development [11].

Drug discovery has also leveraged on nature's resources for many years. Medicinal plants, one of those natural products, have long been used as remedies in traditional medicine and ethnomedicine around the world [12,13]. Till date, many drugs have been produced from plants, many plants are being explored for therapeutics properties, and many other plants are pending investigation [14,15]. Plants with a variety of medicinal qualities, including potential anticancer effects, are concealed among Africa's lush vegetation [16]. *Duranta erecta* is one of the plants found in the African flora which is at the epicenter of this study. The objectives of this study include target and drug discovery for breast cancer treatment using *in silico* approaches, study of the anticancer property of *Duranta erecta* against breast cancer, and evaluation of the mechanism of action *in silico*. Uncountable similar studies have been conducted across the world. A recent study investigated the anticancer potential of *Foeniculum vulgare* Mill., a medicinal plant, against breast cancer [17]. In non-small cell lung cancer, kinase inhibitors targeting PI3Kα were identified based on molecular docking and dynamics approach [11]. Network pharmacology analysis coupled with molecular docking study revealed the mechanism of action of medicinal plants in the treatment of breast cancer and hepatocellular carcinoma [18,19]. Therefore, the present study approach is justified and significant.

2. Materials and Methods

2.1. Cell culture

MCF-7 human breast cancer cell lines (CDC, Atlanta, USA) were obtained from the tissue culture laboratory of the WHO National Polio Reference Lab, Department of Virology, College of Medicine, University College Hospital, Ibadan, Nigeria. The cells were cultured in the minimum essential medium (MEM) with 10% fetal bovine serum (FBS) and maintained in MEM with 2% FBS, and then placed in a humidified incubator at 37 C with the microplates sealed to conserve carbon dioxide (CO_2). The medium includes 1% non-essential amino acids, 0.07% NaHCO3, 100 units/mL Penicillin, 2 mM L-glutamine, vitamin solution, and 100 mg/mL Streptomycin.

2.2. Plant collection and extraction procedure

The plant was collected following the guidelines on good agricultural and collection practices in herbal medicine, and then it was taken to Forest Herbarium Ibadan for the authentication. The plant materials were air-dried in the shade until completely dry. An electric crushing machine was used to grind the materials to powder. Extraction by cold maceration was done by soaking 50 g of the powdered plant materials in methanol (ratio 1:3) for 72 hours. Air-drying was used to concentrate the extracts.

2.3. MTT assay

Cell viability was determined using the 3-(4, 5-dimethyl thiazol-2-yl) 2, 5-diphenyltetrazolium bromide (MTT) assay (Sigma, Chem., St. Louis, MO). The crude extracts' cytotoxic potential was determined using the IC50 values. A hundred microliter (100 µL) of cells was pipetted into a 96-well plate and kept in MEM for 24 hours. The following day, a stock solution of the extracts at 10 mg/mL was prepared with DMSO. The stock solution was ten-fold serially diluted to yield concentrations of 1000, 100, 10, 1, 0.1, and 0.01 µg/mL. The final DMSO concentration in the tested dilutions was less than 1%. The cells were incubated in triplicate with the extracts at 37 C in a carbon dioxide environment for 72 hours. The negative control is the cells without treatment, while vincristine sulfate (100 - 0.001 g/mL) served as the positive control. After 72 hours, the MCF-7 cell lines' maintenance medium was removed, and 25 µL of MTT solution (2 mg/mL in PBS) was added to each well and incubated for 2 hours at 37 C. The yellow MTT that was added to the cell lines was converted to an insoluble purple formazan (an indicator of cell

viability). The MTT solution was removed after 2 hours, and 75 µL of DMSO was added to each well to solubilize the purple formazan. All the tests were carried out in triplicate. Finally, the optical density (OD) was determined by reading the plates with a microplate reader at 490 nm.

2.4. Establishment of Target Database of Breast Cancer

The Cancer Genome Atlas (TCGA) database and the Gene Expression Omnibus (GEO) datasets were investigated for this purpose.

The GEO datasets[20], which include breast cancer data, were derived from www.ncbi.nlm.nih.gov/geo/ with the following parameters: entry type = datasets; organism = Homo sapiens; study type = expression profiling by array. After a thorough screening, four datasets (GSE10797, GSE20437, GSE3744, and GSE9574) were considered for further analyses. The datasets provide information on the mRNA expressions of both breast cancer samples and normal samples. The limma package of R software was used to screen the differentially expressed genes (DEGs) between breast cancer and normal samples. The thresholds used are $|\log FC| \geq 2$ and P-value < 0.05.

TCGA database was accessed through the Gene Expression Profiling Interactive Analysis (GEPIA, http://gepia2.cancer-pku.cn/). GEPIA is a web server that provides interactive gene expression profiling and analysis for cancer and normal genes [21]. The DEGs screening was performed using the limma R package and the same thresholds as for GEO.

The genes obtained from the two databases were combined to form the target database of breast cancer. The duplicates genes were removed, and the genes IDs and names were confirmed using UniProt database https://www.uniprot.org/ [22].

2.5. Screening of Chemical Compounds in the Plants

The chemical constituents in *Duranta erecta*, were retrieved from the PhytoChemical Interactions DataBase (https://www.genome.jp/db/pcidb). The synonyms of the plants names were searched in the world flora online plant list database (https://wfoplantlist.org/plant-list) and used in the previous database for the compounds retrieval. Published literatures were retrieved from the PubMed database (https://pubmed.ncbi.nlm.nih.gov/) for additional compounds.

Thereafter, the drug-likeness of the compounds was evaluated as follows. The compounds obtained were input in PubChem database (http://pubchem.ncbi.nlm.nih.gov/) to retrieve the simplified molecular-input line-entry system (SMILES) structure. The SMILES were subsequently inserted in MolSoft database (https://www.molsoft.com/mprop/) to predict the drug-likeness. The prediction was based on the parameters of the Lipinski's rule, namely the molecular weight (MW), the number of Hydrogen bond donors (NHBD), the number of Hydrogen bond acceptors (NHBA), and the Lipophilicity (LogP). The Lipinski's rule, also known as the "rule of five", states that poor absorption or permeation is more likely when the NHBD is more than 5, the NHBA more than 10, the MW greater than 500 g/mol, and the LogP is greater than 5 [23]. A molecule is more likely to be a good drug (oral drugs) when three criteria out of the four are respected.

When the SMILES was not available in PubChem, the molecule was first drawn in MarvinSketch (https://chemaxon.com/products/marvin) and then the SMILES exported into the MolSoft database. At the end, only the molecules conform with the Lipinski's rule of five were considered for further analysis.

2.6. Establishment of Target Database of Chemical Compounds

The molecules obtained from the previous step were used for target prediction. Their SMILES were input in the SWISS TargetPrediction database (http://www.swisstargetprediction.ch/). *Homo sapiens* was selected as the organism of interest. The results were exported in excel format. The targets of different constituents of the plant were combined, the duplicates eliminated, and the targets database was created for the plant.

2.7. Screening of Common Targets and Construction of Protein-Protein Interaction Networks

The targets database of breast cancer and the targets database of the plant were screened for the common targets. The common targets were obtained by drawing a Venn diagram using the web server https://bioinformatics.psb.ugent.be/webtools/Venn/.

The common targets were introduced into the STRING database version 11.5 (https://string-db.org/) to construct the protein-protein interaction (PPI) network. The species *Homo sapiens* was selected and the minimum confidence of ≥ 0.4 was set for the interaction score evaluation. The PPI network was then exported into Cytoscape 3.9.1 (https://www.cytoscape.org/) for core

targets (hub genes) screening. The hub genes were determined by using the parameter “Degree” of the CytoHubba plug-in.

2.8. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway Enrichment Analyses

The hub genes were used for GO and KEGG enrichment analyses. The analyses were conducted in the enrichment tool ShinyGO version 0.76 (http://ge-lab.org/go/). The following parameters were considered for the analyses: species = Human, false discovery rate (FDR) cut-off = 0.05, and number of pathways to show = 20. The biological process (BP), the cellular component (CC), and the molecular function (MF) categories were considered for the gene ontology analysis.

2.9. Molecular Docking

2.9.1. Ligands Library Preparation

The ligands in our context are the compounds (metabolites) retained after drug-likeness study. For the purpose of the library preparation, the structure data file (SDF) format of the metabolites was downloaded from PubChem. Subsequently, ligands library was prepared in Discovery Studio Visualizer 2021.

2.9.2. Preparation of Targets

The target proteins' (hub genes') structures were downloaded from the RCSB Protein Data Bank database (https://www.rcsb.org/) and saved in PDB format. The proteins IL6 (PDBID: 1ALU), EGFR (PDBID: 7VRE), JUN (PDBID: 1FOS), ESR1 (PDBID: 7UJ8), EZH2 (PDBID: 4MI5), CCNB1 (PDBID: 2B9R), CCNA2 (PDBID: 1H1P), CDK1 (PDBID: 4YC6), AURKA (PDBID: 2DWB), and MMP9 (PDBID: 5TH6) were imported in Discovery Studio Visualizer 2021 and prepared by removing molecules of water and unwanted ligands.

2.9.3. Docking

To bind the proteins and ligands in the desired protein-ligand complexes and obtain the binding energies, PyRx AutoDock Vina software was used. The same software was used to minimize the ligands' energy. Using the PyRx AutoDock program, the PDB files of the ligands and target proteins were converted to PDBQT format. The procedures for sorting boxes, creating grid boxes, and converting molecules to PDBQT format are all performed in AutoDock. To find the ideal binding site, the grid box was maximized. Thereafter, the PDBQT files were converted

into PDB using Cygwin64 Terminal software and then the complexes were formed and visualized in Discovery Studio Visualizer 2021 and the web server protein-ligand interaction profiler.

2.10. Statistical analysis

The cellular growth inhibition by the tested extracts was calculated as the percentage inhibitory activity and further expressed as the IC_{50} value (refers to the concentration of the tested extract to inhibit 50% growth of the cancer cells).

$$\textbf{\% inhibition} = (1 - [At / Ac] \times 100)$$

Where;

At is the absorbance of the test extract and

Ac is the absorbance of the control

Tests were carried out in triplicates and the data were presented as mean ± SD (standard deviation). Microsoft Excel was used for analysis of mean, SD, and percentage inhibition while non-linear regression analysis was used to determine IC_{50} using Graph Pad Prism 6.

3. Results

3.1. Authentication of *Duranta erecta*

The plant was collected for authentication on June 2, 2021. The plant was collected at Sokoto Crescent, Awolowo Stadium, University of Ibadan; the GPS coordinates are latitude 7.439236, and longitude 3.885735. The plant was authenticated at Forest Herbarium Ibadan and the voucher number is FHI 113235.

3.2. *Duranta erecta* suppresses breast cancer proliferation

The methanol extracts of the plant, both leaves (DEL) and fruits (DEF) extracts, showed cytotoxic effects on the MCF-7 cells, as shown in figure 4. DEF showed a better effect (IC_{50} = 9.99 µg/mL, 95% CI = 5.25-19.03 µg/mL) against DEL (IC_{50} = 15.07 µg/mL, 95% CI = 10.65-21.34 µg/mL).

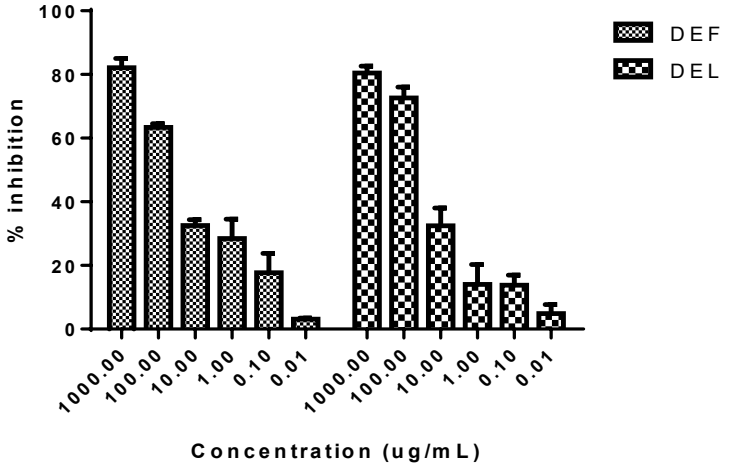


Figure 4: The cytotoxicity effect of *Duranta erecta* on breast cancer cells

3.3. Target Database of Breast Cancer

The four GEO datasets (GSE10797, GSE20437, GSE3744, and GSE9574) yielded 915 DEGs and the TCGA database provided 742 DEGs, making it a total of 1657 DEGs identified. After the duplicates were removed 1318 DEGs were obtained and made up the breast cancer targets database (BC). The DEGs include upregulated and downregulated genes. Basic information about the GEO datasets is in Table 2.

Table 2: GEO Datasets

Accession Number	**Number of Samples**	**Platform**	**Name of the Platform**
GSE10797	66	GPL571	Affymetrix Human Genome U133Plus 2.0 Array
GSE20437	42	GPL96	Affymetrix Human Genome U133A Array
GSE3744	47	GPL570	Affymetrix Human Genome U133Plus 2.0 Array
GSE9574	29	GPL96	Affymetrix Human Genome U133A Array

3.4. Database of the Chemical Compounds Harvested from the Plant

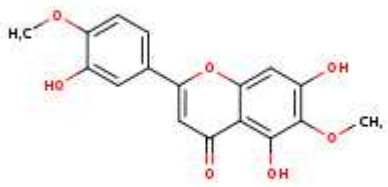
Based on the search strategy of this work, a total of 77 compounds (Appendix 1) were found in *Duranta erecta*. Out of these 77 compounds, 16 compounds were not found in PubMed or other relevant databases for their structures and/or SMILES. The remaining 61 compounds were evaluated for their drug-likeness based on Lipinski's rule of five. Only 39 compounds finally passed the test of drug-likeness (Table 3). The structures of the compounds are depicted on figure 5.

Table 3: *Duranta erecta's* Chemical Compounds and Drug-Likeness

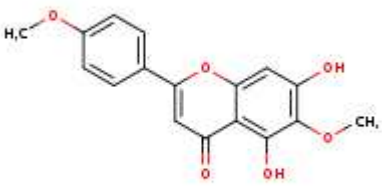
Metabolites	**Pubchem CID**	**Molecular formula**	**Mol. Weight (g/mol)**	**NHBA**	**NHBD**	**MolLogP**
5,7,3'-Trihydroxy-6,4'-dimethoxyflavone (Desmethoxycentaureidin)	5469524	$C_{17}H_{14}O_7$	330.07	7	3	3.12
Pectolinarigenin	5320438	$C_{17}H_{14}O_6$	314.29	6	2	3.58
alpha-Amyrine (Viminalol)	73170	$C_{30}H_{50}O$	426.7	1	1	7.77
Beta-Sitosterol	222284	$C_{29}H_{50}O$	414.7	1	1	8.45
Scutellarein	5281697	$C_{15}H_{10}O_6$	286.24	6	4	2.85
Penduletin 4'-methyl ether (5-hydroxy-3,6,7,4'-tetramethoxyflavone)	5318355	$C_{19}H_{18}O_7$	358.3	7	1	3.37
5,7-Dihydroxy-3'-(2-hydroxy-3-methyl-3-butenyl)-3,6,4'-trimethoxyflavone	11743305	$C_{23}H_{24}O_8$	428.4	8	3	3.95
3,7-Dihydroxy-3'-(2-hydroxy-3-methyl-3-butenyl)-4',5,6-trimethoxyflavone	10410309	$C_{23}H_{24}O_8$	428.4	8	3	3.23
3,7-Dihydroxy-2-[4-hydroxy-3-(4-hydroxy-3-methylbutyl)phenyl]-5,6-dimethoxy-4H-1-benzopyran-4-one	10573935	$C_{22}H_{24}O_8$	416.4	8	4	2.46

3,7-Dihydroxy-2-[3-(4-hydroxy-3-methylbutyl)-4-methoxyphenyl]-5,6-dimethoxy-4H-1-benzopyran-4-one	15511334	$C_{23}H_{26}O_8$	430.4	8	3	2.99
3,7,4'-Trihydroxy-3'-(7-methyl-8-acetoxyoctyl)-5,6-dimethoxyflavone	11306857	$C_{28}H_{34}O_9$	514.22	9	3	4.91
Aliarin 4'-methyl ether	44259768	$C_{23}H_{26}O_8$	430.4	8	3	3.55
(+)-3,13-Clerodadien-16,15-olid-18-oic acid	101736433	$C_{20}H_{28}O_4$	332.2	4	1	3.29
Scoparone (6,7-dimethoxycoumarin)	8417	$C_{11}H_{10}O_4$	206.19	4	0	1.05
Hardwickiic acid	6451316	$C_{20}H_{28}O_3$	316.4	3	1	3.91
Ergosterol peroxide	5351516	$C_{28}H_{44}O_3$	428.6	3	1	6.39
Rosenonolactone	11723309	$C_{20}H_{28}O_3$	316.4	3	0	3.69
Naringenin	932	$C_{15}H_{12}O_5$	272.25	5	3	2.38
Cinnamic acid	444539	$C_9H_8O_2$	148.16	2	1	2.22
4-Methoxycinnamic acid ((E)-p-methoxycinnamic acid)	699414	$C_{10}H_{10}O_3$	178.18	3	1	2.18
Myristoleic acid	5281119	$C_{14}H_{26}O_2$	226.35	2	1	4.99
Ursolic acid	64945	$C_{30}H_{48}O_3$	456.7	3	2	6.46
Oleanolic acid	10494	$C_{30}H_{48}O_3$	456.7	3	2	6.66
Apigenin	5280443	$C_{15}H_{10}O_5$	270.24	5	3	3.22
Beta-amyrin	73145	$C_{30}H_{50}O$	426.7	1	1	7.95
Campenoside	101599955	$C_{25}H_{30}O_{10}$	490.5	10	4	0.85

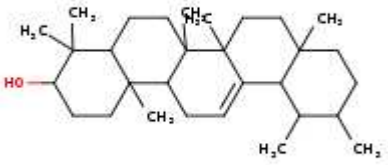
Acacetin	5280442	$C_{16}H_{12}O_5$	284.26	5	2	3.74
Diosmetin	5281612	$C_{16}H_{12}O_6$	300.26	6	3	3.29
Luteolin	5280445	$C_{15}H_{10}O_6$	286.24	6	4	2.78
Quercetin	5280343	$C_{15}H_{10}O_7$	302.23	7	5	1.19
Cleomiscosin A	442510	$C_{20}H_{18}O_8$	386.4	8	2	1.67
Durantin A	11224109	$C_{22}H_{20}O_9$	428.4	9	1	2.11
Repenin A	44583433	$C_{29}H_{24}O_{10}$	532.5	10	2	4
2,4'-dimethoxy-3'-(2-hydroxy-3-methyl-3-butenyl)-acetophenone	11010837	$C_{15}H_{20}O_4$	264.32	4	1	2.54
Flavoglaucin	119037	$C_{19}H_{28}O_3$	304.4	3	2	6.44
Acroptilin	442140	$C_{19}H_{23}ClO_7$	398.8	7	2	0.64
Hyrtial	126946	$C_{26}H_{40}O_3$	400.6	3	0	5.82
α-Cadinol	10398656	$C_{15}H_{26}O$	222.37	1	1	4.27
Betulin	72326	$C_{30}H_{50}O_2$	442.7	2	2	7.13



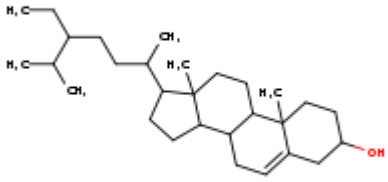
1



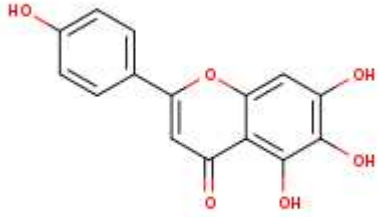
2



3

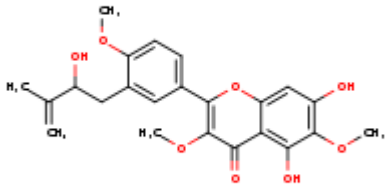


4



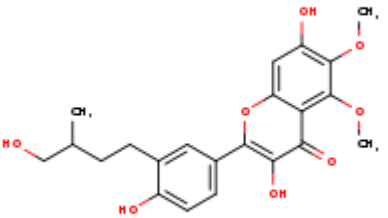
5

6

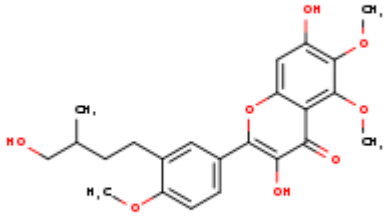


7

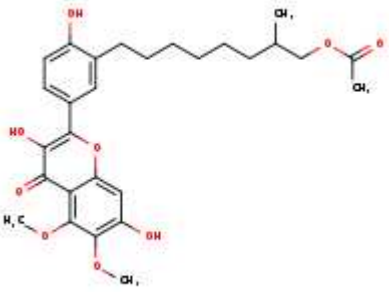
8



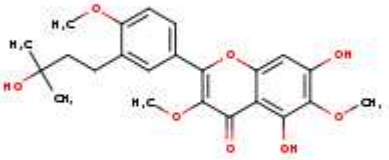
9



10

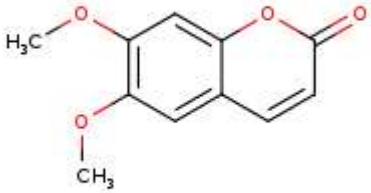


11

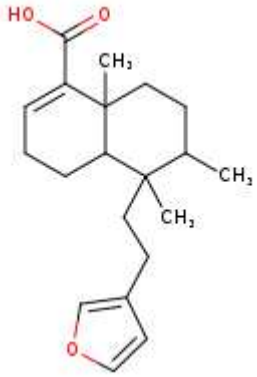


12

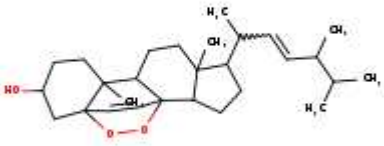
13



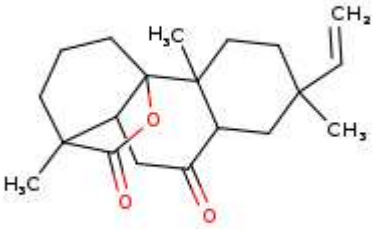
14



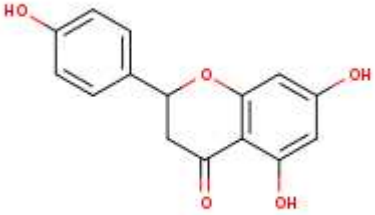
15



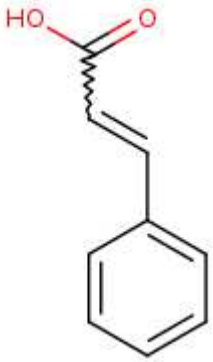
16



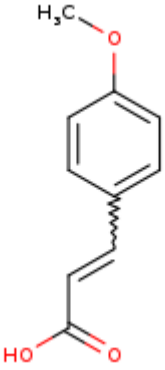
17



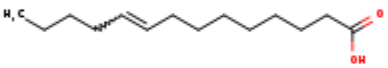
18



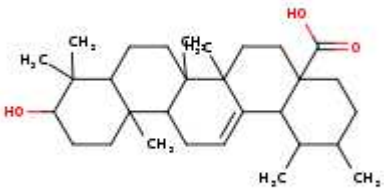
19



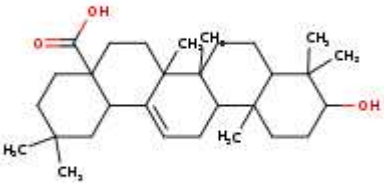
20



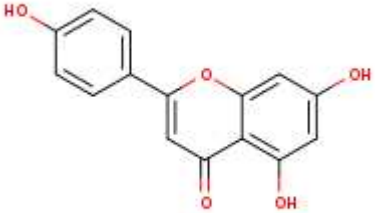
21



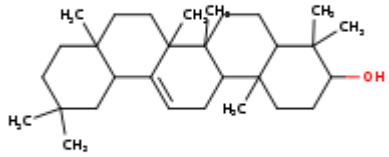
22



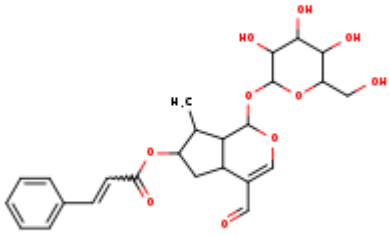
23



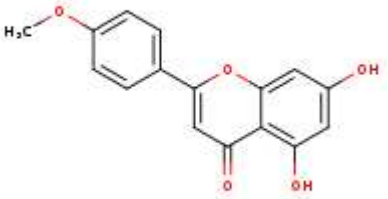
24



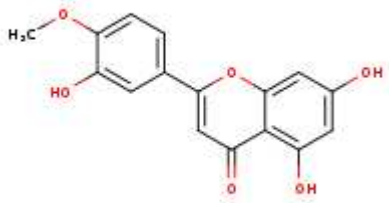
25



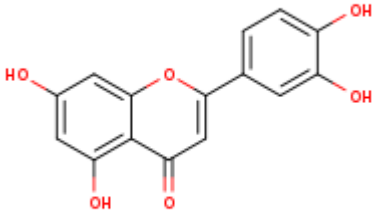
26



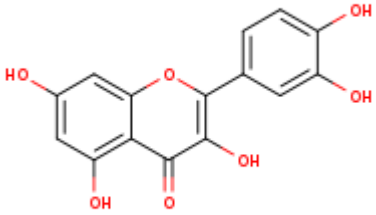
27



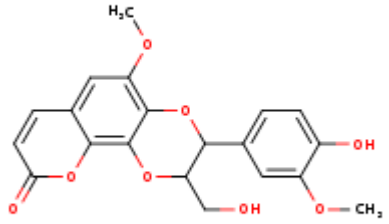
28



29

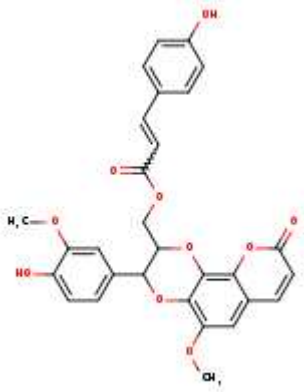


30

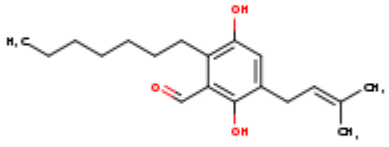


31

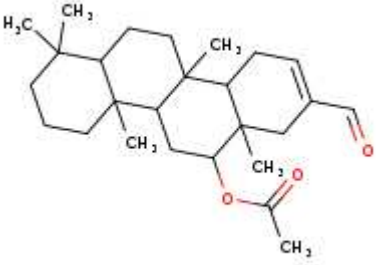
32



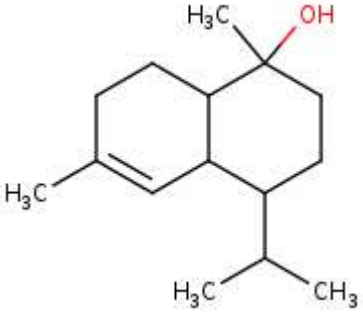
33



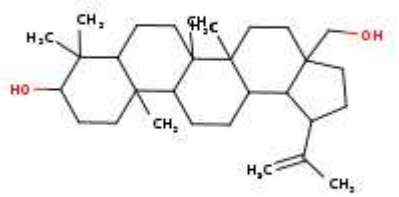
34



35



36



37

Figure 5: The Structures of the 37 Chemical Compounds from *Duranta erecta* based on drug-likeness.

(1) Desmethoxycentaureidin; (2) Pectolinarigenin; (3) Viminalol; (4) Beta-Sitosterol; (5) Scutellarein; (6) Penduletin 4'-methyl ether; (7) 5,7-Dihydroxy-3'-(2-hydroxy-3-methyl-3-butenyl)-3,6,4'-trimethoxyflavone; (8) 3,7-Dihydroxy-3'-(2-hydroxy-3-methyl-3-butenyl)-4',5,6-trimethoxyflavone; (9) 3,7-Dihydroxy-2-[4-hydroxy-3-(4-hydroxy-3-methylbutyl)phenyl]-5,6-dimethoxy-4H-1-benzopyran-4-one; (10) 3,7-Dihydroxy-2-[3-(4-hydroxy-3-methylbutyl)-4-methoxyphenyl]-5,6-dimethoxy-4H-1-benzopyran-4-one; (11) 3,7,4'-Trihydroxy-3'-(7-methyl-8-acetoxyoctyl)-5,6-dimethoxyflavone; (12) Aliarin 4'-methyl ether; (13) (+)-3,13-Clerodadien-16,15-olid-18-oic acid; (14) Scoparone; (15) Hardwickiic acid; (16) Ergosterol peroxide; (17) Rosenonolactone; (18) Naringenin; (19) Cinnamic acid; (20) 4-Methoxycinnamic acid; (21) Myristoleic acid; (22) Ursolic acid; (23) Oleanolic acid; (24) Apigenin; (25) Beta-amyrin; (26) Campenoside; (27) Acacetin; (28) Diosmetin; (29) Luteolin; (30) Quercetin; (31) Cleomiscosin A; (32) Durantin A; (33) Repenin A; (34) Flavoglaucin; (35) Hyrtial; (36) α-Cadinol; (37) Betulin.

3.5. Target Database of Chemical Compounds

The target proteins were predicted in the SWISS TargetPrediction database for the 39 compounds that were judged drug likened. Two compounds, 2,4'-dimethoxy-3'-(2-hydroxy-3-methyl-3-butenyl)-acetophenone and Acroptilin, showed no predicted targets in the database. The remaining 37 compounds yielded 3700 targets. After removing the duplicates, a total number of 809 targets were retained for the target database of chemical compounds (*Duranta erecta*).

3.6. Common Targets and PPI Networks

The Venn diagram shown on figure 6 indicates that BC and *Duranta erecta* share 102 common genes (Table 4).

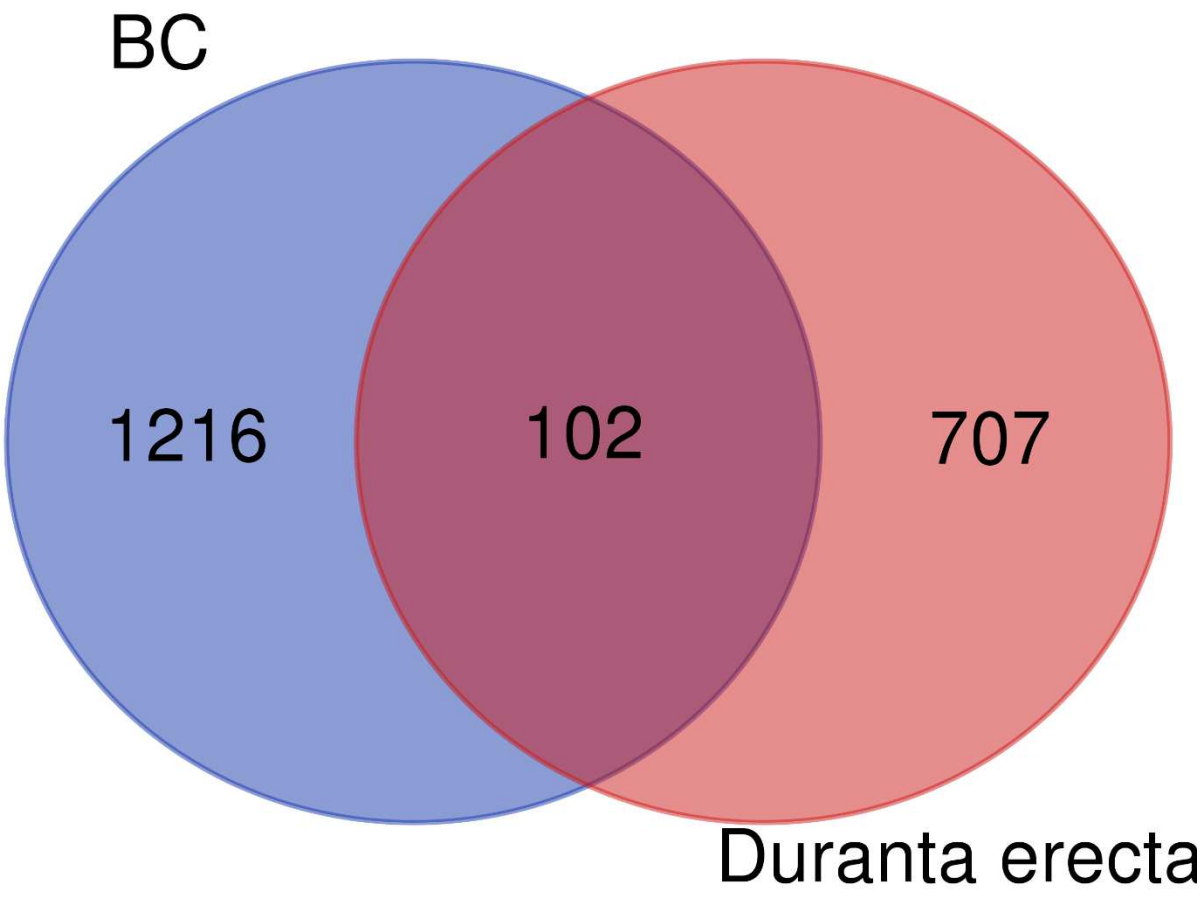


Figure 6: Venn diagram showing 102 common targets between the chemical compounds' targets of *Duranta erecta* and the disease targets of breast cancer

Table 4: The Common Genes between Breast Cancer and *Duranta erecta*

Gene Symbol	UniProt ID	Protein Name	Expression in Breast Cancer	Target Class
ACACB	O00763	Acetyl-CoA carboxylase 2	Downregulated	Enzyme
ADAMTS5	Q9UNA0	A disintegrin and metalloproteinase with thrombospondin motifs 5	Downregulated	Protease
ADRB2	P07550	Beta-2 adrenergic receptor	Downregulated	G protein-coupled receptor
AGTR1	P30556	Type-1 angiotensin II receptor	Downregulated	Family A G protein-coupled receptor
AKR1B10	O60218	Aldo-keto reductase family 1 member B10	Downregulated	Enzyme
AKR1C1	Q04828	Aldo-keto reductase family 1 member C1	Downregulated	Enzyme
AKR1C2	P52895	Aldo-keto reductase family 1 member C2	Downregulated	Enzyme
AKR1C3	P42330	Aldo-keto reductase family 1 member C3	Downregulated	Enzyme
ANPEP	P15144	Aminopeptidase N	Downregulated	Protease
AR	P10275	Androgen receptor	Upregulated	Nuclear receptor
AURKA	O14965	Aurora kinase A	Upregulated	Kinase
AURKB	Q96GD4	Aurora kinase B	Upregulated	Kinase
AVPR2	P30518	Vasopressin V2 receptor	Downregulated	G protein-coupled receptor
BBOX1	O75936	Gamma-butyrobetaine hydroxylase 1	Downregulated	Oxidoreductase
CA12	O43570	Carbonic anhydrase 12	Upregulated	Lyase
CA4	P22748	Carbonic anhydrase 4	Downregulated	Lyase
CCNA2	P20248	Cyclin A2	Upregulated	Cyclin
CCNB1	P14635	Cyclin B1	Upregulated	Cyclin
CCNB2	O95067	Cyclin B2	Upregulated	Cyclin
CCND2	P30279	Cyclin D2	Downregulated	Cyclin
CCNE1	P24864	Cyclin E1	Upregulated	Cyclin
CCNE2	O96020	Cyclin E2	Upregulated	Cyclin
CDC25C	P30307	M-phase inducer phosphatase 3	Upregulated	Phosphatase
CDC7	O00311	Cell division cycle 7-related protein kinase	Upregulated	Kinase

Gene Symbol	UniProt ID	Protein Name	Expression in Breast Cancer	Target Class
CDK1	P06493	Cyclin-dependent kinase 1	Upregulated	Kinase
CENPE	Q02224	Centromere-associated protein E	Upregulated	Motor protein
CES1	P23141	Liver carboxylesterase 1	Downregulated	Hydrolase
CFD	P00746	Complement factor D	Downregulated	Hydrolase
CNR1	P21554	Cannabinoid receptor 1	Downregulated	G protein-coupled receptor
CPA3	P15088	Mast cell carboxypeptidase A	Downregulated	Protease
CTSG	P08311	Cathepsin G	Downregulated	Protease
CXCL8	P10145	C-X-C motif chemokine 8 (Interleukin 8)	Upregulated	Cytokine
CYP26A1	O43174	Cytochrome P450 26A1	Downregulated	Cytochrome P450
CYP26B1	Q9NR63	Cytochrome P450 26B1	Downregulated	Cytochrome P450
DUSP1	P28562	Dual specificity protein phosphatase 1	Downregulated	Phosphatase
DYRK2	Q92630	Dual specificity tyrosine-phosphorylation-regulated kinase 2	Upregulated	Kinase
EDNRB	P24530	Endothelin receptor type B	Downregulated	G protein-coupled receptor
EGFR	P00533	Epidermal growth factor receptor	Downregulated	Kinase
ELOVL6	Q9H5J4	Elongation of very long chain fatty acids protein 6	Upregulated	Transferase
EPAS1	Q99814	Endothelial PAS domain-containing protein 1	Downregulated	Unclassified protein
ESR1	P03372	Estrogen receptor	Upregulated	Receptor
EZH2	Q15910	Histone-lysine N-methyltransferase EZH2	Upregulated	Transferase
F10	P00742	Coagulation factor X	Downregulated	Protease
F3	P13726	Coagulation factor VII/Tissue factor	Downregulated	Surface antigen
FABP4	P15090	Fatty acid-binding protein 4	Downregulated	Carrier protein
FEN1	P39748	Flap endonuclease 1	Downregulated	Hydrolase
GSTM2	P28161	Glutathione S-transferase Mu 2	Downregulated	Transferase

Gene Symbol	UniProt ID	Protein Name	Expression in Breast Cancer	Target Class
HPGD	P15428	15-hydroxyprostaglandin dehydrogenase [NAD(+)]	Downregulated	Oxidoreductase
HSP90B1	P14625	Endoplasmin/ Heat shock protein 90 kDa beta member 1	Upregulated	Chaperon
IL6	P05231	Interleukin-6	Downregulated	Cytokine
JUN	P05412	Transcription factor Jun/ Proto-oncogene c-Jun	Downregulated	Transcription factor
KIF11	P52732	Kinesin-like protein 1	Upregulated	Motor protein
KIT	P10721	Mast/stem cell growth factor receptor Kit	Downregulated	Kinase
MAOA	P21397	Monoamine oxidase A	Downregulated	Oxidoreductase
MAOB	P27338	Monoamine oxidase B	Downregulated	Oxidoreductase
MAPT	P10636	Microtubule-associated protein tau	Downregulated	Unclassified protein
MELK	Q14680	Maternal embryonic leucine zipper kinase	Upregulated	Kinase
MME	P08473	Neprilysin	Downregulated	Protease
MMP1	P03956	Matrix metalloproteinase 1	Upregulated	Protease
MMP12	P39900	Matrix metalloproteinase 12	Upregulated	Protease
MMP13	P45452	Matrix metalloproteinase 13	Upregulated	Protease
MMP9	P14780	Matrix metalloproteinase 9	Upregulated	Protease
MYLK	Q15746	Myosin light chain kinase, smooth muscle	Downregulated	Kinase
NEK2	P51955	Never in mitosis A-related kinase 2	Upregulated	Kinase
NGFR	P08138	Tumor necrosis factor receptor superfamily member 16/ Low-affinity nerve growth factor receptor	Downregulated	Receptor
NR3C2	P08235	Mineralocorticoid receptor/ Nuclear receptor subfamily 3 group C member 2	Downregulated	Nuclear receptor
NR4A1	P22736	Nuclear receptor subfamily 4 group A member 1	Downregulated	Nuclear receptor
NTRK2	Q16620	Neurotrophic tyrosine kinase receptor type 2	Downregulated	Kinase

Gene Symbol	UniProt ID	Protein Name	Expression in Breast Cancer	Target Class
NTRK3	Q16288	Neurotrophic tyrosine kinase receptor type 3	Downregulated	Kinase
OXTR	P30559	Oxytocin receptor	Downregulated	G protein-coupled receptor
PAK2	Q13177	p21-activated kinase 2	Upregulated	Kinase
PBK	Q96KB5	PDZ-binding kinase	Upregulated	Kinase
PDE2A	O00408	cGMP-dependent 3',5'-cyclic phosphodiesterase	Downregulated	Hydrolase
PDE3B	Q13370	cGMP-inhibited 3',5'-cyclic phosphodiesterase 3B	Downregulated	Hydrolase
PDE5A	O76074	cGMP-specific 3',5'-cyclic phosphodiesterase	Downregulated	Hydrolase
PDGFRA	P16234	Platelet-derived growth factor receptor alpha	Downregulated	Kinase
PDK4	Q16654	Pyruvate dehydrogenase kinase isoform 4	Downregulated	Kinase
PER2	O15055	Period circadian protein homolog 2	Downregulated	Unclassified protein
PGR	P06401	Progesterone receptor	Downregulated	Nuclear receptor
PIK3C2G	O75747	Phosphatidylinositol 3-kinase C2 domain-containing subunit gamma	Downregulated	Kinase
PIK3R1	P27986	Phosphatidylinositol 3-kinase regulatory subunit alpha	Downregulated	Kinase
PLA2G2A	P14555	Phospholipase A2 group IIA	Downregulated	Enzyme
PLK1	P53350	Polo-like kinase 1	Upregulated	Kinase
PPARG	P37231	Peroxisome proliferator-activated receptor gamma	Downregulated	Nuclear receptor
PTGER3	P43115	Prostaglandin E2 receptor EP3 subtype	Downregulated	G protein-coupled receptor
PTGS2	P35354	Prostaglandin G/H synthase 2/Cyclooxygenase-2	Downregulated	Oxidoreductase
PTPRG	P23470	Protein-tyrosine phosphatase gamma	Downregulated	Phosphatase
RBP4	P02753	Plasma retinol-binding protein	Downregulated	Secreted protein

Gene Symbol	UniProt ID	Protein Name	Expression in Breast Cancer	Target Class
RET	P07949	Proto-oncogene tyrosine-protein kinase receptor Ret	Upregulated	Kinase
S1PR1	P21453	Sphingosine 1-phosphate receptor 1	Downregulated	G protein-coupled receptor
SLC16A3	O15427	Monocarboxylate transporter 4	Upregulated	Transporter
SQLE	Q14534	Squalene monooxygenase	Upregulated	Oxidoreductase
TACR1	P25103	Tachykinin receptor 1/ Substance-P receptor	Downregulated	G protein-coupled receptor
TAOK1	Q7L7X3	Thousand and one amino acid protein kinase 1	Upregulated	Kinase
TDO2	P48775	Tryptophan 2,3-dioxygenase	Upregulated	Oxidoreductase
TGFBR2	P37173	Transforming growth factor-beta receptor type II	Downregulated	Kinase
THRB	P10828	Thyroid hormone receptor beta	Downregulated	Receptor
TOP2A	P11388	DNA topoisomerase 2-alpha	Upregulated	Topoisomerase
TTK	P33981	Dual specificity protein kinase TTK	Upregulated	Kinase
TYMP	P19971	Thymidine phosphorylase	Upregulated	Transferase
TYMS	P04818	Thymidylate synthase	Upregulated	Transferase
XDH	P47989	Xanthine dehydrogenase/oxidase	Downregulated	Oxidoreductase

The 102 common targets were subsequently imported into STRING database for PPI network construction, as shown on figure 7. The network included 102 nodes and 549 edges. The average node degree was 10.9 and the average local clustering coefficient was 0.598. The statistics also revealed that the expected number of edges is 209, which is much less than the 549 edges observed. This supports the fact that the PPI network has significantly more interactions than expected. This indicates that the proteins in the network are at least partially biologically connected as a group.

Thereafter, the PPI network was exported into Cytoscape for hub genes screening. The top ten targets were selected according to the node degree (Table 5). The network of the hub targets is

illustrated in figure 8. The size of a node is proportional to the node degree score, and the thicker the edge, stronger the connectivity.

Figure 7: Protein-Protein Interactions

Table 5: The Top 10 Hub Genes

Rank	Gene Name	Degree Score
1	IL6	40
2	EGFR	38
3	JUN	35
4	ESR1	31
5	EZH2	30
6	CCNB1	29
7	CCNA2	27
8	CDK1	26
9	AURKA	25
9	MMP9	25

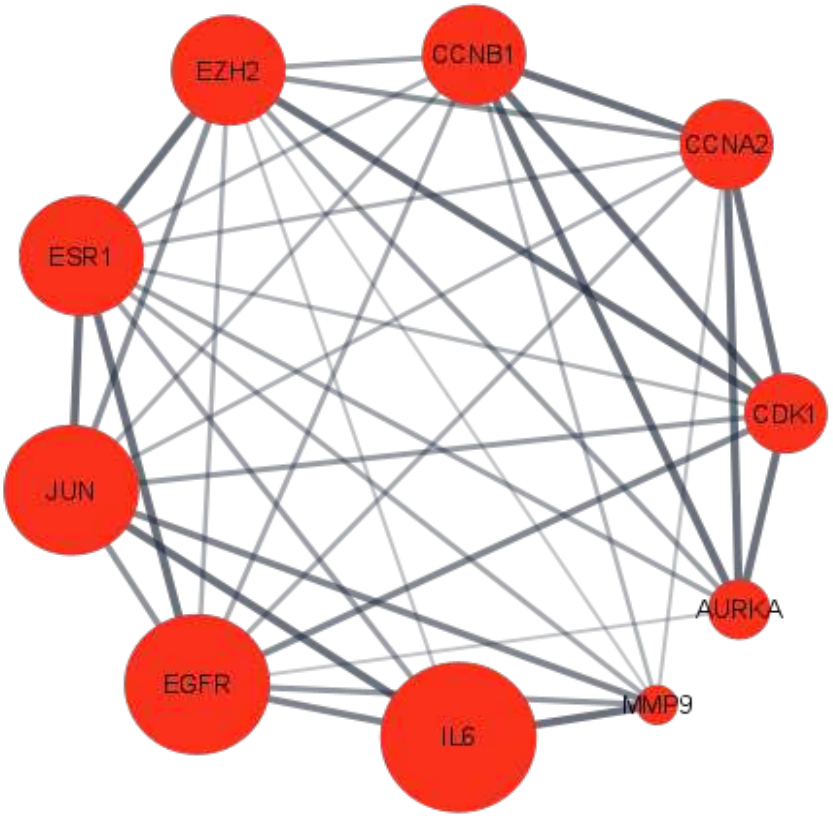


Figure 8: PPI Network of Top Ten Hub Genes

A chemical compound-targets network was also constructed using cytoscape, as shown in figure 9. The compound durantin A has the highest number of predicted targets, 6 hub proteins. Most compounds (11 of them) could bind to 5 targets, and some could bind to 3, 2, and 1 target(s). This result suggests that durantin A might be the most efficacious of all the compounds in *Duranta erecta* against breast cancer. The target ESR1 could bind to 20 compounds, EGFR and CDK1 to 19, MMP9 to 18, CCNB1 to 14, AURKA to 6, CCNA2 to 5, and EZH2, IL6, and JUN could bind to 2 compounds.

3.7. GO enrichment Analysis and KEGG Pathway Enrichment

The GO enrichment analysis of the ten common targets was performed to further investigate the various mechanisms by which *Duranta erecta* can contribute to breast cancer treatment. A total of 1005 items were identified, divided into three categories: 920 biological processes (BP), 50 molecular functions (MF), and 35 cellular components (CC) (Appendix 2, 3, and 4). The top 10 to 20 enriched BP, MF, and CC terms are displayed on figure 10A to 10C. The ten most highly enriched BP terms include animal organ regeneration, regeneration, response to an inorganic substance, positive regulation of cell population proliferation, cellular response to reactive oxygen species, cellular response to oxygen-containing compound, muscle cell proliferation, regulation of DNA-templated transcription, cell cycle G1/S phase transition, and response to an organic cyclic compound. The observation is that the targets are involved in biological processes related to cell proliferation. Highly enriched MF terms include nitric-oxide synthase regulator activity, Ubiquitin-like protein ligase binding, protein kinase binding, cyclin-dependent protein serine/threonine kinase regulator activity, and other terms, confirming the role of the targets in cell proliferation. The CC terms which include pronucleus, cyclin-dependent protein kinase holoenzyme complex, kinase complexes, spindle complexes, also consolidate the fact that the ten hub targets play important role in cell proliferation.

The KEGG pathway enrichment analysis revealed that the targets were highly enriched for 81 pathways (Appendix 5). The top 20 highly enriched pathways were filtered and shown in figure 11. It is noteworthy that the targets are involved in numerous pathways that are found in cancer. Table 6 contains the related information of the top 20 pathways. The diagrams of selected 3 top pathways are shown in figure 12A to 12C. These results suggest that the targets are involved in

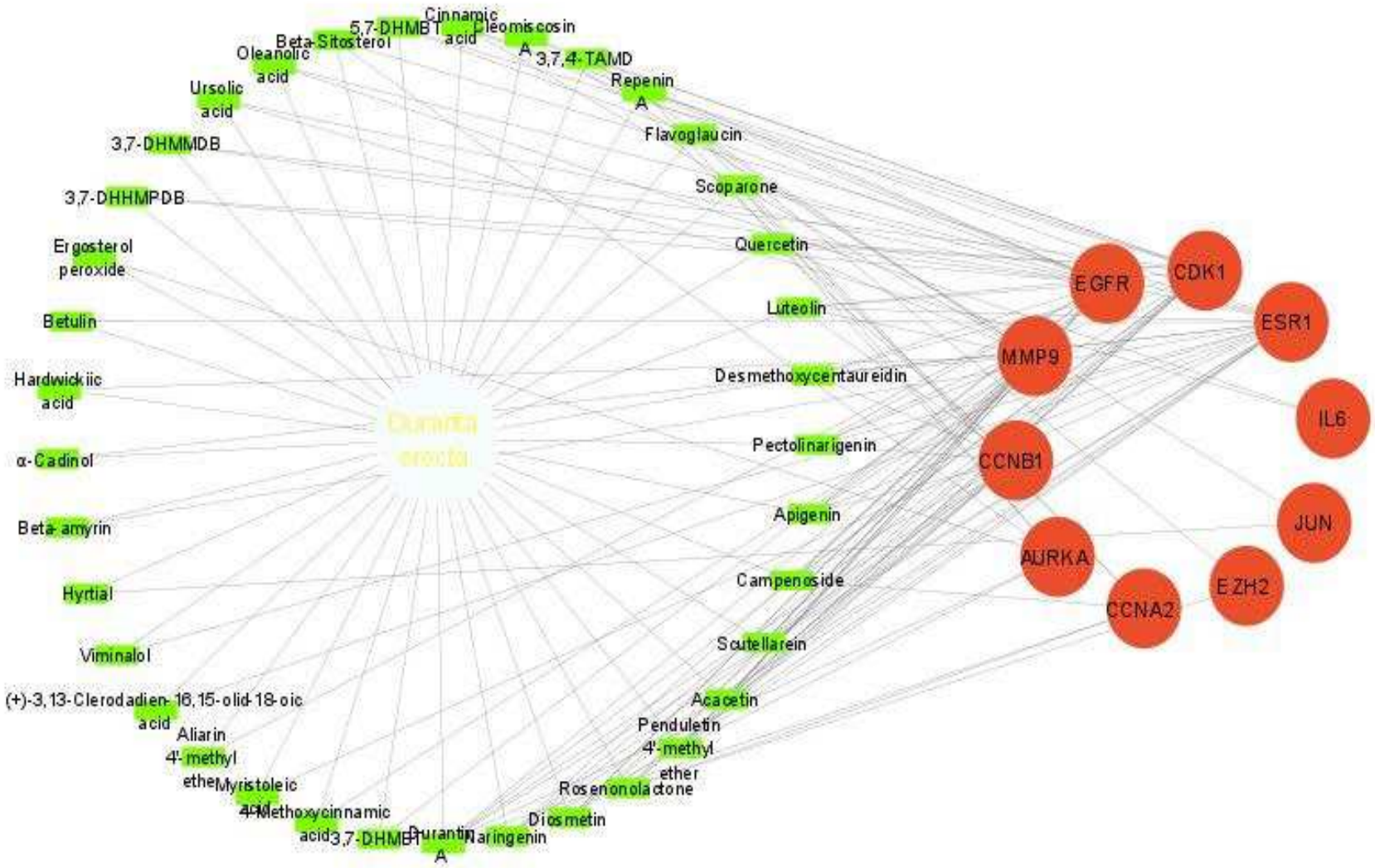


Figure 9: Chemical compound-Target Network

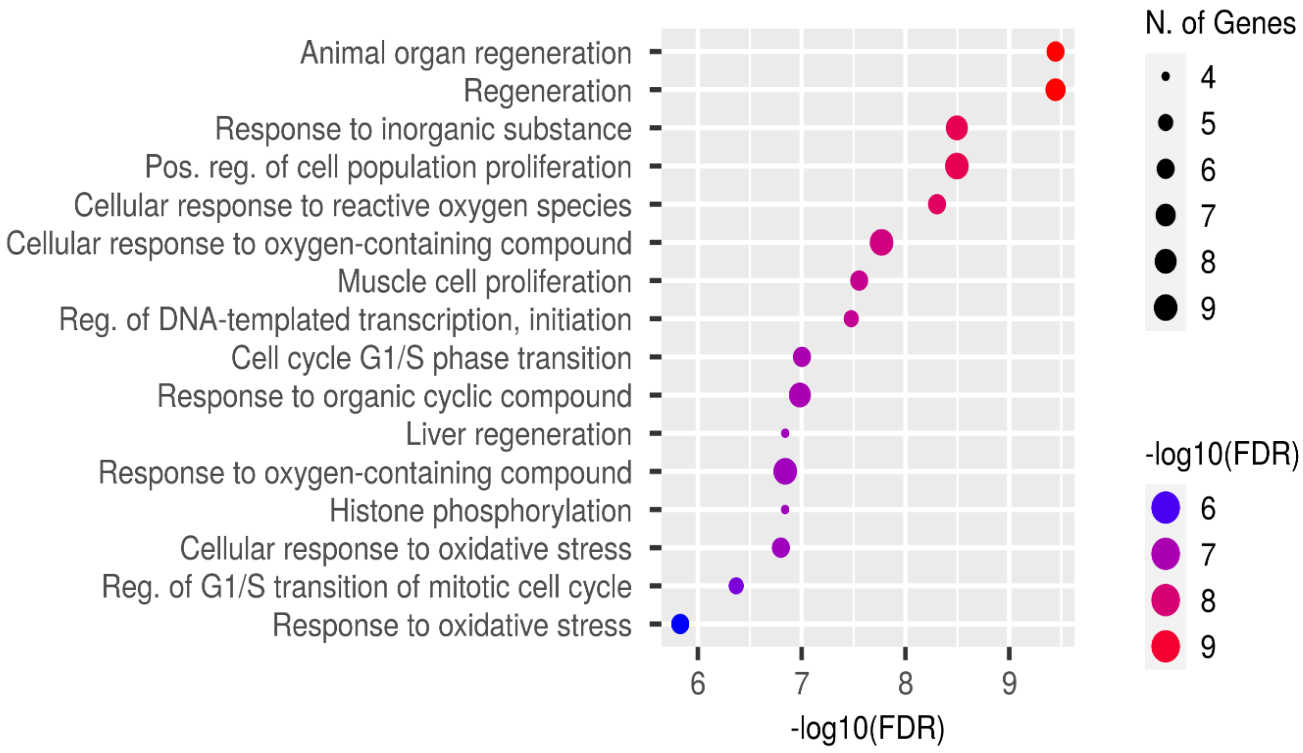


Figure 10A: GO Enrichment Analysis Chart for Biological Process Terms

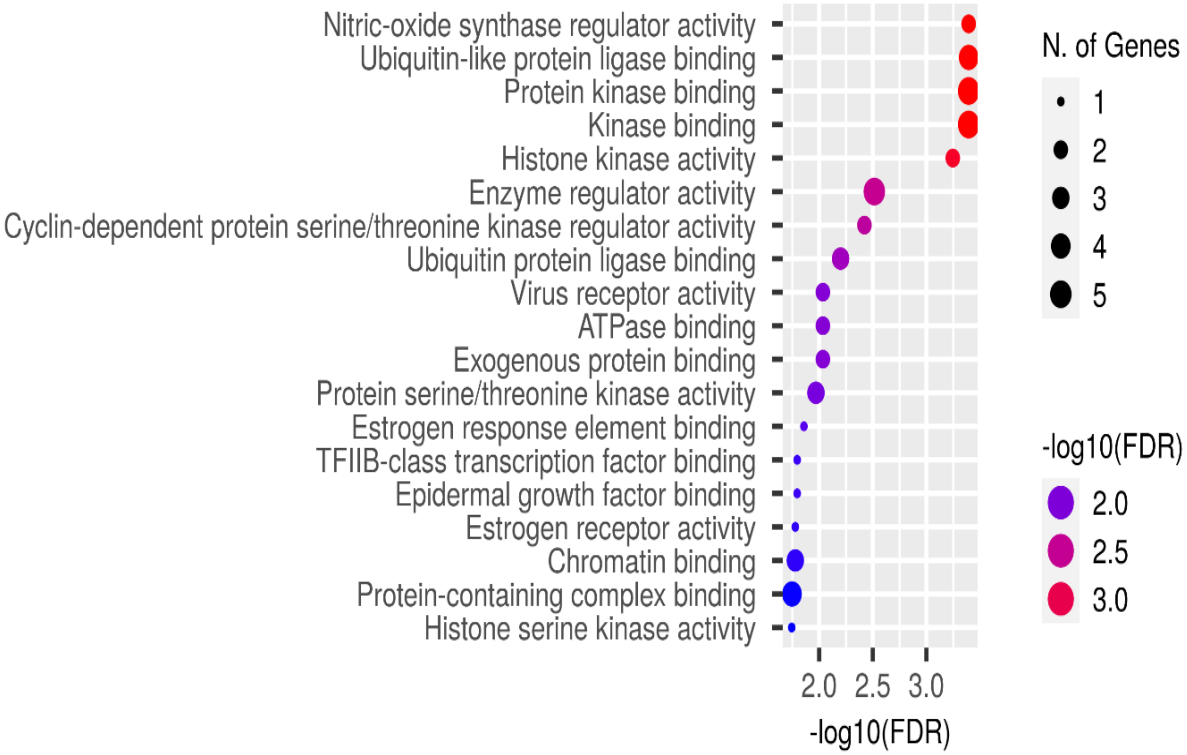


Figure 10B: GO Enrichment Analysis Chart for Molecular Function Terms

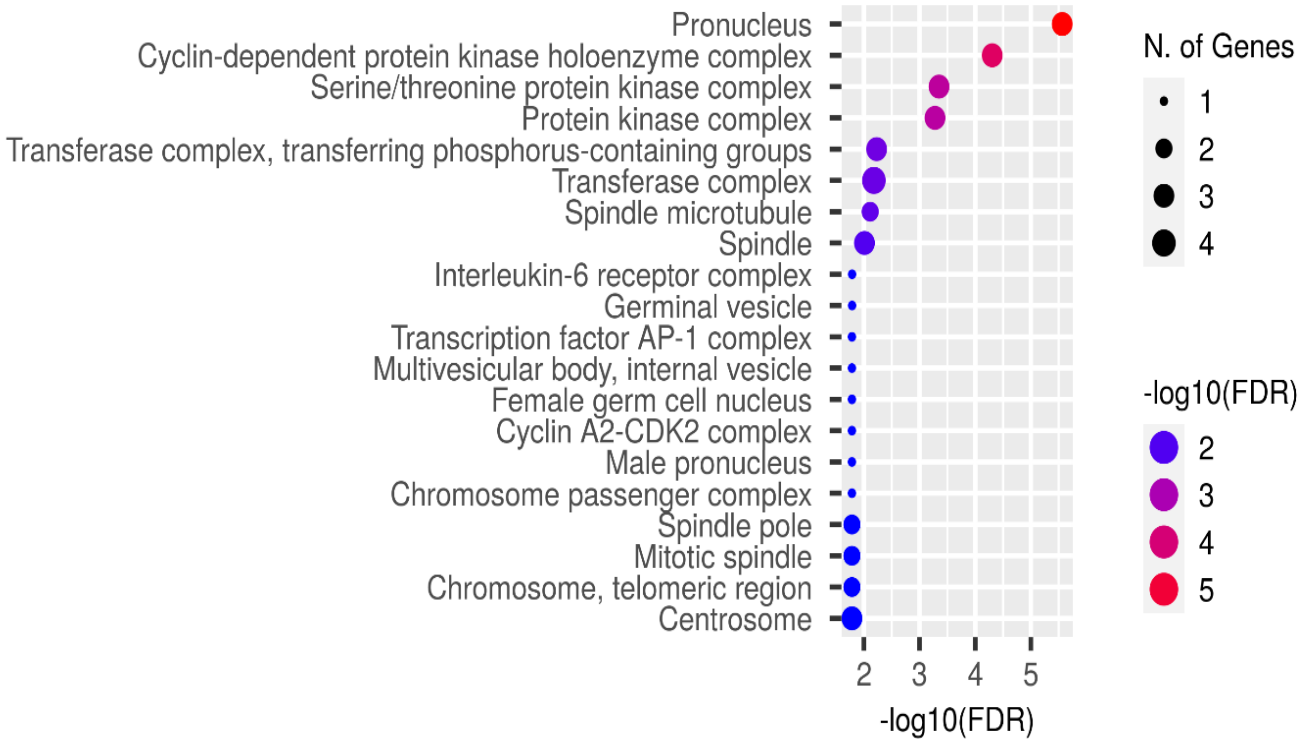


Figure 10C: GO Enrichment Analysis Chart for Cellular Components Terms

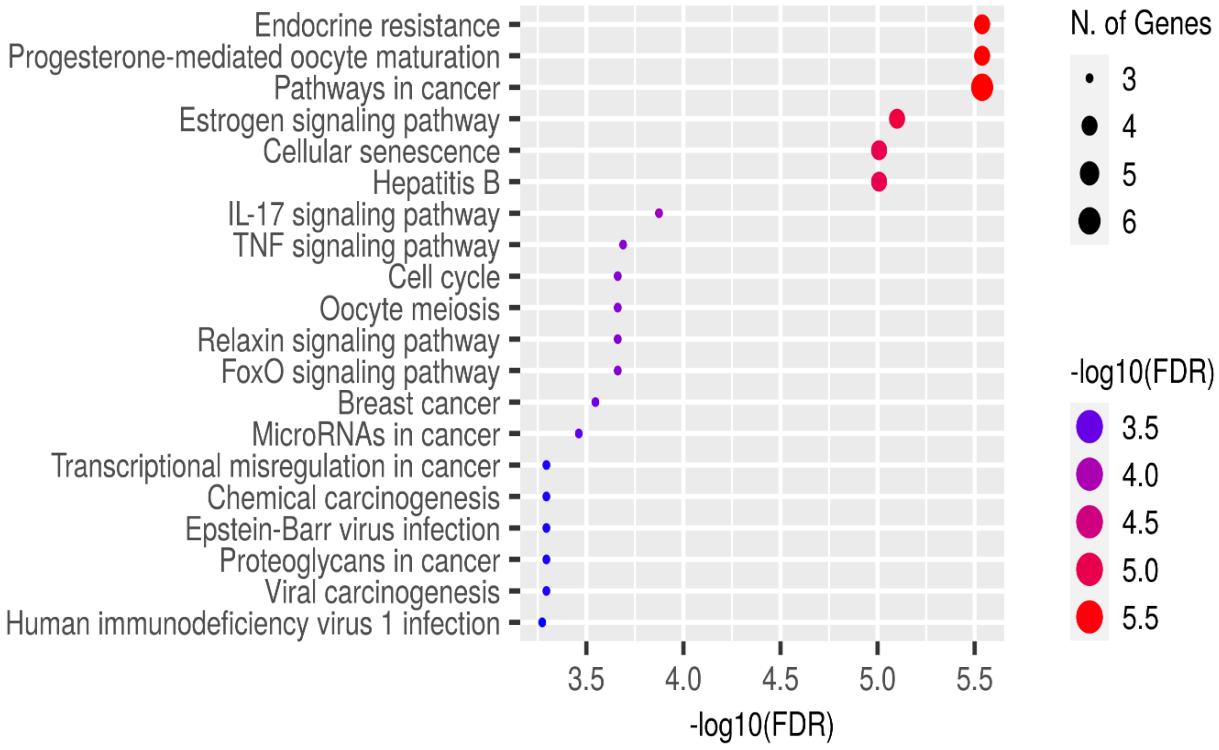


Figure 11: KEGG Enrichment Analysis Chart

Table 6: KEGG enrichment results of key genes

Pathway	**Genes Count**	**Genes**	**Enrichment FDR**
Endocrine resistance	4	ESR1 MMP9 EGFR JUN	2.89E-06
Progesterone-mediated oocyte maturation	4	AURKA CCNB1 CCNA2 CDK1	2.89E-06
Pathways in cancer	6	ESR1 MMP9 IL6 CCNA2 EGFR JUN	2.89E-06
Estrogen signaling pathway	4	ESR1 MMP9 EGFR JUN	7.94E-06
Cellular senescence	4	CCNB1 IL6 CCNA2 CDK1	9.82E-06
Hepatitis B	4	MMP9 IL6 CCNA2 JUN	9.82E-06
IL-17 signaling pathway	3	MMP9 IL6 JUN	0.000133552
TNF signaling pathway	3	MMP9 IL6 JUN	0.00020435
FoxO signaling pathway	3	CCNB1 IL6 EGFR	0.000217896
Cell cycle	3	CCNB1 CCNA2 CDK1	0.000217896
Oocyte meiosis	3	AURKA CCNB1 CDK1	0.000217896
Relaxin signaling pathway	3	MMP9 EGFR JUN	0.000217896
Breast cancer	3	ESR1 EGFR JUN	0.000283868
MicroRNAs in cancer	3	MMP9 EZH2 EGFR	0.000345805
Epstein-Barr virus infection	3	IL6 CCNA2 JUN	0.000507796
Transcriptional misregulation in cancer	3	MMP9 IL6 CCNA2	0.000507796
Viral carcinogenesis	3	CCNA2 CDK1 JUN	0.000507796
Proteoglycans in cancer	3	ESR1 MMP9 EGFR	0.000507796
Chemical carcinogenesis	3	ESR1 EGFR JUN	0.000507796
Human immunodeficiency virus 1 infection	3	CCNB1 CDK1 JUN	0.000533452

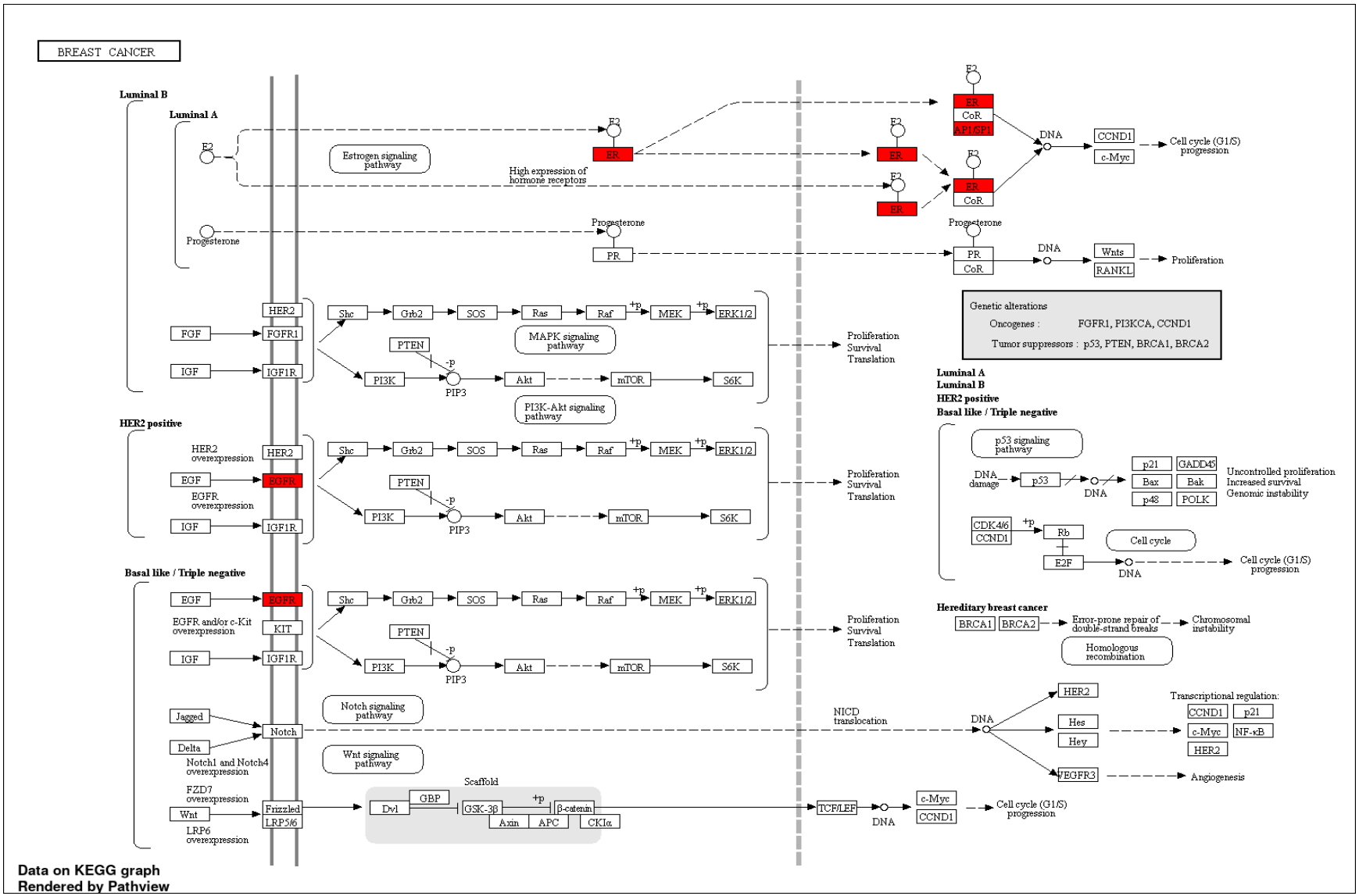


Figure 12A: Diagram of “Breast Cancer” KEGG Pathway. The hub genes involved are in red

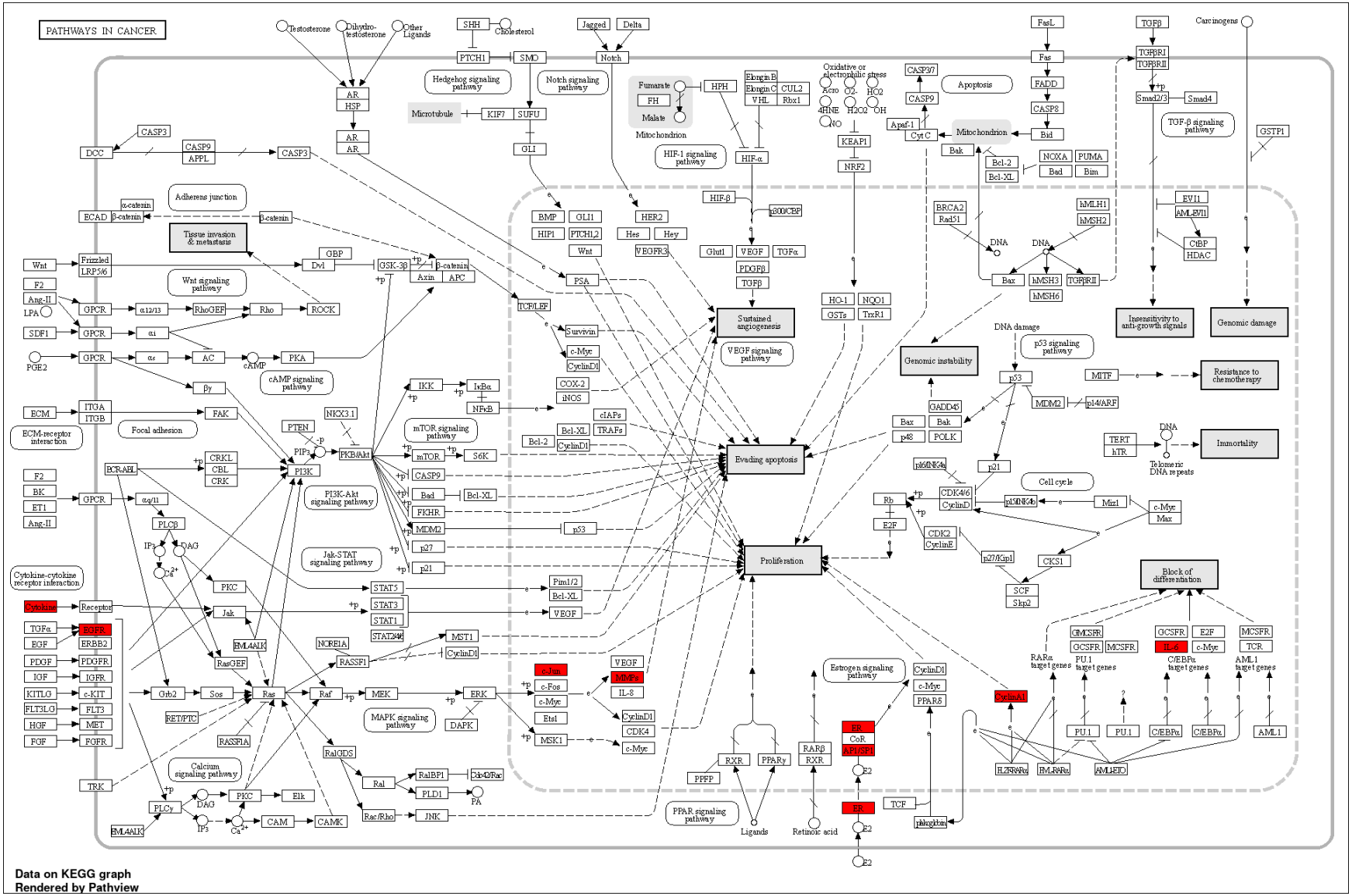


Figure 12B: Diagram of “Pathways in Cancer” KEGG Pathway. The hub genes involved are in red

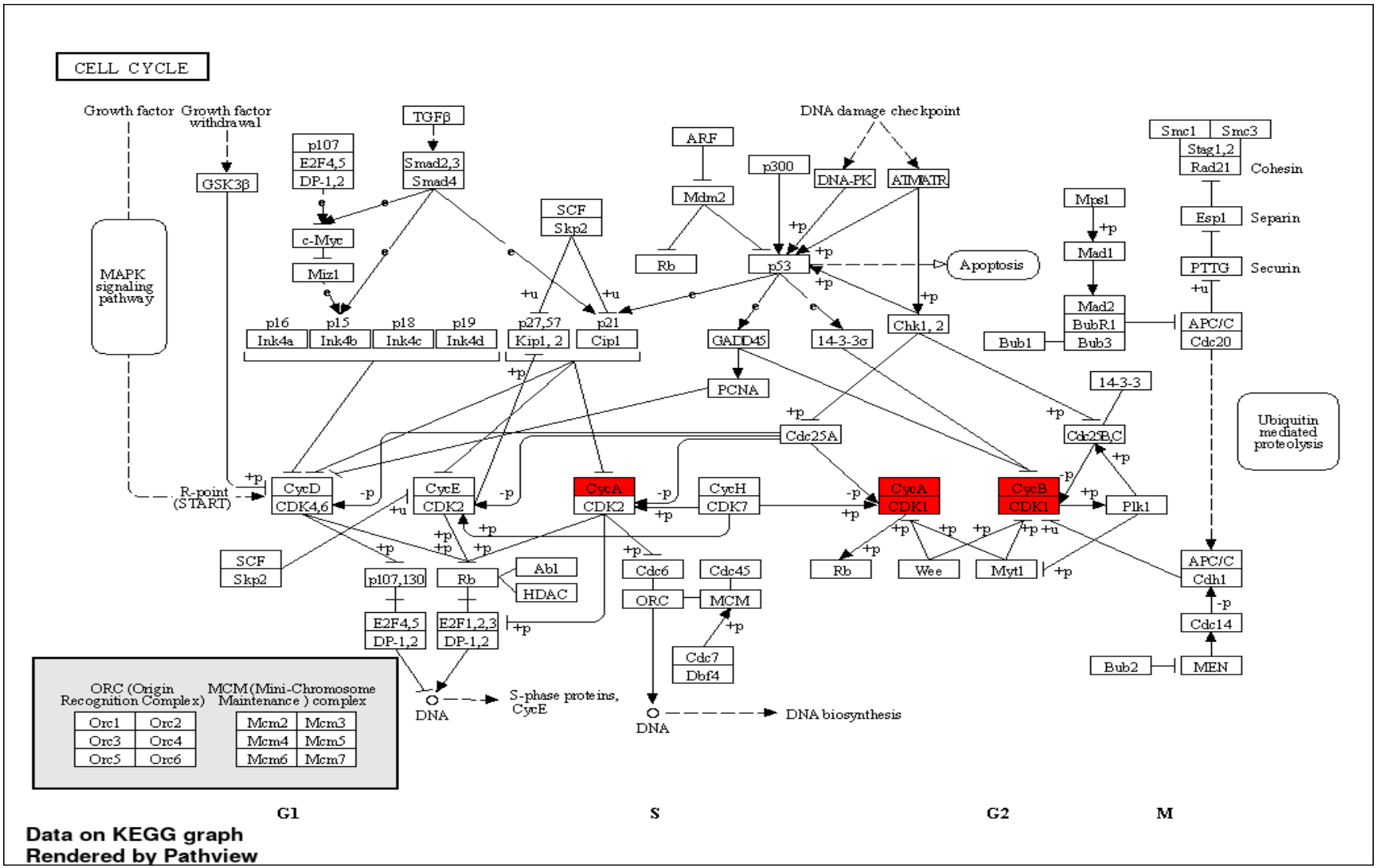


Figure 12C: Diagram of "Cell Cycle" KEGG Pathway. The hub genes involved are in red

Cell proliferation in cancer. Therefore, it is justifiable to say that *Duranta erecta*'s compounds can treat breast cancer by tackling pathways that are involved in cancer cell proliferation through the modulation of the expression of the implicated targets.

3.8. Molecular Docking

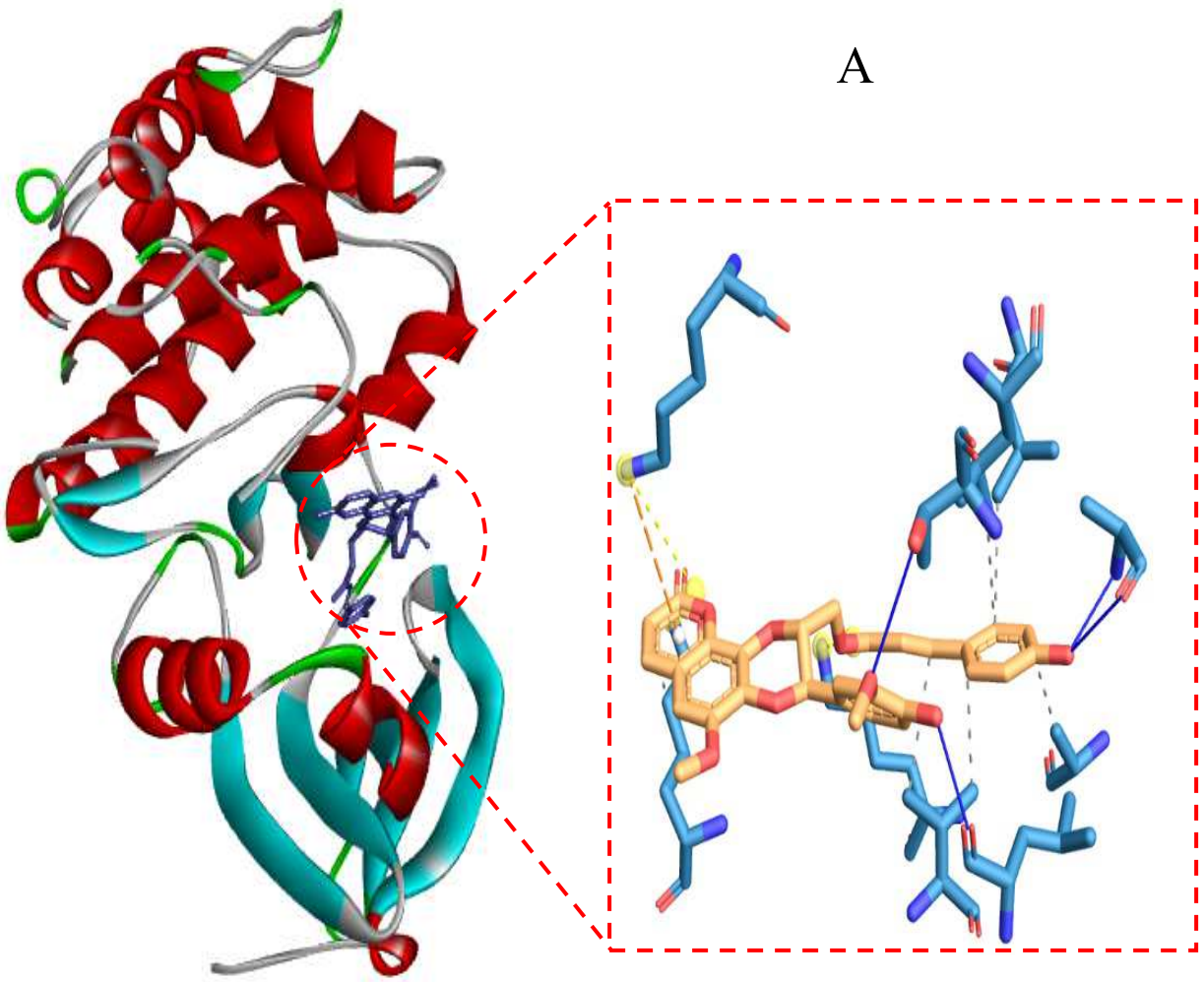
The PPI network's top ten targets were docked with the chemical compounds. According to the findings, they had favorable binding activities because their binding energies were less than 5 kcal/mol [24]. According to the docking results, the compounds from *Duranta erecta* generally performed well with the key targets. Stable receptor-ligand complexes were created as a result of interactions involving hydrogen bonds, hydrophobic bonds, and pi-interactions between the compounds and the targets. The docking results of the compounds with highest binding affinity for each target are found in table 7 while the complete binding energy information is found in appendix 6. The docking models are depicted on figure 13 (A-K). The docking model of repenin A with AURKA was the complex with the highest binding energy (-9.5 kcal/mol). Repenin A is the compound that occupies the first three place in terms of highest binding energy. It is followed by campenoside which occupied to places. Repenin A interacts with AURKA by forming hydrogen bonds with LEU139, ALA213, ALA213, and THR217 of chain A, hydrophobic bonds with LYS143, VAL147, VAL147, ALA160, LEU194, and LEU263 of chain A, and pi-interaction with LYS258 of chain A.

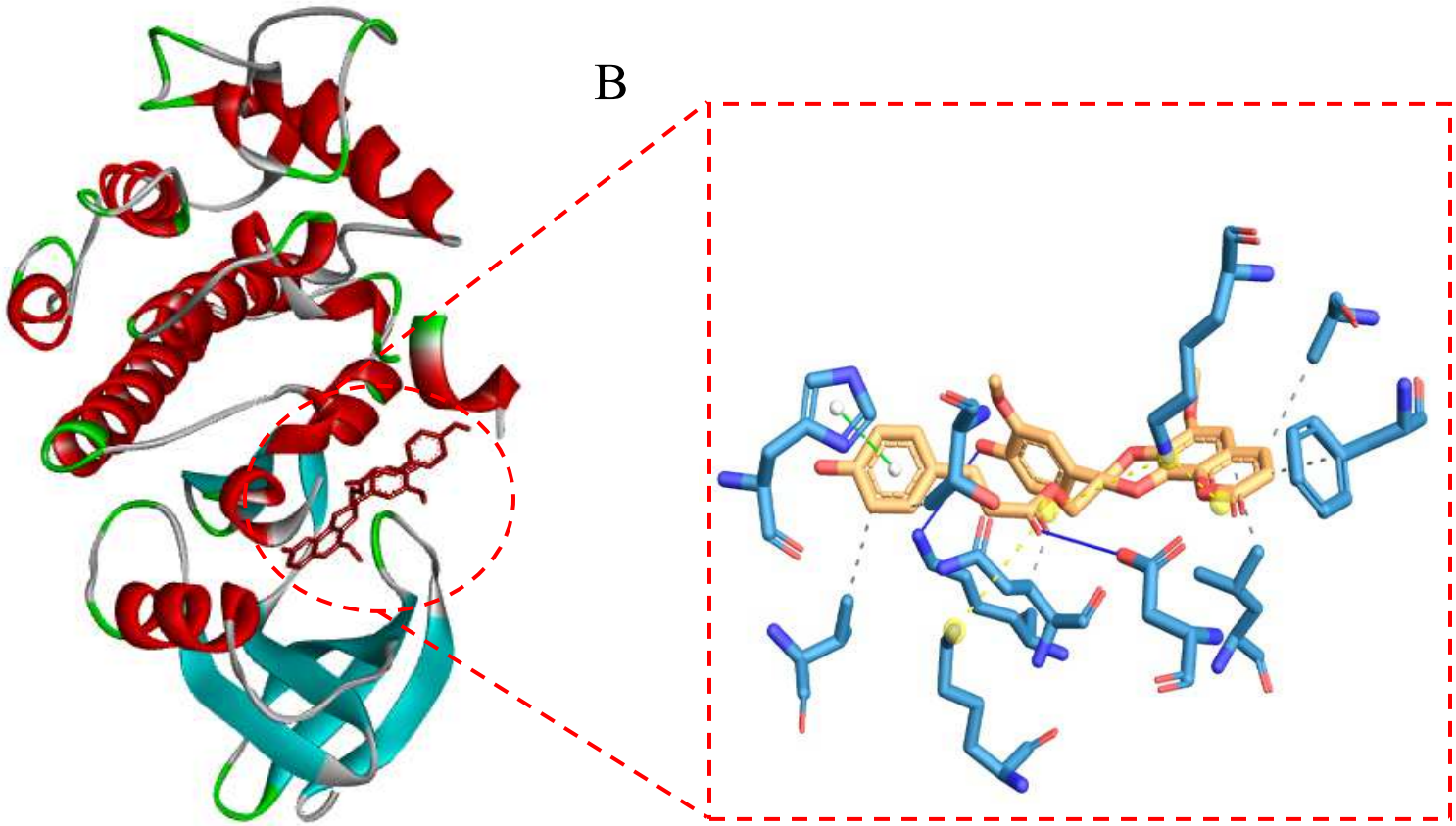
Table 7: Docking Information of Compounds with Highest Binding Affinity for Each Target

Target	Compound	Binding energy (Kcal/mol)	Center coordinates (x, y, z)	Hydrogen bonds	Hydrophobic interactions	Pi interactions
AURKA	Repenin A	-9.5	4.7645 69.3352 65.8161	LEU139A ALA213A ALA213A THR217A	LYS143A VAL147A VAL147A ALA160A LEU194A LEU263A	LYS258A
CDK1	Repenin A	-9	-2.2898 35.0817 -6.254	LYS88A ASP146A	THR14A ALA31A PHE80A GLN132A LEU135A VAL165A	HIS162A
EGFR	Repenin A	-9	-7.3583 59.0362 -23.9963	MET793A GLY796A	PHE723A VAL726A ALA743A LEU792A MET793A LEU844A	PHE723A
MMP9	Scutellarein	-8.9	28.2222 16.6024 37.3648	GLU402A GLU416A ARG424A HIS432A	LEU188A LEU397A VAL398A LEU418A TYR423A	-
ESR1	Clerodadien	-8.8	-0.2608 -8.5307 22.2421	GLU353B	LEU346B LEU346B ALA350B LEU387B LEU391B PHE404B ILE424B	-
CCNA2	Campenoside	-7.8	50.2431 62.556 -0.4011	LEU253B GLN254B	VAL221B GLU224B TYR225B ILE281B ILE281B	-
CCNB1	Campenoside	-7.6	-65.8746 41.523 -11.3405	ALA20A ARG68A ARG68A GLN71A ASN72A ASN72A	LEU17B LEU17B ASN130A ARG135B	-

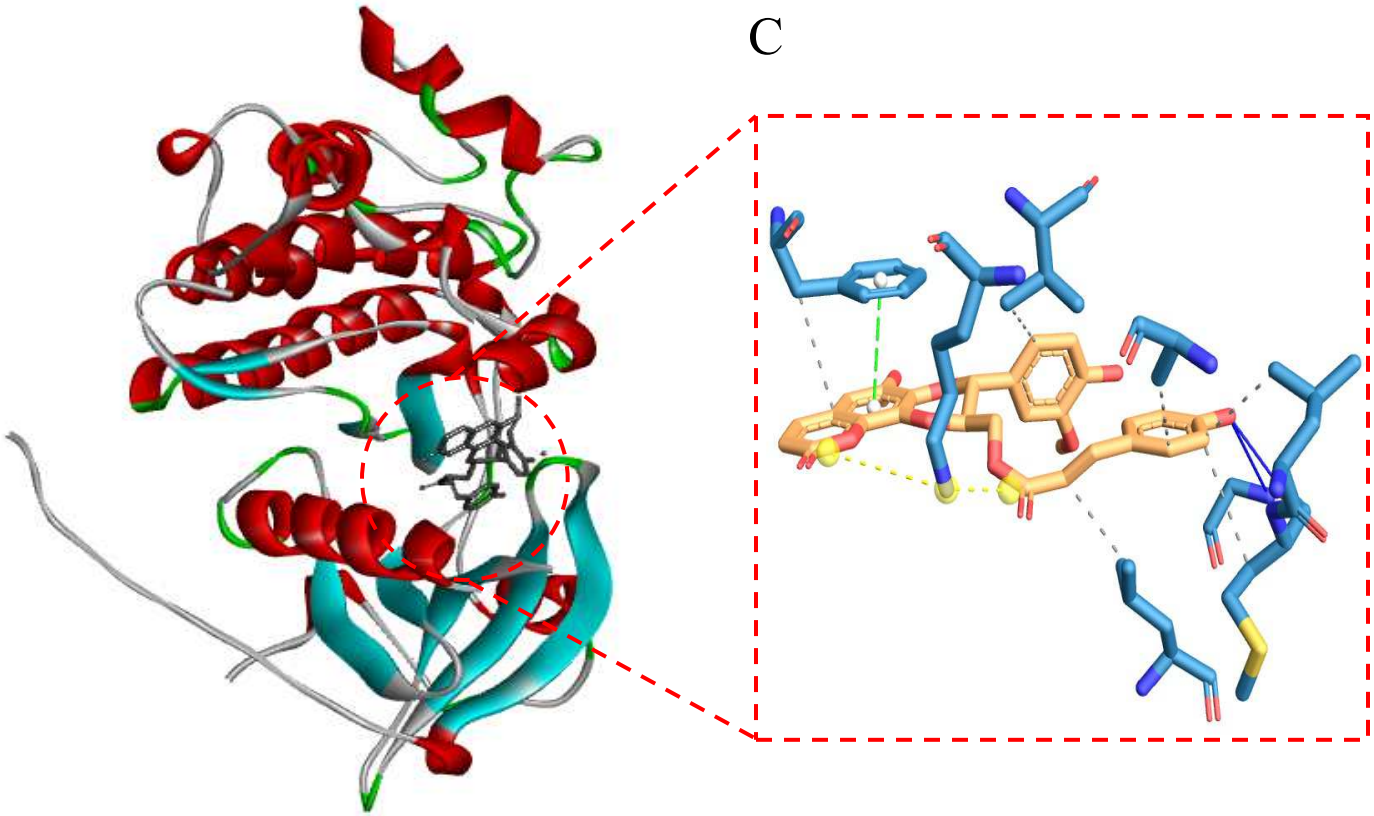
IL6	Oleanolic acid	-7.3	2.6711 -20.0828 8.9102	ARG182A	LEU33A LEU33A LEU33A LYS171A LEU178A LEU178A ARG182A	-
IL6	Ursolic acid	-7.3	2.6711 -20.0828 8.9102	ASN63A	ASN63A GLU93A PRO139A THR143A LEU147A LEU147A	-
EZH2	Rosenonolactone	-6.8	27.8952 21.9713 10.4163	PHE547A GLU549A	PHE547A PHE547A ARG583A VAL704A MET706A	-
JUN	Hyrtial	-6.1	24.9324 5.6737 -14.1387	LYS292F	GLU284F ALA287F ARG288F ARG288F GLU291F GLU291F LYS292F	-

A

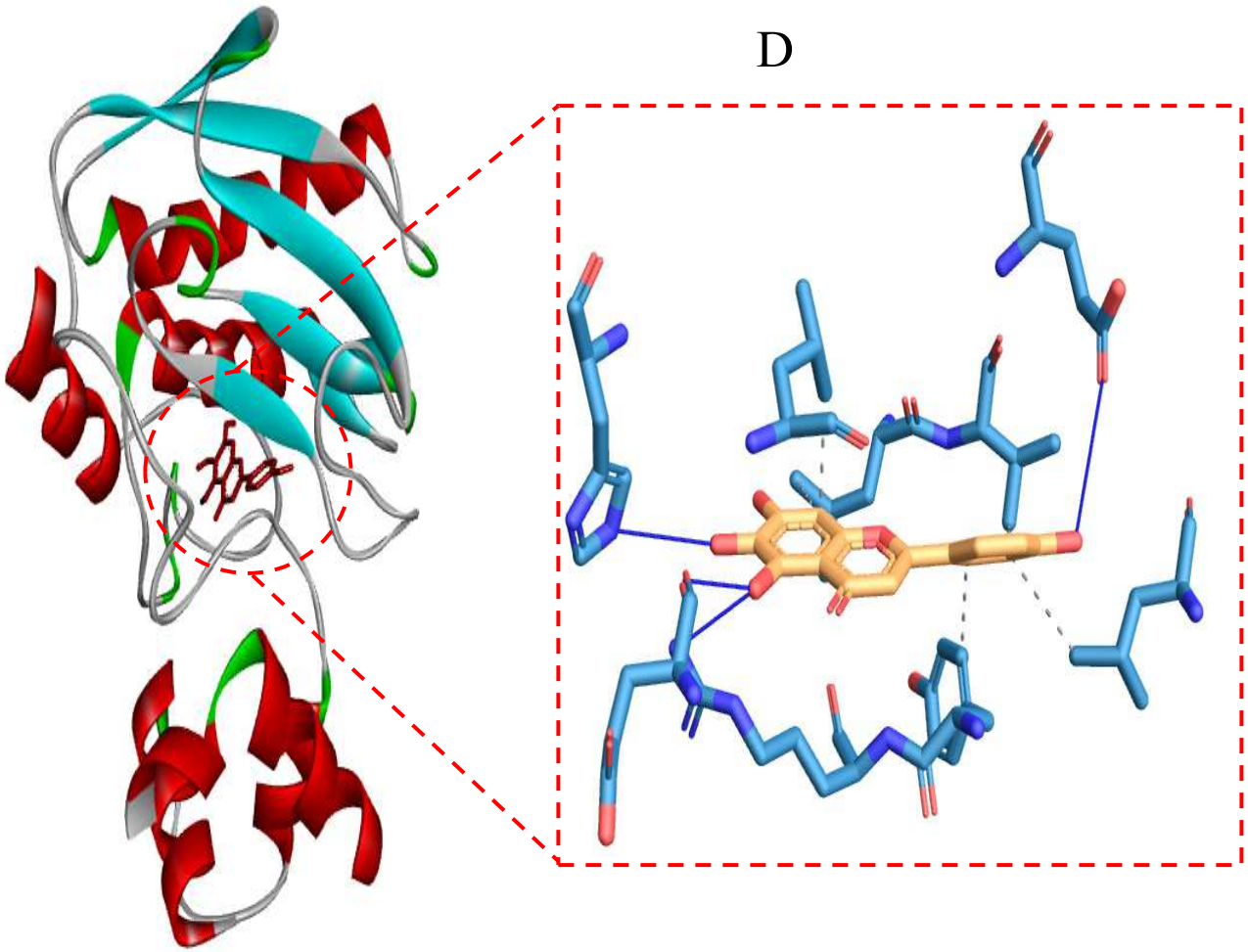



B

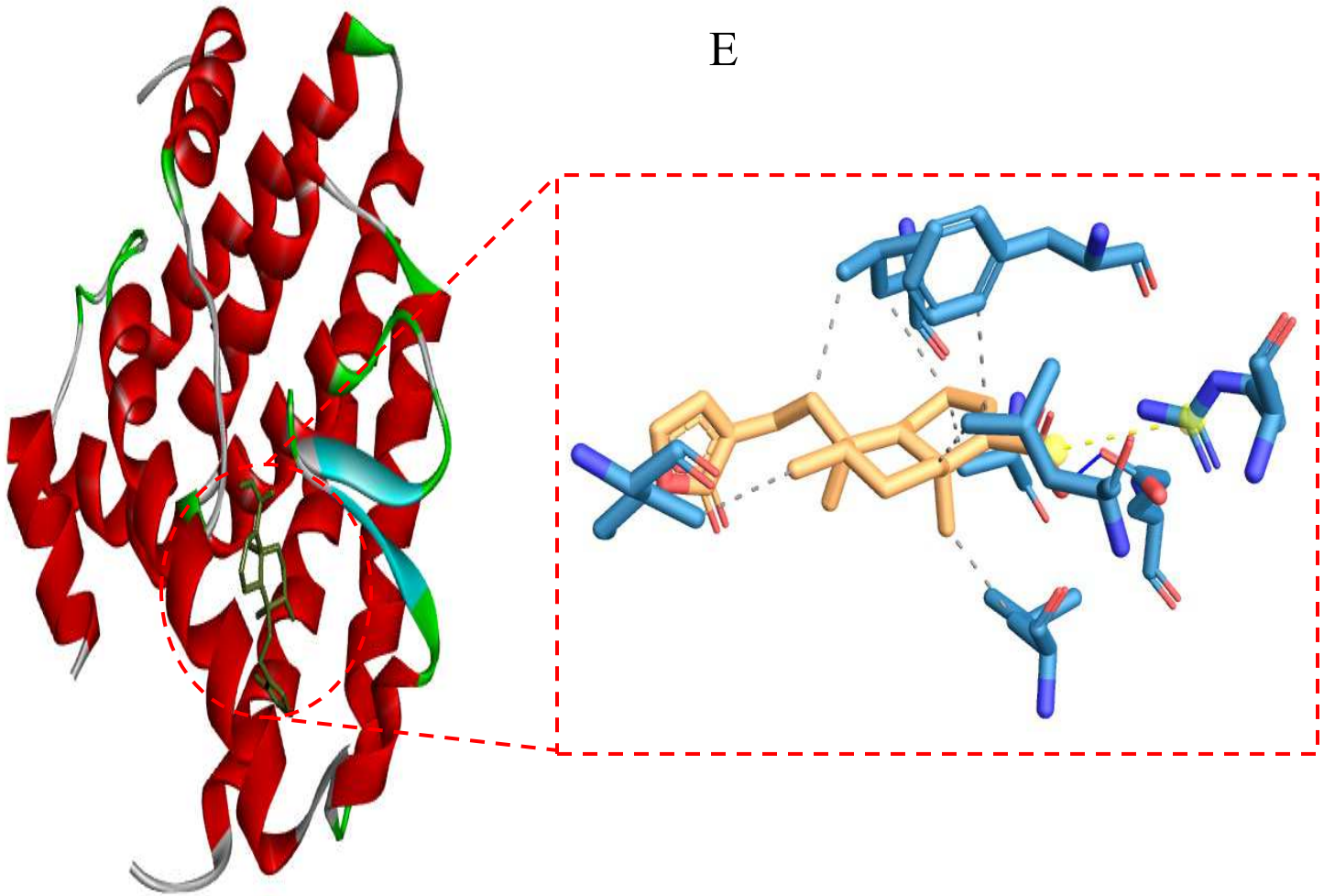
C

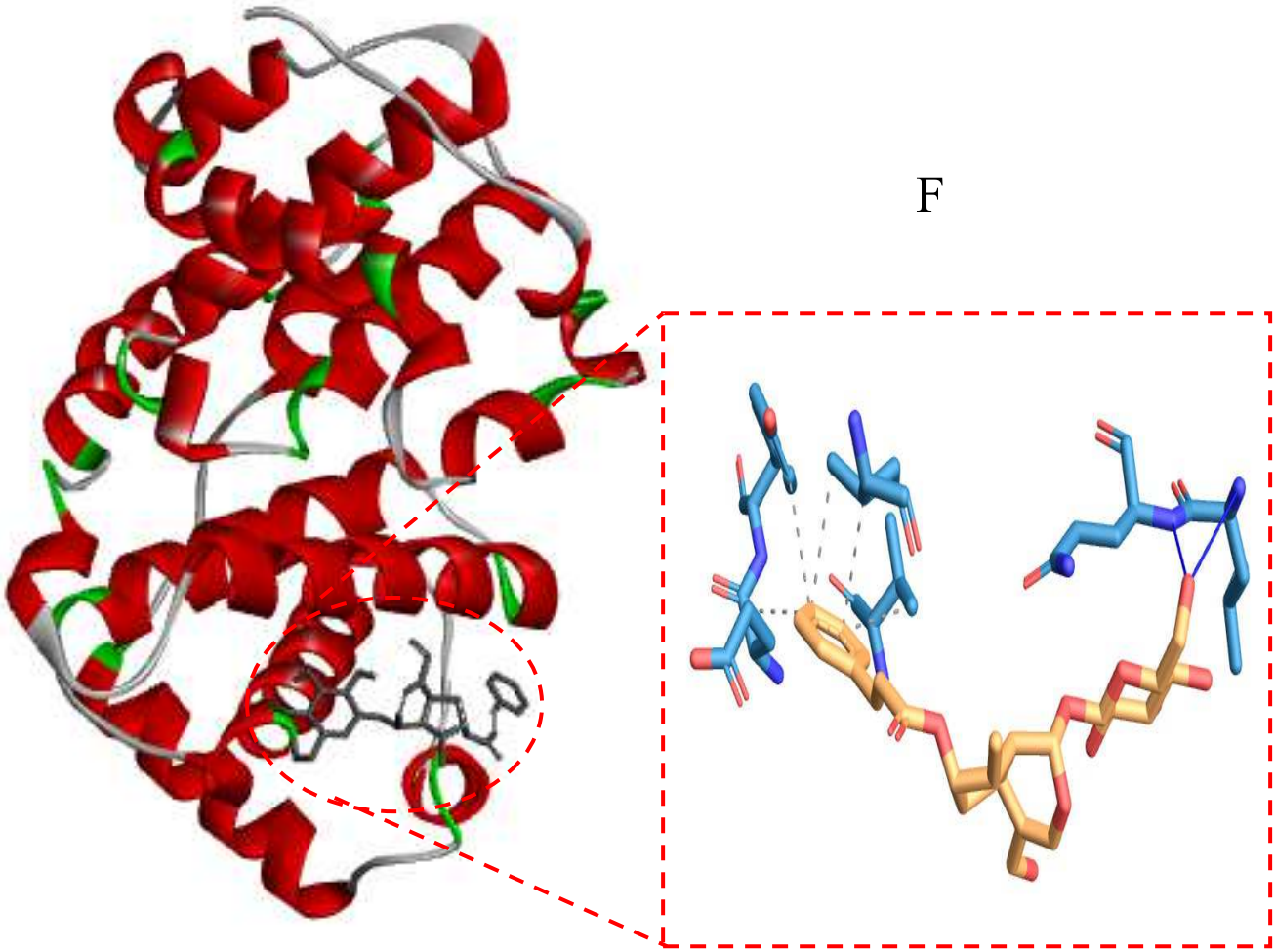


D



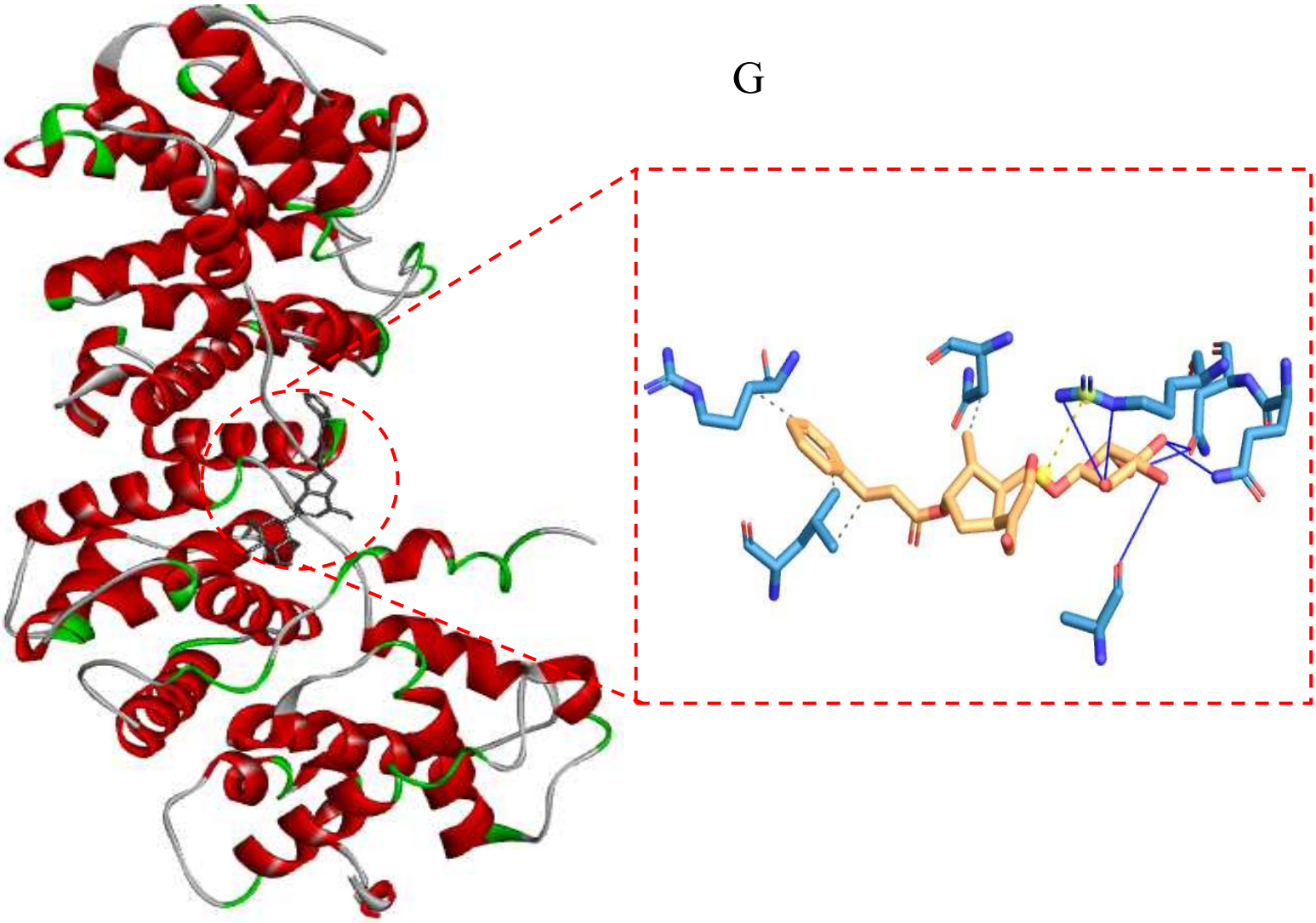
E



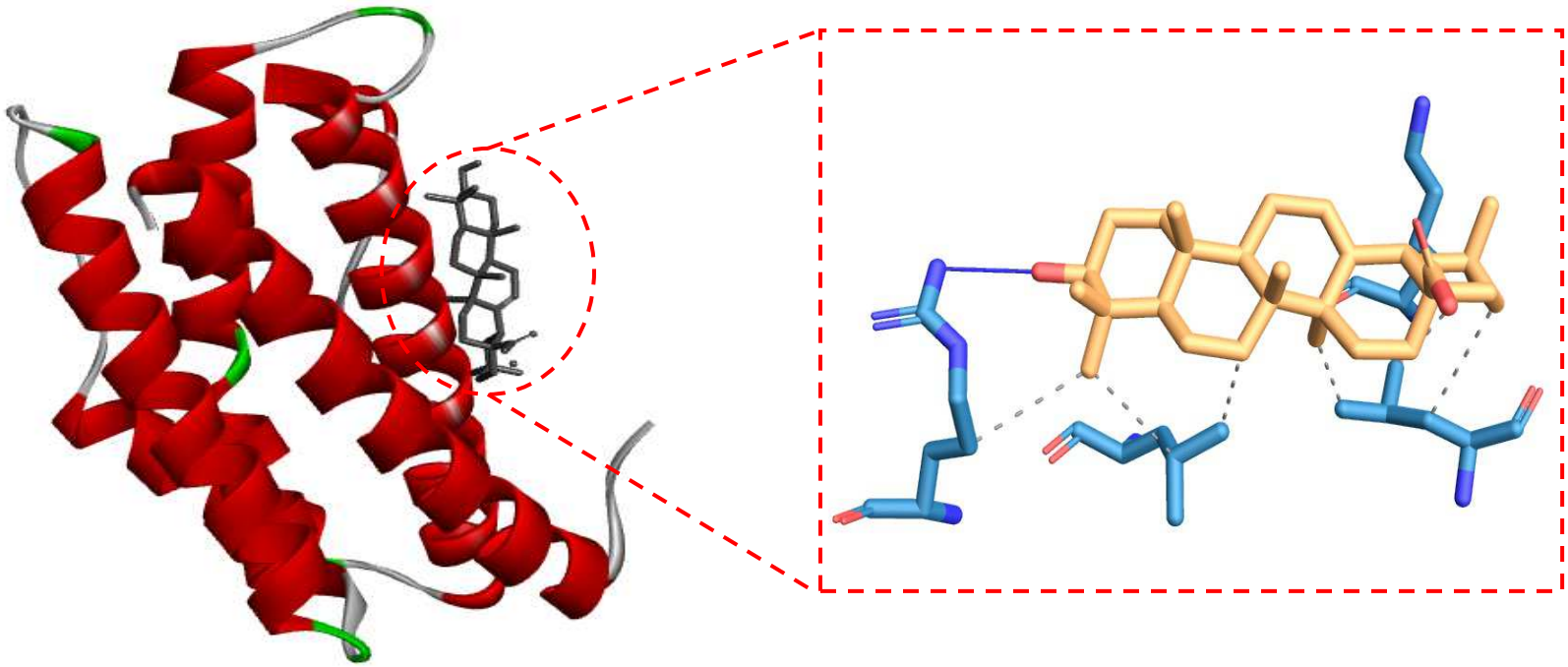


F

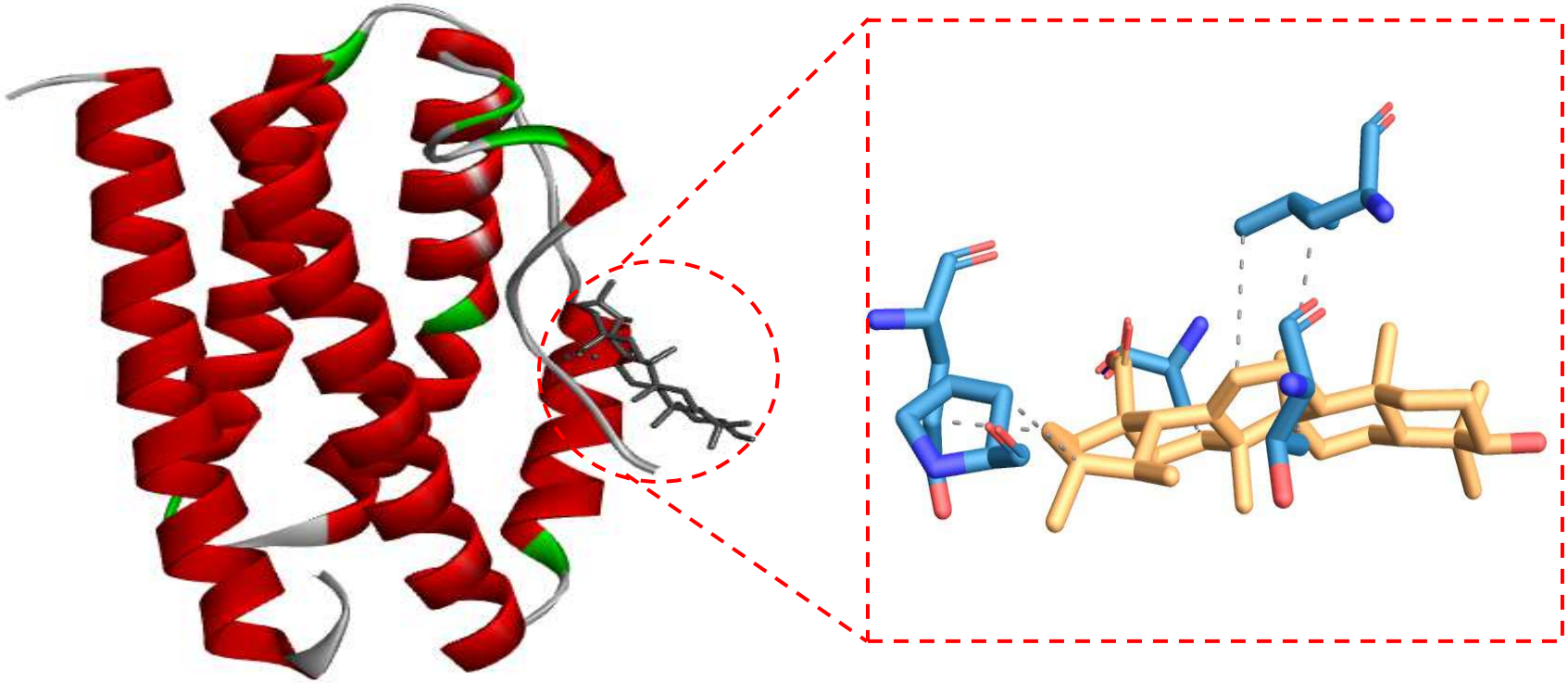
G



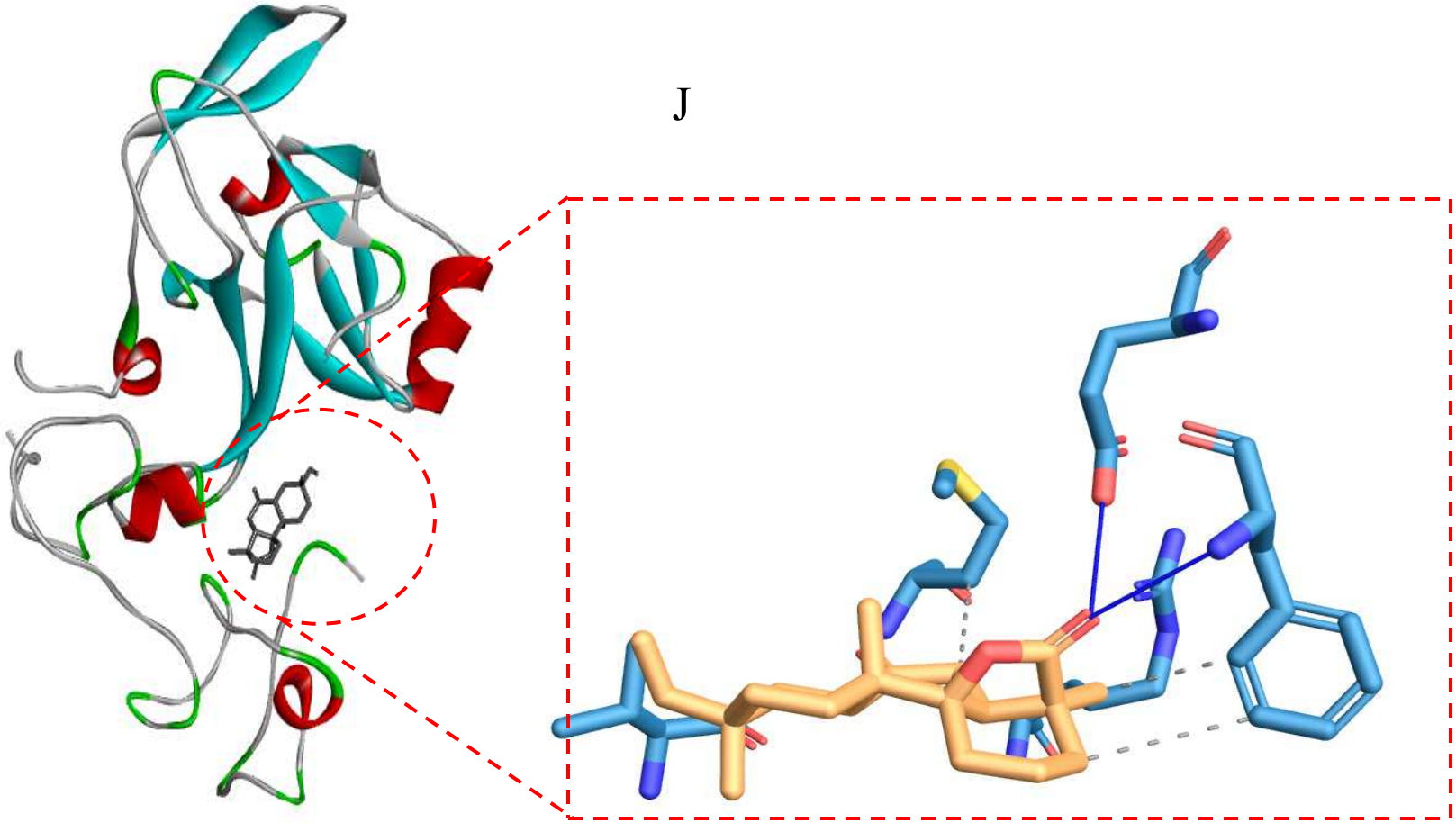
H



I



J



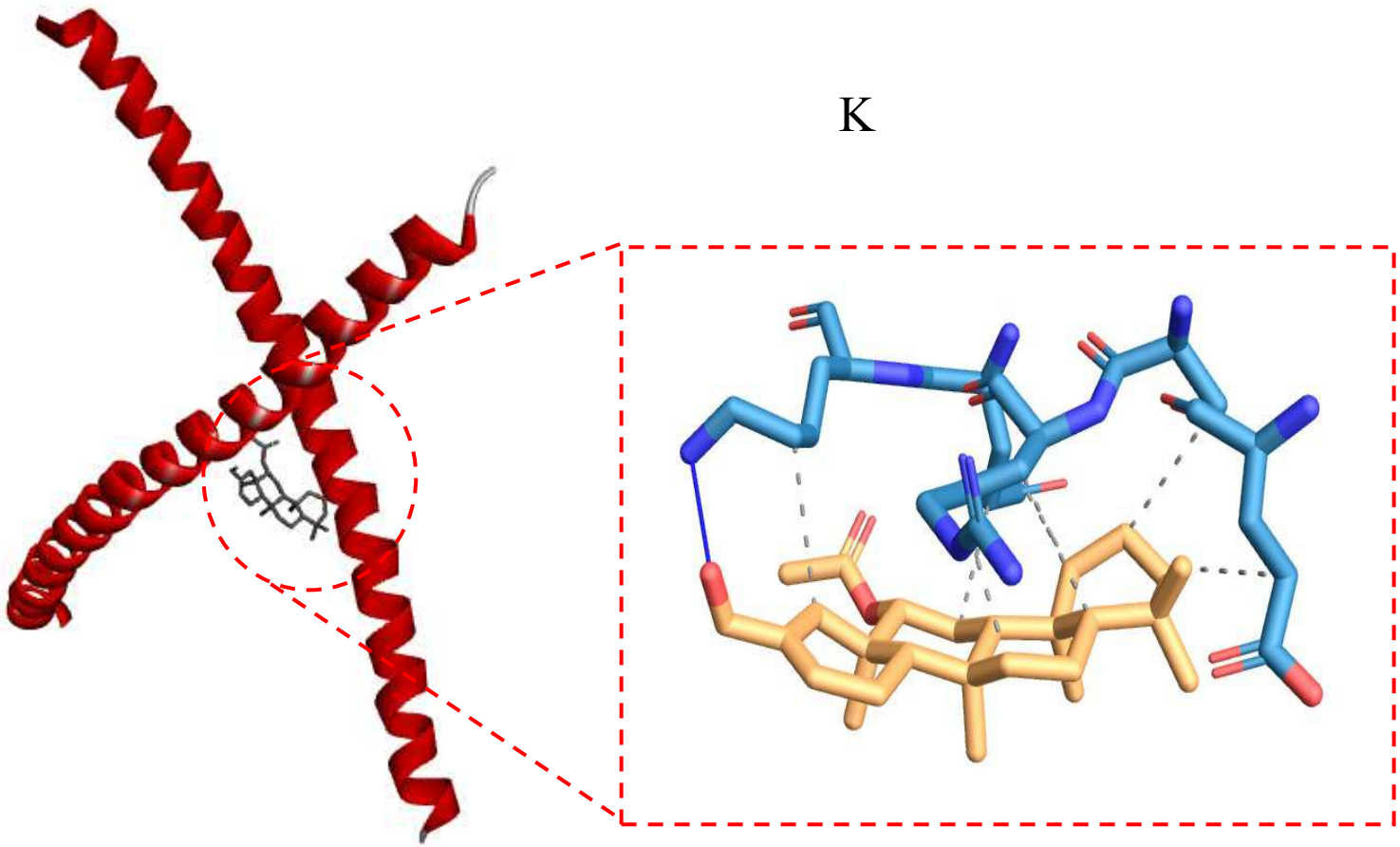
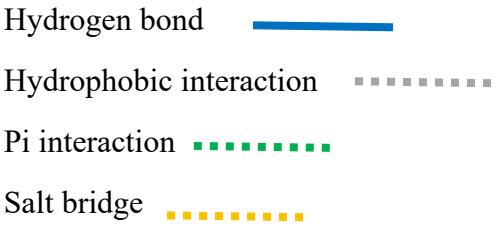

K

Figure 13: The docking models of compounds with highest binding affinity for each target. (A) Repenin A/AURKA complex; (B) Repenin A/CDK1 complex; (C) Repenin A/EGFR complex; (D) Scutellarein/MMP9 complex; (E) Clerodadien/ESR1 complex; (F) Campenoside/CCNA2 complex; (G) Campenoside/CCNB1 complex; (H) Oleanolic acid/IL6 complex; (I) Ursolic acid/IL6 complex; (J) Rosenonolactone/EZH2 complex; (K) Hyrtial/JUN complex.



4. Discussion

The medicinal properties of *Duranta erecta* have been proven for different ailments, including cancer, based on *in vitro*, *in vivo*, and *in silico* experiments [25–30]. The plant has not been investigated extensively in breast cancer and the molecular mechanism of action is unknown. Therefore, this study investigated the anticancer property, with keen interest in the molecular mechanism, of *Duranta erecta* on breast cancer leveraging on computational approaches. In parallel, discovery of candidate drugs from the plant were made.

Leading pharmaceutical companies and researchers have used computer-aided drug discovery (CADD) techniques in preliminary studies to help accelerate the drug discovery and development process while minimizing costs and failures in the final stage. The use of rational drug design as part of CADD provides valuable insights into the binding affinity and molecular interaction between target protein and ligand. Pharmacophore modeling, quantitative structure-activity relationship (QSAR), quantum mechanics, and molecular docking are some of the methods used to identify new inhibitors from chemical databases [31]. In the pipeline of drug discovery, CADD techniques help specifically in hit and lead identification. However, even before lead identification, target identification is key in this process. Target identification is usually done using chemical biology techniques. However, other techniques such as bioinformatics are helping nowadays in target identification [32].

In this study, network pharmacology approach in addition to molecular docking and *in vitro* study were used to identify targets in breast cancer, identify drug candidates from *Duranta erecta* that can act on those targets, hence elucidate the mechanism of action of the plant against breast cancer. The study revealed that *Duranta erecta* extracts exerts cytotoxicity effect on breast cancer cells. It revealed several drug-likened compounds from *Duranta erecta* which can modulate the identified ten hub targets. Among all the compounds, repenin A had the first three highest binding affinity (Table 7). The best binding complex was repenin A/AURKA (-9kcal/mol), followed by repenin A/CDK1 and repenin A/EGFR. Campenoside had the highest binding affinity for two targets, namely CCNA2 (-7.8 kcal/mol) and CCNB1 (-7.6 kcal/mol). The compounds repenin A and campenoside are most likely the most efficacious drug candidates, targeting the mentioned protein targets, based on these findings. They showed the lowest binding energies for more than one target while other compounds were ranked first for only one target. Another thing is that some compounds bound to more targets, five or more, but failed to rank

first for any. For instance, durantin A was predicted to bind to six targets but failed to rank first for any of them. These findings are consistent with previous similar studies that used the same approaches to attain the same objectives; that is using computational approaches for drug discovery [33–35].

As demonstrated by the results of the gene ontology analysis and KEGG pathway analysis (figure 9 and 10), the ten hub targets are involved in the pathways in cancer and favor cell proliferation. Interestingly, sustaining cell proliferation is a hallmark of cancer [36,37]. AURKA has been shown to be an oncogenic gene, involved in the progression of different types of cancer including breast cancer. In clear cell renal cell carcinoma, the gene was shown to progress tumor growth by modulating FOXO3A and SKP2 genes [38]. It contributes to genomic stability by maintaining a functional G2/M checkpoint. A study showed that inhibition of AURKA leads to cell cycle arrest in G2/M phase and subsequent apoptosis [39]. Another study in human colorectal cancer depicted similar findings and showed that AURKA knockdown significantly increases the effectiveness of radiation on cell proliferation [40]. The same results were observed in non-small cell lung carcer [41]. In breast cancer, AURKA is associated with poor prognosis and drive the disease progression [42,43]. This is evidence that AURKA protein is involved in breast cancer and most importantly it druggable [39,44]. Fortunately, the findings of the present study are consistent with previous studies. AURKA was discovered in this study to be a target protein in breast cancer, and it had a strong binding affinity with repenin A and other compounds from *Duranta erecta*.

The proteins CCNA2 and CCNB1 are cyclins and CDK1 is a cyclin dependent kinase. They are also regulators of the cell cycle, and subsequently cell proliferation. Previous studies confirmed that they also play an oncogenic role in breast cancer [45]. The expression levels of the three genes are high in breast cancer and correlate with advanced tumor stage [46,47]. They have therefore been justified as targetable proteins for the management of the disease. Prominent compounds of the studied plant showed stable conformation with the proteins, consolidating the importance of this work.

EGFR and ESR1 are receptors and are also involved in cell growth and proliferation. Thus, their inhibition is therapeutic strategy to constrain the progression of tumor. A few evidence are found in the literature [48]. For instance, Al-Wahaibi and colleagues synthesized indole-2-carboxamides

and prove its antiproliferative and apoptotic activities through EGFR/CDK2 inhibition [49]. The inhibition of EGFR was also associated with tumor microenvironment modulation in non-small cell lung cancer [50]. The inhibition of ESR1 also represses tumor growth in diverse cancer, including breast cancer [51,52]. In this study it was also revealed that the compounds of *Duranta erecta* have strong binding affinity with these two receptors and thus can modulate their activity to constrain breast cancer. Repenin A showed the highest binding affinity for EGFR (-9 kcal/mol) and (+)-3,13-Clerodadien-16,15-olid-18-oic acid showed the highest binding ability for ESR1 (-8.8 kcal/mol).

The other targets determined as core targets in this study were also supported by previous studies as oncogenic driven proteins. That is the case of EZH2 [53,54], IL6 [55,56], JUN [57], and MMP9 [58,59].

Most of the drug candidates found in *Duranta erecta* in this study have proven their medicinal relevance all through the literature. For instance, a study showed that scutellarein, a flavonoid, affects fibroblast differentiation, proliferation, and apoptosis to inhibit pulmonary fibrosis [60]. Another study showed that it reduces tubular epithelial cell death and inflammation in ischemic kidney injury by degrading the COX-2 protein [61]. In colon and ovarian cancers, it inhibits metastasis and proliferation and induces apoptosis [62,63]. Beta-sitosterol and quercetin were found to enhance brain development and quercetin alleviates oxidative damage [64,65]. Quercetin has also shown pharmaceutical applications in cardiovascular diseases and cancer [66,67]. Scoparone, a coumarin compound, is another compound found in *Duranta erecta* which has a track record of medicinal properties in the literature. It alleviates hepatic fibrosis [68], inhibits cancer progression [69], alters sphingolipid metabolism in nonalcoholic steatohepatitis [70], and has osteogenic effects [71].

5. Conclusion

In conclusion, it is evidenced in this study that *Duranta erecta* has anticancer property against breast cancer. This is possible due to the active chemical compounds it contains, namely repenin A, scutellarein, campenoside, durantin A, and others. The compounds have been proven to be viable drug candidates. They repress breast cancer by inhibiting cell proliferation through targeting proteins such as AURKA, CDK1, CCNA2, and EGFR, which were also discovered in this study. This study is the first report of the *in silico* anticancer property of *Duranta erecta* and its possible the mechanism of action. It sets a basis for future studies that can lead to the approval

of new drug candidates for the treatment of breast cancer patients. The *in vitro* and *in vivo* experiments are expected to confirm the findings of this study, and hopefully clinical studies will follow.

References

1. Bray, F., Parkin, D. M., & African Cancer Registry Network. Cancer in sub-Saharan Africa in 2020: a review of current estimates of the national burden, data gaps, and future needs. *Lancet Oncol.* **23**, 719–728 (2022).
2. Mubarik, S. *et al.* Epidemiological and sociodemographic transitions of female breast cancer incidence, death, case fatality and DALYs in 21 world regions and globally, from 1990 to 2017: An Age-Period-Cohort Analysis. *J. Adv. Res.* **37**, 185–196 (2022).
3. Rweyemamu, L. P. *et al.* Breast cancer in East Africa: Prevalence and spectrum of germline SNV/indel and CNVs in BRCA1 and BRCA2 genes among breast cancer patients in Tanzania. *Cancer Med.* (2022) doi:10.1002/cam4.5091.
4. Okifo, F. O. *et al.* Breast cancer treatment and outcomes at Cape Coast Teaching Hospital, Ghana. *Ghana Med. J.* **55**, 190–197 (2021).
5. Knapp, G. C. *et al.* The out-of-pocket cost of breast cancer care at a public tertiary care hospital in Nigeria: an exploratory analysis. *Pan Afr. Med. J.* **41**, 272 (2022).
6. Ntekim, A., Oluwasanu, M. & Odukoya, O. Breast Cancer in Adolescents and Young Adults Less Than 40 Years of Age in Nigeria: A Retrospective Analysis. *Int. J. Breast Cancer* **2022**, 9943247 (2022).
7. Djiwa, T. *et al.* Histo-Molecular Profile of Breast Cancer in Young Women in Togo. *Clin. Pathol. Thousand Oaks Ventura Cty. Calif* **15**, 2632010X221112452 (2022).
8. Nthontho, K. C. *et al.* Pharmacogenetics of Breast Cancer Treatments: A Sub-Saharan Africa Perspective. *Pharmacogenomics Pers. Med.* **15**, 613–652 (2022).
9. Ngwa, W. *et al.* Cancer in sub-Saharan Africa: a Lancet Oncology Commission. *Lancet Oncol.* **23**, e251–e312 (2022).
10. Commissioner, O. of the. Step 1: Discovery and Development. *FDA* (2019).

11. Halder, D., Das, S., R, A. & R S, J. Molecular docking and dynamics based approach for the identification of kinase inhibitors targeting PI3Kα against non-small cell lung cancer: a computational study. *RSC Adv.* **12**, 21452–21467 (2022).
12. Petrovska, B. B. Historical review of medicinal plants' usage. *Pharmacogn. Rev.* **6**, 1–5 (2012).
13. Silva, J., Alvariño, R., Goettert, M. I., Caruncho, H. J. & Alves, C. Editorial: Natural products as drivers in drug development for neurodegenerative disorders. *Front. Pharmacol.* **13**, 932179 (2022).
14. Fasinu, P. S., Okoye, F. B. C., Abiodun, O. O., Kamdem, R. S. T. & Ogbole, O. O. Editorial: Fungal Bioactive Metabolites of Pharmacological Relevance. *Front. Pharmacol.* **13**, 912068 (2022).
15. Ogbole, O. O. *et al.* In vitro antiviral activity of peptide-rich extracts from seven Nigerian plants against three non-polio enterovirus species C serotypes. *Virol. J.* **18**, 161 (2021).
16. Aloke, C., Ohanenye, I. C., Aja, P. M. & Ejike, C. E. C. C. Phytochemicals from medicinal plants from African forests with potentials in rheumatoid arthritis management. *J. Pharm. Pharmacol.* rgac043 (2022) doi:10.1093/jpp/rgac043.
17. Kaur, B. *et al.* An In Silico Investigation to Explore Anti-Cancer Potential of Foeniculum vulgare Mill. Phytoconstituents for the Management of Human Breast Cancer. *Mol. Basel Switz.* **27**, 4077 (2022).
18. Bai, H. *et al.* Network Pharmacology Analysis, Molecular Docking, and In Vitro Verification Reveal the Action Mechanism of Prunella vulgaris L. in Treating Breast Cancer. *Evid.-Based Complement. Altern. Med. ECAM* **2022**, 5481563 (2022).
19. Khan, S. A. & Lee, T. K. W. Investigations of nitazoxanide molecular targets and pathways for the treatment of hepatocellular carcinoma using network pharmacology and molecular docking. *Front. Pharmacol.* **13**, 968148 (2022).

20. Clough, E. & Barrett, T. The Gene Expression Omnibus Database. *Methods Mol. Biol. Clifton NJ* **1418**, 93–110 (2016).
21. Tang, Z., Kang, B., Li, C., Chen, T. & Zhang, Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* **47**, W556–W560 (2019).
22. UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489 (2021).
23. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **23**, 3–25 (1997).
24. Hsin, K.-Y., Ghosh, S. & Kitano, H. Combining Machine Learning Systems and Multiple Docking Simulation Packages to Improve Docking Prediction Reliability for Network Pharmacology. *PLOS ONE* **8**, e83922 (2013).
25. Anita, S. & Wagh, S. R. B. Phytochemical analysis and in-vitro anticancer activity of Duranta erecta L. (Verbenaceae). *Int J Pharm Sci Res* **10**, 2941–2946 (2019).
26. Eke, I. G. & Okpara, G. C. Anti-hyperglycemic and anti-dyslipidemic activities of methanol ripe fruit extract of Duranta erecta L (Verbenaceae) in normoglycemic and hyperglycemic rats. *J. Tradit. Complement. Med.* **11**, 209–216 (2021).
27. Khanal, P. & Patil, B. M. In vitro and in silico anti-oxidant, cytotoxicity and biological activities of Ficus benghalensis and Duranta repens. *Chin. Herb. Med.* **12**, 406–413 (2020).
28. Patil, A. *et al.* GLUT-2 mediated glucose uptake analysis of Duranta repens: In-silico and In-vitro approach. *J. Diabetes Metab. Disord.* **21**, 419–427 (2022).
29. Srivastava, M. & Shanker, K. Duranta erecta Linn: A critical review on phytochemistry, traditional uses, pharmacology, and toxicity from phytopharmaceutical perspective. *J. Ethnopharmacol.* **293**, 115274 (2022).

30. Udobi, M. I. *et al.* Evaluation of the anthelminthic potential of Duranta erecta L. (Verbenaceae) fruits used in Nigerian ethnomedicine as a vermifuge. *J. Ethnopharmacol.* **216**, 57–62 (2018).
31. Gurung, A. B., Ali, M. A., Lee, J., Farah, M. A. & Al-Anazi, K. M. An Updated Review of Computer-Aided Drug Design and Its Application to COVID-19. *BioMed Res. Int.* **2021**, 8853056 (2021).
32. Baig, M. H. *et al.* Computer Aided Drug Design: Success and Limitations. *Curr. Pharm. Des.* **22**, 572–581 (2016).
33. Adebesin, A. O., Ayodele, A. O., Omotoso, O., Akinnusi, P. A. & Olubode, S. O. Computational evaluation of bioactive compounds from Vitis vinifera as a novel β-catenin inhibitor for cancer treatment. *Bull. Natl. Res. Cent.* **46**, 183 (2022).
34. Liu, J. *et al.* Network Pharmacology Prediction and Molecular Docking-Based Strategy to Discover the Potential Pharmacological Mechanism of Huai Hua San Against Ulcerative Colitis. *Drug Des. Devel. Ther.* **15**, 3255–3276 (2021).
35. Nogales, C. *et al.* Network pharmacology: curing causal mechanisms instead of treating symptoms. *Trends Pharmacol. Sci.* **43**, 136–150 (2022).
36. Hanahan, D. Hallmarks of Cancer: New Dimensions. *Cancer Discov.* **12**, 31–46 (2022).
37. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
38. Li, P. *et al.* Aurora-A/FOXO3A/SKP2 axis promotes tumor progression in clear cell renal cell carcinoma and dual-targeting Aurora-A/SKP2 shows synthetic lethality. *Cell Death Dis.* **13**, 606 (2022).
39. Mancini, M. *et al.* Polo-like kinase-1, Aurora kinase A and WEE1 kinase are promising druggable targets in CML cells displaying BCR::ABL1-independent resistance to tyrosine kinase inhibitors. *Front. Oncol.* **12**, 901132 (2022).

40. Liu, F. *et al.* Knockdown of AURKA sensitizes the efficacy of radiation in human colorectal cancer. *Life Sci.* **271**, 119148 (2021).
41. Wang, J. *et al.* Repression of the AURKA-CXCL5 axis induces autophagic cell death and promotes radiosensitivity in non-small-cell lung cancer. *Cancer Lett.* **509**, 89–104 (2021).
42. Kahl, I. *et al.* The cell cycle-related genes RHAMM, AURKA, TPX2, PLK1, and PLK4 are associated with the poor prognosis of breast cancer patients. *J. Cell. Biochem.* **123**, 581–600 (2022).
43. Peng, F. *et al.* Oncogenic AURKA-enhanced N6-methyladenosine modification increases DROSHA mRNA stability to transactivate STC1 in breast cancer stem-like cells. *Cell Res.* **31**, 345–361 (2021).
44. Du, R., Huang, C., Liu, K., Li, X. & Dong, Z. Targeting AURKA in Cancer: molecular mechanisms and opportunities for Cancer therapy. *Mol. Cancer* **20**, 15 (2021).
45. Wang, C., Xie, X., Li, W. & Jiang, D. Expression of KIF2A, NDC80, CDK1, and CCNB1 in breast cancer patients: Their interaction and linkage with tumor features and prognosis. *J. Clin. Lab. Anal.* e24647 (2022) doi:10.1002/jcla.24647.
46. Wang, Y. *et al.* Integrated Profiling Identifies CCNA2 as a Potential Biomarker of Immunotherapy in Breast Cancer. *OncoTargets Ther.* **14**, 2433–2448 (2021).
47. Xing, Z. *et al.* Expression and prognostic value of CDK1, CCNA2, and CCNB1 gene clusters in human breast cancer. *J. Int. Med. Res.* **49**, 300060520980647 (2021).
48. Alharbi, K. S. *et al.* An overview of epithelial growth factor receptor (EGFR) inhibitors in cancer therapy. *Chem. Biol. Interact.* **366**, 110108 (2022).
49. Al-Wahaibi, L. H. *et al.* Synthesis and Biological Evaluation of Indole-2-Carboxamides with Potent Apoptotic Antiproliferative Activity as EGFR/CDK2 Dual Inhibitors. *Pharm. Basel Switz.* **15**, 1006 (2022).
50. Selenz, C. *et al.* EGFR Inhibition Strongly Modulates the Tumour Immune Microenvironment in EGFR-Driven Non-Small-Cell Lung Cancer. *Cancers* **14**, 3943 (2022).

51. Brett, J. O., Spring, L. M., Bardia, A. & Wander, S. A. ESR1 mutation as an emerging clinical biomarker in metastatic hormone receptor-positive breast cancer. *Breast Cancer Res. BCR* **23**, 85 (2021).
52. Ge, Q. *et al.* miR-4324-RACGAP1-STAT3-ESR1 feedback loop inhibits proliferation and metastasis of bladder cancer. *Int. J. Cancer* **144**, 3043–3055 (2019).
53. Gonzalez, M. E. *et al.* EZH2 T367 phosphorylation activates p38 signaling through lysine methylation to promote breast cancer progression. *iScience* **25**, 104827 (2022).
54. Zhao, Y., Hu, Z., Li, J. & Hu, T. EZH2 Exacerbates Breast Cancer by Methylating and Activating STAT3 Directly. *J. Cancer* **12**, 5220–5230 (2021).
55. Siersbæk, R. *et al.* IL6/STAT3 Signaling Hijacks Estrogen Receptor α Enhancers to Drive Breast Cancer Metastasis. *Cancer Cell* **38**, 412-423.e9 (2020).
56. Vasiyani, H. *et al.* DNA damage induces STING mediated IL-6-STAT3 survival pathway in triple-negative breast cancer cells and decreased survival of breast cancer patients. *Apoptosis Int. J. Program. Cell Death* (2022) doi:10.1007/s10495-022-01763-8.
57. Binato, R. *et al.* NRIP1 is activated by C-JUN/C-FOS and activates the expression of PGR, ESR1 and CCND1 in luminal A breast cancer. *Sci. Rep.* **11**, 21159 (2021).
58. Joseph, C. *et al.* Elevated MMP9 expression in breast cancer is a predictor of shorter patient survival. *Breast Cancer Res. Treat.* **182**, 267–282 (2020).
59. Zhang, Q. *et al.* G6PD upregulates Cyclin El and MMP9 to promote clear cell renal cell carcinoma progression. *Int. J. Med. Sci.* **19**, 47–64 (2022).
60. Miao, K. *et al.* Scutellarein inhibits BLM-mediated pulmonary fibrosis by affecting fibroblast differentiation, proliferation, and apoptosis. *Ther. Adv. Chronic Dis.* **11**, 2040622320940185 (2020).

61. Liu, D., Zhang, C., Hu, M. & Su, K. Scutellarein relieves the death and inflammation of tubular epithelial cells in ischemic kidney injury by degradation of COX-2 protein. *Int. Immunopharmacol.* **101**, 108193 (2021).
62. Lang, X. *et al.* Scutellarein induces apoptosis and inhibits proliferation, migration, and invasion in ovarian cancer via inhibition of EZH2/FOXO1 signaling. *J. Biochem. Mol. Toxicol.* **35**, e22870 (2021).
63. Li, Y., Wang, J., Zhong, S., Li, J. & Du, W. Scutellarein inhibits the development of colon cancer via CDC4-mediated RAGE ubiquitination. *Int. J. Mol. Med.* **45**, 1059–1072 (2020).
64. Chandra, R., Singh, S. & Ganguly, C. β-Sitosterol & quercetin enhances brain development in iodine deficient rat models. *Nutr. Health* 2601060221122209 (2022) doi:10.1177/02601060221122209.
65. Wang, J. *et al.* Alleviating effect of quercetin on cadmium-induced oxidative damage and apoptosis by activating the Nrf2-keap1 pathway in BRL-3A cells. *Front. Pharmacol.* **13**, 969892 (2022).
66. Papakyriakopoulou, P. *et al.* Potential Pharmaceutical Applications of Quercetin in Cardiovascular Diseases. *Pharm. Basel Switz.* **15**, 1019 (2022).
67. Rhman, M. A., Devnarain, N., Khan, R. & Owira, P. M. O. Synergism Potentiates Oxidative Antiproliferative Effects of Naringenin and Quercetin in MCF-7 Breast Cancer Cells. *Nutrients* **14**, 3437 (2022).
68. Gao, Y. *et al.* Scoparone alleviates hepatic fibrosis by inhibiting the TLR-4/NF-κB pathway. *J. Cell. Physiol.* (2020) doi:10.1002/jcp.30083.
69. Li, N. *et al.* Scoparone inhibits pancreatic cancer through PI3K/Akt signaling pathway. *World J. Gastrointest. Oncol.* **13**, 1164–1183 (2021).

70. Song, Q., Liu, H., Zhang, Y., Qiao, C. & Ge, S. Lipidomics Revealed Alteration of the Sphingolipid Metabolism in the Liver of Nonalcoholic Steatohepatitis Mice Treated with Scoparone. *ACS Omega* **7**, 14121–14127 (2022).

71. Park, K.-R. *et al.* Effects of Scoparone on differentiation, adhesion, migration, autophagy and mineralization through the osteogenic signalling pathways. *J. Cell. Mol. Med.* **26**, 4520–4529 (2022).