Eksakta : Berkala Ilmiah Bidang MIPA
VOLUME 27 NO 04 2026, pp 578-597
ISSN : Print 1411-3724 — Online 2549-7464
DOI : https://doi.org/10.24036/eksakta/vol27-iss04/707

Eksakta
Berkala Ilmiah Bidang MIPA
http://www.eksakta.ppj.unp.ac.id/index.php/eksakta

Article

Integrative Transcriptomic and Docking Analysis of Coffee Bioactives Targeting CCNA2, AKT1, and CDK2 in TNBC

Ida Neni Haryanti[1,2], Linda Erlina[3*], Ade Arsianti[3], Aryo Tedjo[3]

[1]Master Programs in Biomedical Sciences, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia
[2]Balai Labkesmas Batam, Kepulauan Riau, Indonesia
[3]Department of Medical Chemistry, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia

Article Info

Article history :
Received June 02, 2026
Revised June 15, 2026
Accepted June 20, 2026
Published August 30, 2026

Keywords :
TNBC,
biomarker,
transcriptomics,
molecular docking

Abstract. Triple-negative breast cancer (TNBC) is an aggressive breast cancer subtype characterized by high proliferation rates, poor prognosis, and limited therapeutic options. The identification of reliable molecular targets and effective multitarget therapeutic candidates remains an important challenge in TNBC research. Coffee-derived bioactive compounds have demonstrated potential anticancer properties, but their interactions with key TNBC-associated targets remain insufficiently understood. Therefore, this study aimed to identify molecular biomarkers and therapeutic targets in TNBC and evaluate the potential of major coffee bioactive compounds using an integrative in silico approach. Five Gene Expression Omnibus (GEO) datasets (GSE38959, GSE186102, GSE65194, GSE45827, and GSE7904) were analyzed to identify differentially expressed genes (DEGs) using $|\log2FC| \geq 1$ and adjusted p-value < 0.05. Protein–protein interaction analysis, machine-learning validation, Kaplan–Meier survival analysis, chemogenomic mapping, molecular docking, and ADMET prediction were subsequently performed. A total of 917 overlapping DEGs were identified, with CCNA2 emerging as a key hub gene associated with poor prognosis. Machine-learning validation achieved 97.7% accuracy and an AUC of 0.964. Chemogenomic analysis prioritized AKT1, CDK2, and CCNA2 as therapeutic targets. Molecular docking revealed favorable interactions of caffeic acid and chlorogenic acid, with caffeic acid showing the most consistent multitarget affinity (−5.94 to −6.96 kcal/mol). These findings suggest that caffeic acid is a promising multitarget candidate for TNBC therapy and warrants further experimental validation.

Corresponding Author :
Linda Erlina
Department of Medical Chemistry, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia
Email : linda.erlina22@ui.ac.id

1. Introduction

Breast cancer is the leading cause of cancer-related death among women worldwide. Among its subtypes, triple-negative breast cancer (TNBC) is the most aggressive because it does not express estrogen receptor (ER), progesterone receptor (PR), or HER2 [1]. Consequently, TNBC does not respond to hormonal or HER2-targeted therapies, leaving chemotherapy as the main treatment option. However, chemotherapy is frequently associated with drug resistance and severe adverse effects, contributing to poor clinical outcomes [2]. TNBC accounts for approximately 15–20% of invasive breast cancers worldwide and is often diagnosed at advanced stages, increasing the risk of metastasis and mortality [3-5].

At the molecular level, TNBC is characterized by high genomic instability associated with defects in DNA repair mechanisms. BRCA1 mutation is the most commonly reported genetic factor in TNBC [6]. In addition, genes such as PTEN, ATM, and STK11 are known to play roles in cell-cycle regulation and DNA repair. These molecular alterations make TNBC highly proliferative and more aggressive than other breast cancer subtypes [7]. Recent therapeutic strategies for TNBC have focused on molecular targets involved in cell proliferation and survival. Cyclin-dependent kinase 2 (CDK2) is a key regulator of the G1/S transition and is frequently overexpressed in TNBC, making it an attractive target for inhibiting tumor growth [8-9].

CCNA2 regulates cell-cycle progression through its interaction with CDK2 and plays an important role in S-phase entry. Overexpression of CCNA2 has been associated with increased proliferation, tumor progression, and poor prognosis in breast cancer, highlighting its potential as a TNBC biomarker [10-11]. The PI3K/AKT pathway is another critical driver of TNBC progression. AKT1 promotes cell survival, proliferation, and resistance to apoptosis while interacting functionally with the CCNA2/CDK2 axis to support DNA replication and cell-cycle progression [12]. Dysregulation of this pathway is common in TNBC and represents an important therapeutic target. [13].

Natural bioactive compounds are increasingly explored as anticancer agents. Coffee contains several biologically active compounds, including chlorogenic acid, caffeic acid, trigonelline, and caffeine, which have been reported to induce apoptosis, inhibit proliferation, and modulate cancer-related signaling pathways [14-15]. Among coffee species, *Coffea liberica* (Excelsa coffee) has attracted attention because of its distinctive phytochemical profile and relatively high phenolic content. Previous studies have identified chlorogenic acid, caffeic acid, trigonelline, and caffeine as major constituents, suggesting potential antioxidant and anticancer activities.

Computational studies suggest that coffee-derived compounds may interact with cancer-associated targets such as PI3K/AKT and cell-cycle regulators. Integrative bioinformatics approaches combining transcriptomics and molecular docking provide an efficient strategy for identifying biomarkers and predicting ligand–protein interactions before experimental validation [16-17].

Although several studies have evaluated coffee compounds using gene-expression analysis or molecular docking, few have integrated transcriptomic meta-analysis, machine-learning validation, chemogenomic mapping, and structural bioinformatics to prioritize TNBC multitarget candidates. Therefore, this study integrates five GEO datasets, protein–protein interaction analysis, hub-gene identification, CTD-based chemogenomics, and molecular docking to identify potential therapeutic targets and coffee-derived compounds relevant to TNBC [16], [18].

Recent studies have demonstrated the anticancer potential of coffee-derived phenolic compounds, including caffeic acid and chlorogenic acid, through the modulation of cell-cycle regulation, apoptosis,

and PI3K/AKT signaling pathways. However, most previous investigations focused on single targets or individual computational approaches, and comprehensive studies integrating transcriptomic biomarkers, chemogenomic interactions, and structural bioinformatics remain limited. Therefore, identifying biologically relevant multitarget candidates associated with TNBC progression is necessary to support the development of more effective therapeutic strategies.

We hypothesized that key TNBC-associated biomarkers identified through integrated transcriptomic analysis are functionally linked to cell-cycle and PI3K/AKT signaling pathways and can be modulated by coffee-derived phenolic compounds. Furthermore, caffeic acid and chlorogenic acid are expected to exhibit favorable interactions with selected molecular targets, supporting their potential as multitarget candidates for TNBC therapy.

Therefore, this study integrates five GEO datasets, protein–protein interaction analysis, hub-gene identification, CTD-based chemogenomics, and molecular docking to identify potential therapeutic targets and coffee-derived compounds relevant to TNBC. This study contributes methodologically to biomarker prioritization strategies and highlights potential multitarget candidates that require further experimental validation based on the integration of transcriptomics, chemogenomics, and structural bioinformatics, with a focus on key targets such as CCNA2, CDK2, and AKT1.

2. Experimental Section

2.1. Materials

This study employed an integrative approach to identify potential biomarkers and therapeutic targets in triple-negative breast cancer (TNBC). The analytical framework included transcriptomic analysis, protein–protein interaction (PPI) network construction, functional enrichment analysis, machine learning-based biomarker validation, chemical–gene interaction integration, and molecular docking simulations [16], [18].

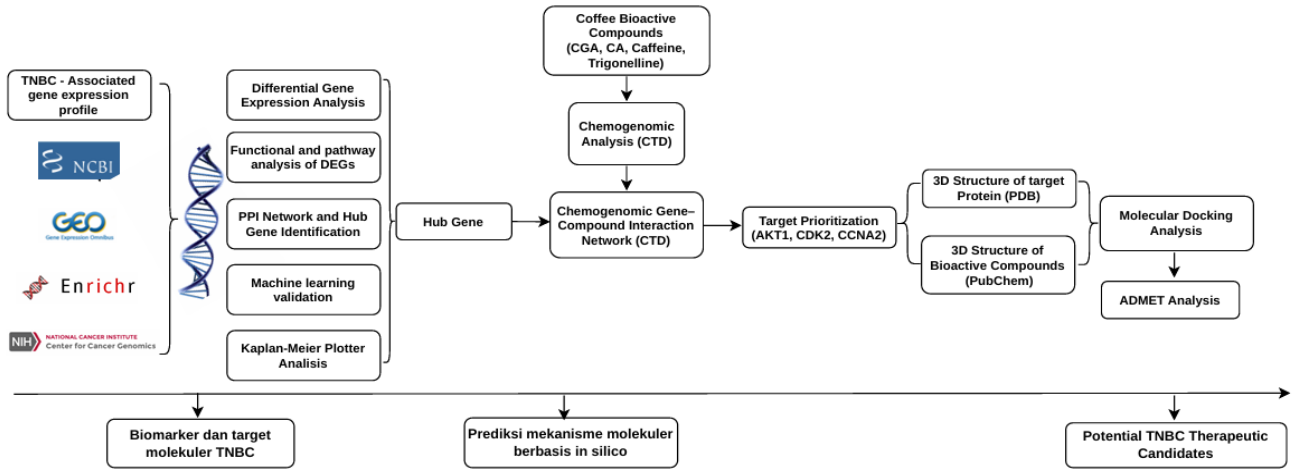

Figure 1. Illustrates the overall workflow of the present study. Transcriptomic datasets associated with TNBC were retrieved from the GEO database and analyzed to identify *Differentially Expressed Genes* (DEGs). Functional enrichment and pathway analyses were subsequently performed to explore the biological processes and signaling pathways associated with the identified DEGs. *Protein–protein interaction* (PPI) network analysis was then conducted to identify hub genes that may play central roles in TNBC progression. The biological and clinical relevance of these hub genes was further evaluated through machine learning validation and Kaplan–Meier survival analysis.

The workflow (Figure 1) began with the identification of differentially expressed genes between TNBC and normal tissue, followed by network analysis to obtain hub genes and biological interpretation through enrichment analysis. The hub genes were then validated using machine

learning approaches and mapped to interactions with coffee bioactive compounds. Priority molecular targets were subsequently analyzed by molecular docking to assess ligand binding affinity and interaction patterns with the proteins. To strengthen the computational interpretation, the study was also designed to support experimental validation through cytotoxicity assays and compound composition analysis.

2.2. Methods

2.2.1 Transcriptomic Data and Preprocessing

Transcriptomic data were obtained from five GEO datasets (http://www.ncbi.nlm.nih.gov/geo/) relevant to TNBC, namely GSE38959, GSE186102, GSE65194, GSE45827, and GSE7904, comprising a total of 218 samples, including 167 TNBC samples and 51 normal samples.

Tabel 1. Characterisation of the GEO Dataset

Dataset	Normal	TNBC
GSE38959	13	30
GSE186102	9	41
GSE65194	11	47
GSE45827	11	41
GSE7904	7	8

The sample distribution for each dataset is presented in Table 1. Differential expression analysis was performed separately for each dataset using the Benjamini–Hochberg false discovery rate (FDR) method, with the criteria $|\log2FC| \geq 1$ and adjusted p-value < 0.05. Genes consistently observed across all datasets were then selected as overlapping DEGs using a Venn diagram in InteractiVenn (https://www.interactivenn.net/) to identify shared genes and minimize the influence of inter-dataset variability. 17 This integration yielded 917 overlapping genes, which were used as the initial candidates for downstream analysis.

2.2.2 Protein–Protein Interaction Network Construction

To understand functional relationships among genes, a protein–protein interaction network was constructed using the STRING database https://string-db.org/ version 12.0). This database integrates protein interaction information from experimental evidence, computational prediction, and literature curation. The overlapping DEGs were used as the primary input for PPI network construction, with a high confidence score threshold (>0.999) to ensure reliable interactions. [19] The overlapping DEGs were used as the primary input for PPI network construction, with a high confidence score threshold (>0.999) to ensure reliable interactions.

Network topology analysis was performed to evaluate interaction patterns among proteins and identify gene groups with important biological roles. Cluster analysis was then carried out using the MCODE algorithm to identify highly connected network modules. The main module obtained was used as the basis for functional analysis and hub gene identification.

2.2.3 Module and Hub Gene Identification

Hub genes were identified using the CytoHubba plugin in Cytoscape version 3.10.4 based on degree centrality and maximal clique centrality (MCC). These methods were used to determine the genes with the highest connectivity and central roles in the PPI network [20]. The ten highest-ranking genes were selected as hub gene candidates: CDK1, CCNB1, CCNA2, BUB1B, MAD2L1, CDC6, MCM3, MCM4, MCM6, and CDT1. These genes are involved in cell-cycle regulation, DNA replication, and mitotic checkpoint control. Genes with high connectivity are assumed to contribute importantly to the biological characteristics of TNBC and may serve as potential biomarkers or therapeutic targets [21-22].

2.2.4 Functional Enrichment Analysis

Functional enrichment analysis was performed to interpret the biological meaning of the identified genes. The analysis was conducted using the ShinyGO and Enrichr platforms (https://maayanlab.cloud/Enrichr/) to map genes to Gene Ontology (GO) categories and relevant signaling pathways [23].

The analysis focused on genes within the main MCODE module with a significance threshold of false discovery rate (FDR) < 0.05. The enrichment analysis emphasized biological processes related to the cell cycle, DNA replication, mitotic checkpoint regulation, and signaling pathways involved in cancer progression. This analysis was used to strengthen the biological interpretation of the hub genes derived from the PPI network.

2.2.5 Machine Learning-Based Biomarker Validation

Biomarker validation was performed to assess the ability of the hub genes to distinguish TNBC samples from normal tissue. The machine learning using orange data mining platform https://orangedatamining.com/ version 3.39 approaches used included Random Forest (RF), Support Vector Machine (SVM), and Neural Network.

The models were evaluated using accuracy, F1-score, Matthews correlation coefficient (MCC), and area under the curve (AUC). Genes showing the best classification performance were considered to have a higher potential as diagnostic biomarkers. This approach was used to ensure that the selected genes were not only central in the network but also relevant for predicting disease status.

2.2.6 Chemical–Gene Interaction Integration

Chemical–gene interaction integration was performed to connect molecular biomarkers with potential therapeutic compounds. The database used was the Comparative Toxicogenomics Database (CTD), a curated resource containing relationships among chemicals, genes, and diseases [24].

The hub genes identified from network analysis and validated by machine learning were used as input for this step. Interactions between these genes and the main coffee bioactive compounds, namely chlorogenic acid, caffeic acid, caffeine, and trigonelline, were then explored. According to the literature, these compounds have demonstrated biological activity in various cancer models [25]. To strengthen target selection, validated hub genes were integrated with CTD (CTD; http://ctdbase.org/) chemical–gene interaction data and mapped to KEGG pathways. Targets were prioritized based on their connectivity with coffee-derived bioactive compounds, involvement in TNBC-related pathways, biological relevance, and the availability of high-quality three-dimensional structures in the Protein Data Bank (PDB) (https://www.rcsb.org/) [26].

Based on these criteria, AKT1, CDK2, and CCNA2 were selected as the principal molecular targets. KEGG analysis indicated that the identified hub genes were mainly enriched in cell-cycle regulation, DNA replication, and cancer-related signaling pathways. CDK2 and CCNA2 were selected due to their central roles in cell-cycle progression and DNA replication, whereas AKT1 was chosen as a key component of the PI3K/AKT pathway, which is functionally linked to Cyclin/CDK

signaling and TNBC cell proliferation. Furthermore, CTD analysis demonstrated interactions between these targets and coffee-derived bioactive compounds, supporting their suitability for subsequent molecular docking studies [26-28].

2.2.7 Target Protein Preparation

Protein preparation was performed for the three main molecular targets, namely AKT1, CDK2, and CCNA2. The three-dimensional protein structures were obtained from the Protein Data Bank and selected based on crystal structure quality, resolution, presence of co-crystallized ligands, and clarity of the active site [29-30].

The protein structures used were AKT1 (PDB ID: 4GV1), CDK2 (PDB ID: 6GUK), and CCNA2 (PDB ID: 7ACK). These structures were selected because of their biological relevance to proliferation, cell-cycle regulation, and cancer cell survival in TNBC. Before molecular docking, the proteins were cleaned of water molecules, co-crystallized ligands, and other non-essential components. Polar hydrogens were then added, Gasteiger charges were assigned, and appropriate atom types were applied to make the receptor structures compatible with docking [31]. After preparation, each protein was saved in .pdbqt format for use in the molecular docking stage.

2.2.8 Test Ligand Preparation

The chemical structures of four coffee bioactive compounds, namely caffeic acid (CID: 689043), caffeine (CID: 231), chlorogenic acid/CGA (CID: 1794427), and trigonelline (CID: 5570), were downloaded from the PubChem (https://pubchem.ncbi.nlm.nih.gov/) database in PDB format. Ligand preparation was performed using AutoDockTools to generate structures ready for docking. The preparation steps included adding polar hydrogens, assigning Gasteiger charges, identifying rotatable bonds, and energy minimization. Afterward, the ligands were converted into .pdbqt format for compatibility with the docking software. This process was important to ensure that the ligands were in a stable state and had conformations suitable for interaction simulations with the target proteins [32].

2.2.9 Molecular Docking Analysis

Molecular docking analysis was performed to predict the binding affinity between coffee-derived bioactive compounds and the target proteins AKT1, CDK2, and CCNA2. Protein and ligand preparation were conducted using AutoDockTools version 1.5.6, whereas docking simulations were performed using AutoDock Vina version 1.2.5. The grid box was positioned at the validated active site of each protein to restrict ligand sampling to the biologically relevant binding pocket [30], [33].

Docking results were evaluated based on binding energy (ΔG), root mean square deviation (RMSD), and amino acid residue interaction patterns within the active site. More negative ΔG values indicate stronger binding affinity, whereas lower RMSD values indicate more reliable docking poses. The pose with the lowest binding energy and the most consistent interaction pattern was selected as the representative binding mode for further analysis [33].

2.2.10 Interaction Analysis and Visualization

The protein–ligand complexes obtained from molecular docking were visually analyzed using PyMOL (https://pymol.org#download) Pymol ™ 1.3.8 , ligplot version 2.3 and Biovia Discovery studio 2025 to evaluate ligand orientation in the active site and identify amino acid residues involved in the interactions. The analysis focused on the target proteins AKT1, CDK2, and CCNA2 with the test ligands chlorogenic acid, caffeic acid, caffeine, and trigonelline.

Interaction evaluation included the identification of hydrogen bonds, hydrophobic interactions, and electrostatic interactions contributing to complex stability. In addition, the analysis assessed ligand positioning relative to the binding site and the involvement of key residues in the target proteins. The

visualization results were used to support the interpretation of the relationship between binding affinity and interaction patterns formed in the protein–ligand complexes.

2.2.11 ADMET Analysis

To evaluate the suitability of the compounds as drug candidates, in silico absorption, distribution, metabolism, excretion, and toxicity (ADMET) analysis was performed. This analysis was intended to provide an initial overview of the pharmacokinetic profile and safety potential of coffee bioactive compounds.

The ADMET results were used to assess whether compounds with good binding affinity also possessed favorable pharmacokinetic characteristics as potential therapeutic agents. Selected compounds were then analyzed using the pkCSM platform (https://biosig.lab.uq.edu.au/pkcsm) to predict pharmacokinetic and toxicity parameters. The parameters examined included intestinal absorption, membrane permeability, tissue distribution, enzyme-mediated metabolism, excretion, and toxicity potential.

2.2.12 Kaplan Meier Plotter

Prognostic analysis was performed using Kaplan–Meier Plotter (https://kmplot.com/analysis/) to evaluate the relationship between target gene expression and overall survival in breast cancer patients. Kaplan–Meier Plotter is a microarray-based database widely used to assess the prognostic value of cancer biomarkers through integration of gene expression and patient clinical data. [34] Prognostic analysis was performed using Kaplan–Meier Plotter to evaluate the relationship between target gene expression and overall survival in breast cancer patients. Kaplan–Meier Plotter is a microarray-based database widely used to assess the prognostic value of cancer biomarkers through integration of gene expression and patient clinical data [34].

3. Results and Discussion

Integrative transcriptomic analysis of five GEO datasets (GSE38959, GSE186102, GSE65194, GSE45827, GSE7904) identified hundreds of differentially expressed genes (DEGs) between triple-negative breast cancer (TNBC) and normal breast tissue using thresholds adjusted $p < 0.05$ and defined $|\log_2 FC|$. Cross-dataset integration via Venn diagram yielded 917 overlapping DEGs (Figure 2), demonstrating high expression consistency and superior representativeness of TNBC biological state compared to dataset-specific DEGs. This integrative approach enhances biomarker identification reliability through reduction of platform-specific bias and improved result reproducibility, aligning with current cancer transcriptomic principles emphasizing cross-dataset consistency for biological validity [35-36].

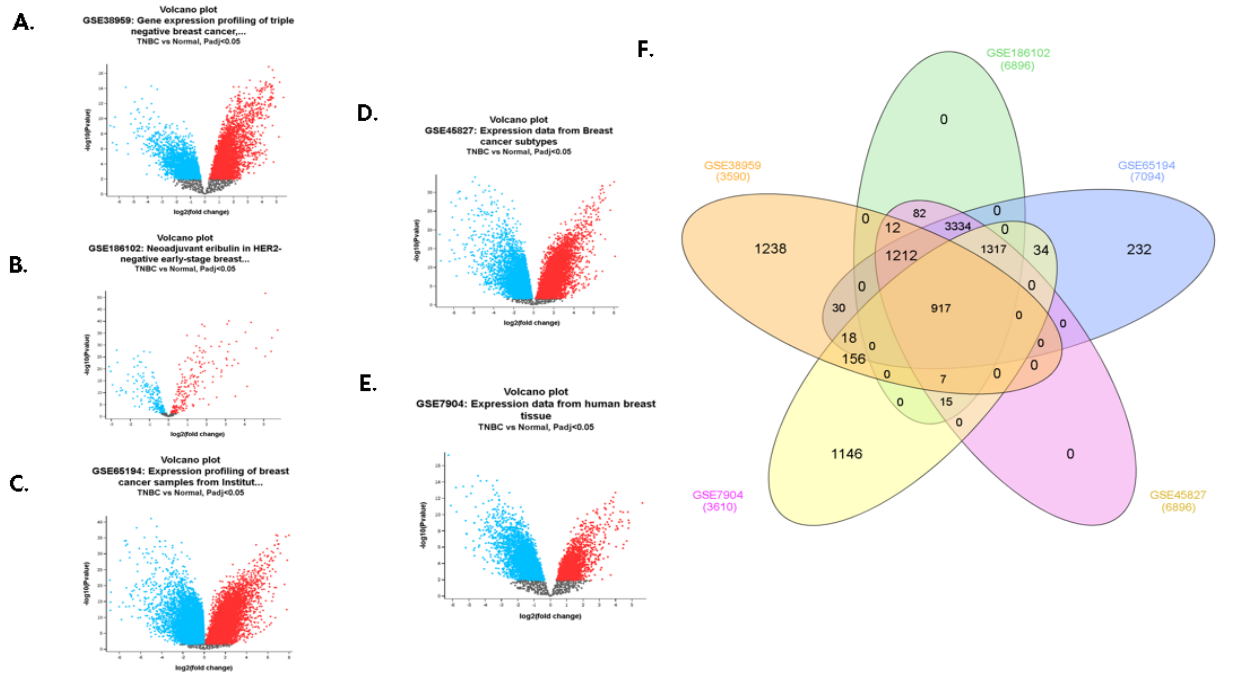

Figure 2. Volcano Plots and Venn Diagram of the Integrated Five GEO datasets. (A) GSE38959, (B) GSE186102, (C) GSE65194, (D) GSE45827, and (E) GSE7904 Show Differentially Expressed Genes Between TNBC and Normal Breast Tissue. (F) The Venn Diagram Shows 917 Overlapping DEGs Across the Five Datasets.

PPI network analysis constructed from 917 overlapping genes showed high connectivity with numerous interprotein interactions reflecting molecular regulation complexity in TNBC. MCODE clustering analysis identified one major module comprising 29 nodes and 362 edges, representing a highly interconnected gene cluster functioning as core regulatory module within TNBC network. In cancer biology, modules with dense connectivity are often associated with essential biological processes such as cell cycle and DNA replication acting as major drivers of tumor proliferation and survival. CytoHubba network (Figure 3) topology analysis identified ten major hub genes: CDK1, CCNB1, CCNA2, BUB1B, MAD2L1, CDC6, MCM3, MCM4, MCM6, and CDT1.

These genes are consistently associated with three major biological processes: cell-cycle regulation (CDK1, CCNB1, CCNA2) controlling G2/M transition and cell division; DNA replication licensing (CDC6, MCM3, MCM4, MCM6, CDT1) ensuring DNA replication occurs only once per cell cycle; and mitotic checkpoint control with chromosomal stability (BUB1B, MAD2L1) maintaining accurate chromosome segregation via spindle assembly checkpoint. CDK1, CCNB1, and CCNA2 overexpression associated with accelerated G2/M transition driving uncontrolled cell proliferation.

MCM complex together with CDC6 and CDT1 plays central role in DNA replication licensing; dysregulation of this system can cause re-replication, DNA damage, and genomic instability which are hallmarks of aggressive cancers like TNBC. BUB1B and MAD2L1 are key spindle assembly checkpoint components preventing chromosome segregation before proper spindle attachment; disruption of this mechanism increases risk of aneuploidy, chromosomal fragmentation, and more aggressive tumor progression [35]. Functional enrichment using GO Biological Process and KEGG through ShinyGO/Enrichr highlighted cell cycle, cancer pathways, DNA replication, mitotic spindle checkpoint, chromosome segregation, and p53 signaling as most prominent pathways (Figure 4) .

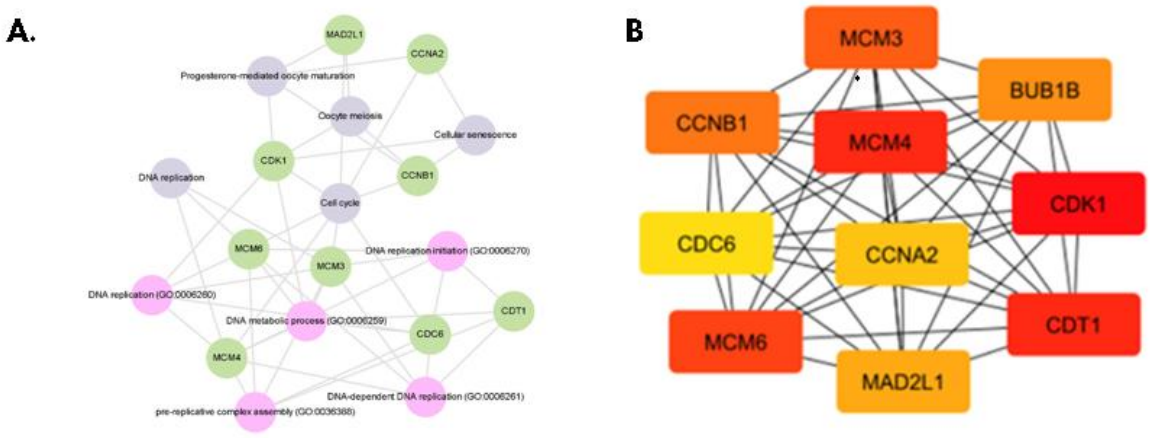

Figure 3. PPI Subnetwork of the 10 Hub Genes: CDK1, CCNB1, CCNA2, BUB1B, MAD2L1, CDC6, MCM3, MCM4, MCM6, and CDT1.

These results are consistent with TNBC profile characterized by strong activation of proliferative pathways and impaired cellular checkpoint control, particularly within p53-dependent axis. p53 pathway dominance reflects genotoxic stress and checkpoint dysfunction enabling cancer cell survival despite genomic damage, since p53 functions as genome guardian halting cell cycle when DNA damage occurs . KEGG mapping showed CDK1, CDK2, CCNA2, CDC6, and MCM family members highly overexpressed in TNBC datasets, confirming their central role in cell-cycle machinery and DNA replication thus overactivation of these pathways drives continuous tumor proliferation. (Figure 5) [35-37].

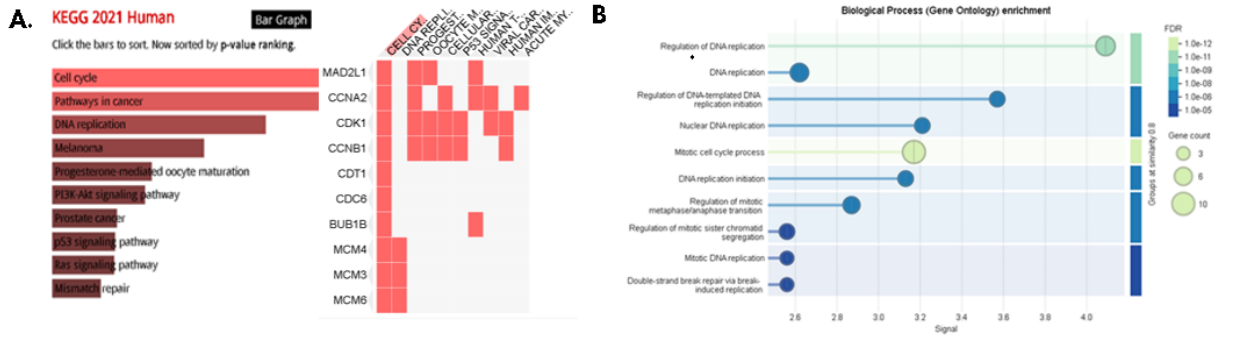

Figure 4. Functional Enrichment Analysis of The Main Module using Shinygo and Enrichr. (A) KEGG Bar Plot and (B) GO Biological Process and KEGG Dot Plot Showing Enrichment of Cell Cycle, DNA Replication, Mitotic Spindle Checkpoint, Chromosome Segregation, Cancer Pathways, and p53 Signaling.

Validation of ten-hub-gene signature consistency and discriminatory power as multigene signature was performed using three machine learning algorithms (Figure 6): Random Forest (RF), Support Vector Machine (SVM), and Neural Network (NN). The classification results showed very high performance: RF and SVM achieved an accuracy of 0.977, F1-score of 0.976, and MCC of 0.945, while the Neural Network yielded an AUC of 0.964, indicating strong capability in distinguishing TNBC from normal tissue.

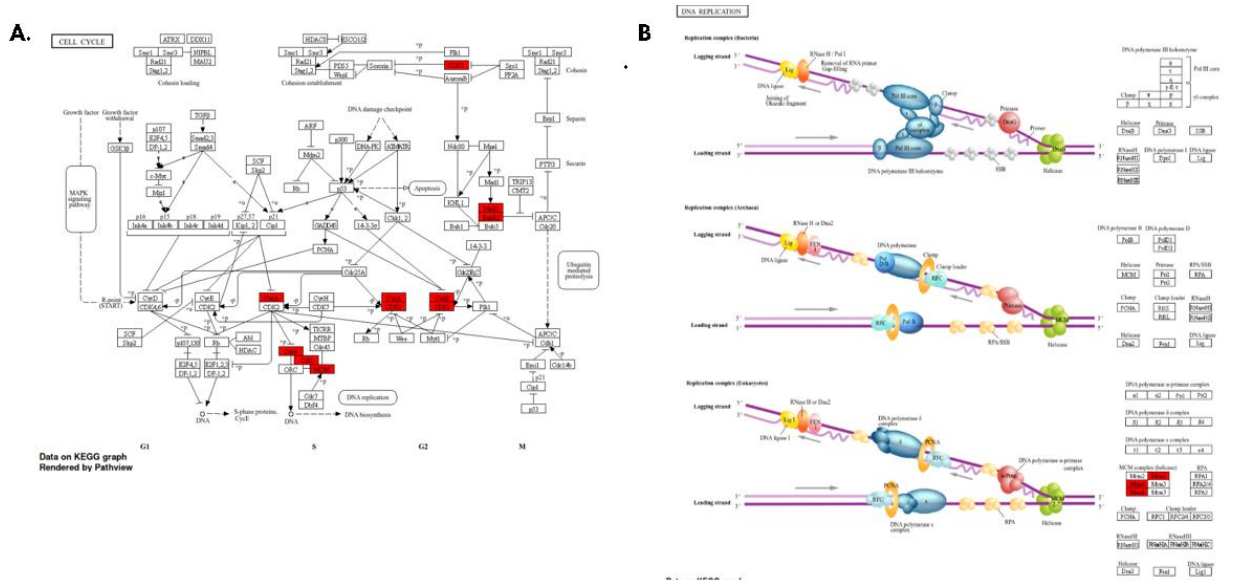

Figure 5. KEGG Pathway Mapping of Cell Cycle and DNA Replication Showing Proteins from the 29-Gene Main Module, including CDK1, CDK2, CCNB1, CCNA2, MAD2L1, MCM3, MCM4, MCM6, CDT1, CDC6, and BUB1B.

Confusion matrix analysis showed that Random Forest achieved the highest sensitivity at 100%, correctly identifying all TNBC samples, although with slightly lower specificity (84.62%). In contrast, SVM and Neural Network both showed sensitivity of 93.33% and specificity of 92.31%, reflecting a well-balanced performance in detecting both TNBC and normal samples. Scatter plot visualization, such as t-SNE or PCA, showed a clear separation between TNBC and control groups, while the heatmap demonstrated a consistent upregulation pattern in most hub genes in TNBC. This approach is consistent with modern cancer bioinformatics, where machine learning-validated multigene signatures are used to improve biomarker identification and cancer subtype classification [36-37].

As an integrative step linking molecular biomarkers with potential therapeutic agents, chemical–gene interaction analysis was performed using the Comparative Toxicogenomics Database (CTD). The integration workflow, shown in Figure 7, revealed that several hub genes were associated with coffee-derived bioactive compounds, particularly chlorogenic acid, caffeic acid, caffeine, and trigonelline. From this mapping, AKT1, CDK2, and CCNA2 were identified as the main targets based on network degree centrality, biological relevance to TNBC progression, and involvement in cancer proliferation and survival pathways. In addition, these proteins have high-quality crystal structures available in the Protein Data Bank (PDB), making them suitable for molecular docking and molecular dynamics simulation.

AKT1 is central kinase on PI3K/AKT/mTOR pathway frequently hyperactivated in TNBC, playing important roles in proliferation, survival, metabolism, metastasis, and therapy resistance [34,43]. PI3K activation induces AKT1 phosphorylation activating downstream proteins mTOR, BAD, and BCL2 thereby promoting cancer cell survival and anti-apoptotic signaling. AKT1 also contributes to cell-cycle regulation through increased cyclin expression and activation of Cyclin A/CDK2 complex involved in G1/S transition and S-phase progression [12].

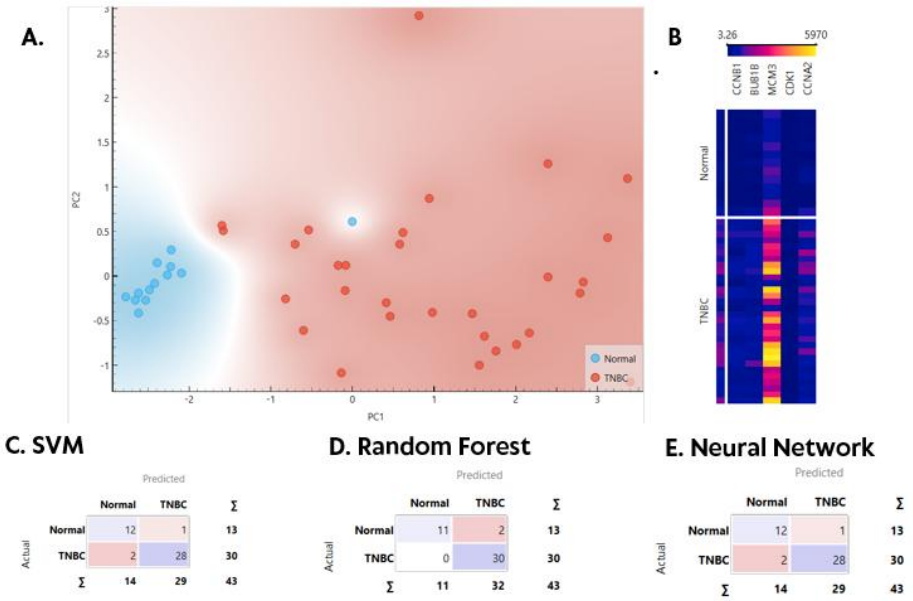

Figure 6. Machine Learning Validation of the 10-Hub-Gene Signature. (A) TNBC and Control Sample Separation, (B) Heatmap of 10 Hub Gene Expression, (C) SVM Performance, (D) Random Forest Performance, and (E) Neural Network Performance.

CDK2 and CCNA2 are major cell-cycle regulators essential for controlling cancer cell proliferation; CDK2 regulates G1/S transition through complex formation with Cyclin A2 (CCNA2), while CCNA2 is involved in S-phase and G2/M regulation supporting DNA replication and cell division. CCNA2 overexpression and CDK2 activation are closely associated with increased proliferation, tumor aggressiveness, genomic instability, and poor prognosis in TNBC. Functional relationship between AKT1, CDK2, and CCNA2 highlights close connection between PI3K/AKT signaling and cell-cycle regulation in maintaining TNBC proliferative phenotype [38].

Based on the integrated network and pathway results, coffee bioactive compounds were further analyzed by molecular docking to evaluate binding affinity and protein–ligand interaction patterns at the active sites of AKT1, CDK2, and CCNA2. The protein–ligand complexes with the best affinity were then evaluated using molecular dynamics simulation to assess their dynamic stability over time. In addition, in silico ADMET analysis was performed to predict the pharmacokinetic profile, drug-likeness, and toxicity potential of coffee bioactive compounds as candidate therapeutic agents against TNBC [39].

The integrative approach suggests that coffee bioactive compounds may modulate the PI3K/AKT pathway and cell-cycle regulation through interactions with AKT1, CDK2, and CCNA2. Such a multitarget strategy is expected to suppress proliferation, induce apoptosis, and inhibit TNBC progression simultaneously by targeting both survival and proliferative signaling pathways [12], [40].

Based on the integrated network and pathway analyses (Figure 8), coffee bioactive compounds were further evaluated by molecular docking to determine binding affinity and protein–ligand interaction patterns at the active sites of AKT1, CDK2, and CCNA2. The complexes with the most favorable affinity were subsequently subjected to molecular dynamics simulation to assess their dynamic stability. In addition, in silico ADMET analysis was performed to predict pharmacokinetic properties, drug-likeness, and toxicity potential.

Docking results showed that AKT1 had the lowest reference binding energy (-9.21 kcal/mol), followed by CDK2 (-8.54 kcal/mol) and CCNA2 (-7.52 kcal/mol), with AKT1 also exhibiting the most favorable RMSD value (1.18 Å). Among the coffee-derived compounds, caffeic acid consistently

showed the best performance across the three proteins, while chlorogenic acid performed optimally against CDK2 (Table 1).

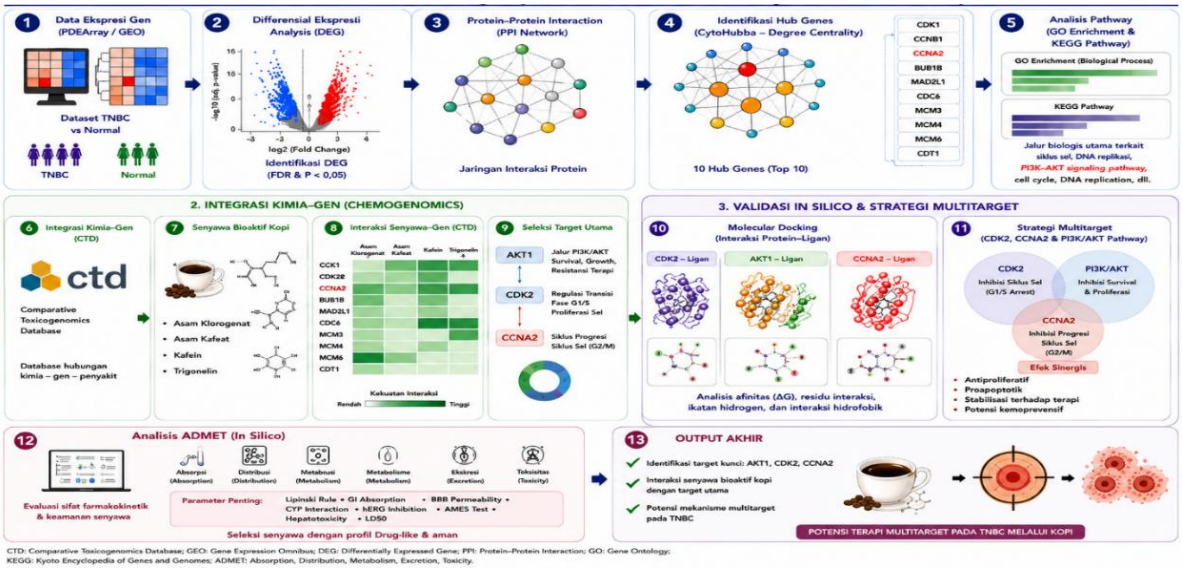

Figure 7. Workflow of Transcriptomic and Chemogenomic Integration for Identifying Potential Molecular Targets of Coffee Bioactive Compounds in TNBC. The Workflow Included DEG Identification, PPI Network Construction, Hub Gene Selection, CTD Integration, and Prioritization of AKT1, CDK2, and CCNA2 for Molecular Docking and Molecular Dynamics Simulation.

Low RMSD values (<3 Å) indicate a stable ligand pose, whereas low inhibition constants (Ki) support the potential for enzymatic inhibition. More negative binding energies (ΔG) indicate stronger protein–ligand stability, and RMSD values below 2 Å suggest conserved and dynamically valid binding poses, thereby strengthening the interpretation of the docking results [29].

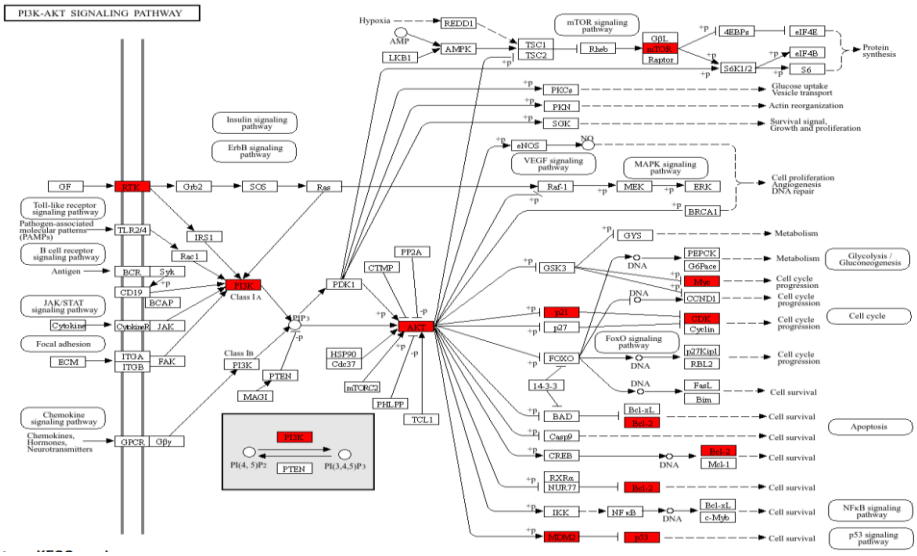

Figure 8. PI3K/AKT Signaling and Cell-Cycle Regulation Involving AKT1, CDK2, and CCNA2 in TNBC. PI3K Activation Triggers AKT1-Mediated Signaling that Promotes Survival, Proliferation, and Cell-Cycle Progression, while the CCNA2/CDK2 Complex Contributes to G1/S Transition and DNA Replication.

The interactions were largely dominated by hydrogen bonds between phenolic or carboxyl groups of the coffee compounds and polar residues in the active site, hydrophobic interactions with hydrophobic pockets of the proteins, and electrostatic or π-stacking interactions with aromatic residues around the ATP-binding site of CDK2. These interaction types are important for ligand-binding stability and specificity and are often associated with competitive inhibition of kinase activity at the ATP-binding site. Table 2 presents the binding energy, RMSD, and docking affinity parameters of chlorogenic acid, caffeic acid, caffeine, and trigonelline for each target protein using AutoDockTools 1.5.6.

The docking score comparison showed that AKT1 had the lowest reference binding energy (-9.21 kcal/mol), followed by CDK2 (-8.54 kcal/mol) and CCNA2 (-7.52 kcal/mol), with AKT1 also exhibiting the most favorable RMSD value (1.18 Å). Among the coffee-derived compounds, caffeic acid consistently showed the best performance across the three proteins, whereas chlorogenic acid showed the strongest affinity for CDK2.

Docking of AKT1 (PDB 4GV1) showed a reference binding energy of -9.21 kcal/mol with an RMSD of 1.18 Å. Among the coffee compounds, caffeic acid (-5.94 kcal/mol, Ki 28 μM), chlorogenic acid (-5.25 kcal/mol), and trigonelline (-4.04 kcal/mol) showed moderate to weak affinity. Caffeic acid had the strongest intermolecular energy (-7.43 kcal/mol), suggesting strong hydrogen-bonding and hydrophobic interactions in the active site of AKT1, which is frequently mutated in TNBC, such as E17K [41-42].

For CDK2 (PDB 6GUK, val2), the reference binding energy was -8.54 kcal/mol with an RMSD of 0.64 Å. Chlorogenic acid showed the strongest performance with -6.49 kcal/mol (Ki 16.32 μM, intermolecular energy -9.77 kcal/mol), followed by caffeic acid (-6.14 kcal/mol) and trigonelline (-4.69 kcal/mol). The high affinity of chlorogenic acid is relevant because CDK2 phosphorylates EZH2 at T416, thereby maintaining the basal-like TNBC phenotype and Erα resistance [9].

Table 2. Molecular Docking Results of Coffee Bioactive Compounds against AKT1, CDK2, and CCNA2.

Protein	PDB	ΔG (kcal/mol)	RMSD	Bioaktive	ΔG (kcal/mol)	RMSD	inhibitor konstat (uM)	intermol energy
AKT1 (val1)	4GV1	-9.21	1.18	Caffeic acid	-5.94	28	44.55	-7.43
				Cafeine	-4.87	23.79	267.85	-4.87
				chlorogenic acid	-5.25	18.35	141.34	-8.53
				Trigonelline	-4.04	22.6	1.09	-4.34
CDK2 (val2)	6GUK	-8.54	0.64	Caffeic acid	-6.14	16.55	31.81	-7.63
				Cafeine	-4.7	19.03	357.96	-4.7
				chlorogenic acid	-6.49	16.32	17.54	-9.77
				Trigonelline	-4.69	20.25	362.79	-4.99
CCNA2 (val1)	7ACK	-7.52	2.96	Caffeic acid	-6.96	39.6	7.9	-8.45
				Cafeine	-4.79	35.52	305.69	-4.79
				chlorogenic acid	-1.79	30.22	51.59	-5.04
				Trigonelline	-4.48	33.94	520.23	-4.78

For CCNA2 (PDB 7ACK, val1), the reference binding energy was -7.52 kcal/mol with an RMSD of 2.96 Å. Caffeic acid showed the best performance (-6.96 kcal/mol, Ki 39.6 μM, intermolecular energy -8.45 kcal/mol), whereas chlorogenic acid showed weak binding (-1.79 kcal/mol). The higher

RMSD suggests greater pose flexibility, yet caffeic acid still appears to have potential for inhibiting cell-cycle progression through CCNA2, which is highly expressed in TNBC. [43]. Overall, AKT1 appears to be the most suitable target based on the best docking score and lowest RMSD, indicating high pose stability, while the higher RMSD of CCNA2 reflects greater site flexibility. Caffeic acid showed the strongest multi-target affinity, with an average binding energy below -6 kcal/mol, suggesting potential as a broad TNBC inhibitor (Table 3).

Based on the docking results for AKT1, CDK2, and CCNA2, caffeic acid was selected as the primary candidate for molecular dynamics analysis, with chlorogenic acid considered as a complementary candidate, particularly for CDK2. Caffeic acid was prioritized because it showed the most consistent interaction profile across all three targets, with comparatively better docking performance, especially against CCNA2, and a tendency to form stable interactions within the binding pocket. Chlorogenic acid was also considered for CDK2 because it exhibited the strongest binding and intermolecular energies for that target.

Candidate selection for molecular dynamics was based not only on docking scores, but also on binding consistency across multiple targets, the strength of noncovalent interactions, and biological relevance to TNBC pathways. As a multitarget compound, caffeic acid may simultaneously modulate proteins involved in proliferation, cell-cycle regulation, and TNBC progression. Therefore, molecular dynamics simulations of the caffeic acid–AKT1, caffeic acid–CDK2, and caffeic acid–CCNA2 complexes may provide a more reliable assessment of complex stability, conformational changes, and binding persistence over time.

Tabel 3. Comparative Docking Scores of Caffeic Acid, Caffeine, Chlorogenic Acid, And Trigonelline using AutoDockTools 1.5.6.

Compound	AKT1 kcal/mol (RMSD)	CDK2 kcal/mol (RMSD)	CCNA2 kcal/mol (RMSD)
Caffeic acid	-5.94 (28)	-6.14 (16.55)	-6.96 (39.6)
Caffeine	-4.87 (23.79)	-4.7 (19.03)	-4.79 (35.52)
CGA	-5.25 (18.35)	-6.49 (16.32)	-1.79 (30.22)
Trigonelline	-4.04 (22.6)	-4.69 (20.25)	-4.48 (33.94)

Visualization of the ligand–protein complexes using PyMOL showed that the test ligands occupied the active pocket (ATP-binding site) of AKT1, CDK2, and CCNA2 with orientations consistent with key active-site residues. This binding position suggests that the coffee bioactive compounds can structurally adapt to the binding domains of the target proteins [29]. For AKT1 (4GV1), caffeic acid showed dominant hydrogen bonding with Ser220, Ser478, and Thr291 in the hinge region, π-π stacking and hydrophobic interactions with Phe438 and Phe593 in the P-loop, and van der Waals interactions with Val164 and Ala177. Stability was further supported by a salt bridge between Asp274 and Arg73, consistent with clinical AKT1 inhibitors targeting the PIP3-binding region. A 2026 study also reported similar AKT–ligand interactions involving Lys189 in TNBC activation [12], [41].

For CDK2 (6GUK), chlorogenic acid formed strong hydrogen bonds with Glu81 and Glu128, along with hydrophobic interactions with Leu83, Ile10, and Leu298 around the ATP-binding pocket. These interactions suggest that chlorogenic acid may inhibit CDK2 activity through ATP-competitive binding [9]. For CCNA2 (7ACK), caffeic acid formed several hydrogen bonds with Tyr15, Asn132, and Asp196, as well as hydrophobic interactions with Ile49 and Val18. Although the RMSD of the

CCNA2 complex was higher than that of the other targets, the residue interaction pattern indicates that caffeic acid can still maintain stable affinity at the active site of CCNA2 [11], [43].

These findings indicate that coffee-derived natural compounds may have the potential to act as multitarget inhibitors against proteins involved in cell-cycle regulation and cell survival in TNBC, namely AKT1, CDK2, and CCNA2. This is consistent with the multitarget inhibition paradigm in TNBC through the CDK2 and PI3K/AKT pathways. AKT1 is a major regulator of PI3K/AKT/mTOR signaling and plays critical roles in proliferation, survival, metastasis, and therapy resistance in TNBC. AKT1 activation is known to enhance anti-apoptotic signaling and tumor progression via downstream pathways such as mTOR and cyclin/CDK signaling [37-38].

AKT1 activates PI3K/AKT/mTOR, the UFL1–AKT loop, and FDX1 phosphorylation to promote proliferation, cuproptosis resistance, and aggressive TNBC progression. CDK2/CCNA2 drives G1/S transition through EZH2 phosphorylation and E2F1–CCNA2 regulation, thereby enhancing invasion and metastasis in TNBC. Inhibition by coffee-derived compounds such as caffeic acid and chlorogenic acid may disrupt these pathways, consistent with the anticancer effects of coffee compounds through antioxidant and anti-inflammatory mechanisms [9], [44]

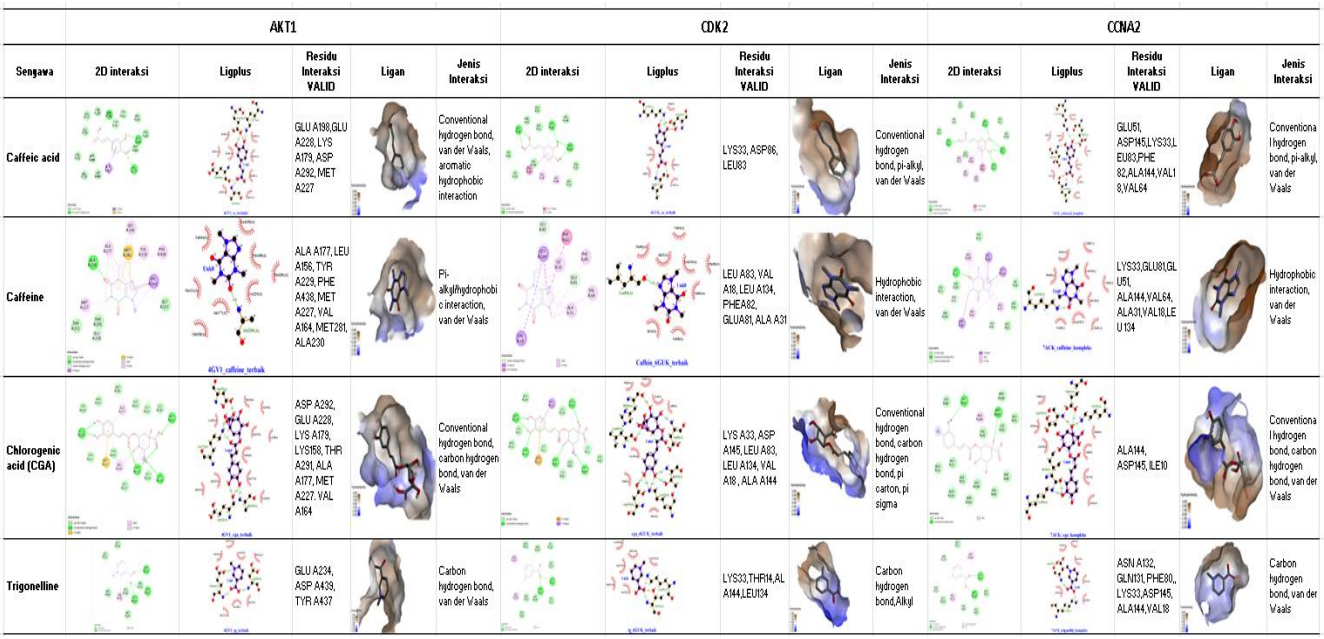

Figure 9. Protein–Ligand Interactions between AKT1, CDK2, and CCNA2 and Coffee Bioactive Compounds Visualized using AutoDockTools, LigPlot, and PyMOL.

The quantitative ADMET analysis (Table 4) indicated that caffeic acid exhibited more favorable pharmacokinetic properties than chlorogenic acid, with higher intestinal absorption (69.4% vs. 36.3%) and greater Caco-2 permeability (0.634 vs. −0.840 log Papp). Both compounds were predicted to be non-mutagenic (AMES negative), non-hepatotoxic, and non-cardiotoxic, with no inhibition of major CYP450 enzymes. These findings suggest that caffeic acid possesses superior oral bioavailability, whereas both compounds exhibit acceptable safety profiles, supporting their potential as multitarget candidates against TNBC.

Tabel 4. Predicted ADMET Properties of Caffeic Acid and Chlorogenic Acid Based

Parameter	Caffeic Acid	Chlorogenic Acid
Absorption		
Water solubility (log mol/L)	-2.33	-2.449
Caco-2 permeability (log Papp $\times 10^{-6}$ cm/s)	0.634	-0.84
Intestinal absorption (%)	69.407	36.377
P-glycoprotein substrate	No	Yes
P-glycoprotein I inhibitor	No	No
Distribution		
VDss (log L/kg)	-1.098	0.581
BBB permeability (log BB)	-0.647	-1.407
CNS permeability (log PS)	-2.608	-3.856
Excretion		
Total clearance (log ml/min/kg)	0.508	0.307
Renal OCT2 substrate	No	No
Toxicity		
AMES toxicity	No	No
hERG I inhibitor	No	No
hERG II inhibitor	No	No
Hepatotoxicity	No	No
Oral acute toxicity LD50 (mol/kg)	2.383	1.973
Oral chronic toxicity LOAEL (log mg/kg_bw/day)	2.092	2.982
Skin sensitisation	No	No
Minnow toxicity (log mM)	2.246	5.741

Kaplan–Meier Plotter analysis (Figure 10) showed that CCNA2 expression was significantly associated with breast cancer prognosis. Patients with high CCNA2 expression had a lower overall survival probability than those with low expression. A hazard ratio (HR) of 1.77, a 95% confidence interval of 1.59–1.96, and a log-rank p-value of $<1 \times 10^{16}$ indicate that CCNA2 overexpression is strongly associated with increased mortality risk in breast cancer patients. In addition, median survival was shorter in the high-expression group, suggesting that elevated CCNA2 expression is associated with more aggressive tumor progression [45].

In contrast to CCNA2, analysis of AKT1 showed no significant association between gene expression and overall survival. An HR of 1.0, a 95% confidence interval of 0.9–1.1, and a log-rank p-value of 0.93 indicate that AKT1 expression alone is not a significant predictor of breast cancer prognosis. This suggests that the biological role of AKT1 may be more strongly influenced by protein activation status and post-translational regulation than by gene expression level alone.

A similar result was observed for CDK2, where high and low expression groups did not differ significantly in survival outcomes. An HR of 1.07, a 95% confidence interval of 0.97–1.18, and a log-rank p-value of 0.2 indicate that CDK2 expression is not sufficiently strong as a standalone prognostic biomarker in breast cancer. Nevertheless, CDK2 remains biologically important in regulating the G1/S transition and cancer cell proliferation [9].

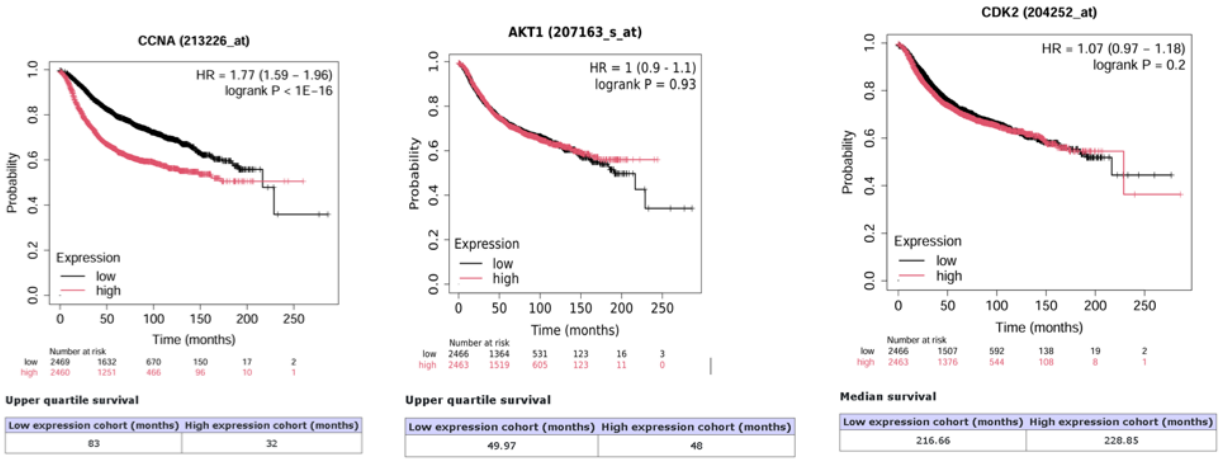

Figure 10. Kaplan–Meier Survival Curves Showing that CCNA2 Overexpression is Associated with Reduced Overall Survival in Breast Cancer Patients, whereas AKT1 and CDK2 show No Significant Prognostic Association.

Kaplan–Meier analysis showed that CCNA2 has the strongest prognostic value compared with AKT1 and CDK2. This finding supports the earlier transcriptomic and network analyses showing that CCNA2 is one of the major hub genes associated with proliferation, tumor progression, and TNBC aggressiveness [45-46]. Overall, the transcriptomic, network, enrichment, machine learning, and docking analyses indicate that TNBC has a strong dependence on interconnected cell-cycle and survival pathways. The selection of AKT1, CDK2, and CCNA2 as the main targets is justified because all three represent a biologically important axis in maintaining the aggressive TNBC phenotype. The docking results provide initial structural evidence that coffee bioactive compounds, especially caffeic acid and chlorogenic acid, may modulate these targets [8], [47-48].

4. Conclusion

Overall, this study demonstrates that TNBC is strongly dependent on proliferative programs involving cell cycle progression, DNA replication, mitotic checkpoint control, and genomic stability. Integration of five transcriptomic datasets yielded 917 consistent overlapping DEGs, supporting a robust molecular signature of TNBC. Network analysis, MCODE clustering, and functional enrichment further revealed that these genes form core modules coordinating cancer cell division. Topological analysis and machine learning validation identified ten major hub genes, with CCNA2 showing the strongest prognostic association with overall survival and emerging as a potential risk biomarker. Although AKT1 and CDK2 were not individually significant prognostic markers, they remain functionally important in PI3K/AKT signaling and G1/S transition. Chemogenomic mapping and molecular docking demonstrated favorable interactions of caffeic acid and chlorogenic acid with these targets. Collectively, these findings support the hypothesis that coffee bioactive compounds may serve as putative multitarget candidates for TNBC therapy.

Among the evaluated compounds, caffeic acid is explicitly recommended as the primary candidate for prioritization in wet-lab testing. This recommendation is based on its consistent multitarget binding profile across AKT1, CDK2, and CCNA2, coupled with favorable predicted ADMET properties, including superior intestinal absorption (69.4%), Caco-2 permeability (0.634 log Papp $\times 10^{-6}$ cm/s), and a non-toxic profile (AMES negative, non-hepatotoxic, non-cardiotoxic). Therefore, caffeic acid should be prioritized for future experimental validation, including cytotoxicity assays, enzymatic kinase inhibition assays, and mechanistic studies in TNBC cell models to confirm

therapeutic potential. Additionally, molecular dynamics simulations of the caffeic acid–AKT1, caffeic acid–CDK2, and caffeic acid–CCNA2 complexes are recommended to assess dynamic stability and binding persistence over time prior to experimental work.

References

[1] Jie, H., Ma, W., & Huang, C. (2025). Diagnosis, prognosis, and treatment of triple-negative breast cancer: a review. *Breast Cancer: Targets and Therapy*, 265-274.

[2] Giordano, A., & Tommonaro, G. (2019). Curcumin and cancer. *Nutrients*, *11*(10), 2376.

[3] Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, *71*(3), 209-249.

[4] Khoirunnisa, S. M., Setiawan, D., Postma, M. J., & de Jong, L. A. (2025). Trends in breast cancer in Indonesia from 2017 to 2020: A national-level analysis by age and disease severity. *Cancer treatment and research communications*, 101000.

[5] Lee, J. (2023). Current treatment landscape for early triple-negative breast cancer (TNBC). *Journal of clinical medicine*, *12*(4), 1524.

[6] Tutt, A. N., Garber, J. E., Kaufman, B., Viale, G., Fumagalli, D., Rastogi, P., ... & Geyer Jr, C. E. (2021). Adjuvant olaparib for patients with BRCA1-or BRCA2-mutated breast cancer. *New England Journal of Medicine*, *384*(25), 2394-2405.

[7] Sharma, P. (2016). Biology and management of patients with triple-negative breast cancer. *The oncologist*, *21*(9), 1050-1062.

[8] Yang, C., Wang, M., Gong, Y., Deng, M., Ling, Y., Li, Q., ... & Zhou, Y. (2023). Discovery and identification of a novel PI3K inhibitor with enhanced CDK2 inhibition for the treatment of triple negative breast cancer. *Bioorganic Chemistry*, *140*, 106779.

[9] Nie, L., Wei, Y., Zhang, F., Hsu, Y. H., Chan, L. C., Xia, W., ... & Hung, M. C. (2019). CDK2-mediated site-specific phosphorylation of EZH2 drives and maintains triple-negative breast cancer. *Nature communications*, *10*(1), 5114.

[10] Asghar, U., Witkiewicz, A. K., Turner, N. C., & Knudsen, E. S. (2015). The history and future of targeting cyclin-dependent kinases in cancer therapy. *Nature reviews Drug discovery*, *14*(2), 130-146.

[11] Wang, Y., Zhong, Q., Li, Z., Lin, Z., Chen, H., & Wang, P. (2021). Integrated profiling identifies CCNA2 as a potential biomarker of immunotherapy in breast cancer. *OncoTargets and therapy*, 2433-2448.

[12] Noorolyai, S., Shajari, N., Baghbani, E., Sadreddini, S., & Baradaran, B. (2019). The relation between PI3K/AKT signalling pathway and cancer. *Gene*, *698*, 120-128.

[13] He, Y., Sun, M. M., Zhang, G. G., Yang, J., Chen, K. S., Xu, W. W., & Li, B. (2021). Targeting PI3K/Akt signal transduction for cancer therapy. *Signal transduction and targeted therapy*, *6*(1), 425.

[14] Mannino, G., Kunz, R., & Maffei, M. E. (2023). Discrimination of green coffee (Coffea arabica and Coffea canephora) of different geographical origin based on antioxidant activity, high-throughput metabolomics, and DNA RFLP fingerprinting. *Antioxidants*, *12*(5), 1135.

[15] Murai, T., & Matsuda, S. (2023). The chemopreventive effects of chlorogenic acids, phenolic compounds in coffee, against inflammation, cancer, and neurological diseases. *Molecules*, *28*(5), 2381.

[16] Aarón, R. H., Sheila, C. M., Julio Emmanuel, G. P., Oscar, J. G., Aurelio, L. M., & Jocksan Ismael, M. C. (2025). In Silico strategies for drug discovery: optimizing natural compounds from foods for therapeutic applications. *Discover Chemistry*, *2*(1), 133.

[17] Paul, J. K., Azmal, M., Haque, A. S. N. B., Talukder, O. F., Meem, M., & Ghosh, A. (2024). Phytochemical-mediated modulation of signaling pathways: a promising avenue for drug discovery. *Advances in Redox Research*, *13*, 100113.

[18] Baião, A. R., Cai, Z., Poulos, R. C., Robinson, P. J., Reddel, R. R., Zhong, Q., ... & Gonçalves, E. (2025). A technical review of multi-omics data integration methods: from classical statistical to deep generative approaches. *Briefings in bioinformatics*, *26*(4), bbaf355.

[19] Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., ... & von Mering, C. (2021). The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic acids research*, *49*(D1), D605-D612.

[20] Chin, C. H., Chen, S. H., Wu, H. H., Ho, C. W., Ko, M. T., & Lin, C. Y. (2014). cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC systems biology*, *8*(Suppl 4), S11.

[21] Xing, Z., Wang, X., Liu, J., Zhang, M., Feng, K., & Wang, X. (2021). Expression and prognostic value of CDK1, CCNA2, and CCNB1 gene clusters in human breast cancer. *Journal of International Medical Research*, *49*(4), 0300060520980647.

[22] Song, S., Wang, Y., & Liu, P. (2022). DNA replication licensing factors: novel targets for cancer therapy via inhibiting the stemness of cancer cells. *International Journal of Biological Sciences*, *18*(3), 1211.

[23] Ge, S. X., Jung, D., & Yao, R. (2020). ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics*, *36*(8), 2628-2629.

[24] Davis, A. P., Grondin, C. J., Johnson, R. J., Sciaky, D., Wiegers, J., Wiegers, T. C., & Mattingly, C. J. (2021). Comparative toxicogenomics database (CTD): update 2021. *Nucleic acids research*, *49*(D1), D1138-D1143.

[25] Davis, A. P., King, B. L., Mockus, S., Murphy, C. G., Saraceni-Richards, C., Rosenstein, M., ... & Mattingly, C. J. (2010). The comparative toxicogenomics database: update 2011. *Nucleic acids research*, *39*(suppl_1), D1067-D1072.

[26] Agu, P. C., Afiukwa, C. A., Orji, O. U., Ezeh, E. M., Ofoke, I. H., Ogbu, C. O., ... & Aja, P. M. (2023). Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management. *Scientific reports*, *13*(1), 13398.

[27] Johnson, J., Chow, Z., Lee, E., Weiss, H. L., Evers, B. M., & Rychahou, P. (2021). Role of AMPK and Akt in triple negative breast cancer lung colonization. *Neoplasia*, *23*(4), 429-438.

[28] Wadhwa, B., Paddar, M., Khan, S., Mir, S., AClarke, P., Grabowska, A. M., ... & Malik, F. (2020). AKT isoforms have discrete expression in triple negative breast cancers and roles in cisplatin sensitivity. *Oncotarget*, *11*(45), 4178.

[29] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., ... & Bourne, P. E. (2000). The protein data bank. *Nucleic acids research*, *28*(1), 235-242.

[30] Huey, R., Morris, G. M., & Forli, S. (2011). Using AutoDock 4 and Vina with AutoDockTools: A Tutorial, Scripps Research Institute.

[31] Trott, O., & Olson, A. J. J. J. O. C. C. (2009). Software news and update AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function. *Effic. Optim. Multithreading*, *31*, 455-461.

[32] Agu, P. C., Afiukwa, C. A., Orji, O. U., Ezeh, E. M., Ofoke, I. H., Ogbu, C. O., ... & Aja, P. M. (2023). Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management. *Scientific reports*, *13*(1), 13398.

[33] El-Hachem, N., Haibe-Kains, B., Khalil, A., Kobeissy, F. H., & Nemer, G. (2017). AutoDock and AutoDockTools for protein-ligand docking: beta-site amyloid precursor protein cleaving enzyme 1 (BACE1) as a case study. In *Neuroproteomics: Methods and Protocols* (pp. 391-403). New York, NY: Springer New York.

[34] Lánczky, A., & Győrffy, B. (2021). Web-based survival analysis tool tailored for medical research (KMplot): development and implementation. *Journal of medical Internet research*, *23*(7), e27633.
[35] Vaniya, B., Patel, N., & Vora, H. (2023). A Study Of Mmp7 Expression In Triple Negative Breast Cancer Patients. *International Association of Biologicals and Computational Digest*, *2*(2), 1-6.
[36] Zhao, B., Xu, Y., Zhao, Y., Shen, S., & Sun, Q. (2020). Identification of potential key genes associated with the pathogenesis, metastasis, and prognosis of triple-negative breast cancer on the basis of integrated bioinformatics analysis. *Frontiers in Oncology*, *10*, 856.
[37] Quan, H., Yin, H., Wang, Z., Lv, Y., Sun, Q., & Yin, T. (2025). Identification of key hub genes and potential therapeutic drugs for nasopharyngeal carcinoma: Insights into molecular mechanisms and treatment strategies. *Brazilian Journal of Otorhinolaryngology*, *91*, 101618.
[38] Deng, Y., Han, Q., Mei, S., Li, H., Yang, F., Wang, J., ... & Zhang, T. (2019). Cyclin-dependent kinase subunit 2 overexpression promotes tumor progression and predicts poor prognosis in uterine leiomyosarcoma. *Oncology Letters*, *18*(3), 2845-2852.
[39] Otto, T., & Sicinski, P. (2017). Cell cycle proteins as promising targets in cancer therapy. *Nature Reviews Cancer*, *17*(2), 93-115.
[40] Manning, B. D., & Toker, A. (2017). AKT/PKB signaling: navigating the network. *Cell*, *169*(3), 381-405.
[41] Hassan, A., & Aubel, C. (2025). The PI3K/Akt/mTOR signaling pathway in triple-negative breast cancer: a resistance pathway and a prime target for targeted therapies. *Cancers*, *17*(13), 2232.
[42] Liu, R., Chen, Y., Liu, G., Li, C., Song, Y., Cao, Z., ... & Liu, Y. (2020). PI3K/AKT pathway as a key link modulates the multidrug resistance of cancers. *Cell death & disease*, *11*(9), 797.
[43] Lu, Y., Su, F., Yang, H., Xiao, Y., Zhang, X., Su, H., ... & Ling, X. (2022). E2F1 transcriptionally regulates CCNA2 expression to promote triple negative breast cancer tumorigenicity. *Cancer Biomarkers*, *33*(1), 57-70..
[44] Makiso, M. U., Tola, Y. B., Ogah, O., & Endale, F. L. (2024). Bioactive compounds in coffee and their role in lowering the risk of major public health consequences: A review. *Food science & nutrition*, *12*(2), 734-764.
[45] Gao, T., Han, Y., Yu, L., Ao, S., Li, Z., & Ji, J. (2014). CCNA2 is a prognostic biomarker for ER+ breast cancer and tamoxifen resistance. *PloS one*, *9*(3), e91771.
[46] Győrffy, B., Surowiak, P., Budczies, J., & Lánczky, A. (2013). Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. *PloS one*, *8*(12), e82241.
[47] Zhou, J. Z., Wen, J. Y., Xu, X. W., Zhao, N., Tang, J. J., Xiao, Y. R., ... & Zhang, Q. (2025). Dual inhibition of AKT and autophagy sensitizes triple negative breast cancer cells to carboplatin. *Translational Oncology*, *58*, 102434.
[48] Zhang, H. P., Jiang, R. Y., Zhu, J. Y., Sun, K. N., Huang, Y., Zhou, H. H., ... & Wang, X. J. (2024). PI3K/AKT/mTOR signaling pathway: an important driver and therapeutic target in triple-negative breast cancer. *Breast Cancer*, *31*(4), 539-551.