



OPEN ACCESS

EDITED BY

Dharmendra Kumar Yadav,
Gachon University, Republic of Korea

REVIEWED BY

Bo-Wei Zhao,
Zhejiang University, China
Saleem Y Bhat,
University of Pennsylvania, United States

*CORRESPONDENCE

Abul Kalam Azad Mandal,
✉ akamandal@rediffmail.com,
✉ akazadmandal@vit.ac.in

RECEIVED 29 October 2025

REVISED 24 November 2025

ACCEPTED 27 November 2025

PUBLISHED 19 December 2025

CITATION

Dharshini LCP and Mandal AKA (2025)
Network-based insights into
miR-30a-5p-mediated regulation and EGCG
targeting in triple-negative breast cancer.
Front. Bioinform. 5:1735106.
doi: 10.3389/fbinf.2025.1735106

COPYRIGHT

© 2025 Dharshini and Mandal. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with
these terms.

Network-based insights into miR-30a-5p-mediated regulation and EGCG targeting in triple-negative breast cancer

Loganathan Chandramani Priya Dharshini and
Abul Kalam Azad Mandal *

Department of Biotechnology, School of Biosciences and Technology, Vellore Institute of
Technology, Vellore, Tamil Nadu, India

Background: Triple-negative breast cancer (TNBC) is defined by the absence of ER, PR, and HER2 expression. This limits the targeted therapies, resulting in poor clinical outcomes. Identifying the molecular targets that can be regulated through miRNAs and natural compounds offers a potential therapeutic platform.

Methods: We combined transcriptomic profiling with miRNA target prediction to identify genes regulated by miR-30a-5p and assess their interaction with the green tea polyphenol, epigallocatechin gallate (EGCG). Differentially expressed genes (DEGs) from TCGA-TNBC datasets and miRNA targets from miRDB, TargetScan, and miRTarBase were screened for common genes. Then, the protein-protein interaction and network topology analyses were performed to identify key hub genes. Molecular docking and simulation were carried out with the four key genes against EGCG.

Results: Data integration yielded 393 overlapping genes and identified ten hub genes- *RRM2*, *KIF11*, *ANLN*, *CDC20*, *CCNA1*, *AGO2*, *YWHAZ*, *DTL*, *SKP2*, and *PCNA*. Pathway enrichment showed that all these hubs are involved in cell cycle and mitotic regulation, which was associated with poor TNBC prognosis. Mutation profiling revealed high alteration rates in *KIF11*, *ANLN*, *CDC20*, and *YWHAZ*, with increased missense mutations and C>T transitions. Molecular docking and simulations identified *YWHAZ* as the most favorable and structurally stable EGCG-binding target, compared to the other three key genes.

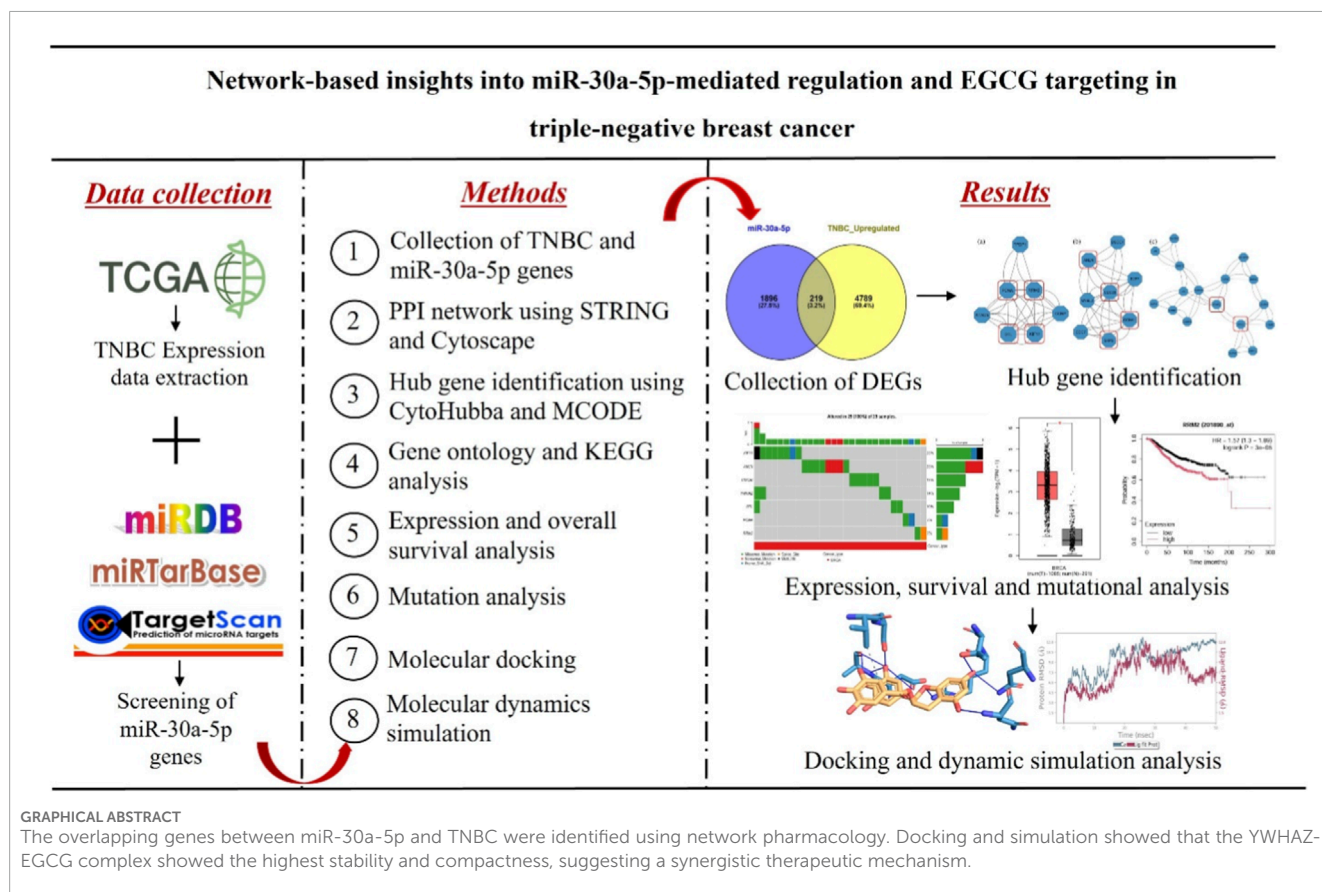
Conclusion: The results emphasizes that EGCG has strong binding affinity towards *YWHAZ*, revealing that miR-30a-EGCG targets TNBC synergistically through cell-cycle-mediated pathways. The findings give rational support for miRNA-guided phytochemical-based TNBC therapeutic development.

KEYWORDS

network pharmacology, TNBC, hsa-miR-30a-5p, EGCG, cell cycle regulation, molecular docking, molecular dynamics simulation

1 Introduction

Triple-negative breast cancer (TNBC), a molecularly diverse subtype of BC, accounts for 15%–20% of all BC cases (Wolff et al., 2013). It encompasses six subtypes and disproportionately affects premenopausal women (<40 years of age), mostly of African American and Hispanic descent from lower socioeconomic backgrounds (Morris et al., 2007). It is defined by heightened risk of early relapse, high metastatic



potential, and dismal prognosis. After initial diagnosis, nearly half of all recurrences occur within 3–5 years, and patients under current treatment regimens have a median overall survival (OS) of only 10.2 months (Benotto et al., 2014). TNBC exhibits a distinct metastatic pattern, preferentially spreading to the brain, lungs, liver, and central nervous system (CNS) (Pradhan et al., 2023). The median time to recurrence ranges from 19 to 40 months, with a staggering 75% mortality rate within 3 months of recurrence (Yin et al., 2020). The distinct molecular phenotype and a lack of well-defined targets make TNBC resistant to endocrine and HER-2-based receptor-targeted therapies (Li et al., 2022). Although chemo-immunotherapy has shown efficacy in individuals exhibiting elevated programmed cell death ligand 1 (PD-L1) expression, major challenges such as therapeutic resistance, persistent metastatic burden, off-target toxicity, and acquired drug resistance continue to limit its success (Chaudhary et al., 2018). These difficulties emphasize the call for the identification of molecular targets and the development of targeted drug delivery to enhance patient outcomes.

MicroRNAs (miRNAs), approximately 19–23 nts, are short non-coding master regulators of post-transcriptional gene expression that drive numerous cellular processes from tumor initiation to metastasis. miRNAs that are dysregulated can drive oncogenesis through complex regulatory networks, eventually leading to therapy resistance (Kim and Croce, 2023). A known tumor suppressor is miR-30a, which inhibits proliferation, invasion, and migration while inducing apoptosis, thereby restraining tumor survival (Jiang et al., 2018). In TNBC, the expression of miR-30a is

significantly decreased, correlating with higher histological grade and lymph node metastasis (Santana et al., 2023). Wang et al. (2024) observed that rosmarinic acid counteracts miR-30a-5p-mediated suppression of B-cell leukemia 2L11 (BCL2L11), increasing programmed cell death in MDA-MB-231-derived BC stem-like cells. Natural compounds are acquiring importance in oncological studies as sources of multi-target therapeutics. Green tea (*Camellia sinensis*) is highly known for its bioactive catechins, comprising 20%–30% of its dry weight (Chacko et al., 2010). Among which, EGCG is the most abundant and pharmacologically active catechin, exhibiting a broad anticancer effect (Aggarwal et al., 2022). It modulates several oncogenic pathways (PI3K/AKT, STAT3, and β -catenin), thereby inhibiting the nuclear translocation of their downstream effectors. This harmonized chaos induces apoptosis while suppressing the anti-apoptotic genes (Marin et al., 2023), fortifying its potential as a targeted anticancer agent.

High-throughput transcriptomic profiling provides a potent approach for unwinding the molecular level complexity of TNBC, allowing for the discovery of biomarkers, gene signatures, regulatory networks, and newer therapeutic targets (Supplitt et al., 2021). Elango et al. (2023) reported CDC123 as a potential target of TNBC through transcriptomic and enrichment analyses, while Kaddoura et al. (2023) reported GATA3 and FOXA1 as important regulators within estrogen-dependent signaling pathways across multiple TNBC datasets.

In this study, we applied an integrative bioinformatics framework to identify miR-30a-5p-regulated molecular targets in

TNBC. Furthermore, to explore the potential synergistic effect, we assessed the binding affinity of EGCG toward key hub genes via molecular docking, followed by molecular dynamics simulations (MDS) to evaluate the interaction dynamics and EGCG's stability within the binding pockets. This multi-layered approach provides mechanistic insight into miRNA-phytochemical synergy and supports the rational development of multi-target therapeutic strategies of TNBC.

2 Materials and methods

2.1 Screening and processing the DEGs in TNBC

To delineate differentially expressed genes (DEGs), RNASeq gene expression data of BRCA were retrieved from the TCGA database. Differential gene expression analysis was carried out using the R package “DESeq2”, and the Wald test was applied for significance testing. Then, the genes are ranked by adjusted p -values based on the Benjamini-Hochberg algorithm. Genes with $\log_2 \text{FC} > 1$ and $\text{adj. } p < 0.05$ were considered upregulated, while those with $\log_2 \text{FC} < -1$ and $\text{adj. } p < 0.05$ were considered downregulated genes. The functional annotation of the identified DEGs was conducted using the g:Profiler, a web-based interface (Cui et al., 2024; Zhao et al., 2022). Finally, visualization of DEGs was performed using the “ggplot2” package and presented as a volcano plot. This DEG profiling workflow and code were obtained from the methodology described by Krishnamoorthy and Karuppasamy (2024). The preprocessing pipeline for TNBC DEGs is available at the following GitHub repository (https://github.com/vitmilab/lab-projects/tree/main/DEG_Identification). The workflow of the study is depicted in Figure 1.

2.2 Selection of miR-30a-5p target genes

The experimentally verified targets of miR-30a-5p were obtained from databases such as miRTarBase, miRDB, and TargetScan. miRTarBase (<https://ngdc.cncb.ac.cn/databasecommons/database/id/167>) focuses exclusively on experimentally validated miRNA-target interactions supported by high-throughput evidence (Chen and Wang, 2020; Zhao et al., 2024). miRDB (<https://mirdb.org/>) employs the MirTarget computational tool to analyze miRNA-target interactions and their regulatory effects based on high-throughput sequencing (HTS) data (McGeary et al., 2019). TargetScan (https://www.targetscan.org/vert_80/) identifies potential miRNA targets by recognizing conserved motifs aligning the miRNA seed regions (Raudvere et al., 2019).

2.3 Acquisition of the common targets of miR-30a-5p and TNBC DEGs

The common targets between “upregulated DEGs vs. miR-30a-5p targets” and “downregulated DEGs vs. miR-30a-5p targets” were identified through “Venny 2.1 (<https://bioinfogp.cnb.csic.es/tools/>)

venny/) by entering the miR-30a-5p targets and disease genes in the query box (Oliveros, 2007).

2.4 PPI network construction

Protein-protein interactions (PPIs) among identified DEGs were assessed using the STRING database v12.0 (<https://string-db.org/>) (Scklarczyk et al., 2023; von Mering et al., 2004). Predicted interactions for the target genes were extracted using a medium confidence score ≥ 0.4 , as calculated by STRING's probabilistic scoring system (Equation 1). The confidence score shows the estimated likelihood that a predicted interaction reflects a true biological association, derived from combined evidence including experiments, co-expression, curated databases, and text mining. A score of ≥ 0.4 was selected to balance network connectivity and specificity, as using stricter cutoffs of ≥ 0.7 shattered the PPI network and removed valid interactions.

$$S = 1 - \prod_i (1 - S_i) \quad (1)$$

2.5 Visualization of protein-interaction network (PIN) and key gene identification

Protein-interaction network (PIN) from STRING was imported and visualized through Cytoscape 3.10.3 (<https://cytoscape.org/>). Topological analysis was performed using the cytoHubba plugin to identify the top 10 hub genes, through six centrality measures, including maximal centrality clique (MCC), maximum neighbour component (MNC), degree, closeness, betweenness, and stress centralities (Equations 2–7) (Chin et al., 2014). For each gene, the ranks from all six algorithms were computed to obtain the mean rank, and genes with the highest composite scores were defined as hub genes.

$$MCC(v) = \sum_{C \in S(v)} (|C| - 1)! \quad (2)$$

Here, the maximal clique's collection containing v is represented by $S(v)$, while the product of all positive integers less than $|C|$ is denoted by $(|C| - 1)!$. Absence of an edge between the neighbour node v , then $MCC(v)$ equals degree.

$$MNC(v) = |V(MC(V))| \quad (3)$$

In this case, the maximum connected component of the $G[N(V)]$ is denoted by $MC(V)$ and $G[N(V)]$ is the induced subgraph of G by $N(V)$.

$$Deg(v) = |N(V)| \quad (4)$$

Here, the node is represented by V and the collections of its neighbours are denoted by $N(V)$.

$$Clo(v) = \sum_{w \in V} \frac{1}{dist(v, w)} \quad (5)$$

$$BC(v) = \sum_{s \neq t \in V \cap C(v)} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad (6)$$

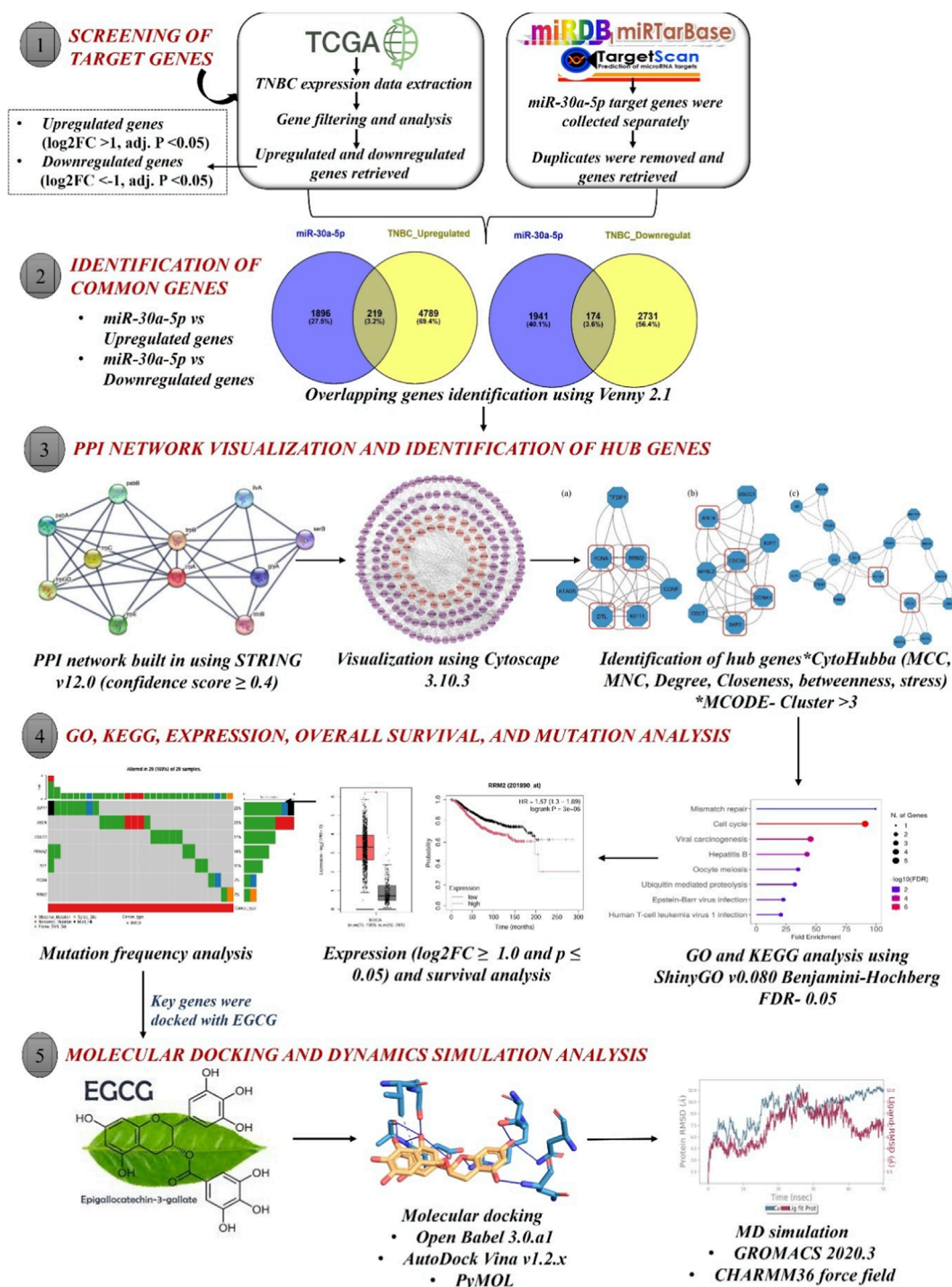


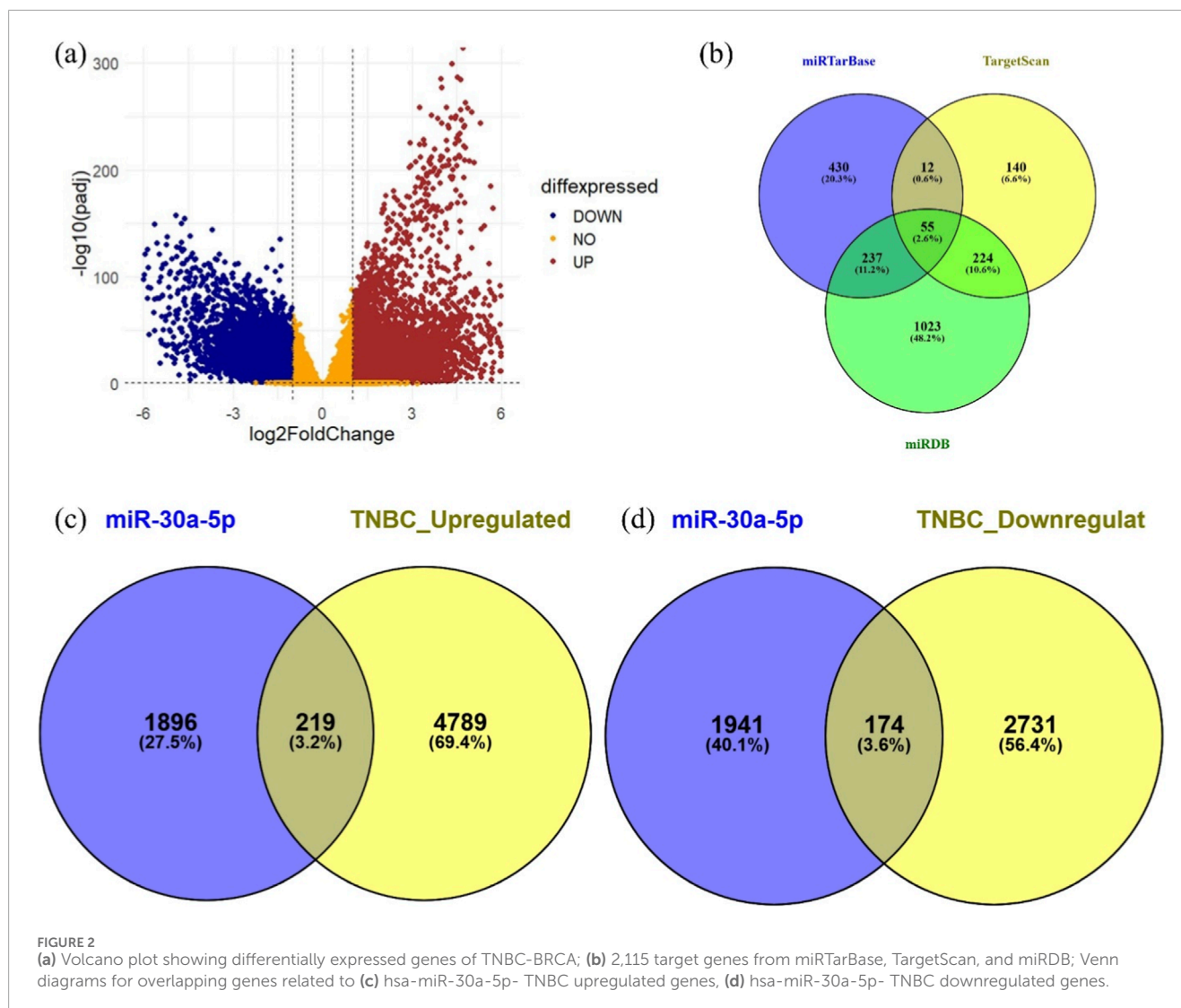
FIGURE 1
Graphical illustration depicting the overall experimental workflow.

In this case, the number of shortest paths between nodes s and t is represented by σ_{st} .

$$Str(v) = \sum_{s \neq t \in V(v)} \sigma_{st}(v) \quad (7)$$

Here, the number of shortest paths from the node s to t is represented by $\sigma_{st}(v)$, which uses the node v .

Gene clusters that are tightly linked within the PIN were extracted from the Molecular Complex Detection (MCODE) plugin. Core targets identified by cytoHubba were cross-referenced with the MCODE clusters with scores > 3 , using default parameters, to further refine the selection of biologically significant genes (Shannon et al., 2003).



2.6 Gene ontology (GO) and pathway enrichment analysis

The biological significance and pathways associated with the key genes were determined using ShinyGO v0.80 (<https://bioinformatics.sdstate.edu/go/>). The significance was assessed using Benjamini-Hochberg false discovery rate (FDR) of 0.05 (Ge et al., 2019). Subsequently, the top 10 GO pathways involved in biological process (BP), molecular function (MF), and cellular component (CC), along with Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, were selected. Fold enrichment of each term was obtained directly from the ShinyGO enrichment output. Fold enrichment shows the overrepresented genes in a particular pathway and is calculated as observed/expected overlap (k/E), where k is the number of submitted genes found in the GO/KEGG term, and E is the expected overlap, which was calculated using the formula: $E = \frac{K}{N} \times n$, where n is the number of genes submitted (10 genes), K is the number of genes annotated to that term, and N is the background universe automatically selected by the enrichment tool.

2.7 Expression and survival analysis

The expression levels between TNBC vs. normal breast tissues were assessed using the Gene Expression Profiling Interactive Analysis 2 (GEPIA2) server (<http://gepia2.cancer-pku.cn/#index>) (Tang et al., 2019), with $\log_2 \text{FC} \geq 1.0$ and $p \leq 0.05$. Furthermore, through Kaplan-Meier (KM) plotter (<https://kmplot.com/analysis/>), the overall survival (OS) analysis in TNBC patients was evaluated.

2.8 Mutation analysis

The mutation frequency was analyzed through the Gene Set Cancer Analysis (GSCA) database (<https://guolab.wchscu.cn/GSCA/#/>). This database provides a comprehensive depot of genomic alterations, including single-nucleotide variants (SNVs), in a variety of cancer forms, including breast invasive carcinoma (BRCA). SNV data corresponding to the hub genes were extracted

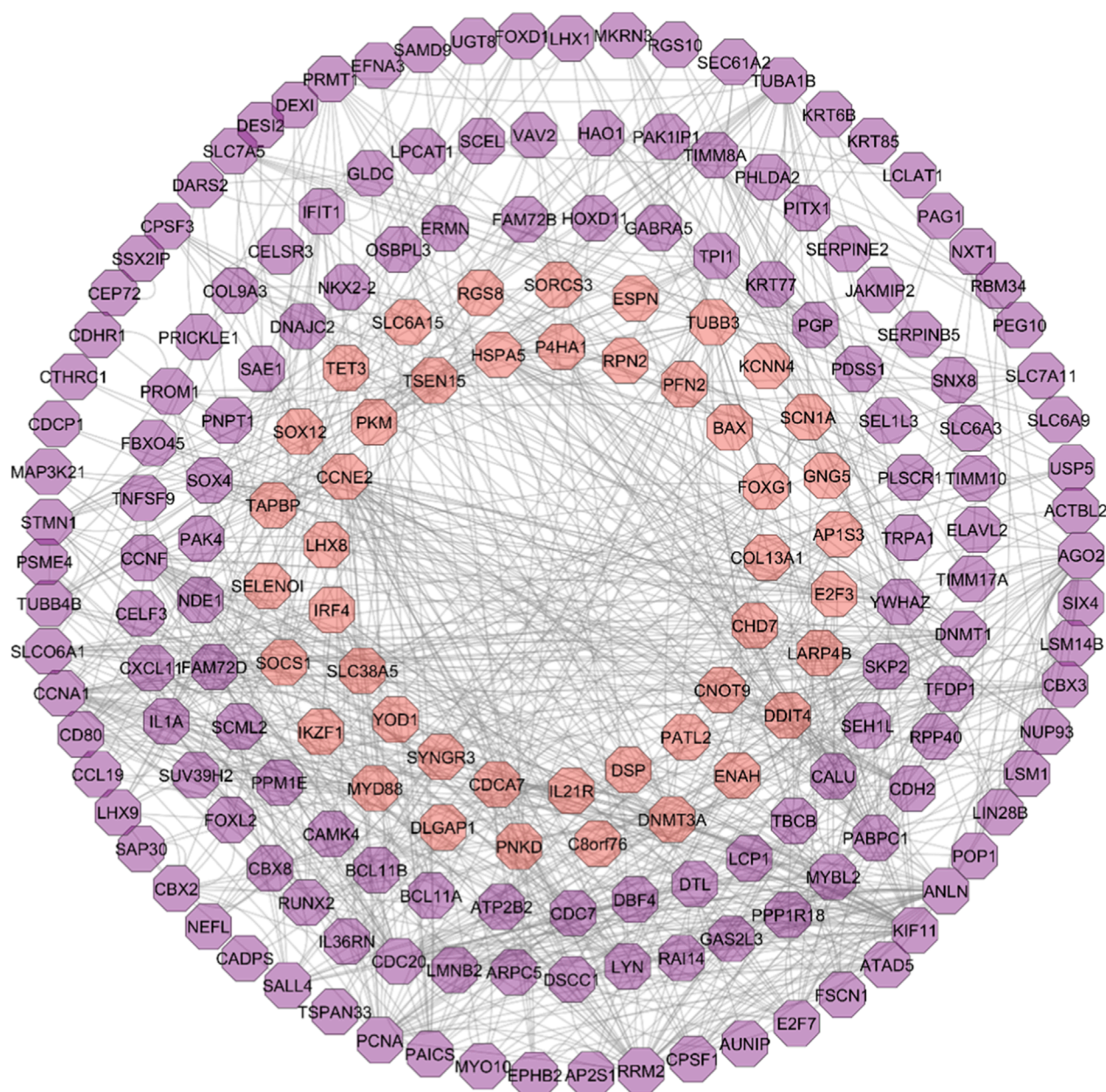


FIGURE 3

Visualization of protein-interaction network using Cytoscape. Purple colour indicates the upregulated genes (219 genes) and orange indicates downregulated genes (174 genes).

and evaluated to determine the mutation frequency and variant types observed in BRCA samples.

2.9 Molecular docking

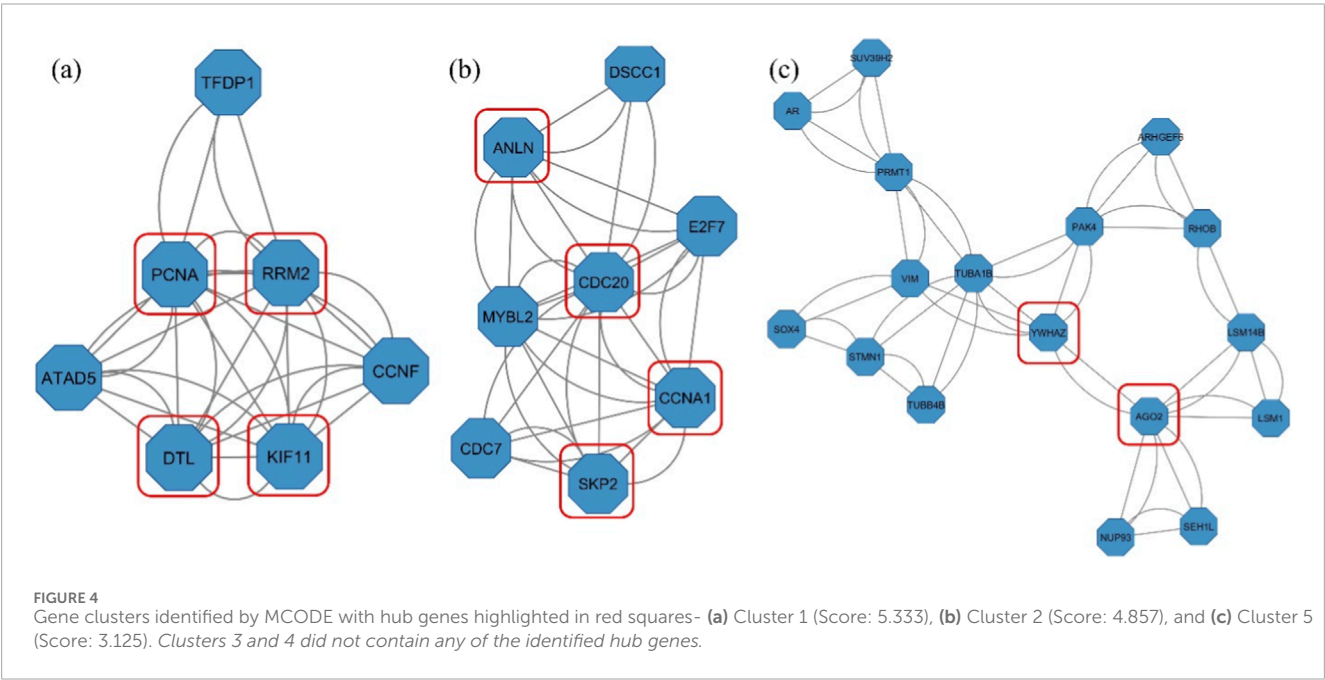
The target proteins' 3D structure was obtained from RCSB Protein Data Bank (<https://www.rcsb.org/>) in PDB format. AutoDock Tools (ADT) 1.5.7 was used for protein preparation, where water molecules, ligands, and heteroatoms were removed. Subsequently, polar and Kollman charges were added, and the processed structures were saved in PDBQT format. The

ligand (EGCG) structure was downloaded from PubChem (CID: 65,064) in SDF format and converted to 3D using Open Babel 3.0.a1, and PDBQT files were generated using AutoDock Tools. A grid box was defined for each target protein to encompass the known active/binding site. Molecular docking was performed using AutoDock Vina v1.2.x. The output PDBQT files were converted and visualized using PyMOL to analyse interactions and binding conformations. The binding affinities (kcal/mol) of the top-ranked poses were recorded. Binding energy represents the predicted free energy of ligand-protein binding $\Delta G_{binding}$, where more negative values indicate stronger and more

TABLE 1 Top 10 genes ranked by cytoHubba based on six topological parameters.

S.No	MCC	MNC	Degree	Closeness	Betweenness	Stress
1	<i>DTL</i> *	<i>DTL</i> *	<i>ANLN</i> *	<i>CFL2</i>	<i>LYN</i>	<i>SNX1</i>
2	<i>RRM2</i> *	<i>RRM2</i> *	<i>RRM2</i> *	<i>ANLN</i> *	<i>ANLN</i> *	<i>LEPR</i>
3	<i>MYBL2</i>	<i>ANLN</i> *	<i>AGO2</i> *	<i>VIM</i>	<i>AGO2</i> *	<i>TIMM8A</i>
4	<i>E2F7</i>	<i>PCNA</i> *	<i>CDC20</i> *	<i>AR</i>	<i>CDC20</i> *	<i>PCNA</i> *
5	<i>CDC20</i> *	<i>CDC20</i> *	<i>YWHAZ</i> *	<i>AGO2</i> *	<i>YWHAZ</i> *	<i>AGO2</i> *
6	<i>KIF11</i> *	<i>KIF11</i> *	<i>KIF11</i> *	<i>CDC20</i> *	<i>ITSN1</i>	<i>YWHAZ</i> *
7	<i>CDC7</i>	<i>CCNA1</i> *	<i>CCNA1</i> *	<i>YWHAZ</i> *	<i>CDH2</i>	<i>CDH2</i>
8	<i>CCNA1</i> *	<i>CDH2</i>	<i>CDH2</i>	<i>CDH2</i>	<i>BDNF</i>	<i>RPP40</i>
9	<i>CCNF</i>	<i>DNMT1</i>	<i>DNMT1</i>	<i>BDNF</i>	<i>DNMT1</i>	<i>BDNF</i>
10	<i>SKP2</i> *	<i>SKP2</i> *	<i>BDNF</i>	<i>DNMT1</i>	<i>RUNX2</i>	<i>DNMT1</i>

From the 12 topological parameters in cytoHubba- MCC, MNC, degree, Closeness, Betweenness, and Stress were selected to identify the most relevant hub genes (Repeated times ≥ 2). These shortlisted genes were further validated by cross-referencing with the gene cluster score >3 identified by the MCODE, plugin. Genes highlighted in*represent key hub genes identified through this integrative analysis.



favorable interactions. Using the PLIP server, the interactions (hydrogen bonding, π - π stacking, and hydrophobic interactions) were evaluated.

2.10 Molecular dynamics (MD) simulation

GROMACS 2020.3 package with the CHARMM36 force field was used for MD simulations to ascertain the structural stability and integrity of the four gene-ligand complexes

(Van Der Spoel et al., 2005). Ligand parameter and topology files were generated through the CHARMM General Force Field (CGenF). Each protein-ligand complex was placed in a dodecahedron box using a simple point charge (SPC) water model, and the system's neutrality was ensured by introducing three chlorine counterions. The system's energy minimization was carried out using the steepest descent algorithm to eliminate unfavourable weak Van der Waals contacts. The algorithms, like the Particle Mesh Ewald (PME) method and the linear constraint solver (LINCS), were applied to constrain the electrostatic and hydrogen bond

TABLE 2 Gene ontology (GO) analysis of the hub genes.

Pathways	No. of genes	Gene ID	Fold enrichment
Biological process			
Regulation of synapse maturation	2	<i>YWHAZ CDC20</i>	305.08
Mitotic cell cycle phase transition	6	<i>CCNA1 DTL SKP2 CDC20 RRM2 ANLN</i>	28.841
Regulation of mitotic cell cycle phase transition	4	<i>DTL CDC20 RRM2 ANLN</i>	25.636
Mitotic cell cycle process	8	<i>ANLN CCNA1 KIF11 DTL SKP2 CDC20 RRM2 PCNA</i>	22.187
Mitotic cell cycle	8	<i>ANLN CCNA1 KIF11 DTL SKP2 CDC20 RRM2 PCNA</i>	18.378
Regulation of cell cycle process	5	<i>DTL CDC20 RRM2 KIF11 ANLN</i>	14.648
Cell cycle process	8	<i>ANLN CCNA1 KIF11 DTL SKP2 CDC20 RRM2 PCNA</i>	13.856
Regulation of cell cycle	7	<i>CCNA1 DTL SKP2 CDC20 RRM2 KIF11 ANLN</i>	13.204
Cell cycle	8	<i>ANLN CCNA1 KIF11 DTL SKP2 CDC20 RRM2 PCNA</i>	9.450
Cellular component			
Cyclin-dependent protein kinase holoenzyme complex	2	<i>CCNA1 PCNA</i>	84.744
Cullin-RING ubiquitin ligase complex	3	<i>CDC20 SKP2 DTL</i>	36.127
Ubiquitin ligase complex	3	<i>CDC20 SKP2 DTL</i>	20.927
Transferase complex	5	<i>CDC20 CCNA1 SKP2 DTL PCNA</i>	12.462
Intracellular protein-containing complex	4	<i>CDC20 SKP2 AGO2 DTL</i>	10.353
Catalytic complex	7	<i>CDC20 CCNA1 SKP2 AGO2 DTL PCNA RRM2</i>	9.449
Microtubule cytoskeleton	5	<i>PCNA CCNA1 KIF11 DTL CDC20</i>	7.939
Cytoskeleton	6	<i>ANLN PCNA CCNA1 KIF11 DTL CDC20</i>	5.377
Nucleoplasm	8	<i>ANLN CDC20 PCNA DTL SKP2 AGO2 CCNA1 YWHAZ</i>	3.995
Nuclear lumen	8	<i>ANLN CDC20 PCNA DTL SKP2 AGO2 CCNA1 YWHAZ</i>	3.680
Molecular function			
Purine-specific mismatch base pair DNA N-glycosylase activity	1	<i>PCNA</i>	762.7
Ribonucleotide-diphosphate reductase activity thioredoxin disulfide as acceptor	1	<i>RRM2</i>	762.7
Oxidoreductase activity acting on CH and CH2 groups disulfide as acceptor	1	<i>RRM2</i>	762.7
DNA polymerase processivity factor activity	1	<i>PCNA</i>	762.7
Dinucleotide insertion or deletion binding	1	<i>PCNA</i>	762.7
Ribonucleotide-diphosphate reductase activity	1	<i>RRM2</i>	762.7
Ubiquitin ligase activator activity	1	<i>CDC20</i>	572.025
Protein C-terminus binding	3	<i>PCNA CDC20 AGO2</i>	33.814
Enzyme activity	5	<i>AGO2 CDC20 PCNA KIF11 YWHAZ</i>	5.114

TABLE 3 Top enriched pathways of the hub genes.

S. No	KEGG pathways	No. of genes	Gene id	Fold enrichment
1	Mismatch repair	1	<i>PCNA</i>	99.482
2	Cell cycle	5	<i>PCNA SKP2 YWHAZ CCNA1 CDC20</i>	90.797
3	Viral carcinogenesis	4	<i>SKP2 YWHAZ CCNA1 CDC20</i>	45.308
4	Hepatitis B	3	<i>PCNA YWHAZ CDC20</i>	42.372
5	Oocyte meiosis	2	<i>YWHAZ CDC20</i>	34.932
6	Ubiquitin mediated proteolysis	2	<i>SKP2 CDC20</i>	32.226
7	Epstein-Barr virus infection	2	<i>SKP2 CCNA1</i>	22.654
8	Human T-cell leukemia virus 1 infection	2	<i>CCNA1 CDC20</i>	20.613

TABLE 4 Summary of docking results showing the PDB ID of target proteins, predicted binding affinities (kcal/mol), number of hydrogen bonds, and key interacting residues for EGCG.

Gene name	PDB id	Binding affinity	No. of hydrogen bonds	Key interacting residues
KIF11	3K5E	−8.5	11	Asn 98, Asp 187, Asn 190, Arg 192, Val 194, Lys 260, Asn 262, Asp 322, Ser 323, Arg 327, Thr 328
ANLN	2Y7B	−8.9	7	Asn 1,033, Asn 1,036, Thr 1,060, Arg 1,062, Asp 1,068, Arg 1,069, Ser 1,074
CDC20	6F08	−7	4	Ser 45, Lys 120, Arg 127, Asn 173
YWHAZ	4GGC	−10.1	8	Asp 191, Trp 234, Ser 275, Trp 276, Arg 316, Trp 363, Ser 404, Ser 448

interactions. The system was equilibrated using canonical NVT (number of particles, volume, and temperature) and isobaric NPT (number of particles, pressure, and temperature) ensembles to stabilize its thermodynamic conditions. Berendsen thermostat used to heat the system to 300 K gradually, with a time lapse of 0.1 ps and a pressure of 1 bar. Then, the MD simulation was carried out for 100 ns with a 2-fs integration time step, and the trajectory data (RMSD, RMSF, Rg, and SASA) were analyzed based on the [Formula \(8-10\)](#). SASA is calculated from the atomic coordinates of a native protein numerically ([Ghahremanian et al., 2022](#)).

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_1^m - x_1^1)^2 + (y_1^m - y_1^1)^2 + (z_1^m - z_1^1)^2} \tag{8}$$

Where x^m, y^m, z^m are the initial coordinates and x^1, y^1, z^1 are the trajectory coordinates at frame t, and N is the number of atoms.

$$RMSF = \sqrt{\frac{1}{T} \sum_{i=1}^T (x_i - \bar{x})^2} \tag{9}$$

Where N is the trajectory frame numbers and \bar{x} is the time-averaged position.

$$Rg = \sqrt{\frac{1}{N} \sum_{i=1}^N |r(i) - r_{center}|^2} \tag{10}$$

Where $r(i)$ is the coordinates of the atom i , and r_{center} is the center of mass, and N is the number of protein atoms.

3 Results

3.1 Identification of DEGs

TNBC-specific transcriptomic data were obtained from the TCGA-BRCA dataset. The dataset was preprocessed, and the DEGs were identified using the “DESeq2 package” in R, ensuing an initial finding of 43,881 DEGs. Subsequently, the data were filtered based on log2 FC and adj. p value, resulting in a 5008 upregulated and 2,905 downregulated refined gene list within the TNBC dataset.

Functional annotation of these DEGs was performed using g:Profiler to traverse through their biological importance. Expression of DEGs was visually interpreted through a volcano plot ([Figure 2a](#)), markedly showing upregulated genes in red and downregulated genes in blue, giving a clear depiction of the expression landscape.

An integrated target gene set of 2,115 genes was compiled through the integrative curation of 2,704 hsa-miR-30a-5p-associated genes, sourced from three well-established miRNA target prediction databases: miRTarBase (734 genes), TargetScan (431 genes), and miRDB (1,539 genes) ([Figure 2b](#)). Redundancy was systematically eliminated to ensure maximal data fidelity in the gene set.

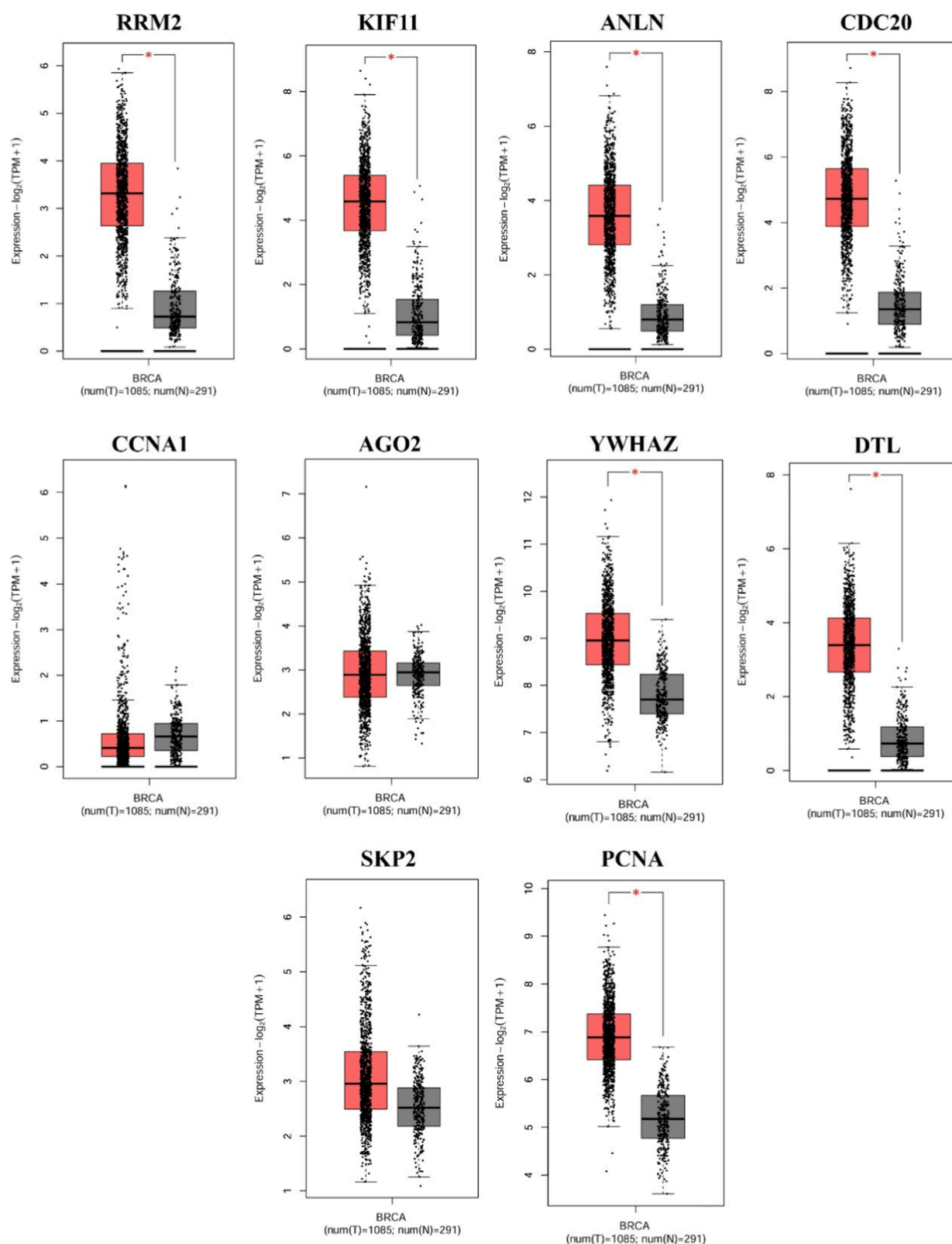


FIGURE 5

Expression validation of hub genes between normal and TNBC samples using GEPIA2. Box plots illustrate the expression levels of each hub gene in TNBC samples (left) compared to normal breast tissue (right) (*indicates $p < 0.05$).

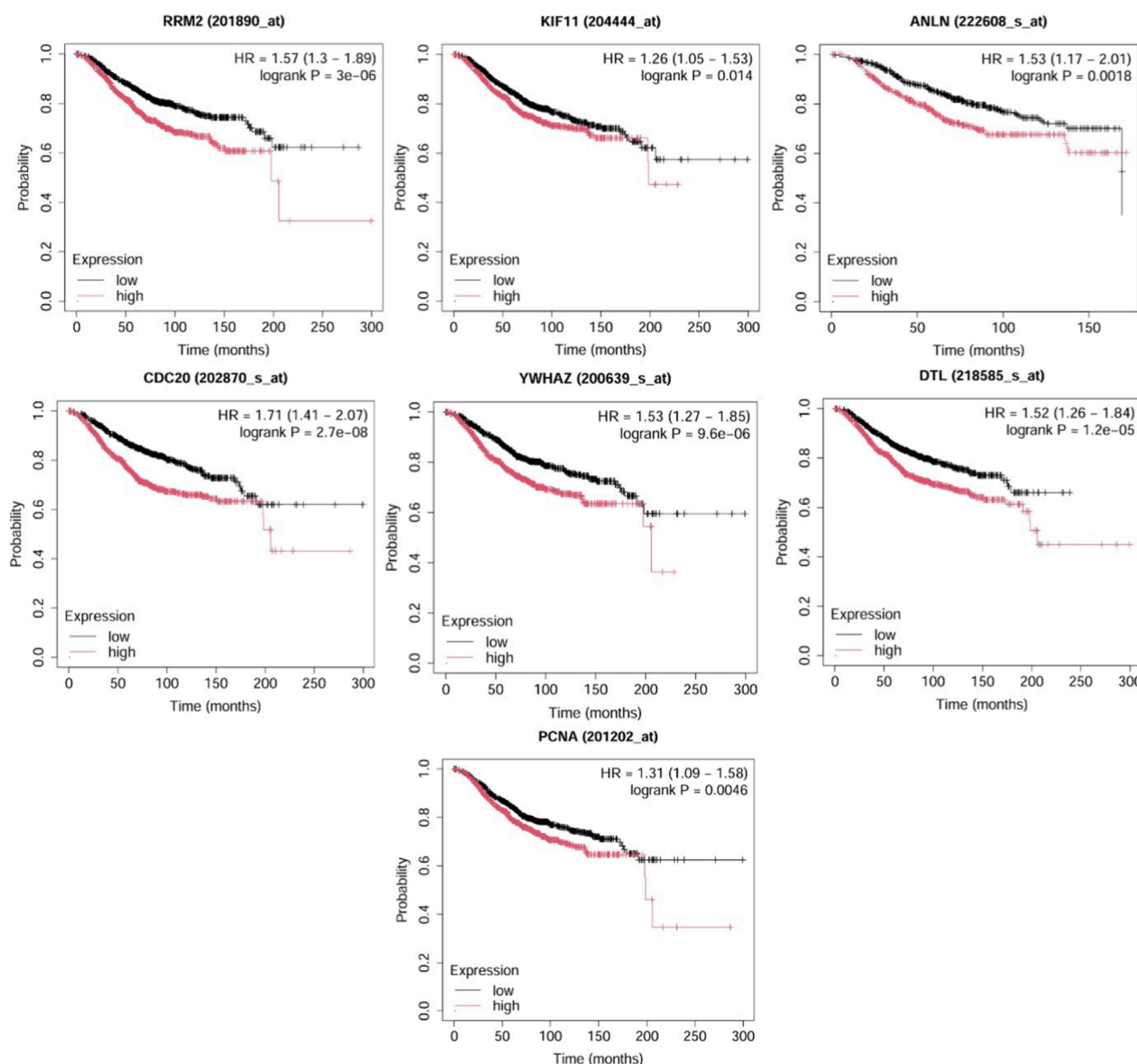


FIGURE 6

Overall survival analysis of the hub genes in TNBC patients using KM Plotter. High expression of each gene is associated with reduced overall survival, with $p < 0.05$ and $HR > 1$ considered statistically significant.

The convergence between miR-30a-5p target genes and TNBC-related genes was examined using InteractiVenn (Figures 2c,d). This integrative analysis revealed 393 overlapping genes, comprising 219 upregulated and 174 downregulated genes. All these genes are differentially expressed in TNBC and predicted to contain validated or putative miR-30a-5p binding. These represent functionally relevant targets of miR-30a-5p within the TNBC.

3.2 Hub gene identification from PPI network via CytoHubba and MCODE

The PPI analysis of the DEGs from the STRING database with a confidence score of ≥ 0.4 generated 392 nodes and 622

edges. The network exhibited an average node degree of 3.17 and an average local clustering coefficient of 0.351. Enrichment analysis yielded a highly significant PPI enrichment p -value of $2.49e-13$ ($p < 0.05$), indicating that the observed interactions are unlikely to be random and reflect biologically meaningful associations.

The resulting PIN from STRING was imported into Cytoscape for further topological assessments (Figure 3). The network topology showed a clustering coefficient of 0.207 and a characteristic path length of 4.187, indicating that the network holds sturdy connections, toughness, and flexibility.

CytoHubba plugin comprised the top 10 hub genes involved in TNBC survival and progression based on the ranking obtained

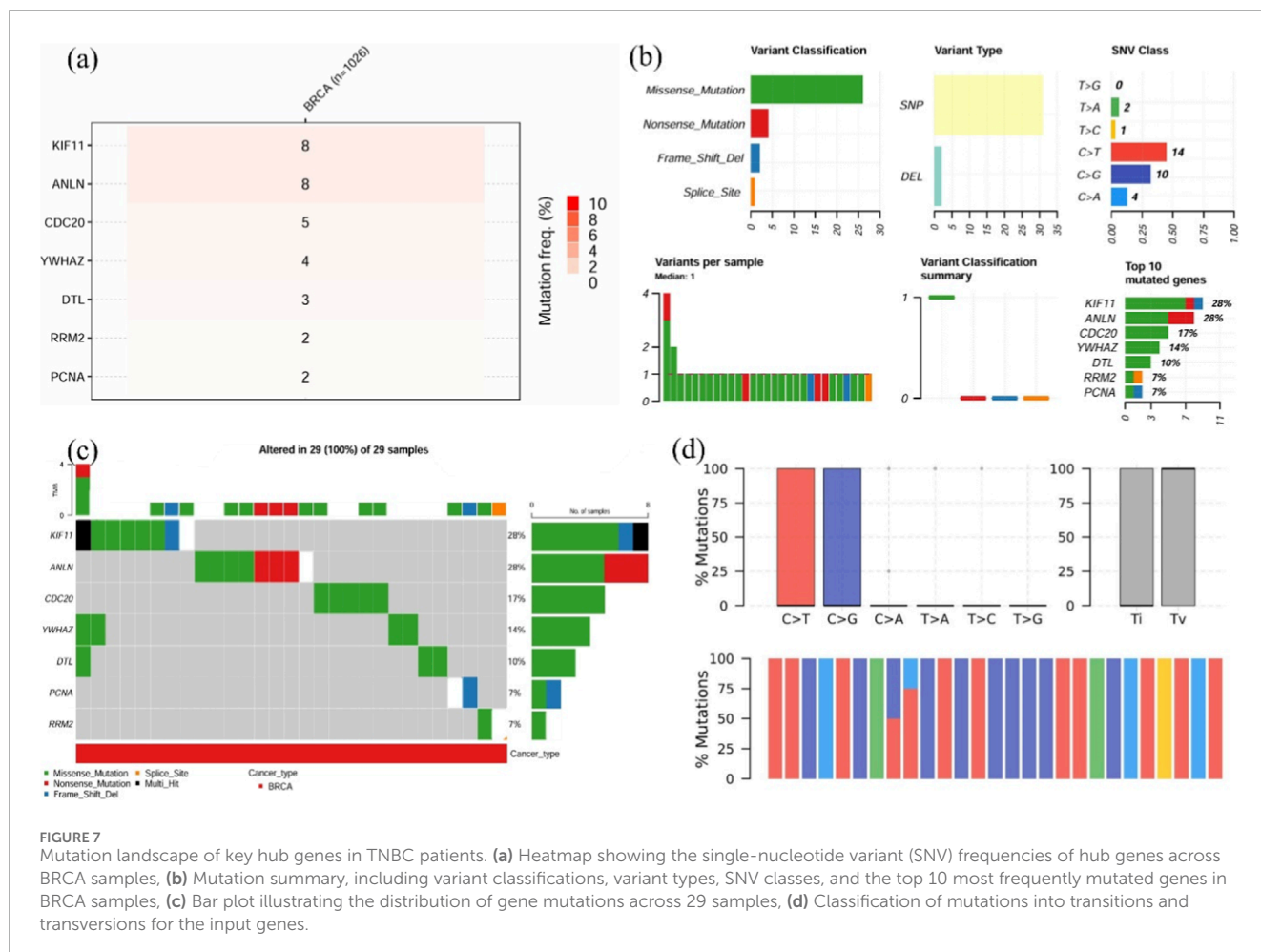


FIGURE 7 Mutation landscape of key hub genes in TNBC patients. **(a)** Heatmap showing the single-nucleotide variant (SNV) frequencies of hub genes across BRCA samples, **(b)** Mutation summary, including variant classifications, variant types, SNV classes, and the top 10 most frequently mutated genes in BRCA samples, **(c)** Bar plot illustrating the distribution of gene mutations across 29 samples, **(d)** Classification of mutations into transitions and transversions for the input genes.

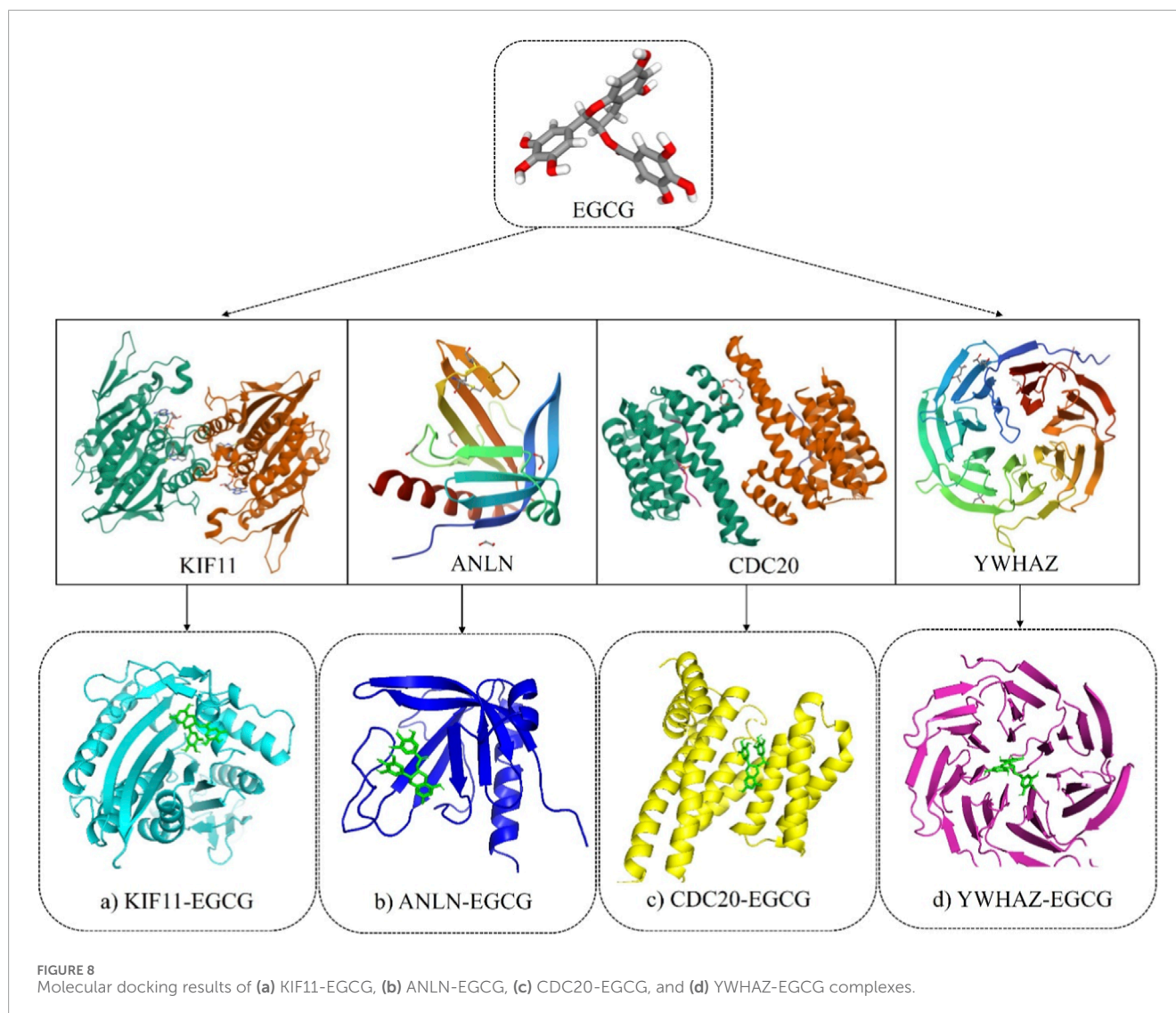
from the six topological parameters. Furthermore, the MCODE plugin detected significant gene clusters within the PPI network. The clusters identified through MCODE were cross-referenced with the results from cytoHubba. Notably, *RRM2*, *KIF11*, *ANLN*, *CDC20*, *CCNA1*, *AGO2*, *YWHAZ*, *DTL*, *SKP2*, and *PCNA* consistently appeared across multiple cytoHubba rankings (Table 1) and were predominantly located within the MCODE clusters with a score >3 (Figure 4).

3.3 Functional annotation through GO and KEGG

GO enrichment analysis was conducted using ShinyGO v0.80 to gain deeper insights into the biological significance of the ten key target genes regulated by miR-30a-5p in TNBC (Table 2). The GO BP analysis revealed significant enrichment of 437 pathways, with top-ranking terms including regulation of synapse maturation, mitotic cell cycle phase transition, regulation, processes, cell cycle, and its regulation. In the CC category, 67 enriched pathways were identified, with the hub genes predominantly localized to functionally critical structures and protein complexes. These included numerous complexes such as cyclin-dependent protein kinase (CDK) holoenzyme, cullin-RING ubiquitin ligase,

transferase, intracellular protein-containing complex, microtubule cytoskeleton, nucleoplasm, and nuclear lumen. Within the MF, 80 significantly enriched terms were observed. The most prominent among these were purine-specific mismatch base pair DNA N-glycosylase activity, ribonucleotide-diphosphate reductase activity thioredoxin disulfide as acceptor, DNA polymerase processivity factor activity, ribonucleotide-diphosphate reductase activity, and enzyme activity associated with nucleotide metabolism and DNA replication. Primarily, the GO pathway enrichment analysis indicated that the genes are linked with cell cycle regulation, emphasizing the importance of hub genes in cell cycle regulation.

From the KEGG analysis, out of 31 distinct pathways, the most significantly enriched ten pathways are detailed in Table 3. Hub genes are prominently involved in the cell cycle pathway, followed by viral carcinogenesis, mismatch repair, hepatitis B, and oocyte meiosis pathways. The dominant involvement of these genes in cell cycle-related pathways strongly suggests that miR-30a-5p may exert its therapeutic potential in TNBC by modulating core signaling cascades. The fold enrichment and total number of genes were reported in Tables 2, 3 for Gene Ontology and KEGG pathways, respectively.



3.4 Hub genes expression validation in TNBC

The expression analysis of the ten hub genes was analyzed using GEPIA2 to assess their expression levels between normal and tumor samples. Expression levels were represented as $-\log_2(\text{TPM}+1)$ to normalize the data. Among these, seven genes, namely, *RRM2*, *KIF11*, *ANLN*, *CDC20*, *YWHAZ*, *DTL*, and *PCNA*, exhibited significantly elevated expression in tumor samples ($p \leq 0.05$). Whereas *CCNA1*, *AGO2*, and *SKP2* did not show statistically significant differences, indicating relatively stable expression between normal and tumor tissues (Figure 5).

3.5 Overall survival analysis

The prognostic significance of seven key upregulated hub genes, namely, *RRM2*, *KIF11*, *ANLN*, *CDC20*, *YWHAZ*, *DTL*, and *PCNA*, was evaluated for overall survival in TNBC patients using

Kaplan-Meier survival analysis. From the above genes, *CDC20* exhibited the strongest prognostic effect with a hazard ratio of 1.71 (confidence interval (CI): 1.41–2.07; $p = 2.7\text{e-}08$). This was followed by *RRM2* (HR = 1.57, CI: 1.3–1.89, $p = 3\text{e-}06$), *YWHAZ* (HR = 1.53, $p = 9.6\text{e-}06$), *ANLN* (HR = 1.53, $p = 0.0018$), and *DTL* (HR = 1.52, $p = 1.2\text{e-}05$). Additionally, *PCNA* (HR = 1.31, $p = 0.0046$) and *KIF11* (HR = 1.26, $p = 0.014$) also exhibited significant associations with reduced survival probability (Figure 6).

3.6 Mutation analysis of the hub genes

Mutation analysis revealed that *KIF11* and *ANLN* had increased mutations (20% each), followed by *CDC20* (17%) and *YWHAZ* (14%). Missense mutation was found to be the abundant variant classification, with SNP and C>T as the variant type and SNV category, respectively. *CDC20* and *YWHAZ* carried missense mutations, whereas *ANLN* contained both missense and nonsense mutations. In contrast, frame-shift mutations

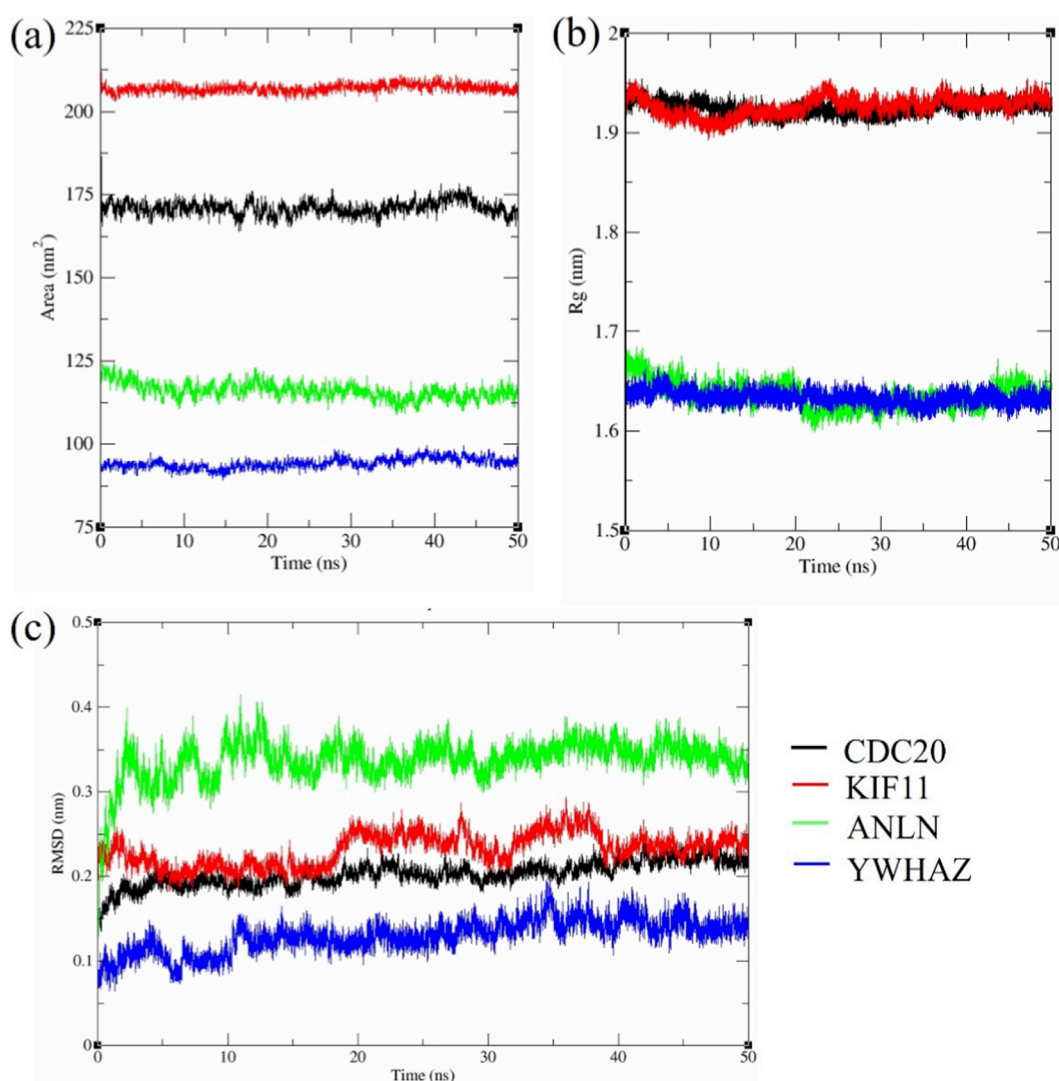


FIGURE 9
MD simulation analysis of KIF11, ANLN, CDC20, and YWHAZ with EGCG (a) RMSD, (b) SASA, (c) Radius of gyration.

were observed in *KIF11*, alongside missense and nonsense mutations (Figures 7a,b). Investigation of 29 BRCA samples showed that 100% of the samples displayed changes in at least one prognostic gene, with 28% of the cases in *KIF11* (Figure 7c). Moreover, comparable transitions and transversions with C>T and C>G being the most occurring conversions (Figure 7d). The increased burden of mutation diversity among the four prognostic genes suggests their role in disease progression and therapeutic targeting.

3.7 Molecular docking analysis

EGCG exhibited the strongest affinity for YWHAZ (−10.1 kcal/mol), forming eight hydrogen (H_2) bonds in the positions of Asp 191, Trp 234, Ser 275, Trp 276, Arg 316, Trp 363, Ser 404, and Ser 448. The next most favorable interaction was with ANLN (−8.9 kcal/mol), forming seven H_2 bonds, followed by

KIF11 (−8.5 kcal/mol) stabilized by eleven polar contacts. CDC20 showed the weakest binding of −7.0 kcal/mol with only four H_2 bonds. Overall, the binding hierarchy was YWHAZ > ANLN > KIF11 > CDC20, suggesting that YWHAZ have the most stable and complementary binding pocket for EGCG (Figure 8). Table 4 gives a summary of docking results with the number of hydrogen bonds.

3.8 Molecular dynamics simulation

The RMSD parameter (Figure 9a) examines the structural stability across all four protein-EGCG docked complexes. Notably, YWHAZ displayed good structural stability with slight deviation (-0.12 ± 0.02 nm), suggesting that the native scaffold was preserved by EGCG binding without persuading extensive conformational alterations. ANLN, in contrast, showed the highest fluctuations (-0.34 ± 0.03 nm), while CDC20 (-0.20 ± 0.03 nm) and KIF11

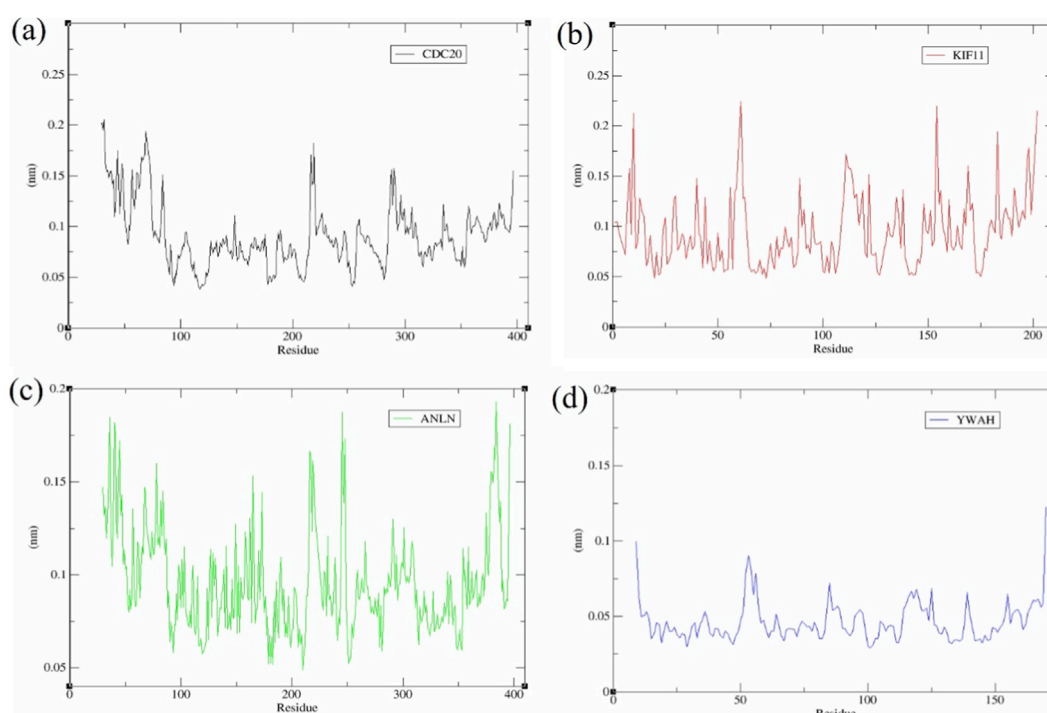


FIGURE 10
RMSF analysis of (a) CDC20, (b) KIF11, (c) ANLN, and (d) YWHAZ with EGCG.

($\sim 0.24 \pm 0.03$ nm) occupied intermediate positions. From the RMSD profile, it was observed that YWHAZ has a structurally rigid environment for EGCG binding, whereas CDC20 and KIF11 hold on to substantial backbone flexibility, leading to transitory shifts.

Consistent with the RMSD hierarchy, the solvent-accessible surface area (SASA) values (Figure 9b) revealed YWHAZ as the most compact complex ($\sim 95 \pm 4$ nm²), followed by ANLN ($\sim 120 \pm 5$ nm²), while CDC20 ($\sim 170 \pm 5$ nm²) and KIF11 ($\sim 205 \pm 4$ nm²) displayed substantially larger solvent exposure. The lower the SASA, the greater the compactness. YWHAZ shows a conformationally restrained state having low SASA, aligning with RMSD data. But CDC20 and KIF11 display greater flexibility and more solvent-exposed binding surfaces due to their increased SASA.

The Radius of Gyration (Rg) remained stable after rapid equilibrium within the first few ns (Figure 9c). Time-averaged Rg values strengthened the compactness hierarchy: YWHAZ (1.64 ± 0.02 nm) and ANLN (1.65 ± 0.02 nm) maintained tightly packed structures, whereas CDC20 (1.91 ± 0.02 nm) and KIF11 (1.93 ± 0.02 nm) were more extended. Low RMSD, SASA, and rigid Rg for YWHAZ highlight the increased stability, while it was vice-versa for CDC20 and KIF11 with their dynamic scaffolds.

RMSF demonstrated the structural grading of YWHAZ with the slightest mobility (baseline ~ 0.03 – 0.06 nm, peaks ~ 0.10 – 0.12 nm) (Figure 10d), reflecting a uniformly firm scaffold. ANLN showed intermediate flexibility (~ 0.12 – 0.18 nm at termini and loop regions) (Figure 10c), while CDC20 and KIF11 displayed distinct peaks (~ 0.20 – 0.27 nm) (Figures 10a,b), consistent with higher global flexibility.

4 Discussion

In recent years, TNBC's pathological burden has accelerated significantly, defined by its aggressive phenotype, high histological grade, and rapid cellular proliferation. Unlike other subtypes of BC, TNBC lacks well-defined molecular targets, resulting in confined treatment options and contributing to poor clinical prognosis (Geyer et al., 2017). The “triple-negative” status of TNBC was defined through immunohistochemical (IHC) profiling reveals the absence of HER2 expression, with less than 1% ER and PR expression (Zagami and Carey, 2022). Researchers are nowadays gaining attention towards phytochemicals and microRNAs in cancer studies. A recent study by Yang et al. (2025) observed that CYCS and MYL12B as the potential targets of EGCG in nasopharyngeal carcinoma through IHC and docking approaches. Bioinformatics analysis demonstrated that EGCG affects NFκB1, Bcl-2, HIF-1α, MMP, etc., thereby reducing the risk of ovarian cancer (Xinqiang et al., 2017). MicroRNAs are widely known for post-transcriptional regulation, where miR-30a impacts prognosis associated with survival in ovarian cancer through hub gene regulation (Lu et al., 2020). Similarly, in osteosarcoma, miR-30a-3p suppresses tumor growth via PTEN upregulation (Zhong et al., 2017). All these analyses illustrate the power of integrative bioinformatics for focusing mechanistic studies.

Based on the above experimental evidence and challenges in TNBC treatment, there is a growing interest in combinatorial therapeutic approaches aimed at overcoming tumor heterogeneity and drug resistance. Our study combines EGCG and miR-30a-5p as a synergistic effect in TNBC treatment. EGCG, a bioactive

polyphenol known for the suppression of cell cycle progression, apoptotic induction, and modulation of oncogenic pathways (Romano and Martel, 2021). Notably, miR-30a-5p enhanced the anticancer efficacy of small molecule drugs in BC by regulating gene expression involved in proliferation and survival (Jiang et al., 2018). Put together, EGCG and miR-30a provide a synergistic and potentially effective therapy for targeting several oncogenic pathways associated with TNBC.

In this study, transcriptomic profiling from the TCGA database yielded 5,008 upregulated and 2,905 downregulated TNBC genes. A total of 2,115 unique target genes of miR-30a-5p were identified through well-known miR-target prediction databases. Comparative intersection analysis uncovered 219 overlapping genes between miR-30a-5p targets and upregulated TNBC genes, and 174 overlapping genes with the downregulated set, yielding a refined pool of 393 common genes selected for further investigation. These genes were subjected to PPI network construction, resulting in a highly interconnected network of 392 nodes and 622 edges, visualized using Cytoscape. Network topology was examined using cytoHubba and MCODE plugins to identify key regulatory hubs. Intriguingly, *RRM2*, *KIF11*, *ANLN*, *CDC20*, *CCNA1*, *AGO2*, *YWHAZ*, *DTL*, *SKP2*, and *PCNA* emerged as top-ranked hub genes across multiple cytoHubba algorithms and were predominantly clustered within MCODE modules with scores exceeding 3. These findings underscore their potential as central molecular effectors in miR-30a-5p-mediated regulation and TNBC progression, providing valuable therapeutic targets.

Functional enrichment through GO and KEGG analyses demonstrated that the hub genes are involved in cell cycle processes, including their progression, regulatory mechanisms, the mitotic checkpoint complex, and the cyclin A2-CDK2 complex. These results collectively highlight the biological significance of the predicted target genes in maintaining cell cycle fidelity. Supporting this, Xiong et al. (2019) reported that upregulation of miR-30a-5p downregulated cyclin B1, cyclin D1, and c-myc, responsible for the cell cycle. Such evidence strongly suggests that miR-30a-5p can exert tumor-suppressive effects in TNBC by targeting essential regulators of the cell cycle pathway.

Expression analysis revealed that 7 out of 10 genes- *RRM2*, *KIF11*, *ANLN*, *CDC20*, *YWHAZ*, *DTL*, and *PCNA* were significantly upregulated in TNBC patients ($p < 0.05$). These genes are intricately involved in fundamental cellular pathways, including DNA replication, cell cycle regulation, and mitotic progression, thereby emphasizing their potential roles in sustaining the proliferative and aggressive phenotype of TNBC. Specifically, *RRM2*, *PCNA*, and *DTL* are associated with DNA synthesis and repair (Wilson et al., 2021; Zhang et al., 2021; Tian et al., 2024), while *KIF11* and *CDC20* regulate mitotic spindle formation and cell cycle transitions (Wang et al., 2025; Sitrý-Shevah et al., 2024). *ANLN* is essential for cytokinesis (Maryam and Chin, 2021), and *YWHAZ* is implicated in modulating signal transduction pathways linked to cell survival (Pedroza et al., 2021). Further, in TNBC patients, the overall survival analysis revealed that elevated expression of the seven genes was linked with poorer outcomes ($p < 0.05$, HR > 1). Therefore, targeted regulation of these overexpressed oncogenic genes may effectively modulate critical tumorigenic pathways and improve patient prognosis. Mutation frequency analysis indicated that among the 7 identified hub genes, *KIF11*, *DTL*, *CDC20*, and

YWHAZ exhibited the highest mutation rates, predominantly in the form of SNPs as the variant type. These genes emerged as the principal key hub genes within the study, stating that the gene modulation via miR-30a-5p may contribute to the downregulation of the key genes.

KIF11 (kinesin family member 11 or kinesin-5 (E.g.,5)) is a vital ATP-driven motor protein that promotes bipolar spindle formation, which drives anti-parallel microtubules separation during cell division (Valensin et al., 2009). Overexpression of *KIF11* has been linked with poor clinical outcomes in numerous malignancies (Liu et al., 2016). In docetaxel-resistant TNBC cells, *KIF11* knockdown inhibited the proliferation of CD44+/CD24-cancer stem-like cells and removed their self-renewal capacity (Jiang et al., 2017). Similarly, *KIF11* inhibition has been shown to reduce breast cancer growth, migration, and invasion (Zhou et al., 2019).

ANLN (Anilin), an important cytokinesis regulator, has a dynamic subcellular localization. During interphase, it resides within the nucleus and relocates during telophase to the cytoplasm. This helps generate the formation of the contractile ring and development of the cleavage furrow through interactions with myosin, F-actin, RhoA, and septin (Piekny and Maddox, 2010). *ANLN* regulates proliferation through cell cycle modulation. Silencing *ANLN* in non-small cell lung cancer (NSCLC) and breast cancer cell lines resulted in polynucleated cells and markedly inhibited cell proliferation (Zhou et al., 2015).

CDC20 (Cell Division Cycle 20), a crucial activator of the anaphase-promoting complex/cyclosome (APC/C), is necessary for appropriate mitotic progression. It is a potential therapeutic target, since dysregulation of its activity can result in mitotic abnormalities and tumorigenesis. Therefore, several *CDC20* inhibitors are currently under clinical investigation (Wang et al., 2015). While elevated *CDC20* expression has been found in BC (Karra et al., 2014), its function in TNBC is yet unknown. Song et al. (2021) observed that reduced *CDC20* expression showed a direct impact on cell migration and invasiveness in MDA-MB-231 cells.

YWHAZ (Tyrosine 3 monooxygenase/tryptophan 5-monooxygenase activation protein zeta or 14-3-3ζ), an important factor in signal transduction and tumor progression. Growing evidence suggests that *YWHAZ* is often overexpressed across numerous cancer types and acts as an oncogene by promoting cellular processes such as proliferation, migration, and invasion through interactions with ErbB2 and p85. Upstream microRNAs, such as miR-193b, miR-451, and miR-30c, control its expression (Gan et al., 2020; Wang et al., 2017). The importance of miRNA-mediated regulation was emphasized by a recent study that found that miR-136-5p elevation inhibited BC by directly targeting *YWHAZ* (Kong et al., 2023).

Docking results demonstrated that EGCG had a substantial binding affinity with *YWHAZ* (−10.1 kcal/mol), a validated target of miR-30a-5p, suggesting a potential synergistic interaction that could improve TNBC therapeutic efficacy. These results indicate the binding predictions, which require further experimental validation for functional confirmation. Previous literature stated that EGCG-protein binding energies typically range from −6 to −11.42 kcal/mol (Wang et al., 2019), consistent with our findings. For instance, Fanai et al. (2025) demonstrated the comparable binding affinity of EGCG with PPARGC1A (−9.55 kcal/mol), FOXO

(−10.21 kcal/mol), and SIRT3 (−11.42 kcal/mol). Another study by Kamaraj (2025) showed the binding affinities of EGCG with ESR1 (−8.2 kcal/mol), MMP2 (−7.4 kcal/mol), MMP9 (−8.3 kcal/mol), MMP13 (−8.8 kcal/mol), and STAT1 (−7.3 kcal/mol). All the above evidence supports the biological plausibility of our modelled interactions. In line with the docking results, MD modelling showed that the YWHAZ-EGCG complex takes the most compact and stable structure. In contrast, ANLN displayed intermediate stabilization, and CDC20/KIF11 maintained dynamic solvent-exposed conformations (Paranthaman et al., 2025).

5 Conclusion

The present study integrates miRNA-based regulation with phytochemical intervention to target the risk caused by TNBC. By focusing on hsa-miR-30a-5p and EGCG, the study identified gene targets implicated in cell cycle and mitotic pathways, revealing multi-target therapeutic strategies. These findings provide support for co-delivery strategies using nanocarrier-based formulations, aiming to harness the synergistic effects of miRNAs and natural bioactive compounds that pave the way for newer treatment approaches for TNBC.

Data availability statement

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Author contributions

LD: Formal Analysis, Data curation, Methodology, Writing – original draft, Software. AM: Writing – review and editing, Validation, Supervision, Conceptualization.

References

- Aggarwal, V., Tuli, H. S., Tania, M., Srivastava, S., Ritzer, E. E., Pandey, A., et al. (2022). Molecular mechanisms of action of epigallocatechin gallate in cancer: recent trends and advancement. *Sem. Cancer Biol.* 80, 256–275. doi:10.1016/j.semcancer.2020.05.011
- Bonotto, M., Gerrattana, L., Poletto, E., Driol, P., Giangreco, M., Russo, S., et al. (2014). Measures of outcome in metastatic breast cancer: insights from a real-world scenario. *Oncologist* 19 (6), 608–615. doi:10.1634/theoncologist.2014-0002
- Chacko, S. M., Thambi, P. T., Kuttan, R., and Nishigaki, I. (2010). Beneficial effects of green tea: a literature review. *Chin. Med.* 5, 13. doi:10.1186/1749-8546-5-13
- Chaudhary, L. N., Wilkinson, K. H., and Kong, A. (2018). Triple-negative breast cancer: who should receive neoadjuvant chemotherapy? *Surg. Oncol. Clin. N. Am.* 27 (1), 141–153. doi:10.1016/j.soc.2017.08.004
- Chen, Y., and Wang, X. (2020). miRDB: an online database for prediction of functional microRNA targets. *Nucleic Acids Res.* 48 (D1), D127–D131. doi:10.1093/nar/gkz757
- Chin, C.-H., Chen, S.-H., Wu, H.-H., Ho, C.-W., Ko, M.-T., and Lin, C.-Y. (2014). cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* 8 (Suppl. 4), S11. doi:10.1186/1752-0509-8-S4-S11
- Cui, S., Yu, S., Huang, H.-Y., Lin, Y.-C.-D., Zhang, B., Xiao, J., et al. (2024). miRTarBase 2025: updates to the collection of experimentally validated microRNA-target interactions. *Nucleic Acid. Res.* 53 (D1), D147–D156. doi:10.1093/nar/gkac1072
- Elango, R., Rashid, S., Vishnubalaji, R., Al-Sarraf, R., Akhtar, M., Ouararhni, K., et al. (2023). Transcriptome profiling and network enrichment analyses identify subtype-specific therapeutic gene targets for breast cancer and their microRNA regulatory networks. *Cell Death Dis.* 14 (415), 415. doi:10.1038/s41419-023-05908-8
- Fanai, H. L., Chand, J., Ahmad, S. F., Attia, S. M., and Emran, T. B. (2025). System biology-based assessment of the molecular mechanism of epigallocatechin gallate in Parkinson's disease: via network pharmacology, *in-silico* evaluation & *in-vitro* studies. *Sci. Rep.* 15, 25678. doi:10.1038/s41598-025-11592-9
- Gan, Y., Ye, F., and He, X. (2020). The role of YWHAZ in cancer: a maze of opportunities and challenges. *J. Cancer* 11 (8), 2252–2264. doi:10.7150/jca.41316
- Ge, S. X., Jung, D., and Yao, R. (2019). ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* 36 (8), 2628–2629. doi:10.1093/bioinformatics/btz931
- Geyer, F. C., Pareja, F., Weigelt, B., Rakha, E., Ellis, I. O., Schnitt, S. J., et al. (2017). The spectrum of triple-negative breast disease: High- and low-grade lesions. *A. J. Pathol.* 187 (10), 2139–2151. doi:10.1016/j.ajpath.2017.03.016

Funding

The authors declare that no financial support was received for the research and/or publication of this article.

Acknowledgements

The authors thank the Vellore Institute of Technology administration, Vellore, for providing resources to complete the work.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Ghahremanian, S., Rashidi, M. M., Raeisi, K., and Toghraie, D. (2022). Molecular dynamics simulation approach for discovering potential inhibitors against SARS-CoV-2: a structural review. *J. Mo. Lipids*. 354, 118901. doi:10.1016/j.molliq.2022.118901
- Jiang, M., Zhuang, H., Xia, R., Gan, L., Wu, Y., Ma, J., et al. (2017). KIF11 is required for proliferation and self-renewal of docetaxel resistant triple negative breast cancer cells. *Oncotarget* 8 (54), 92106–92118. doi:10.18632/oncotarget.20785
- Jiang, L., Zhang, H., and Tang, J. (2018). MiR-30a: a novel biomarker and potential therapeutic target for cancer. *J. Oncol.* 2018, 5167829. doi:10.1155/2018/5167829
- Kaddoura, R., Alqutami, F., Asbaita, M., and Hachim, M. (2023). *In silico* analysis of publicly available transcriptomic data for the identification of triple-negative breast cancer-specific biomarkers. *Life* 13 (2), 422. doi:10.3390/life13020422
- Kamaraj, B. (2025). Exploring the therapeutic potential of epigallocatechin-3-gallate (Green tea) in periodontitis using network pharmacology and molecular modelling approach. *Int. J. Mol. Sci.* 26 (18), 9144–. doi:10.3390/ijms26189144
- Karra, H., Repo, H., Ahonen, I., Lyytyniemi, E., Pitkänen, R., Lintunen, M., et al. (2014). Cdc20 and securin overexpression predict short-term breast cancer survival. *Brit. J. Cancer*. 110 (12), 2905–2913. doi:10.1038/bjc.2014.252
- Kim, T., and Croce, C. M. (2023). MicroRNA: trends in clinical trials of cancer diagnosis and therapy strategies. *Exp. Mol. Med.* 55, 1314–1321. doi:10.1038/s12276-023-01050-9
- Kong, Z., Han, Q., Zhu, B., Wan, L., and Feng, E. (2023). Circ_0069094 regulates malignant phenotype and paclitaxel resistance in breast cancer cells via targeting the miR-136-5p/YWHAZ axis. *Thorac. Cancer* 14 (19), 1831–1842. doi:10.1111/1759-7714.14928
- Krishnamoorthy, H. R., and Karuppusamy, R. (2024). Deciphering the prognostic landscape of triple-negative breast cancer: a focus on immune-related hub genes and therapeutic implications. *Biotechnol. Appl. Biochem.* 72, 825–845. doi:10.1002/bab.2700
- Li, Y., Zhang, H., Merker, Y., Chen, L., Liu, N., Leonov, S., et al. (2022). Recent advances in therapeutic strategies for triple-negative breast cancer. *J. Hematol. Oncol.* 15, 121. doi:10.1186/s13045-023-01497-3
- Liu, L., Liu, X., Mare, M., Dumont, A. S., Zhang, H., Yan, D., et al. (2016). Overexpression of Eg5 correlates with high grade astrocytic neoplasm. *J. Neurooncol.* 126, 77–80. doi:10.1007/s11060-015-1954-3
- Lu, W., Wu, Y., Lu, X., Zhu, T., Ren, Z. L., and Yu, Z. (2020). Bioinformatics analysis of prognostic value and prospective pathway signal of miR-30a in ovarian cancer. *J. Ovarian Res.* 13, 120. doi:10.1186/s13048-020-00722-8
- Marin, V., Burgos, V., Perez, R., Maria, D. A., Pardi, P., and Paz, C. (2023). The potential role of epigallocatechin-3-gallate (EGCG) in breast cancer treatment. *Int. J. Mol. Sci.* 24 (13), 10737. doi:10.3390/ijms241310737
- Maryam, A., and Chin, Y. R. (2021). ANLN enhances triple-negative breast cancer stemness through TWIST1 and BMP2 and promotes its spheroid growth. *Front. Mol. Biosci.* 8, 700973. doi:10.3389/fmolb.2021.700973
- McGeary, S. E., Lin, K. S., Shi, C. Y., Pham, T. M., Bisaria, N., Kelley, G. M., et al. (2019). The biochemical basis of microRNA targeting efficacy. *Science* 366 (6472), eaav1741. doi:10.1126/science.aav1741
- Morris, G. J., Naidu, S., Topham, A. K., Guiles, F., Xu, Y., McCue, P., et al. (2007). Differences in breast carcinoma characteristics in newly diagnosed African-American and Caucasian patients: a single-institution compilation compared with the national cancer Institute's surveillance, epidemiology, and end results database. *Cancer* 110 (4), 876–884. doi:10.1002/cncr.22836
- Oliveros, J. C. (2007). An interactive tool for comparing lists with Venn's diagrams. Available online at: <https://bioinfogp.cnb.csic.es/tools/venny/index.html>.
- Paranthaman, P., Krishnamoorthy, H. R., Veluchamy, S., Karuppusamy, R., and Veerappapillai, S. (2025). Unraveling the complexities in TNBC gene interaction networks towards better therapeutics. *Biotechnol. Appl. Biochem.* 2025. doi:10.1002/bab.70039
- Pedroza, D. A., Ramirez, M., Rajamanickam, V., Subramani, R., Margolis, V., Gurbuz, T., et al. (2021). miRNome and functional network analysis of PGRMC1 regulated miRNA target genes identify pathways and biological functions associated with triple negative breast cancer. *Front. Oncol.* 11 (2021), 710337. doi:10.3389/fonc.2021.710337
- Piekny, A. J., and Maddox, A. (2010). The myriad roles of anillin during cytokinesis. *Sem. Cell Dev. Biol.* 21 (9), 881–891. doi:10.1016/j.semcdb.2010.08.002
- Pradhan, R., Dey, A., Taliyan, R., Puri, A., Kharavtekar, S., and Dubey, S. K. (2023). Recent advances in targeted nanocarriers for the management of triple negative breast cancer. *Pharmaceutics* 15 (1), 246. doi:10.3390/pharmaceutics15010246
- Raudvere, U., Kolsberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., et al. (2019). g:Profiler: a web server for functional enrichment analysis and conversion of gene lists (2019 update). *Nucleic Acids Res.* 47 (W1), W191–W198. doi:10.1093/nar/gkz369
- Romano, A., and Martel, F. (2021). The role of EGCG in breast cancer prevention and therapy. *Mini Rev. Med. Chem.* 21 (7), 883–898. doi:10.2174/138955752099920211194445
- Santana, T. A. B. d.S., de Oliveira Passamai, L., de Miranda, F. S., Borin, T. F., Borges, G. F., Luiz, W. B., et al. (2023). The role of miRNAs in the prognosis of triple-negative breast cancer: a systematic review and meta-analysis. *Diagnostics* 13 (1), 127. doi:10.3390/diagnostics13010127
- Sklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., et al. (2023). The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* 51 (D1), D638–D646. doi:10.1093/nar/gkac1000
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13 (11), 2498–2504. doi:10.1101/gr.1239303
- Sitry-Shevah, D., Miniowitz-Shemtov, S., Dan, T. L., and Hershko, A. (2024). The mitotic checkpoint complex controls the association of Cdc20 regulatory protein with the ubiquitin ligase APC/C in mitosis. *Proc. Natl. Acad. Sci. U.S.A.* 121 (37), e2413089121. doi:10.1073/pnas.2413089121
- Song, C., Lowe, V. J., and Lee, S. (2021). Inhibition of Cdc20 suppresses the metastasis in triple negative breast cancer (TNBC). *Breast Cancer* 28 (5), 1073–1086. doi:10.1007/s12282-021-01242-z
- Supplitt, S., Karpinski, P., Sasiadek, M., and Laczmanska, I. (2021). Current achievements and applications of transcriptomics in personalized cancer medicine. *Int. J. Mol. Sci.* 22 (3), 1422. doi:10.3390/ijms22031422
- Tang, Z., Kang, B., Li, C., Chen, T., and Zhang, Z. (2019). GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acid. Res.* 47 (W1), W556–W560. doi:10.1093/nar/gkz430
- Tian, J., Wen, M., Goa, P., Feng, M., and Wei, G. (2024). RUVBL1 ubiquitination by DTL promotes RUVBL1/2-β-catenin-mediated transcriptional regulation of NHEJ pathway and enhances radiation resistance in breast cancer. *Cell Death Dis.* 15 (4), 259. doi:10.1038/s41419-024-06651-4
- Valensin, S., Ghiron, C., Lamanna, C., Kremer, A., Rossi, M., Ferruzzi, P., et al. (2009). KIF11 inhibition for glioblastoma treatment: reason to hope or a struggle with the brain? *BMC Cancer* 9, 196. doi:10.1186/1471-2407-9-196
- Van Der Spoel, D., Lindhal, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. C. (2005). GROMACS: fast, flexible, and free. *J. Comput. Chem.* 26, 1701–1718. doi:10.1002/jcc.20291
- Von Mering, C., Jensen, L. J., Snel, B., Hooper, S. D., Krupp, M., Foglierini, M., et al. (2004). STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acid. Res.* 33 (Database Issue), D433–D437. doi:10.1093/nar/gki005
- Wang, L., Zhang, J., Wan, L., Zhou, X., Wang, Z., and Wei, W. (2015). Targeting Cdc20 as a novel cancer therapeutic strategy. *Pharmacol. Ther.* 151, 141–151. doi:10.1016/j.pharmthera.2015.04.002
- Wang, W., Zhang, L., Wang, Y., Ding, Y., Chen, T., Wang, Y., et al. (2017). Involvement of miR-451 in resistance to paclitaxel by regulating YWHAZ in breast cancer. *Cell Death Dis.* 8, e3071. doi:10.1038/cddis.2017.460
- Wang, W., Xiong, X., Li, X., Zhang, Q., Yang, W., and Du, L. (2019). *In silico* investigation of the anti-tumor mechanisms of epigallocatechin-3-gallate. *Molecules* 24 (7), 1445. doi:10.3390/molecules24071445
- Wang, W., Zhang, Y., Huang, X., Li, D., Lin, Q., Zhuang, H., et al. (2024). The role of the miR-30a-5p/BCL2L1 pathway in rosmarinic acid-induced apoptosis in MDA-MB-231-derived breast cancer stem-like cells. *Front. Pharmacol.* 15, 1445034. doi:10.3389/fphar.2024.1445034
- Wang, B., Bao, L., Li, X., Sun, G., Yang, W., Xie, N., et al. (2025). Identification and validation of the important role of KIF11 in the development and progression of endometrial cancer. *J. Transl. Med.* 23 (48), 48. doi:10.1186/s12967-025-06081-6
- Wilson, E. A., Sultana, N., Shah, K. N., Elford, H. L., and Faridi, J. S. (2021). Molecular targeting of RRM2, NF-κB, and mutant TP53 for the treatment of triple-negative breast cancer. *Small Mol. Ther.* 20 (4), 655–664. doi:10.1158/1535-7163.mct-20-0373
- Wolff, A. C., Hammond, M. E. H., Hicks, D. G., Dowsett, M., McShane, L. M., Allison, K. H., et al. (2013). Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American society of clinical oncology/college of American pathologists clinical practice guideline update. *J. Clin. Oncol.* 31 (31), 3997–4013. doi:10.1200/jco.2013.50.9984
- Xinqiang, S., Mu, Z., Lei, C., and Mun, L. Y. (2017). Bioinformatics analysis on molecular mechanism of green tea compound epigallocatechin-3-gallate against ovarian cancer. *Clin. Transl. Sci.* 10 (4), 302–307. doi:10.1111/cts.12470
- Xiong, J., Wei, B., Ye, Q., and Liu, W. (2019). MiR-30a-5p/UBE3C axis regulates breast cancer cell proliferation and migration. *Biochem. Biophys. Res. Commun.* 516 (3), 1013–1018. doi:10.1016/j.bbrc.2016.03.069
- Yang, Y., Luo, W., Feng, Z., Chen, X., Li, J., Zuo, L., et al. (2025). An integrative analysis combining bioinformatics, network pharmacology and experimental methods identified key genes of EGCG targets in Nasopharyngeal carcinoma. *Discov. Oncol.* 16, 742. doi:10.1007/s12672-025-02365-x
- Yin, L., Duan, J.-J., Bian, X.-W., and Yu, S.-C. (2020). Triple-negative breast cancer molecular subtyping and treatment progress. *Breast Cancer Res.* 22, 61. doi:10.1186/s13058-020-01296-5
- Zagami, P., and Carey, L. A. (2022). Triple negative breast cancer: pitfalls and progress. *Npj Breast Cancer* 8, 95. doi:10.1038/s41523-022-00468-0
- Zhang, S., Zhou, T., Wang, Z., Yi, F., Li, C., Guo, W., et al. (2021). Post-translational modifications of PCNA in control of DNA synthesis and DNA damage

tolerance-the implications in carcinogenesis. *Int. J. Biol. Sci.* 17 (14), 4047–4059. doi:10.7150/ijbs.64628

Zhao, B., Su, X., Hu, P., Ma, Y., Zhou, X., and Hu, L. (2022). A geometric deep learning framework for drug repositioning over heterogeneous information networks. *Brief. Bioinfo.* 23, bbac384. doi:10.1093/bib/bbac384

Zhao, B., Su, X., Yang, Y., Li, D., Li, G., Hu, P., et al. (2024). A heterogeneous information network learning model with neighborhood-level structural representation for predicting lncRNA-miRNA interactions. *Comput. Struct. Biotechnol. J.* 23, 2924–2933. doi:10.1016/j.csbj.2024.06.032

Zhong, B., Guo, S., Zhang, W., Zhang, C., Wang, Y., and Zhang, C. (2017). Bioinformatics prediction of miR-30a targets and its inhibition of cell proliferation of osteosarcoma by up-regulating the expression of PTEN. *BMC Med. Genomics.* 10, 64. doi:10.1186/s12920-017-0300-3

Zhou, W., Wang, Z., Shen, N., Pi, W., Jiang, W., Huang, J., et al. (2015). Knockdown of ANLN by lentivirus inhibits cell growth and migration in human breast cancer. *Mol. Cell Biochem.* 398 (1–2), 11–19. doi:10.1007/s11010-014-2200-6

Zhou, J., Chen, W., Yang, L., Wang, J., Sun, J. Y., Zhang, W. W., et al. (2019). KIF11 functions as an oncogene and is associated with poor outcomes from breast cancer. *Cancer Res. Treat.* 51 (3), 1207–1221. doi:10.4143/crt.2018.460