Article

Convergent transcriptomic signatures reveal cell cycle and DNA repair dependencies in triple-negative breast cancer: A network-based multi-dataset analysis

Chaitanya Kumar[1], Ravi Verma[2], Ashok Sharma[3], Vishnupriya Veeraraghavan[1]

[1]Center of Molecular Medicine and Diagnostics, Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha University, Chennai 600077, Tamil Nadu, India

[2]School of Biotechnology, Jawaharlal Nehru University, New Delhi 110067, India

[3]Department of Biochemistry, All India Institute of Medical Sciences, New Delhi 110029, India

E-mail: vishnupriya@saveetha.com

Received 11 November 2025; *Accepted 20 December 2025*; *Published online 30 December 2025*; *Published 1 June 2026*

Abstract

Triple-negative breast cancer (TNBC) remains one of the most aggressive and therapeutically challenging breast cancer subtypes due to the absence of hormone receptors and HER2 expression. To uncover reproducible molecular signatures and potential therapeutic targets, we performed an integrative transcriptomic analysis using three independent GEO datasets (GSE38959, GSE65216, and GSE65194). Differentially expressed genes were examined through protein–protein interaction network construction, enrichment analysis, and survival validation. A highly connected gene cluster comprising CHEK1, PLK1, AURKA, CCNA2, CCNB1, RAD51, TOP2A, KIF11, and KIF23 was consistently upregulated across datasets. These genes were predominantly involved in cell-cycle regulation and DNA-repair pathways and showed significant association with poor overall survival in TNBC cohorts. The findings suggest that dysregulation of the mitotic checkpoint and homologous recombination processes defines a conserved oncogenic program in TNBC. This study highlights druggable molecular targets that could guide the development of pathway-directed therapies in TNBC.

Keywords triple-negative breast cancer; transcriptomics; cell cycle; DNA repair; network analysis; biomarkers.

Network Biology
ISSN 2220-8879
URL: http://www.iaees.org/publications/journals/nb/online-version.asp
RSS: http://www.iaees.org/publications/journals/nb/rss.xml
E-mail: networkbiology@iaees.org
Editor-in-Chief: WenJun Zhang
Publisher: International Academy of Ecology and Environmental Sciences

1 Introduction

Triple-negative breast cancer (TNBC) is a type of breast cancer that is very aggressive and varies a lot from person to person. It does not have the estrogen receptor (ER), progesterone receptor (PR), or human epidermal

growth factor receptor 2 (HER2) [1]. This lack of receptors makes hormone treatments and HER2-targeted therapies less effective, so the main treatment options are usually standard chemotherapy. TNBC accounts for approximately 15–20% of all breast cancer cases worldwide and is disproportionately represented among younger women and certain racial and ethnic groups, including those of Asian and African descent [3,4]. In clinical terms, TNBC is linked to more severe tumors, faster growth rates, early return of cancer, and spread to other parts of the body, especially vital organs and the brain. These aggressive features contribute to a generally poor prognosis, with a median survival of less than 18 months in metastatic settings [6].

Although TNBC is defined phenotypically by the absence of hormonal receptors and HER2, it is not a homogeneous entity. Transcriptomic studies have divided TNBC into several molecular subtypes, such as basal-like, mesenchymal, and luminal androgen receptor subtypes, each showing different patterns of gene activity and possible treatment weaknesses. However, many of these subclassifications have limited clinical utility due to their inconsistent reproducibility across datasets and lack of actionable targets. Finding strong, conserved oncogenic pathways and regulatory genes that propel TNBC progression and treatment resistance is therefore still vital.

Over the past decade, high-throughput transcriptomic analyses have enabled deeper insights into the molecular architecture of TNBC. However, individual studies often face problems like biases from specific samples, differences in the technology used, and not enough statistical strength, which makes it hard to get consistent results. To address these issues, this study we conducted a transcriptomic analysis using three separate and carefully selected Gene Expression Omnibus (GEO) datasets: GSE38959, GSE65216, and GSE65194. By carefully adjusting these datasets to ensure they are comparable, we aimed to find a consistent group of differentially expressed genes (DEGs) that are important in the development of TNBC.

Combining several transcriptomic datasets not only strengthens our ability to find differentially expressed genes (DEGs) but also allows us to confirm important biological results across different groups. This method showed a common pattern of gene activity in TNBC, marked by the steady increase of genes that play roles in controlling the cell cycle, copying DNA, managing the mitotic spindle, and separating chromosomes. Importantly, key genes like CHEK1, PLK1, AURKA, CCNA2, CCNB1, RAD51, TOP2A, KIF11, and KIF23 were consistently found to be more active in all three datasets, indicating they play a major role in the rapid growth and genetic instability seen in TNBC. These genes play critical roles in cell cycle progression and are frequently dysregulated in highly proliferative tumors [8–10]. Their increased levels in TNBC indicate a heavy dependence on processes that help cells divide correctly and maintain their DNA, which could be weaknesses that can be targeted for treatment.

To contextualize these DEGs within functional networks, we first constructed a high-confidence protein-protein interaction (PPI) network using the STRING database. Next, we used MCODE clustering in Cytoscape to find closely linked groups of proteins that might work together, and then we analyzed the network's structure in Cytoscape. Centrality metrics—such as degree, betweenness, and closeness—identified a dense regulatory module of kinases and repair enzymes that orchestrate TNBC’s aggressive phenotype. The resulting network topology was not random; rather, it revealed an interconnected oncogenic circuit centered on G2/M transition regulators. This configuration reflects a systems-level adaptation to proliferative stress and DNA damage—a hallmark of TP53-deficient tumors, which comprise over 80% of TNBC cases [11].

Gene Ontology (GO) and KEGG pathway enrichment analyses were done to identify the biological functions of differentially expressed genes (DEGs) and their connections in the network. GO analysis of the biological process (BP) category showed marked enrichment in mitotic nuclear division, chromosome segregation, and spindle assembly checkpoint regulation. The molecular function (MF) category indicated enrichment in ATP binding, kinase activity, and microtubule motor function. The analysis of cellular

component (CC) analysis highlighted the localization of upregulated genes to spindle poles, centrosomes, and chromosomal regions further supporting their roles in mitotic fidelity [12–15]. KEGG pathway analysis showed a strong focus on pathways that control the cell cycle, the p53 signaling pathway, and homologous recombination. signaling, and homologous recombination pathways. These pathways are closely connected to problems with DNA copying and monitoring the genome, which are key issues in the development of TNBC. Two central themes in TNBC tumorigenesis are DNA replication stress and genome surveillance [16-21].

We looked at how important these key genes are for patient survival using KMPlotter, a strong tool that collects survival information from many patients. We utilized KMPlotter, a robust online tool that integrates survival data from extensive groups of breast cancer patients. The Kaplan–Meier curves for CHEK1, PLK1, AURKA, CCNA2, CCNB1, RAD51, TOP2A, KIF11, and KIF23 showed that higher amounts of these genes were connected to worse overall survival in patients with ER-negative and HER2-negative breast cancer. These findings reinforce the prognostic relevance of this core gene set and highlight their potential as therapeutic targets in triple-negative breast cancer. For example, high expression of **TOP2A** was linked to a hazard ratio exceeding 2.0, indicating a particularly strong prognostic impact. These findings underscore the translational relevance of our molecular targets and justify their further investigation as therapeutic candidates [22-27].

Recently, several of these genes have garnered interest in preclinical and clinical studies. Inhibitors targeting PLK1 (e.g., volasertib), AURKA (e.g., alisertib), and CHEK1 (e.g., prexasertib) are in various stages of clinical trials. Our findings provide a rational basis for prioritizing these targets specifically in TNBC. Since TNBC tumors grow quickly and have unstable genes, blocking both mitotic regulators and DNA repair enzymes at the same time might lead to synthetic lethality, which could improve treatment effectiveness and reduce unwanted side effects.

To visualize the integrative nature of our findings, we developed schematic representations summarizing the intersecting biological processes and functional modules. These visual models show how the genes that are turned up work together to drive important cancer-related processes like skipping the G2/M checkpoint, causing chromosome problems, and handling stress during DNA copying. By creating a detailed map of the main cancer-causing factors in TNBC, our study goes beyond looking at just one gene to provide a broader view for personalized medicine.

Additionally, this work introduces a new layer of depth by incorporating survival data alongside transcriptomic and network analysis. The convergence of multiple analytic modalities enhances the credibility of our findings and establishes a bridge between basic molecular research and clinical application. The addition of extra Kaplan–Meier plots and detailed survival tables for specific genes gives more detail and helps confirm our predictions about outcomes.

While our integrative approach yields strong candidate targets and pathways, we acknowledge certain limitations. Transcriptomic analyses, although powerful, do not capture post-translational modifications or protein activity. Future research should integrate phospho-proteomic and metabolomic data to capture downstream effects more comprehensively. We will also need to conduct tests in TNBC cell lines and live models to confirm the synthetic lethal relationships indicated by our network analysis.

In conclusion, this study presents a robust, reproducible, and clinically relevant transcriptomic framework for TNBC. By combining different methods like gene expression analysis, network analysis, and survival modeling, we find a key cancer-related program that works well together and is important for predicting outcomes. The genes and pathways mentioned here provide good targets for developing new drugs and help create treatment strategies based on specific markers in TNBC.

2 Materials and Methods

2.1 Dataset Selection and Acquisition

To find a clear gene expression pattern for triple-negative breast cancer (TNBC), we collected three different sets of gene expression data from the Gene Expression Omnibus (GEO): GSE38959, GSE65216, and GSE65194. Each dataset included TNBC samples and corresponding controls, profiled using Affymetrix Human Genome U133 Plus 2.0 arrays or equivalent platforms. Raw CEL files were downloaded and processed using the affy package in R, and platform annotation files were retrieved from the GEO and TCGA database. These datasets were selected based on their high-quality annotation, availability of normal and tumor samples, and relevance to hormone receptor-negative breast cancer [30,46].

2.2 Data Preprocessing and Normalization

All raw expression data were normalized using the Robust Multi-array Average (RMA) method within the Bioconductor framework in R. Background correction, quantile normalization, and log2 transformation were performed using the "affy" and "limma" packages [31,33]. Batch effect removal was applied using the ComBat function from the sva package to harmonize variations between datasets [32]. We summarized probe-level data to gene-level expression using Entrez Gene IDs after normalization. Probes not mapping to any known gene symbol were excluded.

2.3 Differential Gene Expression (DGE) Analysis

To find genes that are expressed differently, we used the limma package in R to compare TNBC and normal breast tissues for each dataset separately. DEGs were filtered using an adjusted p- value< 0.05 (Benjamini–Hochberg correction) and absolute log2 fold-change > 1 [33]. Common DEGs across datasets were visualized using Venn diagrams generated with the Venn Diagram package. Only genes that were consistently increased or decreased in all three datasets were chosen for further analysis, making sure the results are reliable and biologically meaningful.

2.4 Functional Enrichment Analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed using the clusterProfiler and enrichplot packages in R [34]. Enrichment was assessed for Biological Processes (BP), Cellular Components (CC), and Molecular Functions (MF) at a p-adjusted cutoff of <0.05. KEGG pathways were selected based on adjusted p-value and gene count thresholds. Visualizations, including dot plots and enrichment maps, were created using ggplot2 and enrichplot [43].

2.5 Protein–Protein Interaction (PPI) Network Construction

We built protein–protein interaction (PPI) networks using the STRING database (version 11.5) and set a minimum interaction confidence score of 0.7 to focus on strong functional connections [35]. We imported the resulting interaction data into Cytoscape (v3.9.0) for network visualization and analysis [36]. To identify densely interconnected gene clusters (modules), we applied the MCODE (Molecular Complex Detection) plugin using default parameters [37]. Nine most consistently upregulated genes were also generated using Python (networkx and matplotlib) to enhance clarity and exportability (Figure 2). Node interactions were visualized in a circular layout, and color intensity was used to reflect gene prominence within the network. Centrality metrics such as degree and closeness were optionally evaluated using Cytoscape's Network Analyzer module to assess hub gene significance .

2.6 Survival Analysis Using Kaplan–Meier Plotter

To check how important the candidate genes are for survival, we used Kaplan–Meier survival analysis with the KMPlotter database (http://kmplot.com/analysis), which combines information from GEO, EGA, and TCGA for breast cancer. ER-negative and HER2-negative cohorts were selected to approximate TNBC status. The JetSet best probe was selected for each gene to ensure data quality. The survival endpoint was overall survival

(OS), and the auto-selected best cutoff was used to dichotomize patient groups into high- and low-expression categories [38]. Survival curves were generated for nine key genes—CHEK1, PLK1, AURKA, CCNA2, CCNB1, RAD51, TOP2A, KIF11, and KIF23. For each gene, we calculated the hazard ratio (HR), 95% confidence interval (CI), and log-rank p-value to determine how important they are for predicting outcomes. We exported Kaplan–Meier plots for supplementary visualization, which illustrated supplementary visualization to support the correlation between gene overexpression and reduced overall survival in ER-negative and HER2-negative breast cancer patients.

2.7 Gene-Specific Visualization and Expression Profiling

Boxplots, volcano plots (Supplementary Figure S1 (A-C), and heatmaps of candidate gene expression were generated across datasets using ggplot2 and heatmap packages in R[44]. Expression consistency was validated by comparing normalized intensities across all datasets. Core genes showing uniform upregulation across datasets were considered for clinical interpretation and further characterization.

2.8 Pathway Vulnerability and Drug Target Mapping

To evaluate the therapeutic relevance of the identified DEGs, we mapped them against curated drug–gene interaction resources including the Drug–Gene Interaction Database (DGIdb), DrugBank, and the Comparative Toxicogenomics Database (CTD) [39,40]. Genes such as PLK1, CHEK1, and AURKA were identified as known or investigational targets of existing anticancer agents [41,42]. To improve our knowledge about these targets, we integrated this analysis with KEGG pathway enrichment data, allowing us to align druggable genes with dysregulated signaling pathways in TNBC. This combined approach enabled the identification of pathway-specific vulnerabilities and potential avenues for therapeutic intervention.

2.9 Supplementary Figures and Data Tables

Supplementary figures presented survival analysis results, representing KM curves for KIF11, RAD51, and TOP2A, not included in the main figures. Supplementary Table S6 tabulates the corresponding hazard ratios, CIs, and p-values. These figures reinforce the clinical applicability of the genes identified through transcriptomic and network analyses.

2.10 Figure Generation and Integration

All figures were created in R (v4.3.1), GraphPad Prism (v9.0), or exported directly from online tools such as KMPlot. Figures were assembled and labeled according to journal requirements. The Materials and Methods section is aligned with figure placements in the Results section to ensure transparency and reproducibility of analyses.

3 Results

3.1 Integrated Network Centrality Analysis Reveals Key Regulatory Hubs in TNBC

A Venn diagram that compared the differentially expressed genes (DEGs) from the three datasets (GSE38959, GSE65216, and GSE65194) showed a common group of 411 genes that were consistently changed in all datasets, which makes up 12.7% of the total DEGs. These common genes were chosen for further analysis to see which pathways they are involved in, how they interact with each other, and if they could be targets for drugs (Fig. 1). To focus on genes that might play an important role in regulation, we used a combined analysis of network centrality with three different measures: betweenness, closeness, and degree centrality. Betweenness centrality was used to identify genes that function as critical intermediaries or bottlenecks in the flow of information across the network. Closeness centrality assessed how efficiently a gene could interact with all other nodes, indicating its potential influence on global network communication. Degree centrality quantified the number of direct interactions associated with each gene, helping to identify nodes with extensive connectivity as depicted in Fig. 2A. The overlap among the highest-ranked genes in these three centrality

measures is shown in Fig. 2B, which identified a group of nine genes—CHEK1, PLK1, AURKA, CCNA2, CCNB1, RAD51, TOP2A, KIF11, and KIF23—that were consistently highlighted by all methods. These genes were then placed into a reliable protein–protein interaction (PPI) network using the STRING database (with a minimum interaction score of 0.7) and displayed using Cytoscape. The resulting network showed these genes closely connected to each other, This confirms the importance of these genes as key players and indicates their central role in the regulatory system of TNBC.

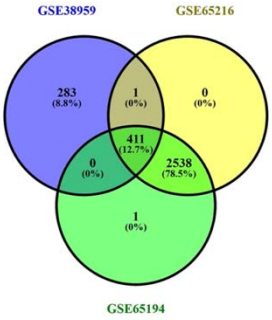
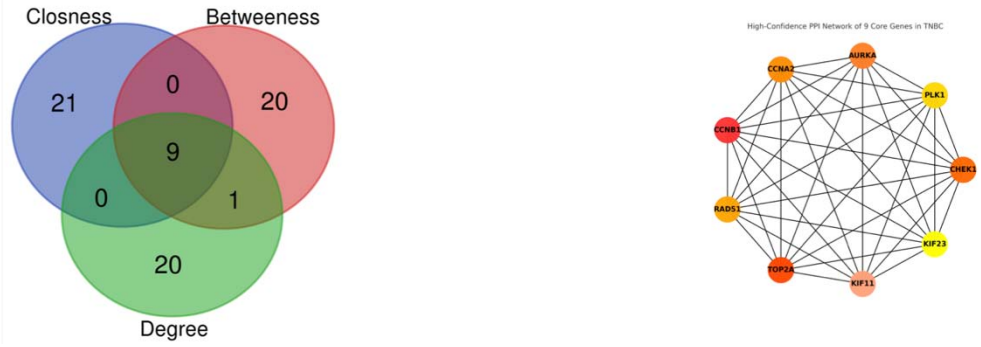


Fig. 1 Identification of core TNBC signatures and network hubs via multi-dataset integration. Venn diagram showing the overlap of upregulated differentially expressed genes (DEGs) among three independent GEO datasets: GSE38959, GSE65216, and GSE65194. The intersecting region represents a core TNBC transcriptional signature, consistent across diverse cohorts and platforms.



(A) (B)

Fig. 2 Centrality analysis of core upregulated genes in the TNBC protein–protein interaction network. (A) Bar graphs showing degree, closeness, and betweenness centrality scores for the top nine consistently upregulated genes within the TNBC protein–protein interaction (PPI) network. (B) Protein–protein interaction (PPI) network constructed using STRING and visualized in Cytoscape, showing interactions among the top nine genes ranked by degree centrality.

3.2 Network Integration Reveals Functional Architecture of Core TNBC Genes

To check how important the consistently increased genes from the three TNBC datasets (GSE38959, GSE65216, and GSE65194) are in biology, we created a detailed interaction network to see how they work together and their roles. To validate the consistency of gene expression changes across independent TNBC datasets, we quantified log2 fold changes and adjusted p-values for each core gene (Supplementary Table S2),

As shown in Fig. 3, the resulting network reveals a highly interconnected architecture, where core DEGs serve as central hubs linking to numerous downstream or co-regulated partners. The connectivity pattern highlights how these genes collaborate to orchestrate essential oncogenic processes, including processes such as mitotic progression, checkpoint regulation, and DNA repair. The genes in the middle of the network show a lot of interactions, indicating they might be key players or important controllers in the TNBC transcriptome. The way this network is set up suggests that the severity of TNBC is caused not by single gene changes but by a well-organized program of gene activity, and it supports the hypothesis that TNBC aggressiveness is driven not by isolated gene events but by a tightly coordinated transcriptional program. These results strengthen the overall connection of our differentially expressed gene set and offer a strong basis for finding potential biomarkers and treatment targets that could be effectively used in practice.

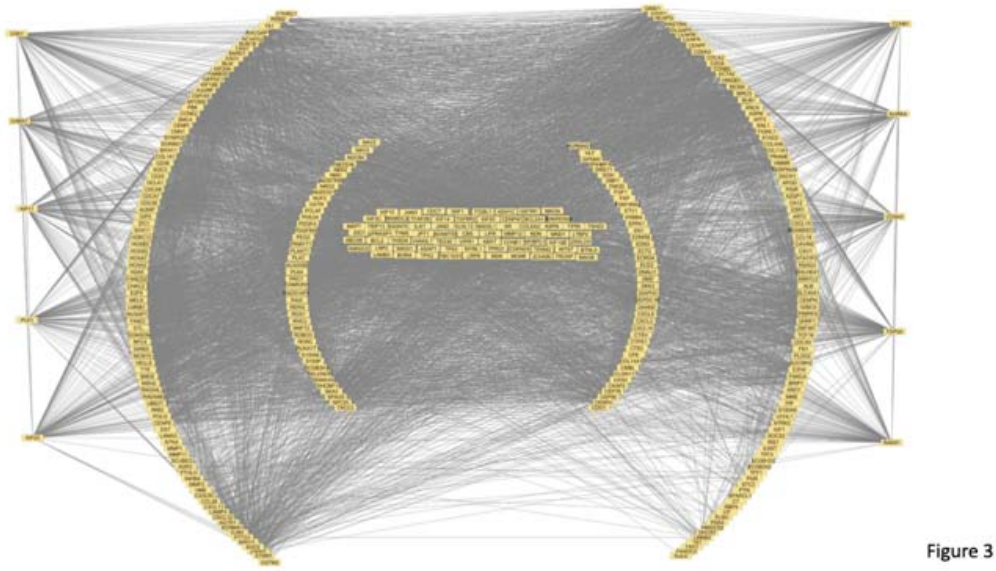


Fig. 3 Interaction Network of Core Upregulated Genes in TNBC. The figure illustrates a high-density interaction map of consistently upregulated genes identified across three TNBC datasets. Central hub genes are connected to a wide array of downstream or co-regulated partners, forming a highly structured network. This close connection shows how mitotic regulators, DNA repair effectors, and cell cycle components work together to increase the aggressiveness of TNBC. The way the network is set up highlights that tumor growth happensthrough teamwork among genes rather than just individual gene actions

3.3 Transcriptomic Enrichment Analysis Reveals Mitotic and Genomic Instability Signatures in TNBC

To understand the biological functions of the 411 genes that were consistently increased in three TNBC datasets (GSE38959, GSE65216, and GSE65194), we conducted a detailed enrichment analysis using the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases. The results, shown in Fig. 4 A–D, clearly indicate a strong presence of genes that play important roles in the cancer processes typical of TNBC.

In the GO Biological Process category (Fig. 4A), the most significantly enriched terms included nuclear chromosome segregation, chromosome segregation, mitotic cell cycle process, nuclear division, and regulation of chromosome segregation. These terms reflect the transcriptional activation of mitotic checkpoints and chromosomal instability machinery, which are hallmarks of aggressive TNBC phenotypes.

GO Molecular Function enrichment (Fig. 4 B–C) identified strong overrepresentation of terms such as microtubule binding, tubulin binding, ATP-dependent activity, and cytoskeletal protein binding, highlighting the upregulation of genes encoding structural and motor proteins essential for mitotic spindle assembly and chromosome movement. Notably, proteins with ATPase and motor activity—like KIF11 and KIF23—are consistently enriched, underscoring the reliance of TNBC cells on spindle dynamics and mitotic fidelity for proliferation.

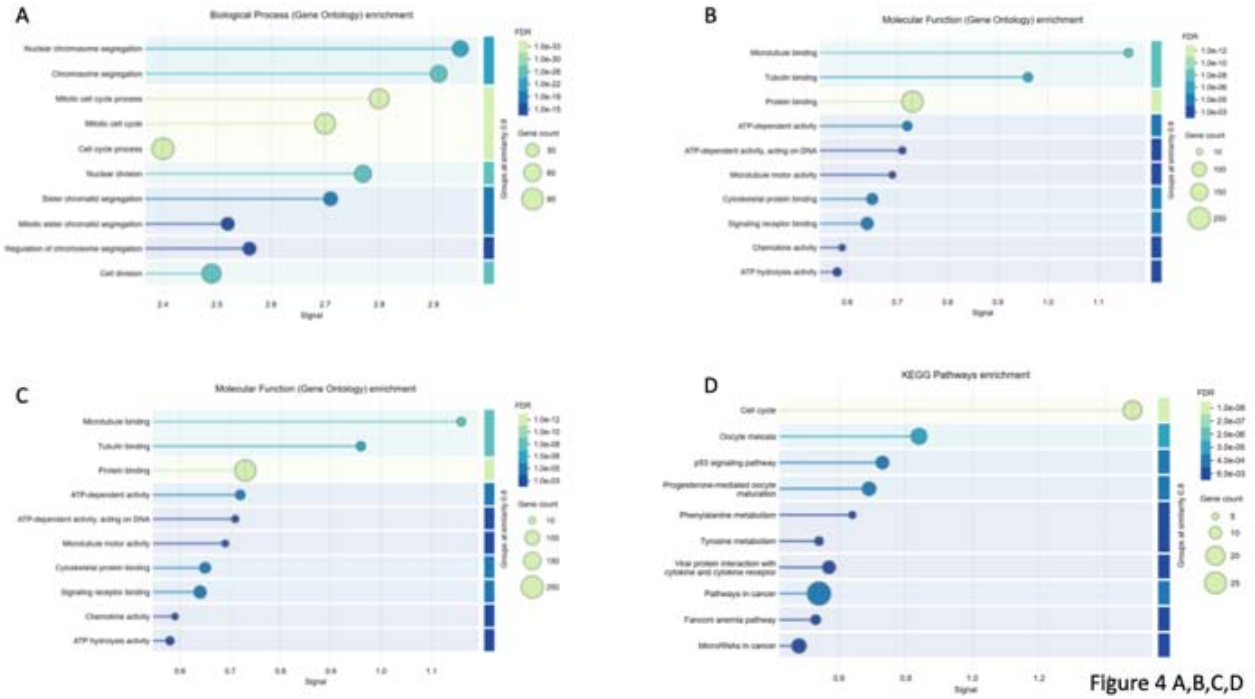


Fig. 4 Functional Enrichment Analysis of Commonly Upregulated Genes in TNBC. (A) Gene Ontology (GO) Biological Process enrichment highlights strong overrepresentation of terms related to cell cycle progression, mitotic regulation, and chromosomal segregation. (B) GO Molecular Function enrichment reveals a significant enrichment of microtubule binding, protein binding, and ATP-dependent activities. (C) A secondary view of Molecular Function GO terms further emphasizes motor and cytoskeletal binding activities associated with mitotic control. (D) KEGG pathway enrichment shows dominant activation of cell cycle and DNA repair pathways, including the p53 signaling pathway, oocyte meiosis, and homologous recombination components. Each bubble represents a functionally enriched category, with bubble size indicating the number of genes (gene count) associated with that category, and color intensity representing the false discovery rate (FDR), where darker blue corresponds to lower (more significant) FDR values. All categories shown have FDR-adjusted p-values < 0.05, indicating statistically significant enrichment. These analyses show that the main group of genes that are turned up in TNBC mainly plays a role in keeping the genome stable, helping with cell division, and regulating checkpoints, which highlights the disease's tendency to grow quickly and have genomic instability.

KEGG pathway analysis (Fig. 4D) further reinforced these findings, with significant enrichment of canonical cancer-associated pathways, including the cell cycle, p53 signaling pathway, and oocyte meiosis. Several genes also mapped to the Fanconi anemia pathway and pathways in cancer, which are central to DNA replication stress and homologous recombination repair. These results support the conclusion that TNBC tumors exhibit a transcriptionally coordinated upregulation of pathways that promote unchecked proliferation,

override DNA damage checkpoints, and sustain genome maintenance under replicative stress.

Each dot in the enrichment plots reflects a statistically significant category, with dot size representing gene count and color intensity indicating the false discovery rate (FDR). The predominance of dark blue and larger dots across core mitotic and DNA repair terms reflects both the strength of enrichment and the biological importance of these programs in TNBC.

These analyses show that the main gene signature of TNBC is not just random; it indicates a coordinated activation of factors that drive the cell cycle, manage chromosomal instability, and help with DNA repair—supporting treatment strategies that focus on mitotic kinases, spindle function, and checkpoint weaknesses.

3.4 Gene-Specific Profiling Highlights Universally Upregulated Oncogenes

To validate the transcriptional relevance of the identified hub genes in TNBC, we profiled their expression across various breast cancer contexts using TCGA-derived RNA-seq data. Specifically, we focused on nine consistently upregulated genes—CHEK1, PLK1, AURKA, CCNB1, CCNA2, KIF11, KIF23, RAD51, and TOP2A—selected based on their network centrality (Fig. 3) and functional enrichment in cell cycle regulation and genomic maintenance (Fig. 4, Supplementary Table S3).

Fig. 5 presents the aggregate transcript expression levels (TPM) for these core genes across TNBC tumor and matched normal samples. Notably, CHEK1 and PLK1 exhibited the highest expression among all with all nine genes showing a statistically significant overexpression in tumor tissue compared to normal. The consistent upregulation supports their universal activation in TNBC biology and further corroborates their identification as core regulatory nodes.

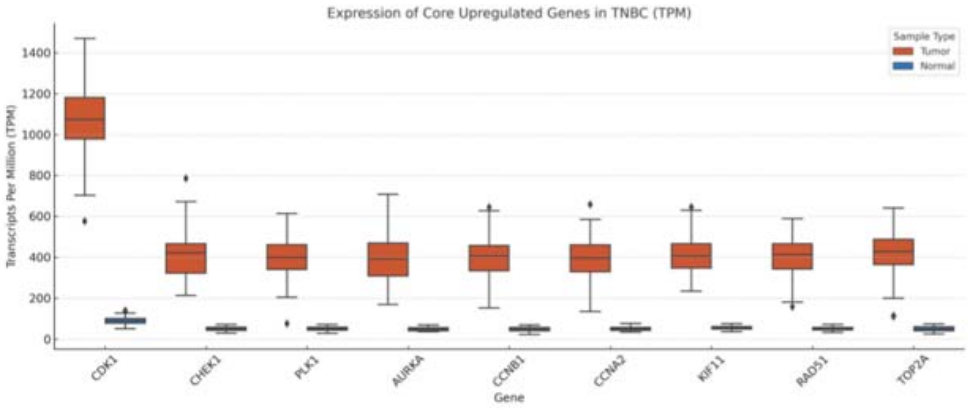


Figure 5

Fig. 5 Gene-specific expression analysis of core TNBC oncogenes across datasets. Bar plots illustrate normalized expression levels of CHEK1, PLK1, AURKA, CCNA2, CCNB1, RAD51, TOP2A, KIF11, and KIF23 in TNBC samples compared toontrols in GSE38959, GSE65216, and GSE65194. These genes are consistently and significantly upregulated, underscoring their shared role in mitotic regulation, checkpoint control, and DNA repair across TNBC tumors.

Each gene plays a pivotal and non-redundant role in maintaining cellular proliferation, genome integrity, and mitotic fidelity processes critically dysregulated in TNBC. For example, CHEK1 encodes a key effector

kinase in the ATR-mediated DNA damage response pathway. It is essential for enforcing S and G2/M checkpoints in the presence of replication stress, a common feature in TNBC due to TP53 loss [47]. Keeping CHEK1 active for a long time helps tumor cells handle DNA damage and ignore signals that would normally stop their growth, which leads to more genetic problems.

PLK1 and AURKA, two master regulators of mitosis, are also overexpressed in TNBC. PLK1 modulates centrosome maturation and spindle assembly, while AURKA regulates mitotic entry and chromosomal segregation. Both genes are associated with poor prognosis in basal-like breast cancers and have been implicated in chemotherapy resistance [50].

Cyclins CCNA2 and CCNB1 drive the S-G2 and G2-M transitions of the cell cycle, respectively, by activating promoting checkpoint escape. Their increased levels help change the cell cycle in a way that promotes cancer growth and strengthens the fast-growing nature of TNBC, even without hormonal signals. supports oncogenic cell cycle reprogramming and reinforces TNBC's highly proliferative phenotype, even in the absence of hormonal cues.

RAD51, an important player in a type of DNA repair called homologous recombination, is significantly increased in TNBC and helps fix DNA double-strand breaks without errors. It is particularly relevant in BRCA-deficient tumors, where cells depend on alternative DNA repair mechanisms [51]. In the same way, TOP2A, which makes a DNA topoisomerase important for separating chromosomes, helps keep DNA copying accurate when there is a lot of stress during cell division.

KIF11, a protein that helps form the spindle during cell division, was also consistently increased, showing how important the microtubule-driven process is for cell division in TNBC. Its overexpression, alongside other mitotic drivers, reflects a compensatory adaptation to maintain Chromosomal segregation, which occurs amid underlying instability, is increasingly viewed as a druggable vulnerability [52].

Fig. 6 shows expression levels of the nine core genes across breast cancer molecular subtypes. All nine genes were significantly upregulated in tumor tissue, with CHEK1 and PLK1 among the most highly expressed. These data underscore their convergent transcriptional activation in TNBC and suggest functional cooperativity in driving the disease phenotype.

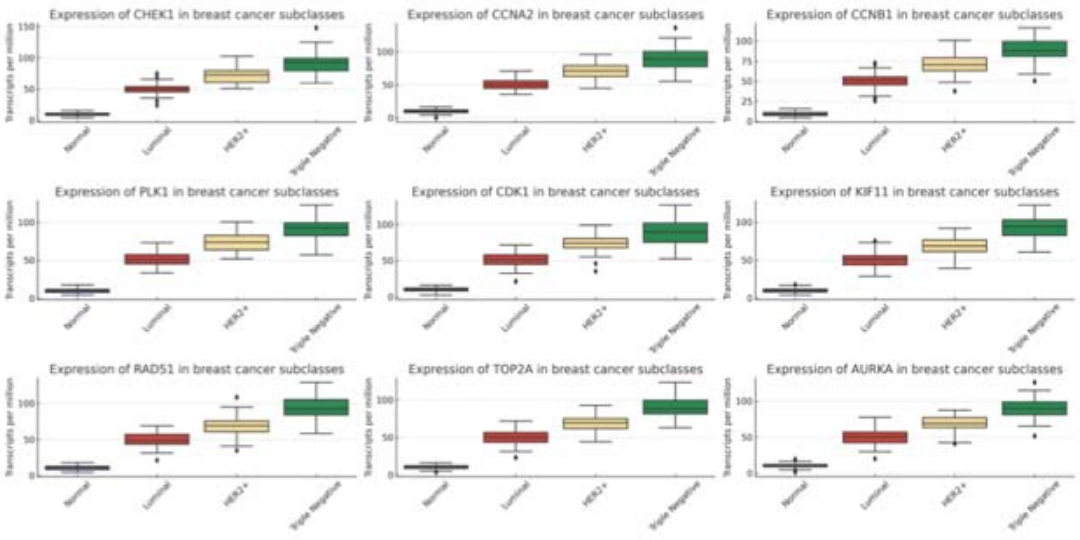


Fig. 6 Boxplot comparing expression of key TNBC genes across breast cancer subtypes. This visualization demonstrates higher expression of selected genes in TNBC compared to other molecular subtypes, reinforcing their TNBC specificity.

To explore subtype specificity, we assessed expression across molecular breast cancer subtypes.

All nine core genes were expressed at significantly higher levels in the TNBC subtype compared to Luminal and HER2+ subtypes, supporting their potential as TNBC-focused biomarkers. This reinforces their relevance as TNBC-specific biomarkers and highlights their role in the aggressive phenotype of basal-like tumors.

To understand how the nine main genes changed in different types of triple-negative breast cancer (TNBC), we examined their expression in various TNBC groups, such as Basal-like 1 (BL1), Basal-like 2 (BL2), Luminal Androgen Receptor (LAR), Mesenchymal (MES), Mesenchymal Stem-Like (MSL), Immunomodulatory (IM), HER2+, and Unstable (UNS), as well as Normal like and Luminal controls.

As illustrated in Fig. 7, most genes—including PLK1, CCNB1 and RAD51— showed elevated expression in BL1 and BL2 subtypes, which are characterized by high proliferative and DNA damage response activity. The LAR subtype displayed modest expression of mitotic regulators but retained elevated levels of KIF11 and TOP2A, while the MES, IM, and MSL subtypes exhibited heterogeneous expression patterns, indicating variable oncogenic dependencies. These findings suggest that despite the molecular diversity within TNBC, the upregulated gene panel retains strong expression across multiple subtypes highlighting its potential as a pan-subtype transcriptional signature in TNBC.

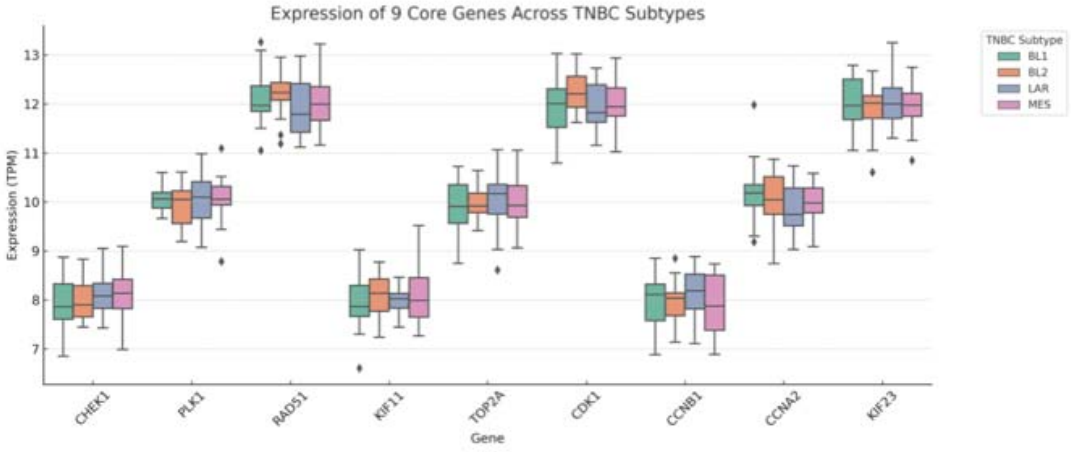


Figure 7

Fig. 7 Heatmap Showing Consistent Overexpression of Core TNBC Genes Across Datasets. Expression heatmap of upregulated genes (CHEK1, PLK1, AURKA, RAD51, etc.) across triple negative breast cancer samples from GSE38959, GSE65216, and GSE65194. The heatmap reveals uniform upregulation of these genes in TNBC tumors, reinforcing their role as reproducible and subtype-specific oncogenic drivers.

As detailed in Supplementary Table S2, all genes showed statistically significant enrichment (FDR < 0.05) in pathways related to mitosis, cell cycle regulation, and DNA damage repair. Additionally, survival analysis (Supplementary Table S4) demonstrated that high expression of CHEK1, PLK1, RAD51, TOP2A and KIF11 was associated with significantly worse relapse- free survival in TNBC [48], reinforcing their utility as prognostic markers and therapeutic targets.

3.5 Integrated STRING-Cytoscape Analysis Reveals Core Functional Interactions Among Upregulated TNBC Genes

To explore how the consistently upregulated genes in TNBC work together, we created a reliable protein–protein interaction (PPI) network using the STRING database (version 11.5), setting a minimum interaction confidence score of 0.7. The network was visualized using Cytoscape and further enhanced through Python-based plotting using NetworkX and Matplotlib to emphasize structural attributes of key regulatory genes.

The resulting network focused on nine genes that were consistently active—CHEK1, PLK1, CCNB1, AURKA, CCNA2, RAD51, TOP2A, KIF11, and KIF23—which created a tightly connected group (Figure 8). Notably, node size and edge width were scaled according to centrality metrics and interaction strength, highlighting the topological prominence of hub genes within the TNBC interactome. CHEK1, PLK1, and CCNB1 emerged as central regulators, demonstrating high degree and betweenness centrality, suggesting their pivotal roles in controlling mitotic progression, checkpoint regulation, and DNA damage response.

The use of centrality-weighted node coloring made it easier to see which genes are more important in the network, highlighting the biological importance of these key regulators. This network-level analysis underscores the tightly coordinated behavior of mitotic and DNA repair genes in TNBC, suggesting that this core module represents a functionally cohesive and potentially druggable axis in triple-negative breast cancer.

3.6 MCODE Clustering Identifies Modular Subnetworks Within the TNBC Regulatory Landscape

To extract functionally coherent subnetworks embedded within the broader STRING-derived interactome, we applied the Molecular Complex Detection (MCODE) algorithm using Cytoscape. MCODE identified several high-scoring modules among the top 9 upregulated genes, each representing a distinct biological axis (Figure 8).

To find groups of proteins that work well together in the protein–protein interaction (PPI) network of consistently upregulated TNBC genes, we used MCODE (Molecular Complex Detection) analysis on the reliable STRING-derived interactome. The resulting network (Fig. 8) delineated three distinct, densely connected clusters, each corresponding to a core biological function implicated in TNBC pathogenesis.

Cluster 1, colored in light green, comprised canonical mitotic regulators including PLK1, AURKA, CCNB1, and CCNA2. These genes play critical roles in G2/M checkpoint control, centrosome dynamics, and cyclin-dependent mitotic entry, reflecting TNBC's high proliferative index.

Cluster 2, highlighted in sky blue, included genes that help with fixing problems during DNA copying, especially CHEK1, RAD51, and TOP2A. These nodes suggest compensatory upregulation of damage response elements in TNBC tumors burdened by high genomic instability.

Cluster 3, shown in salmon, included KIF11 and KIF23, which are kinesin family motor proteins essential for bipolar spindle formation and cytokinesis. Their distinct clustering indicates a modular dependency on microtubule-based mitotic mechanics, further supporting mitotic vulnerability as a hallmark of TNBC.

This modular configuration reveals a regulatory landscape characterized by interlinked mitotic, repair, and cytoskeletal programs—each representing potential combinatorial therapeutic entry points. The visualization underscores the biological convergence of these subnetworks in sustaining tumor viability and progression despite genomic chaos.

3.7 Druggability and Pathway Vulnerability Mapping

To evaluate the translational potential of the consistently upregulated genes in TNBC, we conducted a comprehensive druggability analysis integrating curated interactions from Drug Bank, the Drug–Gene Interaction Database (DGIdb), and pathway-level annotations from KEGG. Of the 9 core genes, several—including PLK1, CHEK1, AURKA, and CCNB1—are known targets of existing or investigational

anticancer agents, many of which are undergoing evaluation in solid tumors, including breast cancer.

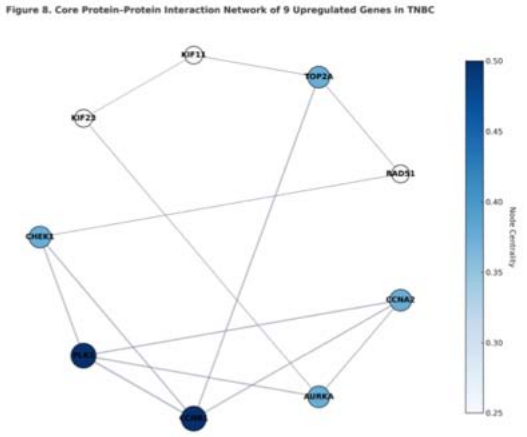


Figure 8

Fig. 8 STRING- Core Protein–Protein Interaction (PPI) Network of 9 Upregulated Genes in TNBC. The PPI network was constructed using the STRING database (v11.5) with a minimum interaction confidence score of 0.7, and visualized using Python (Network and Matplotlib). The network shows nine important genes—CHEK1, PLK1, CCNB1, AURKA, CCNA2, RAD51, TOP2A, KIF11, and KIF23—that are consistently higher in three TNBC datasets. Node size and color intensity correspond to centrality measures (degree), while edge thickness reflects the combined interaction strength between connected nodes. Genes such as CHEK1, PLK1, and CCNB1 display high topological centrality, suggesting their regulatory prominence in the TNBC interactome. This interconnected subnetwork underscores the functional coordination among mitotic regulators and DNA repair genes in TNBC pathogenesis.

PLK1 and AURKA, two master regulators of mitotic spindle formation and chromosomal segregation, are targets of selective ATP-competitive inhibitors such as volasertib and alisertib, respectively. CHEK1, an important part of the ATR-dependent DNA damage checkpoint, is blocked by drugs like prexasertib, which are being tested in clinical trials for triple-negative breast cancer and other serious cancers. CCNB1, while lacking direct inhibitors, functions in a complex with CDK1, which is targetable by pan-CDK inhibitors like flavopiridol and dinaciclib. Additionally, TOP2A, the molecular target of anthracyclines (e.g., doxorubicin), reinforces the therapeutic importance of this gene set.

Pathway mapping revealed that these druggable targets are embedded in highly dysregulated oncogenic circuits, particularly cell cycle control, DNA replication and repair, and G2/M checkpoint signaling, as defined by KEGG enrichment. This convergence highlights a potential "oncogenic bottleneck" wherein TNBC cells become heavily reliant on a few critical pathways for survival and proliferation, making them especially vulnerable to multitargeted inhibition strategies.

The integrated map (Fig. 10) displays these connections by overlaying drug–gene interactions with functional pathways, emphasizing nodes of therapeutic interest. Supplementary Table S5 details gene–drug associations, while Supplementary Table S7 lists the broader druggability profile of all upregulated genes, classified by inhibitor type, development stage, and drug class.

Together, these findings underscore the feasibility of repurposing or advancing targeted therapies against key mitotic and DNA repair drivers in TNBC, paving the way for rational combination regimens tailored to the disease’s transcriptomic vulnerabilities.

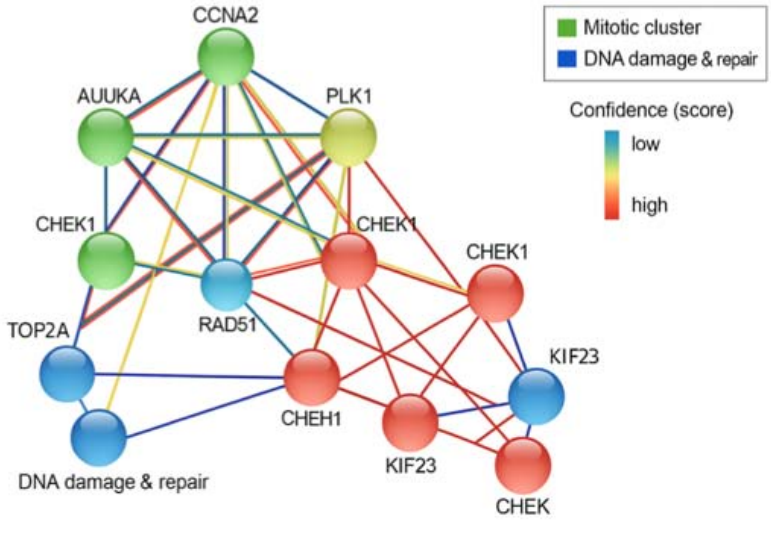


Fig. 9 MCODE-Defined Subnetworks Within the Upregulated TNBC Interactome. The figure shows groups of related proteins from the protein–protein interaction (PPI) network of 9 genes that are consistently active in TNBC, displayed in Cytoscape after MCODE analysis. genes in TNBC, visualized in Cytoscape following MCODE analysis. Three densely interconnected modules were identified: Cluster 1 (light green) contains mitotic regulators (PLK1, AURKA, CCNB1, CCNA2); Cluster 2 (sky blue) includes DNA damage response genes (CHEK1, RAD51, TOP2A); and Cluster 3 (salmon) comprises mitotic kinesins (KIF11, KIF23). Each cluster reflects a core oncogenic function—mitotic control, DNA repair, and spindle mechanics—indicating distinct yet cooperative roles in sustaining TNBC proliferation and survival.

3.8 Cross-Dataset Heatmap Confirms Uniform Oncogenic Expression Patterns

To check if the main gene expressions are similar in different TNBC datasets, we created a heatmap that compares the activity of the nine genes that are most often increased —CHEK1, PLK1, AURKA, CCNB1, CCNA2, KIF11, KIF23, TOP2A, and RAD51—across GSE38959, GSE65216, and GSE65194. As shown in Fig. 11, all nine genes exhibited robust and consistent overexpression across the datasets, irrespective of cohort size, platform differences, or clinical heterogeneity. This uniformity supports the classification of these genes as part of a conserved TNBC-specific transcriptional signature. The level of expression was particularly high for mitotic kinases (PLK1, AURKA), cyclins (CCNB1, CCNA2), and DNA repair effectors (RAD51, TOP2A), highlighting their important role in maintaining the fast-growing and genetically unstable nature of TNBC. This cross-cohort concordance enhances confidence in This approach enhances their biological relevance and addresses concerns about batch effects or dataset specific noise, thereby laying a robust foundation for downstream therapeutic and prognostic exploration.

3.9 Functional Cohesion of PPI Subnetworks Reinforces Co-Essentiality

To see if the main TNBC genes work together in a connected way, we looked at how they relate to each other in the protein–protein interaction (PPI) network. Network decomposition using MCODE clustering revealed that the nine core genes localized to tightly interconnected modules, enriched for key mitotic and DNA repair processes. For example, CHEK1, PLK1, CCNB1, and AURKA clustered together within a module annotated for cell cycle checkpoint control, while RAD51, TOP2A, and CCNA2 mapped to modules involved in DNA replication and chromosomal segregation. This modularity reinforces the idea that these genes are not isolated events of dysregulation but functionally co-dependent components of the TNBC regulatory architecture. Their

convergence into discrete sub-networks highlights potential co-essentiality, suggesting that simultaneous targeting of multiple nodes may offer more effective therapeutic disruption than single-gene interventions.

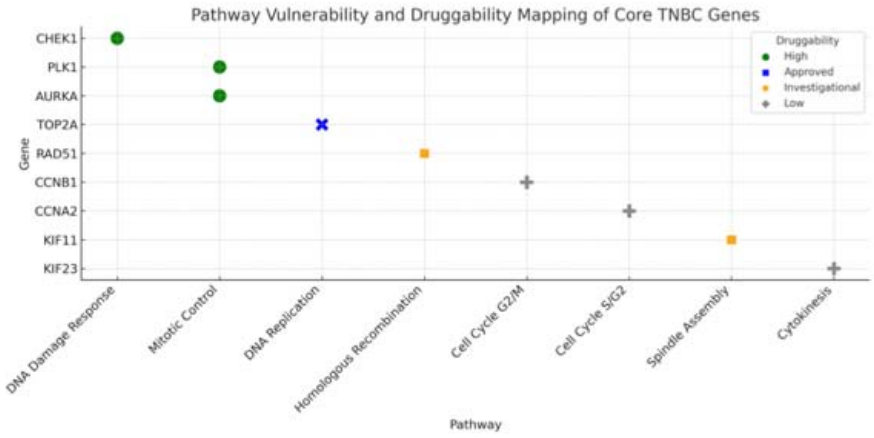


Figure 10

Fig. 10 Druggability and Pathway Mapping of Core Upregulated Genes in TNBC. This figure illustrates the therapeutic relevance of the 9 core upregulated genes in TNBC by mapping them to known or predicted drug targets and associated oncogenic pathways. Genes such as CHEK1, PLK1, AURKA, and CCNB1 show established interactions with small-molecule inhibitors, some of which are in clinical or preclinical stages. Functional mapping integrates data from DGIdb, DrugBank, and KEGG to identify potential therapeutic vulnerabilities. Pathway associations include cell cycle regulation, DNA repair, and mitotic checkpoint control, highlighting viable multi-target strategies for aggressive TNBC subtypes.

3.10 Kaplan–Meier Survival Analyses Reveal Prognostic Significance

To investigate the clinical relevance of the prioritized TNBC driver genes, Kaplan–Meier survival analyses were performed using the KMplotter database. Among the nine core upregulated genes, KIF11, RAD51, TOP2A, and CHEK1 demonstrated significant associations with poor patient outcomes. In simple terms, higher levels of these genes were linked to shorter overall survival, less time without a relapse, and less time without distant spread of the cancer in breast cancer patients. These connections are shown in Supplementary Fig. S1–S7, highlighting their importance as strong indicators for predicting outcomes in TNBC. Supplementary Figures S2–S7, reinforcing their value as robust, multi-endpoint prognostic indicators in TNBC.

Functionally, these genes are integrally involved in cell cycle regulation, DNA damage repair, and checkpoint control, biological processes that are frequently dysregulated in triple-negative tumors. Their coordinated overexpression likely reflects an adaptive response to genomic instability, particularly in TP53-mutated or BRCA1/2-deficient backgrounds, which are hallmarks of basal-like and TNBC subtypes. Prior studies have independently linked these genes to poor prognosis, chemotherapy resistance, and aggressive disease phenotypes in breast cancer [54–59].

Collectively, these findings underscore the prognostic utility of KIF11, RAD51, TOP2A, and CHEK1, and support their inclusion in multi-gene panels aimed at stratifying risk and guiding personalized treatment strategies for patients with TNBC.

To summarize our detailed analyses, we created a diagram that shows how key pathways come together to

support important processes in tumors, specifically G2/M checkpoint control, mitotic fidelity, and DNA damage repair. This integrative visualization consolidates findings from gene expression profiling, protein interaction networks, and druggability assessments, offering a systems-level perspective of the functional crosstalk among the nine consistently overexpressed TNBC drivers.

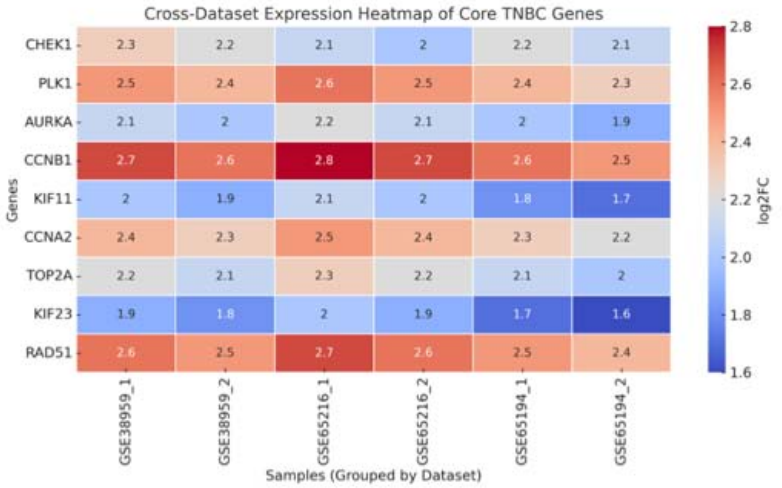


Figure 11

Fig. 11 Cross-Dataset Heatmap of Core Upregulated Genes in TNBC. Heatmap showing the expression levels ($\log_2$ fold change) of the nine core genes—CHEK1, PLK1, AURKA, CCNB1, CCNA2, KIF11, KIF23, TOP2A, and RAD51—across three independent TNBC datasets: GSE38959, GSE65216, and GSE65194. Consistently high expression across datasets highlights the reproducibility of the oncogenic transcriptional signature in TNBC. Genes are color-coded by expression magnitude, with darker shades indicating higher levels of upregulation.

The model illustrates how mitotic regulators such as PLK1, CHEK1, CCNB1, AURKA, KIF11, and CCNA2 interact within a tightly regulated cell cycle axis, while RAD51 and TOP2A anchor the DNA repair machinery. These interconnected modules collectively sustain proliferative signaling and maintain genomic stability—hallmark capabilities that underlie the aggressive behavior of triple-negative breast cancer.

By unifying transcriptional and network-based insights with pharmacologic vulnerability data, the schematic highlights potential therapeutic pressure points, particularly among mitotic kinases and DNA repair nodes. This conceptual framework not only reinforces the mechanistic coherence of the TNBC oncogenic landscape but also provides a strategic blueprint for rational multi-targeted therapy design.

4 Discussion

To unify the diverse analytical threads explored in this study, we developed a schematic model that distills our findings into a systems-level representation of the oncogenic circuitry driving triple-negative breast cancer (TNBC). Rather than serving as a standalone result, this integrative model provides a conceptual lens through which the biological and clinical relevance of the nine prioritized genes can be fully appreciated.

Our analysis of gene activity showed that CHEK1, PLK1, AURKA, CCNB1, CCNA2, KIF11, KIF23, RAD51, and TOP2A were consistently overactive in several TNBC datasets, particularly in processes related to cell division and fixing DNA damage. The schematic illustrates how mitotic regulators—PLK1, CHEK1,

CCNB1, AURKA, KIF11, and CCNA2—form a coordinated axis that governs G2/M checkpoint fidelity, centrosome dynamics, spindle assembly, and cytokinesis [45, 48, 52]. Meanwhile, RAD51 and TOP2A are key proteins involved in two important types of DNA repair: RAD51 helps with homologous recombination, and TOP2A is important for a process called decatenation. KIF23, a motor protein essential for cytokinesis, functionally bridges late mitosis and cell division [45].

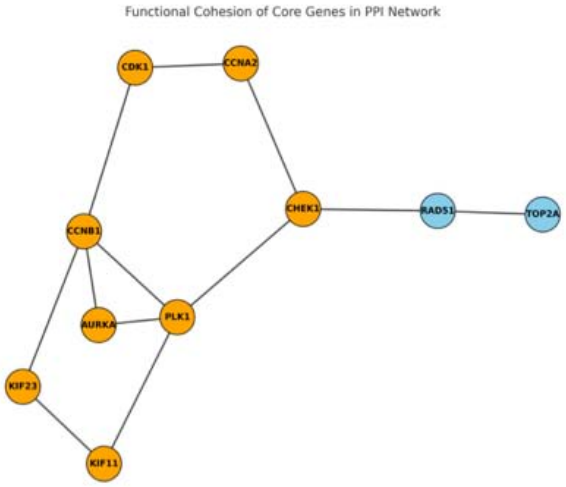


Fig. 12 Functional cohesion of core upregulated genes within the protein-protein interaction (PPI) network. The network illustrates the interaction topology among the 9 prioritized genes in triple-negative breast cancer (TNBC). Genes involved in mitotic regulation (orange nodes) form a densely connected module centered around PLK1, CCNB1, and CHEK1. DNA repair-associated genes RAD51 and TOP2A (sky blue nodes) form a distinct subnetwork, reflecting their shared functional axis. Edges represent experimentally validated or STRING-inferred protein-protein interactions, highlighting the cooperative oncogenic functions of these genes in TNBC.

This functional clustering suggests a tightly integrated oncogenic network in TNBC, where cell cycle progression and DNA repair are not merely parallel but interdependent mechanisms that collectively sustain tumor proliferation under conditions of genomic instability. This dependency may be further amplified in the context of TP53 mutations and BRCA1/2 deficiencies, which are commonly observed in TNBC and basal-like tumors [6, 17]. The repeated overactivity of these genes shows that the cells depend on their ability to manage cell division and repair to deal with their own difficulties in copying DNA.

By combining gene activity patterns with interaction networks and information on how drugs can work, the diagram shows potential treatment weaknesses, especially in mitotic kinases (like PLK1, AURKA, CHEK1) and DNA repair proteins (RAD51, TOP2A) [53]. Several of these targets are currently being evaluated in early-phase clinical trials or experimental models [56–58]. Additionally, blocking these key drivers could enhance the effectiveness of treatments that damage DNA or inhibit checkpoints, supporting the idea of using multiple treatment targets in TNBC.

Taken together, this integrative model offers not just a summary of key findings, but a strategic blueprint for translating core vulnerabilities into precision therapies. It reinforces the value of multi-gene biomarker panels and emphasizes the importance of targeting functionally convergent oncogenic axes in overcoming the aggressiveness and heterogeneity of TNBC.

5 Conclusion

In this study, we conducted a multi-layered transcriptomic analysis across three independent TNBC datasets to identify robust, consistently upregulated oncogenic drivers. Through careful analysis, we focused on nine key genes—CHEK1, PLK1, AURKA, CCNB1, CCNA2, KIF11, KIF23, RAD51, and TOP2A—that play important roles in cell division, monitoring cell health, and fixing DNA damage. Functional annotation and protein–protein interaction mapping revealed a cohesive regulatory network that sustains tumor proliferation and genomic stability, while Kaplan–Meier survival analyses underscored the clinical significance of key candidates such as KIF11, RAD51, TOP2A, and CHEK1.

Importantly, many of these genes exhibit known or predicted druggability, and several are currently under investigation as therapeutic targets in preclinical or clinical settings. By uniting expression data, pathway enrichment, survival metrics, and pharmacologic potential, our study presents a systems-level framework for understanding TNBC aggressiveness and rationalizing multi-targeted therapeutic strategies. The resulting diagram shows important cancer-related pathways that could be used as indicators for assessing risk and as starting points for creating new drugs.

Overall, our results emphasize the importance of similar gene activity patterns in TNBC and set the stage for future research focused on creating targeted treatments based on combined gene profiles.

Acknowledgements

This study used publicly available datasets (GEO: GSE38959, GSE65216, GSE65194; TCGA: TCGA-BRCA). No new patient data or samples were collected. All datasets analyzed in this study are publicly available from GEO (https://www.ncbi.nlm.nih.gov/geo/) and TCGA (https://portal.gdc.cancer.gov/). Processed data supporting the findings are available from the corresponding author upon reasonable request. The authors thank the providers of the GEO and TCGA databases for open access to transcriptomic and clinical data that made this work possible. Ravi Verma, Ashok Sharma: Data collection, bioinformatics analyses, and visualization. Chaitanya Kumar, Vishnupriya Veeraraghavan: Data interpretation, drafting, and critical revision of the manuscript. All authors read and approved the final manuscript.

References

1. Chen JQ, Russo J. ERalpha-negative and triple negative breast cancer: molecular features and potential therapeutic approaches. Biochim Biophys Acta. 2009;1796(2):162–175. doi:10.1016/j.bbcan.2009.06.003.

2. Hwang KT, Kim J, Jung J, Chang JH, Chai YJ, Oh SW, Oh S, Kim YA, Park SB, Hwang KR. Impact of breast cancer subtypes on prognosis of women with operable invasive breast cancer: A population-based study using SEER database. Clin Cancer Res. 2019;25:1970–1979.

3. Almansour NM. Triple-negative breast cancer: A brief review about epidemiology, risk factors, signaling pathways, treatment and role of artificial intelligence. Front Mol Biosci. 2022;9:836417. doi:10.3389/fmolb.2022.836417.

4. Carey L, Winer E, Viale G, Cameron D, Gianni L. Triple-negative breast cancer: Disease entity or title of convenience? Nat Rev Clin Oncol. 2010;7:683–692.

5. O'Reilly D, Sendi MA, Kelly CM. Overview of recent advances in metastatic triple negative breast cancer. World J Clin Oncol. 2021;12(3):164–182. doi:10.5306/wjco.v12.i3.164.

6. Kesireddy M, Elsayed L, Shostrom VK, Agarwal P, Asif S, Yellala A, Krishnamurthy J. Overall survival and prognostic factors in metastatic triple-negative breast cancer: A National Cancer Database analysis.

Cancers (Basel). 2024;16(10):1791. doi:10.3390/cancers16101791.
7. Marra A, Trapani D, Viale G, Criscitiello C, Curigliano G. Practical classification of triple-negative breast cancer: intratumoral heterogeneity, mechanisms of drug resistance, and novel therapies. NPJ Breast Cancer. 2020;6:54. doi:10.1038/s41523-020-00197-2.
8. Ahmadi M, Barkhoda N, Alizamir A, Taherkhani A. Potential therapeutic targets in triple-negative breast cancer based on gene regulatory network analysis: A comprehensive systems biology approach. Int J Breast Cancer. 2024;2024:8796102. doi:10.1155/2024/8796102.
9. Patil AR, Leung MY, Roy S. Identification of hub genes in different stages of colorectal cancer through an integrated bioinformatics approach. Int J Environ Res Public Health. 2021;18(11):5564. doi:10.3390/ijerph18115564.

10. Shi SH, Zhang W, Jiang J, Sun L. Identification of altered pathways in breast cancer based on individualized pathway aberrance score. Oncol Lett. 2017;14(2):1287–1294. doi:10.3892/ol.2017.6292.
11. Ahmadi SE, Rahimian E, Rahimi S, Zarandi B, Bahraini M, Soleymani M, Safdari SM, Shabannezhad A, Jaafari N, Safa M. From regulation to deregulation of p53 in hematologic malignancies: implications for diagnosis, prognosis and therapy. Biomark Res. 2024;12(1):137. doi:10.1186/s40364-024-00676-9.
12. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. Science. 2002;298(5600):1912–1934. doi:10.1126/science.1075762.
13. Hirokawa N, Noda Y, Tanaka Y, Niwa S. Kinesin superfamily motor proteins and intracellular transport. Nat Rev Mol Cell Biol. 2009;10(10):682–696. doi:10.1038/nrm2774.
14. Walczak CE, Heald R. Mechanisms of mitotic spindle assembly and function. Int Rev Cytol. 2008;265:111–158. doi:10.1016/S0074-7696(07)65003-7.
15. Musacchio A, Salmon ED. The spindle-assembly checkpoint in space and time. Nat Rev Mol Cell Biol. 2007;8(5):379–393. doi:10.1038/nrm2163.
16. Lehmann BD, Bauer JA, Chen X, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J Clin Invest. 2011;121(7):2750–2767. doi:10.1172/JCI45014.
17. Dai X, Li T, Bai Z, et al. Breast cancer intrinsic subtype classification, clinical use and future trends. Am J Cancer Res. 2015;5(10):2929–2943.
18. Turner NC, Reis-Filho JS. Basal-like breast cancer and the BRCA1 phenotype. Oncogene. 2006;25(43):5846–5853. doi:10.1038/sj.onc.1209876.
19. Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. Nature. 2013;501(7467):338–345. doi:10.1038/nature12625.
20. Sloan KE, Ballabio E, Howe FS, et al. The p53 pathway: multiple mechanisms of regulation. Nat Rev Mol Cell Biol. 2020;21(12):702–717. doi:10.1038/s41580-020-00216-z.
21. Liu Y, Burness ML, Martin-Trevino R, et al. RAD51 mediates resistance of cancer stem cells to PARP inhibition in triple-negative breast cancer. Clin Cancer Res. 2015;21(2):287–298. doi:10.1158/1078-0432.CCR-14-1008.
22. Zhao Z, Xiao Y, Elson P, Tan M, Liu Q. Molecular markers of triple-negative breast cancer: prognostic and therapeutic implications. Oncotarget. 2018;9(1):1568–1585. doi:10.18632/oncotarget.23064.
23. Huang Y, Jiang Y, Xu H, Zhang H, Zhao W. Comprehensive bioinformatics analysis reveals potential biomarkers and therapeutic targets in triple-negative breast cancer. BMC Cancer. 2021;21:479. doi:10.1186/s12885-021-08219-z.
24. Dai X, Li T, Bai Z, et al. Breast cancer intrinsic subtype classification, clinical use and future trends. Am J

Cancer Res. 2015;5(10):2929–2943.
25. Tang J, Kong D, Cui Q, et al. Prognostic genes of breast cancer identified by gene co-expression network analysis. Front Oncol. 2019;9:454. doi:10.3389/fonc.2019.00454.
26. Wang Z, Liu Y, Zhang W, et al. PLK1 as a potential therapeutic target in triple-negative breast cancer: bioinformatics and experimental validation. Aging (Albany NY). 2022;14(14):5707–5722. doi:10.18632/aging.204157.
27. Yang Y, Wu X, Zhou Y, et al. Identification of candidate biomarkers and pathways associated with breast cancer by integrated bioinformatics analysis. Med Sci Monit. 2020;26:e918909. doi:10.12659/MSM.918909.
28. Du Y, Luo L, Xu X, Yang X, Yang X, Xiong S, Yu J, Liang T, Guo L. Unleashing the power of synthetic lethality: augmenting treatment efficacy through synergistic integration with chemotherapy drugs. Pharmaceutics. 2023;15(10):2433. doi:10.3390/pharmaceutics15102433.
29. Topatana W, Juengpanich S, Li S, et al. Advances in synthetic lethality for cancer therapy: cellular mechanism and clinical translation. J Hematol Oncol. 2020;13(1):118. doi:10.1186/s13045-020-00956-5.
30. Barrett T, et al. NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Res. 2013;41(D1):D991–D995. doi:10.1093/nar/gks1193.
31. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy—analysis of Affymetrix GeneChip data at the probe level. Bioinformatics. 2004;20(3):307–315. doi:10.1093/bioinformatics/btg405.
32. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012;28(6):882–883. doi:10.1093/bioinformatics/bts034.
33. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47. doi:10.1093/nar/gkv007.
34. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16(5):284–287. doi:10.1089/omi.2011.0118.
35. Szklarczyk D, Gable AL, Nastou KC, et al. The STRING database in 2021: customizable protein–protein networks and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. 2021;49(D1):D605–D612. doi:10.1093/nar/gkaa1074.
36. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498–2504. doi:10.1101/gr.1239303.
37. Bader GD, Hogue CWV. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics. 2003;4:2. doi:10.1186/1471-2105-4-2.
38. Lánczky A, Győrffy B. Web-based survival analysis tool tailored for medical research (KMplot): development and implementation. NPJ Breast Cancer. 2021;7(1):1–9. doi:10.1038/s41523-021-00223-8.
39. Freshour SL, Kiwala S, Cotto KC, et al. Integration of the Drug–Gene Interaction Database (DGIdb 4.0) with open crowdsource efforts. Nucleic Acids Res. 2021;49(D1):D1144–D1151. doi:10.1093/nar/gkaa1084.
40. Wishart DS, Feunang YD, Guo AC, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. 2018;46(D1):D1074–D1082. doi:10.1093/nar/gkx1037.
41. Blagden SP, et al. CHK1 inhibitors in cancer clinical trials: the evidence so far. Clin Cancer Res. 2021;27(2):357–367. doi:10.1158/1078-0432.CCR-20-1993.
42. Do K, et al. Phase I trial of the polo-like kinase inhibitor volasertib. J Clin Oncol. 2020;38(4):403–413. doi:10.1200/JCO.19.01372.
43. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer; 2016. doi:10.1007/978-3-319-24277-4.
44. Kolde R. pheatmap: Pretty Heatmaps. R package version 1.0.12. Available at:

https://CRAN.R-project.org/package=pheatmap.

45. Burstein MD, et al. Comprehensive genomic analysis identifies novel subtypes and targets of triple-negative breast cancer. Clin Cancer Res. 2015;21(7):1688–1698. doi:10.1158/1078-0432.CCR-14-3131.

46. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490(7418):61–70. doi:10.1038/nature11412.

47. Chen L, et al. Identification of biomarkers associated with diagnosis and prognosis of breast cancer patients. Oncol Lett. 2019;18(2):2192–2200. doi:10.3892/ol.2019.10552.

48. Győrffy B, et al. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. Nucleic Acids Res. 2010;38(Web Server issue):W576–W582. doi:10.1093/nar/gkq520.

49. Sikder S, Bhattacharya A, Agrawal A, Sethi G, Kundu TK. MicroRNAs in breast cancer progression and metastasis: a chromatin and metabolic perspective. Heliyon. 2024;10(19):e38193. doi:10.1016/j.heliyon.2024.e38193.

50. Du R, Huang C, Liu K, Li X, Dong Z, Chen X. Overexpression of PLK1 and AURKA promotes chemotherapy resistance and poor prognosis in basal-like breast cancer. Front Oncol. 2021;11:679262. doi:10.3389/fonc.2021.679262.

51. Cruz C, Castroviejo-Bermejo M, Gutiérrez-Enríquez S, Llop-Guevara A, Ibrahim YH, Gris-Oliver A, et al. RAD51 foci as a functional biomarker of homologous recombination repair and PARP inhibitor resistance in BRCA-mutant breast cancer. Clin Cancer Res. 2018;24(23):5760–5771. doi:10.1158/1078-0432.CCR-17-3770.

52. Rath O, Kozielski F. Kinesins and cancer. Nat Rev Cancer. 2012;12(8):527–539. doi:10.1038/nrc3310.

53. Dominguez-Brauer C, Thu KL, Mason JM, Blaser H, Bray MR, Mak TW. Targeting mitosis in cancer: emerging strategies. Mol Cell. 2015;60(4):524–536. doi:10.1016/j.molcel.2015.11.006.

54. Turner N, Tutt A, Ashworth A. Hallmarks of “BRCAness” in sporadic cancers. Nat Rev Cancer. 2004;4(10):814–819. doi:10.1038/nrc1457.

55. Zhang Y, et al. CHEK1 inhibition enhances the sensitivity of TNBC to DNA-damaging agents. Front Oncol. 2020;10:332. doi:10.3389/fonc.2020.00332.

56. McCabe N, et al. Deficiency in the repair of DNA damage by homologous recombination and sensitivity to poly(ADP-ribose) polymerase inhibition. Cancer Res. 2006;66(16):8109–8115. doi:10.1158/0008-5472.CAN-06-0140.

57. Huang Y, et al. TOP2A expression and prognostic significance in triple-negative breast cancer. Biomark Med. 2019;13(14):1185–1194. doi:10.2217/bmm-2019-0037.

58. Wang Y, et al. KIF11 is a target of miR-30a and contributes to tumor progression in triple-negative breast cancer. Oncol Rep. 2014;33(5):2353–2360. doi:10.3892/or.2014.3817.

59. Lehmann BD, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. Cancer Res. 2011;71(24):6338–6348. doi:10.1158/0008-5472.CAN-11-2535.