

Investigation of Key Molecular Mechanisms of 6PPDQ in Breast Cancer Development and Progression through Network Toxicology and Computational Simulation

Haili Lu

Tongji University

Xi Li

Hubei University of Medicine

Jing Han

Tongji University

Hong Quan

quanhong-quan@163.com

Tongji University

Research Article

Keywords: 6PPDQ, Breast cancer, Network toxicology, Molecular docking, Molecular mechanism

Posted Date: March 27th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9012385/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

The rapid expansion of tire manufacturing has led to widespread environmental contamination by 6PPD-quinone (6PPDQ). Although 6PPDQ exhibits well-documented toxicological properties, its mechanistic links to breast cancer pathogenesis remain unexplored. This study represents the first attempt to examine these associations using network toxicology approaches.

Methods

Putative 6PPDQ targets were retrieved from PharmMapper and SwissTargetPrediction. Breast cancer datasets (GSE3744, GSE15852, GSE21422) from GEO were used for differential expression analysis and diagnostic modeling, with GSE52194 for validation. WGCNA identified disease modules and hub genes; GeneCards curated breast cancer gene sets. Intersection of 6PPDQ targets and hub genes yielded cross-targets for GO/KEGG enrichment and PPI network analysis. Topological algorithms (DMNC, MCC, Degree, EPC) prioritized key cross-targets. Molecular docking validated 6PPDQ-protein binding interactions.

Results

Thirty-six cross-target genes were successfully identified. Functional enrichment analysis revealed predominant involvement in cholesterol biosynthesis inhibition, hormone metabolic processes, and carboxylic acid binding, alongside significant pathway enrichment in PI3K-Akt signaling, breast cancer-specific cascades, transcriptional misregulation in cancer, AMPK signaling, IL-17 signaling, and xenobiotic metabolism via cytochrome P450. Integration of two machine learning algorithms with four Cytoscape topological approaches multi-dimensionally pinpointed two core genes: ESR1 and HMGCR. The risk prediction model constructed upon these genetic markers demonstrated robust predictive capability. Molecular docking analyses confirmed strong binding affinities between 6PPDQ and both core targets, with calculated binding energies of -7.7 kcal/mol for ESR1 and - 7.9 kcal/mol for HMGCR.

Conclusions

This investigation establishes ESR1 and HMGCR as pivotal targets mediating 6PPDQ-induced breast carcinogenesis. Our findings suggest that 6PPDQ promotes breast cancer development through multifaceted mechanisms encompassing metabolic reprogramming with cholesterol synthesis suppression, disruption of hormonal homeostasis, tumor microenvironment modulation, and activation of oncogenic signaling pathways. These discoveries furnish novel mechanistic insights into 6PPDQ-

mediated breast cancer pathogenesis and establish a theoretical foundation for therapeutic target development.

1. Introduction

Breast cancer (BC) represents the most frequently diagnosed malignancy among women globally, characterized by intricate pathological progression involving sophisticated networks of gene interactions and signaling pathway dysregulation[1–3]. While targeted therapeutic interventions and immunotherapeutic approaches have substantially enhanced patient survival outcomes, persistent challenges including tumor heterogeneity, acquired drug resistance, and metastatic recurrence continue to impede clinical management[3, 4]. Against this backdrop, the potential contribution of environmental pollutants to carcinogenesis has progressively captured scientific attention.

6PPD-quinone (6PPDQ), an emerging environmental contaminant, has garnered considerable interest following demonstrations of its pronounced toxicity in ecotoxicological investigations[5, 6]. Nevertheless, the precise molecular mechanisms through which 6PPDQ influences human breast cancer pathobiology remain completely unknown. The present investigation employs an integrative approach combining network toxicology, bulk RNA sequencing (bulkRNA-seq), and molecular docking simulations to systematically dissect the molecular regulatory networks underlying 6PPDQ-mediated breast cancer development and progression, with the objective of uncovering novel mechanisms driving pollutant-associated tumor advancement.

Breast cancer etiology encompasses both inherited genetic susceptibility and environmental exposure components. Beyond established genetic risk factors such as BRCA1/2 mutations[7], accumulating epidemiological evidence demonstrates that chronic exposure to endocrine-disrupting compounds, polycyclic aromatic hydrocarbons, and other environmental pollutants can perturb hormonal equilibrium within breast tissue, provoke DNA damage and oxidative stress, consequently elevating disease susceptibility[8, 9]. However, contemporary research predominantly concentrates on conventional pollutants, with inadequate attention directed toward emerging contaminants including 6PPDQ[10]. As an oxidative derivative of the rubber antioxidant 6PPD, 6PPDQ has been increasingly detected in urban stormwater runoff in recent years and has demonstrated lethal toxicity toward aquatic species, notably coho salmon[11–13]. Despite relatively low environmental residue concentrations, bioaccumulation through trophic transfer combined with chronic low-dose exposure may exert detrimental effects on metabolically active, hormone-responsive breast tissue.

Current toxicological evidence indicates that 6PPDQ elicits cytotoxicity through induction of reactive oxygen species (ROS) generation, mitochondrial impairment, and apoptotic cell death. The quinone moiety within its molecular architecture participates in redox cycling, continuously generating ROS that subsequently damages macromolecular structures and activates pro-oncogenic signaling cascades including NF- κ B and MAPK pathways[14–16]. However, systematic investigations addressing 6PPDQ-specific targets in breast cancer, associated regulatory networks, and impacts on malignant phenotypes

remain notably absent. Traditional toxicological methodologies prove insufficient for comprehensive analysis of complex multi-target, multi-pathway regulatory networks, whereas integration of computational biology with high-throughput omics technologies offers an effective complementary strategy.

This investigation adopts a "prediction-validation-mechanistic analysis" investigative framework: initially screening putative 6PPDQ targets through network toxicology to construct target-pathway regulatory networks; subsequently employing bulkRNA-seq technology to characterize transcriptomic alterations in breast cancer cells following 6PPDQ exposure, identifying differentially expressed genes (DEGs) and enriched pathways; finally, conducting molecular docking simulations to validate binding interactions between 6PPDQ and core targets, thereby elucidating molecular mechanisms governing regulation of malignant breast cancer phenotypes.

By focusing on the potential carcinogenic properties of the emerging pollutant 6PPDQ, this multi-omics integrative analysis aims to address critical knowledge gaps regarding environmental pollutant-breast cancer associations, thereby providing scientific evidence supporting development of more precise preventive strategies.

2. Materials and Methods

2.1 Data Acquisition and Preprocessing

Bulk RNA sequencing data from three independent breast cancer cohorts (GSE3744, GSE15852, and GSE21422; total $n = 152$) were retrieved from the NCBI Gene Expression Omnibus (GEO) database. Batch effect correction was performed using the sva R package (version 3.44.0) prior to dataset integration. An additional transcriptome dataset (GSE52194; $n = 20$) was designated as an external validation cohort. All Affymetrix microarray data underwent $\log_2$ transformation before merging to ensure dimensional consistency. Batch annotations rigorously documented sample collection timing and sequencing batch information. Ensembl gene identifiers (version 104) served as the primary anchor for cross-platform integration.

Toxicological profiling of 6PPDQ was conducted using ProTox 3.0 (Prediction Of Toxicity Of Chemicals; <https://tox.charite.de/>), with detailed results provided in Supplementary File 1. Putative human protein targets of 6PPDQ were retrieved from PharmMapper (<http://www.lilab-ecust.cn/pharmmapper/>) and SwissTargetPrediction (<http://www.swisstarget-prediction.ch/>) databases.

2.2 Identification of 6PPDQ-Breast Cancer Cross-Targets

Differential expression analysis was executed using the limma package (version 3.52.4), applying thresholds of $|\log_2 \text{fold-change}| > 0.585$ and false discovery rate (FDR) < 0.05 . Weighted gene co-expression network analysis (WGCNA, version 1.72) was applied to the merged training dataset to identify biologically meaningful co-expression modules. Optimal soft-thresholding power $\beta = 3$ was

selected (scale-free topology $R^2 = 0.80$), and the topological overlap matrix (TOM) was calculated with minimum module size set at 80 genes. Module-trait associations were evaluated through Pearson correlation analysis, with genes satisfying $|\text{module membership}| > 0.6$ and $|\text{gene significance}| > 0.6$ criteria designated as hub genes.

2.3 Functional Enrichment and Protein-Protein Interaction Network Analysis

To investigate potential molecular mechanisms underlying 6PPDQ-induced breast cancer development, functional enrichment and protein-protein interaction (PPI) network analyses were performed on candidate target genes.

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were conducted utilizing the DAVID database (<https://david.ncifcrf.gov>). GO enrichment encompassed three ontological categories: Biological Process (BP), Cellular Component (CC), and Molecular Function (MF). Fisher's exact test assessed enrichment significance of target genes across GO terms and KEGG pathways, with $P < 0.05$ serving as the significance threshold for identifying functionally enriched biological processes and signaling pathways. Protein-protein interaction relationships among candidate targets were extracted from the STRING (<https://string-db.org/>). Retrieved interaction data were imported into Cytoscape software for network visualization, facilitating interpretation of inter-target relationships and underlying regulatory mechanisms.

2.4 Machine Learning-Based Core Target Prioritization

To identify the most influential genes from the 36-gene signature, an ensemble machine learning approach was implemented incorporating least absolute shrinkage and selection operator (LASSO) regression and support vector machine recursive feature elimination (SVM-RFE). Genes consistently retained by both algorithms were designated as core targets.

2.5 Core Target Pathway Mapping

Classical signal transduction cascades involving the final gene sets were analyzed through the KEGG pathway database (<https://www.kegg.jp/kegg/kegg2.html>) and graphically mapped onto breast cancer-specific pathway diagrams to generate mechanistic hypotheses.

2.6 Molecular Docking Simulations

Three-dimensional structural data for target proteins and small molecule compounds were retrieved from PubChem (<https://pubchem.ncbi.nlm.nih.gov/>) and RCSB Protein Data Bank (<https://www.rcsb.org/>), respectively. Molecular docking was performed using CB-Dock2 software (<https://cadd.labshare.cn/cb-dock2/index.php>) with grid dimensions of $100 \times 100 \times 100$, 50 independent docking runs, and maximum iteration count of 20,000.

3. Results

3.1 Differentially Expressed Gene Screening

Batch correction was applied to included breast cancer samples and control group sequencing data. Principal component analysis results before and after batch correction are presented in Figs. 1A and 1B. Differential expression analysis between breast cancer and normal control groups was subsequently conducted using the limma package, applying filtering criteria of $|\log_2FC| > 0.585$ and $FDR < 0.05$. Results demonstrated that compared to the normal control group, 594 genes exhibited significant differential expression in the breast cancer group, comprising 282 upregulated and 312 downregulated genes (Fig. 1C).

To identify gene modules closely associated with breast cancer pathogenesis, weighted gene co-expression network analysis (WGCNA) was performed on training dataset. The pickSoftThreshold function initially evaluated a range of soft-thresholding powers β . Based on scale-free topology fit index assessment, optimal soft-thresholding power $\beta = 3$ was selected (Fig. 1D) for weighted adjacency matrix construction, subsequently transformed into the topological overlap matrix (TOM) to quantify inter-gene co-expression similarity. Hierarchical clustering analysis was then performed on the TOM dissimilarity matrix, incorporating dynamic tree cutting algorithm with minimum module gene count of 60, deepSplit parameter of 3, and module merging dissimilarity threshold of 0.25. This analysis identified nine distinct gene co-expression modules, with module-gene dendrogram correspondence illustrated in Fig. 1E. Notably, the turquoise module displayed strong positive disease correlation, yielding 2,129 hub genes highly associated with breast cancer (Fig. 1F). These hub genes demonstrated elevated connectivity within their respective modules and significant phenotype associations, suggesting important roles in breast cancer pathogenesis.

The GeneCards database was subsequently queried to curate breast cancer-associated targets, identifying 639 breast cancer-related genes. Venn diagram intersection analysis was performed among the 2,129 WGCNA hub genes, 594 differentially expressed genes, and 639 GeneCards genes. Genes appearing in at least two screening approaches were defined as breast cancer-related targets, ultimately identifying 529 disease-associated genes meeting these criteria (Fig. 1G).

(A, B) Principal component analysis plots of included breast cancer samples before and after batch correction; (C) Heatmap of differential expression analysis results for breast cancer mRNA sequencing data following batch correction; (D) Soft-thresholding selection: left panel displays scale-free topology fit index R^2 , right panel displays average connectivity, with red line indicating $R^2 = 0.8$ threshold; (E) Gene clustering dendrogram: colored bars represent distinct co-expression modules; (F) Module-trait correlation heatmap: values indicate correlation coefficients, parenthetical numbers indicate P values, color intensity represents correlation strength; (G) Venn diagram: illustrating intersection of DEGs, GeneCards targets, and WGCNA hub genes, with genes appearing in at least two screening methods defined as breast cancer-related targets, ultimately yielding 529 core candidate targets.

3.2 Cross-Target Gene Confirmation

To systematically identify 6PPDQ molecular targets, human genes associated with 6PPDQ were retrieved from both PharmMapper and SwissTargetPrediction databases. Following data integration and duplicate removal, 258 potential 6PPDQ-related target genes were obtained (Fig. 2A). Intersection with breast cancer-related genes subsequently identified 36 cross-target genes common to both sets (Fig. 2B).

(A) Screening of 6PPDQ-associated targets; (B) Intersection analysis between breast cancer-related genes and 6PPDQ targets, identifying 36 cross-target genes.

3.3 Cross-Target Acquisition

GO and KEGG functional enrichment analyses were performed on the 36 cross-targets using the DAVID database to investigate specific molecular mechanisms involved in 6PPDQ-induced breast cancer. GO analysis revealed predominant involvement in biological processes including regulation of lipid metabolic process, cellular ketone metabolic process, and hormone metabolic process; cellular component localization in pronucleus; and molecular functions encompassing carboxylic acid binding, organic acid binding, and monocarboxylic acid binding (Fig. 3A). KEGG enrichment analysis demonstrated significant pathway involvement in PI3K-Akt signaling, breast cancer, transcriptional misregulation in cancer, AMPK signaling, PPAR signaling, IL-17 signaling, xenobiotic metabolism by cytochrome P450, and fatty acid degradation (Fig. 3B).

(A) GO functional enrichment results for 36 cross-targets; (B) KEGG pathway enrichment results for 36 cross-targets.

3.4 Diagnostic Model Construction

Based on the 36 cross-targets, 113 machine learning methods and their combinations were employed to train diagnostic predictive models. The batch-corrected integrated dataset served as the training cohort, while three pre-correction breast cancer sequencing datasets (GSE3744, GSE15852, and GSE21422) functioned as independent validation sets. Additionally, a fourth sequencing dataset (GSE52194) was retrieved from GEO as an independent test cohort. As shown in Fig. 4A, the "Lasso+Stepglm[forward]" algorithm achieved optimal diagnostic performance, maintaining high efficacy across both validation and test sets. Model evaluation through confusion matrices corroborated strong diagnostic performance (Figs. 4F-I). Expression analysis of modeling genes under the "Lasso+Stepglm[forward]" algorithm is presented in Fig. 4J. Individual gene diagnostic performance for breast cancer is illustrated in Fig. 4K, demonstrating consistently high diagnostic efficacy across modeling genes.

(A) Performance heatmap of diagnostic models across training, validation, and test cohorts under various algorithm combinations; (B-E) Diagnostic ROC curves for validation and test cohorts under Lasso+Stepglm[forward] algorithm; (F-I) Confusion matrices for validation and test cohorts under Lasso+Stepglm[forward] algorithm; (J) Box plots of differential expression analysis for breast cancer modeling genes; (K) Individual gene diagnostic ROC curves for breast cancer prediction.

3.5 SHapley Additive exPlanations Analysis

SHapley Additive exPlanations (SHAP) represents a machine learning interpretability framework grounded in game theory Shapley values. This approach calculates marginal contributions of individual features to model predictions, ensuring equitable attribution allocation and additivity while providing prediction-level interpretability. To enhance risk prediction model interpretability and reliability, SHAP analysis was applied to modeling features. Feature importance rankings based on SHAP values for "Control" and "Treatment" groups are presented in Fig. 5A. The breast cancer prediction model constructed from 15 modeling genes demonstrated robust performance across five algorithms, with Random Forest achieving optimal AUC (0.929; 95% CI 0.825–1.000) (Fig. 5B). The SHAP summary plot (Fig. 5C) indicated that RBP4 contributed highest explanatory power (mean |SHAP|=0.119), followed by ADH1C (0.069), with remaining genes showing progressively decreasing contributions. The waterfall plot (Fig. 5D) illustrated how modeling genes collectively shifted individual sample predictions from baseline (0.647) to final values (1.0), with elevated RBP4 expression contributing most substantially to increased prediction probability, followed by ADH1C, thereby validating modeling gene predictive reliability.

(A) Grouped SHAP feature importance bar chart for modeling genes in breast cancer; (B) Validation set ROC curves comparing multiple algorithms for systematic evaluation of machine learning model discriminative performance; (C) SHAP summary plot displaying modeling gene contributions; (D) SHAP force plot illustrating progression from baseline 0.674 to final sample prediction ($f(x) = 1$) through sequential feature contributions.

3.6 Core Target Screening

To further delineate core targets mediating 6PPDQ action in breast cancer, a protein-protein interaction (PPI) network of 36 cross-targets was constructed (detailed in Figure S1). This network was imported into Cytoscape, and four topological algorithms (MNC, MCC, Degree, and EPC) from the CytoHubba plugin were applied to extract top-ranked genes from each approach (Figs. 6A-D). Intersection of multi-dimensional topological screening identified 7 core proteins (Fig. 6E).

LASSO algorithm screening subsequently identified 19 core targets (Fig. 6F). SVM-RFE analysis retained 22 variables at peak model performance, yielding 22 feature targets (Fig. 6G). Intersection of genes from both machine learning approaches identified 13 core targets (Fig. 6H). Final intersection between topological algorithm-derived core proteins and machine learning-derived core targets yielded 2 definitive core targets: ESR1 and HMGCR (Fig. 6I).

(A-D) Top 10 genes from DMNC, MCC, Degree, and EPC algorithms, respectively, with node colors transitioning from red to yellow indicating decreasing scores; (E) Venn diagram of four algorithm results identifying 7 consensus core targets; (F) Lasso regression-based core target screening; (G) Support vector machine-based core target screening; (H) Intersection of two machine learning approaches yielding 13 core targets; (I) Final intersection identifying 2 core targets.

3.7 Risk Prediction Model Construction and Validation

Using two screened 6PPDQ-induced breast cancer core signature genes (ESR1 and HMGCR), a risk prediction nomogram was constructed. Analysis revealed HMGCR upregulation as a breast cancer risk factor, while ESR1 downregulation similarly represented a risk factor (Fig. 7A). Nomogram validation through calibration curves demonstrated strong agreement between predicted and observed probabilities, indicating excellent predictive performance (Fig. 7B). Decision curve analysis (DCA) assessed clinical net benefit, revealing moderate predictive efficacy (Fig. 7C). Clinical impact curves showed close correspondence between predicted and actual curves, further supporting favorable clinical utility (Fig. 7D).

(A) Nomogram based on 2 core cross-target genes; (B) Calibration curve; (C) Decision curve; (D) Clinical impact curve.

3.8 Molecular Docking Prediction of 6PPDQ-Core Target Interactions

To validate direct binding between 6PPDQ and breast cancer core targets, molecular docking simulations were performed for 6PPDQ with ESR1 and HMGCR. 6PPDQ demonstrated favorable binding affinity for both targets, with calculated binding energies of -7.7 kcal/mol (ESR1) and -7.9 kcal/mol (HMGCR), respectively (absolute values exceeding 7 kcal/mol), suggesting stable complex formation. The 6PPDQ-ESR1 docking pose localized to a pocket comprising residues MET343, LEU346, THR347, ASN348, LEU349, ALA350, ASP351, GLU353, LEU354, TRP383, LEU384, LEU387, MET388, LEU391, ARG394, PHE404, MET421, ILE424, GLY521, MET522, HIS524, LEU525, TYR526, MET528, LYS529, CYS530, VAL533, VAL534, PRO535, LEU536, and LEU539 (Fig. 8A). The 6PPDQ-HMGCR docking pose occupied a region containing residues ALA525, ARG590, ALA654, MET655, GLY656, MET657, ASN658, MET659, SER661, LYS662, GLU665, VAL683, SER684, ASP690, LYS691, LYS692, GLY765, GLN766, ASP767, GLY803, THR804, VAL805, GLY806, GLY807, GLY808, and THR809 (Fig. 8B), for detailed molecular docking information, please refer to Table S1. These findings provide preliminary molecular-level validation of direct 6PPDQ-core target protein interactions.

(A) 6PPDQ-ESR1 molecular docking; (B) 6PPDQ-HMGCR molecular docking.

3.9 Core Target Signaling Pathway Investigation

The KEGG PATHWAY Database (<https://www.kegg.jp/kegg/kegg2.html>) was utilized to explore signaling pathways involving the 2 core differentially expressed genes in cancer contexts. Results positioned HMGCR downstream of the AMPK signaling pathway, participating in cholesterol synthesis inhibition, with potential regulatory roles in 6PPDQ-induced breast cancer. The proposed regulatory mechanism is illustrated in Fig. 9.

4. Discussion

This investigation integrates network toxicology, bulkRNA-seq, and molecular docking to systematically elucidate molecular associations between the environmental pollutant 6PPDQ and breast cancer

development and progression. Core findings establish ESR1 and HMGCR as pivotal mediators, revealing their promotion of malignant progression through multi-dimensional mechanisms encompassing metabolic reprogramming, hormonal signaling disruption, and tumor microenvironment remodeling.

Multi-algorithm topological screening identified ESR1 as the primary core target in 6PPDQ-induced breast cancer. Molecular docking revealed a binding energy of -7.7 kcal/mol, indicating stable molecular interaction. ESR1 encodes estrogen receptor α (ER α), a critical transcription factor governing breast tissue development and hormonal responsiveness[17, 18]. Epidemiological evidence demonstrates that environmental endocrine disruptors can interfere with ER α signaling homeostasis through estrogen mimicking or antagonism, thereby elevating breast cancer susceptibility[19]. Notably, as a quinone-containing compound, 6PPDQ's benzoquinone moiety generates reactive oxygen species (ROS) through redox cycling[20, 21], and ROS has been established to induce aberrant ER α pathway activation or receptor degradation[21].

Recent investigations demonstrate that 6PPDQ and its precursor 6PPD activate estrogen-related receptor γ (ERR γ), inducing hepatic lipid accumulation through ERR γ signaling. Although ERR γ and ER α differ in ligand-binding domains, they share co-activators and downstream targets[22], suggesting 6PPDQ may disrupt estrogen signaling networks through non-canonical pathways. The differential ESR1 expression pattern in breast cancer tissue, combined with molecular docking results, supports the hypothesis that 6PPDQ may directly bind ER α or indirectly modify receptor function through oxidative stress, consequently disrupting hormone-sensitive regulation of breast epithelial cells. This mechanism parallels previously reported patterns where environmental pollutants including polychlorinated biphenyls (PCBs) and phthalates promote breast cancer through ER α pathway modulation[23, 24].

HMGCR, the second core target identified, serves as the rate-limiting enzyme of the mevalonate (MVA) pathway, governing cellular de novo cholesterol synthesis. Cholesterol metabolic reprogramming has emerged as a critical metabolic hallmark of breast cancer malignant progression[25, 26]. Tumor cells satisfy rapid proliferation requirements for membrane biosynthesis and signaling molecules (steroid hormones, isoprenylated modifiers) through HMGCR upregulation. Molecular docking demonstrated a -7.9 kcal/mol binding energy between 6PPDQ and HMGCR, suggesting potential direct binding or allosteric modulation of HMGCR enzymatic activity.

Cholesterol metabolism intimately connects with breast cancer hormone dependence[27]. Cholesterol serves as an estrogen synthesis precursor, and HMGCR-mediated cholesterol synthesis enhancement can indirectly promote local tumor estrogen production, establishing autocrine growth stimulation loops[28, 29]. Furthermore, elevated HMGCR expression correlates with breast cancer stem cell (BCSC) maintenance[30]. Clendening et al. demonstrated that HMGCR overexpression induces stem cell-like phenotypes in breast cancer cells, enhancing tumorigenic capacity and therapeutic resistance[30]. AMPK pathway enrichment in this investigation supports this mechanism: AMPK, functioning as an energy metabolism sensor, phosphorylates and inhibits HMGCR activity, thereby suppressing cholesterol synthesis[31]; 6PPDQ potentially interferes with AMPK-HMGCR regulatory axes, relieving metabolic

suppression and promoting tumor cell proliferation[32]. Additionally, the demonstrated therapeutic potential of statins (HMGCR inhibitors) in breast cancer adjuvant therapy corroborates the viability of HMGCR-targeted anti-tumor strategies.

KEGG enrichment analysis revealed significant enrichment of 6PPDQ-breast cancer cross-targets in PI3K-Akt and IL-17 signaling pathways. The PI3K-Akt/mTOR cascade represents a central regulator of cell survival, proliferation, and metabolism, with aberrant activation present in approximately 40% of breast cancers[33]. Environmental pollutants including PCBs and benzo[a]pyrene promote breast cancer cell proliferation and chemoresistance through PI3K-Akt pathway activation[34, 35]. This investigation suggests 6PPDQ may activate PI3K-Akt through ROS-dependent mechanisms: quinone redox activity induces oxidative stress, potentially inhibiting PTEN phosphatase activity or activating receptor tyrosine kinases, resulting in sustained Akt phosphorylation[36]. IL-17 pathway enrichment suggests 6PPDQ may promote breast cancer through tumor immune microenvironment modulation. IL-17, primarily secreted by Th17 and $\gamma\delta$ T cells[36], activates NF- κ B and STAT3 transcription factors through autocrine or paracrine mechanisms, promoting proliferation and anti-apoptotic responses[37]. Research demonstrates that IL-17 directly stimulates breast cancer cell proliferation and enhances invasiveness through epithelial-mesenchymal transition (EMT) induction[38, 39]. 6PPDQ-induced oxidative stress and inflammatory responses may recruit IL-17-producing cells to infiltrate tumor tissues, establishing pro-tumorigenic inflammatory microenvironments. Moreover, IL-17 and PI3K-Akt pathways exhibit positive feedback regulation: IL-17 receptor activation upregulates PI3K expression, while Akt activation enhances IL-17 transcriptional activity, creating signal amplification effects[40].

The "metabolic reprogramming-hormonal dysregulation-immune modulation" triad mechanism revealed in this investigation reflects the complexity of multi-target, multi-level network perturbation through which 6PPDQ promotes breast cancer. HMGCR-mediated cholesterol synthesis enhancement satisfies tumor cell membrane requirements while potentially disrupting endocrine balance through steroid hormone synthesis. Additionally, cholesterol metabolites (e.g., 27-hydroxycholesterol) possess immunomodulatory functions capable of suppressing anti-tumor immune responses. Conversely, the AhR/cyp1b1 pathway activated by 6PPDQ (confirmed in zebrafish models) exhibits dual roles in breast cancer: AhR ligands such as dioxins antagonize ER α to inhibit tumor growth, whereas most environmental pollutants (polycyclic aromatic hydrocarbons, phthalates) promote cell migration and angiogenesis through AhR activation.

Notably, human 6PPDQ exposure occurs through diverse pathways, including inhalation of tire wear particle-containing air, consumption of contaminated water, and food chain bioaccumulation. Recent studies have detected 6PPDQ in human blood, urine, and cerebrospinal fluid, confirming systemic bioavailability[41]. Although direct epidemiological evidence linking 6PPDQ to human breast cancer remains absent, molecular mechanisms predicted through computational toxicology in this investigation, combined with demonstrated toxicological characteristics including DNA damage, oxidative stress, and endocrine disruption induction, suggest that chronic low-dose exposure may elevate breast cancer risk through cumulative effects.

5. Limitations

This study has several limitations that should be considered. First, the analysis relied on currently available public RNA-seq datasets, with a combined sample size of 152 cases (for training/validation) plus 20 for independent testing. While this provides initial insights, the findings require validation in larger, more diverse cohorts. Second, the mechanistic links identified between 6PPDQ and breast cancer (e.g., targeting of ESR1 and HMGCR) are primarily supported by bioinformatics predictions and molecular docking simulations. These proposed interactions and regulatory pathways warrant direct experimental validation in breast cancer cell lines and animal models to establish causality. Third, the dose-response relationship of 6PPDQ exposure and its metabolic activation in biological systems remain to be elucidated through controlled in vitro/in vivo studies and large-scale epidemiological investigations. Finally, although the diagnostic model based on the identified targets shows promising performance, its clinical utility and generalizability need to be prospectively validated in independent patient populations.

6. Conclusions

This investigation employed network toxicology integrated with multi-omics analysis to systematically elucidate, for the first time, molecular associations between 6PPDQ, a transformation product of tire rubber antioxidants, and breast cancer development and progression. ESR1 and HMGCR were identified as core targets mediating 6PPDQ-induced breast cancer, with findings revealing malignant progression promotion through a cascade mechanism of "hormonal signaling disruption-cholesterol metabolic reprogramming-tumor microenvironment remodeling". Molecular docking confirmed strong binding affinities between 6PPDQ and both targets, providing structural foundations for direct interaction. Functional enrichment analysis indicated involvement of PI3K-Akt, AMPK, and IL-17 signaling pathways in 6PPDQ carcinogenic effects, constituting a complex regulatory network.

Declarations

Clinical trial number

Not applicable.

Ethics, Consent to Participate, and Consent to Publish declarations

Not applicable.

Funding

Not yet determined.

Author Contribution

LH and LX contributed equally. LH, LX, HJ and QH participated in the conception and design of the study. LH and LX organized the database and statistical analysis. LH, LX and HJ divided the work and participated in the picture drawing. LH and LX wrote the first draft of the manuscript. QH participated in the revision of the manuscript. All authors read and agreed to the final manuscript and authorship arrangement.

Acknowledgement

LH and LX contributed equally. LH, LX, HJ and QH participated in the conception and design of the study. LH and LX organized the database and statistical analysis. LH, LX and HJ divided the work and participated in the picture drawing. LH and LX wrote the first draft of the manuscript. QH participated in the revision of the manuscript. All authors read and agreed to the final manuscript and authorship arrangement.

Data Availability

The data used in this study were obtained from the GEO ([https:// www.ncbi.nlm.nih.gov/geo/](https://www.ncbi.nlm.nih.gov/geo/)) database, both of which are available in publicly available databases. This study complies with its data use and publication rules.

References

1. Menon G, Alkabban FM, Ferguson T. Breast Cancer. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing; 2025.
2. Valente S, Roesch E. Breast cancer survivorship. J SURG ONCOL. 2024;130(1):8–15.
3. Traves KP, Cokenakes SEH. Breast Cancer Treatment. AM FAM PHYSICIAN. 2021;104(2):171–8.
4. Pesapane F, Nicosia L, Cassano E. Updates on Breast Cancer. CANCERS. 2023;15(22):5392.
5. Zhao HN, Thomas SP, Zylka MJ, Dorrestein PC, Hu W. Urine Excretion, Organ Distribution, and Placental Transfer of 6PPD and 6PPD-Quinone in Mice and Potential Developmental Toxicity through Nuclear Receptor Pathways. ENVIRON SCI TECHNOL. 2023;57(36):13429–38.
6. Li G, Dan L, Zhai W, Liu C, Chen M, Xu Q, Huang Y. Respiratory toxicity mechanism of 6PPD and 6PPD-quinone: An integrated study based on network toxicology and molecular docking. ECOTOX ENVIRON SAFE. 2025;301:118494.
7. Rudolph A, Chang-Claude J, Schmidt MK. Gene-environment interaction and risk of breast cancer. BRIT J CANCER. 2016;114(2):125–33.

8. Santaliz Casiano A, Lee A, Teteh D, Madak Erdogan Z, Treviño L. Endocrine-Disrupting Chemicals and Breast Cancer: Disparities in Exposure and Importance of Research Inclusivity. *Endocrinology*. 2022;163(5):bqac34.
9. White AJ, Chen J, McCullough LE, Xu X, Cho YH, Teitelbaum SL, Neugut AI, Terry MB, Hibshoosh H, Santella RM, et al. Polycyclic aromatic hydrocarbon (PAH)-DNA adducts and breast cancer: modification by gene promoter methylation in a population-based study. *Cancer causes control: CCC*. 2015;26(12):1791–802.
10. Prosser RS, Salole J, Hang S. Toxicity of 6PPD-quinone to four freshwater invertebrate species. *Environmental pollution (Barking, Essex: 1987)* 2023, 337:122512.
11. Li C, Zhang Y, Yin S, Wang Q, Li Y, Liu Q, Liu L, Luo X, Chen L, Zheng H, et al. First insights into 6PPD-quinone formation from 6PPD photodegradation in water environment. *J HAZARD MATER*. 2023;459:132127.
12. Deng M, Ji X, Peng B, Fang M. In Vitro and In Vivo Biotransformation Profiling of 6PPD-Quinone toward Their Detection in Human Urine. *ENVIRON SCI TECHNOL*. 2024;58(21):9113–24.
13. Yi J, Ruan J, Yu H, Wu B, Zhao J, Wang H, Chen R, Yang Q, Chen J, Sun D. Environmental fate, toxicity, and mitigation of 6PPD and 6PPD-Quinone: Current understanding and future directions. *Environmental pollution (Barking, Essex: 1987)* 2025, 375:126352.
14. Qi Y, Qiu A, Wei X, Huang Y, Huang Q, Huang W. Effects of 6PPD-Quinone on Human Liver Cell Lines as Revealed with Cell Viability Assay and Metabolomics Analysis. *TOXICS* 2024, 12(6):389.
15. Zhang W, Wu B, Hu C, Rong W, Huang Y, Hu S, Qin Y. EGFR and CYP signaling disruption underlies 6PPD-quinone hepatotoxicity: Insights from a network and machine learning approach. *Toxicology*. 2025;517:154235.
16. Zhou H, Wu Z, Wang X, Jiang L, Sun H, Li H, Yan Z, Wang Y, Yao X, Zhang C, et al. 6PPD-quinone exposure induces oxidative damage and physiological disruption in *Eisenia fetida*: An integrated analysis of phenotypes, multi-omics, and intestinal microbiota. *J HAZARD MATER*. 2025;493:138334.
17. Bidard F, Mayer EL, Park YH, Janni W, Ma C, Cristofanilli M, Bianchini G, Kalinsky K, Iwata H, Chia S, et al. First-Line Camizestrant for Emerging ESR1-Mutated Advanced Breast Cancer. *N Engl J Med*. 2025;393(6):569–80.
18. Dustin D, Gu G, Fuqua SAW. ESR1 mutations in breast cancer. *CANCER-AM CANCER SOC*. 2019;125(21):3714–28.
19. Toy W, Shen Y, Won H, Green B, Sakr RA, Will M, Li Z, Gala K, Fanning S, King TA, et al. ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. *NAT GENET*. 2013;45(12):1439–45.
20. Li Y, Zeng J, Liang Y, Zhao Y, Zhang S, Chen Z, Zhang J, Shen X, Wang J, Zhang Y et al. A Review of N-(1,3-Dimethylbutyl)-N'-phenyl-p-Phenylenediamine (6PPD) and Its Derivative 6PPD-Quinone in the Environment. *TOXICS* 2024, 12(6):394.
21. Redman ZC, Begley JL, Hillestad I, DiMento BP, Stanton RS, Aguaa AR, Pirrung MC, Tomco PL. Reactive Oxygen Species and Chromophoric Dissolved Organic Matter Drive the Aquatic

- Photochemical Pathways and Photoproducts of 6PPD-quinone under Simulated High-Latitude Conditions. *ENVIRON SCI TECHNOL*. 2023;57(49):20813–21.
22. Zhang S, Su X, Zhang J, He T, Tang L, Zhao Z, Cao L. Tire rubber derivative 6PPD and 6PPD-Q induce lipid accumulation in hepatocytes through ER α pathway. *J Environ Sci*. 2025;156:173–84.
 23. Baker KM, Bauer AC. Green Tea Catechin, EGCG, Suppresses PCB 102-Induced Proliferation in Estrogen-Sensitive Breast Cancer Cells. *INT J BREAST CANCER* 2015, 2015:163591.
 24. Fiocchetti M, Bastari G, Cipolletti M, Leone S, Acconcia F, Marino M. The Peculiar Estrogenicity of Diethyl Phthalate: Modulation of Estrogen Receptor α Activities in the Proliferation of Breast Cancer Cells. *TOXICS* 2021, 9(10):237.
 25. Jiang X, Wang M, Cui G, Wu Y, Wei Z, Yu S, Wang A, Zou W, Pan Y, Li X, et al. Tetramethylpyrazine attenuates the cancer stem cell like-properties and doxorubicin resistance by targeting HMGCR in breast cancer. *Phytomedicine: Int J phytotherapy phytopharmacology*. 2025;136:156344.
 26. Yang Y, Liu Y, Zou T, Liu J, Zhou X, Tao R, Liu S. HMGCR: a malignancy hub - frontiers in cancer diagnosis and therapy. *FRONT ONCOL*. 2025;15:1698320.
 27. Murdock DJ, Sanchez RJ, Mohammadi KA, Fazio S, Geba GP. Serum cholesterol and the risk of developing hormonally driven cancers: A narrative review. *CANCER MED-US*. 2023;12(6):6722–67.
 28. Valko-Rokytovská M, Očenáš P, Salayová A, Kostecká Z. Breast Cancer: Targeting of Steroid Hormones in Cancerogenesis and Diagnostics. *Int J Mol Sci*. 2021;22(11):5878.
 29. Simigdala N, Gao Q, Pancholi S, Roberg-Larsen H, Zvelebil M, Ribas R, Folkard E, Thompson A, Bhamra A, Dowsett M, et al. Cholesterol biosynthesis pathway as a novel mechanism of resistance to estrogen deprivation in estrogen receptor-positive breast cancer. *Breast cancer research: BCR*. 2016;18(1):58.
 30. Marks MP, Giménez CA, Isaja L, Vera MB, Borzone FR, Pereyra-Bonnet F, Romorini L, Videla-Richardson GA, Chasseing NA, Calvo JC, et al. Role of hydroxymethylglutaryl-coenzyme A reductase in the induction of stem-like states in breast cancer. *J CANCER RES CLIN*. 2024;150(2):106.
 31. Teng Y. The Potential Therapeutic Applications of Natural Products in the Oxidative Stress-Related MVA Pathway: Focus on HMGCR. *Antioxid (Basel Switzerland)*. 2025;14(8):1001.
 32. Lee MKS, Cooney OD, Lin X, Nadarajah S, Dragoljevic D, Huynh K, Onda D, Galic S, Meikle PJ, Edlund T, et al. Defective AMPK regulation of cholesterol metabolism accelerates atherosclerosis by promoting HSPC mobilization and myelopoiesis. *MOL METAB*. 2022;61:101514.
 33. Miricescu D, Totan A, Stanescu-Spinu I, Badoiu SC, Stefani C, Greabu M. PI3K/AKT/mTOR Signaling Pathway in Breast Cancer: From Molecular Landscape to Clinical Aspects. *Int J Mol Sci*. 2020;22(1):173.
 34. Dzobo K, Hassen N, Senthebane DA, Thomford NE, Rowe A, Shipanga H, Wonkam A, Parker MI, Mowla S, Dandara C. Chemoresistance to Cancer Treatment: Benzo- α -Pyrene as Friend or Foe? *Molecules*. 2018;23(4):930.

35. Yang X, Liang W, Feng Z, Li G, Chen X, Zhang J. Molecular mechanisms of polychlorinated biphenyls in breast cancer: insights from network toxicology and molecular docking approaches. *FRONT PHARMACOL*. 2025;16:1604993.
36. Tiburcio D, Parsell M, Shapiro H, Adolphe S, Naranjo O, George S, Toborek M. Endocrine disruption to metastasis: How phthalates promote breast carcinogenesis. *ECOTOX ENVIRON SAFE*. 2025;303:118874.
37. Wang R, Yang L, Zhang C, Wang R, Zhang Z, He Q, Chen X, Zhang B, Qin Z, Wang L, et al. Th17 cell-derived IL-17A promoted tumor progression via STAT3/NF- κ B/Notch1 signaling in non-small cell lung cancer. *ONCOIMMUNOLOGY*. 2018;7(11):e1461303.
38. Song Y, Yang M, Zhang H, Sun Y, Tao Y, Li H, Zhang J, Li Y, Yang J. IL-17 Affects the Progression, Metastasis, and Recurrence of Laryngeal Cancer via the Inhibition of Apoptosis through Activation of the PI3K/AKT/FAS/FASL Pathways. *J IMMUNOL RES* 2020, 2020:2953191.
39. Huang Q, Han J, Fan J, Duan L, Guo M, Lv Z, Hu G, Chen L, Wu F, Tao X, et al. IL-17 induces EMT via Stat3 in lung adenocarcinoma. *AM J CANCER RES*. 2016;6(2):440–51.
40. Gu F, Li Q, Gao Q, Jiang J, Zhu K, Huang X, Pan J, Yan J, Hu J, Wang Z, et al. IL-17 induces AKT-dependent IL-6/JAK2/STAT3 activation and tumor progression in hepatocellular carcinoma. *MOL CANCER*. 2011;10:150.
41. Zhang S, Tang J, Qiu Z, Huo X, Liu D, Zeng X. Environmental and Human Health Risks of 6PPD and 6PPDQ: Assessment and Implications. *TOXICS* 2025, 13(10):873.

Figures

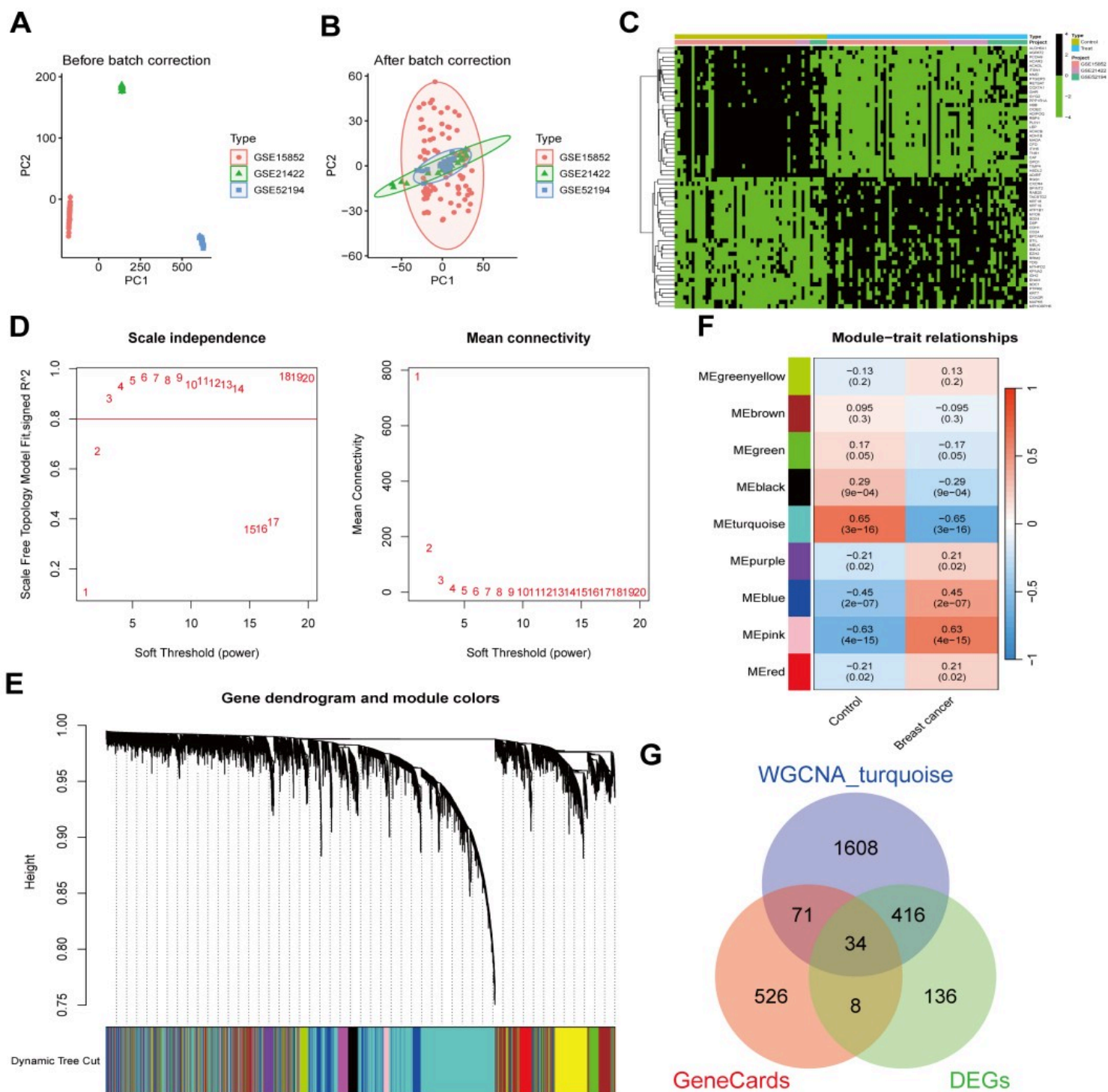


Figure 1

Screening of Breast Cancer Differentially Expressed Genes.

(A, B) Principal component analysis plots of included breast cancer samples before and after batch correction; (C) Heatmap of differential expression analysis results for breast cancer mRNA sequencing data following batch correction; (D) Soft-thresholding selection: left panel displays scale-free topology fit index R^2 , right panel displays average connectivity, with red line indicating $R^2 = 0.8$ threshold; (E) Gene clustering dendrogram: colored bars represent distinct co-expression modules; (F) Module-trait

correlation heatmap: values indicate correlation coefficients, parenthetical numbers indicate P values, color intensity represents correlation strength; (G) Venn diagram: illustrating intersection of DEGs, GeneCards targets, and WGCNA hub genes, with genes appearing in at least two screening methods defined as breast cancer-related targets, ultimately yielding 529 core candidate targets.

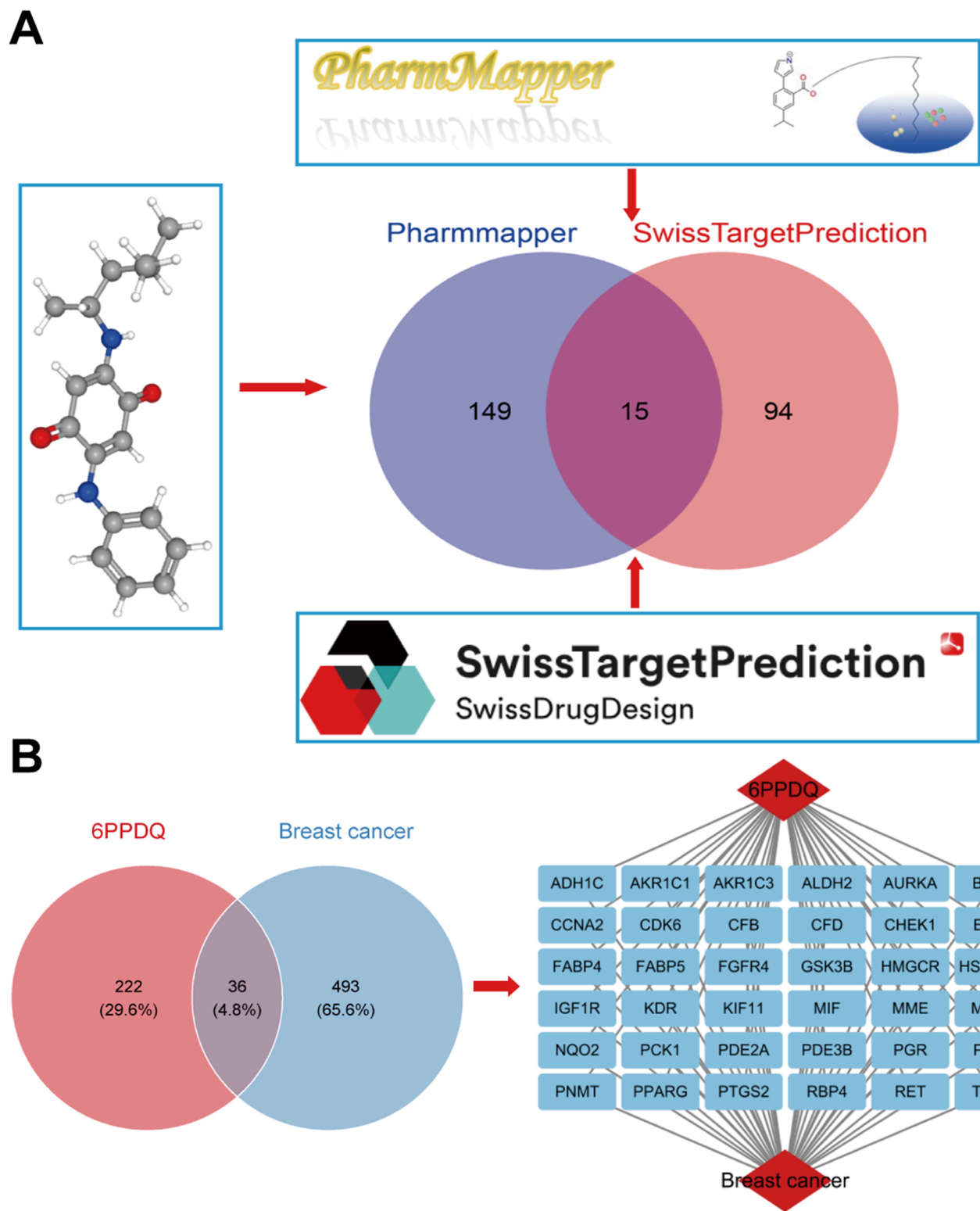


Figure 2

Confirmation of Cross-Target Genes.

(A) Screening of 6PPDQ-associated targets; (B) Intersection analysis between breast cancer-related genes and 6PPDQ targets, identifying 36 cross-target genes.

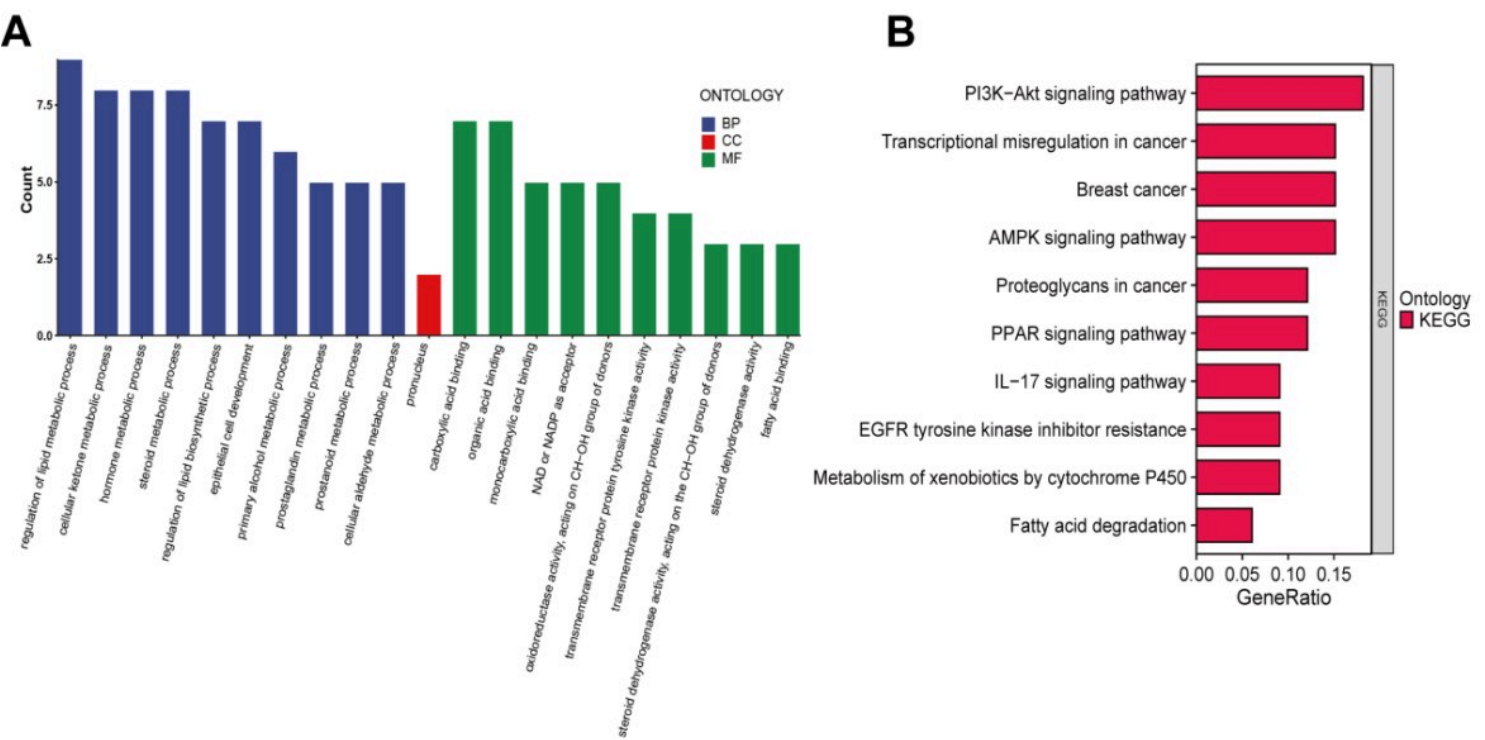


Figure 3

Enrichment Analysis.

(A) GO functional enrichment results for 36 cross-targets; (B) KEGG pathway enrichment results for 36 cross-targets.

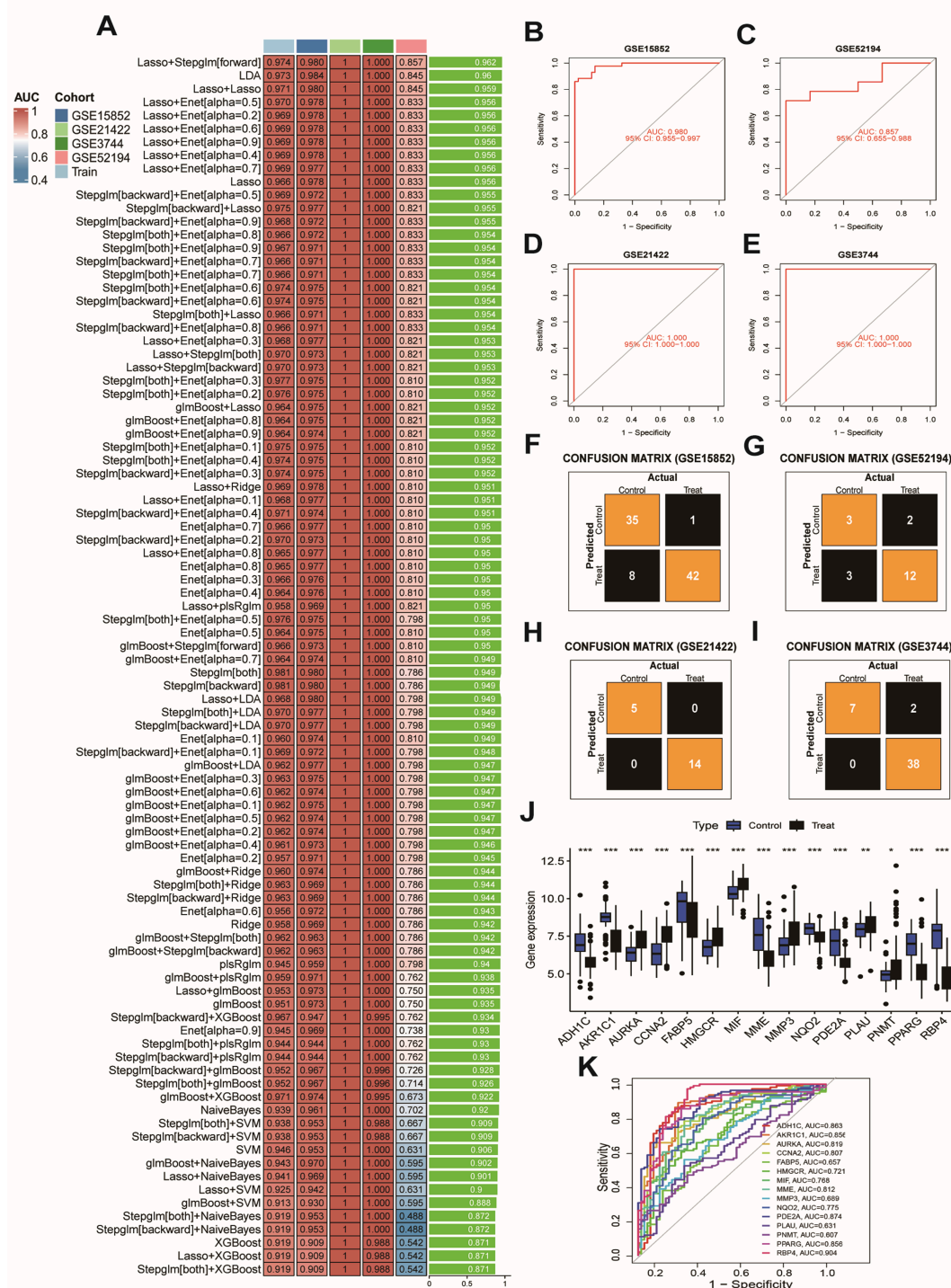


Figure 4

Diagnostic Model Based on 113 Machine Learning Screening Methods.

(A) Performance heatmap of diagnostic models across training, validation, and test cohorts under various algorithm combinations; (B-E) Diagnostic ROC curves for validation and test cohorts under Lasso+Stepglm[forward] algorithm; (F-I) Confusion matrices for validation and test cohorts under

Lasso+Stepglm[forward] algorithm; (J) Box plots of differential expression analysis for breast cancer modeling genes; (K) Individual gene diagnostic ROC curves for breast cancer prediction.

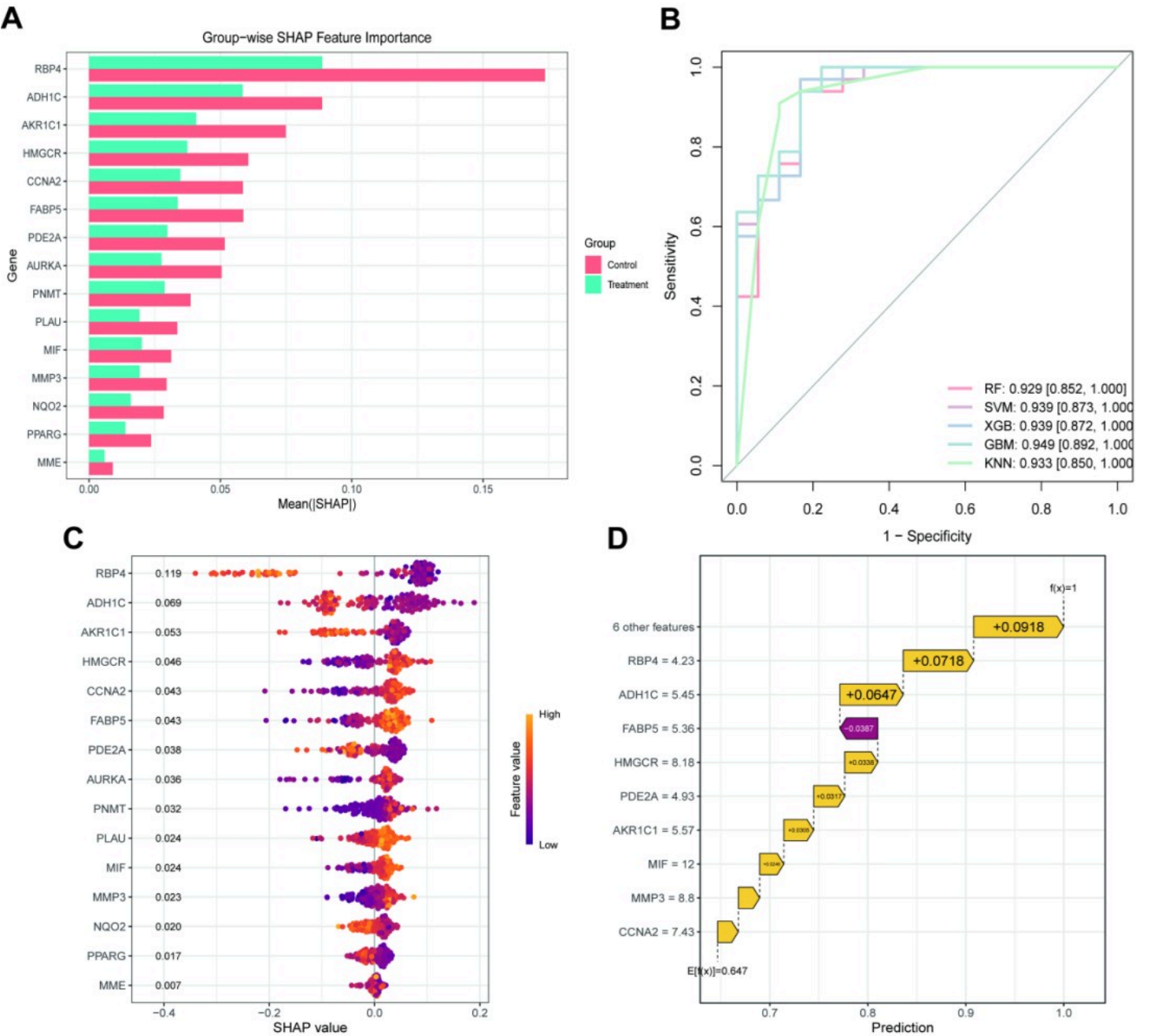


Figure 5

SHAP Analysis.

(A) Grouped SHAP feature importance bar chart for modeling genes in breast cancer; (B) Validation set ROC curves comparing multiple algorithms for systematic evaluation of machine learning model discriminative performance; (C) SHAP summary plot displaying modeling gene contributions; (D) SHAP force plot illustrating progression from baseline 0.674 to final sample prediction ($f(x) = 1$) through sequential feature contributions.

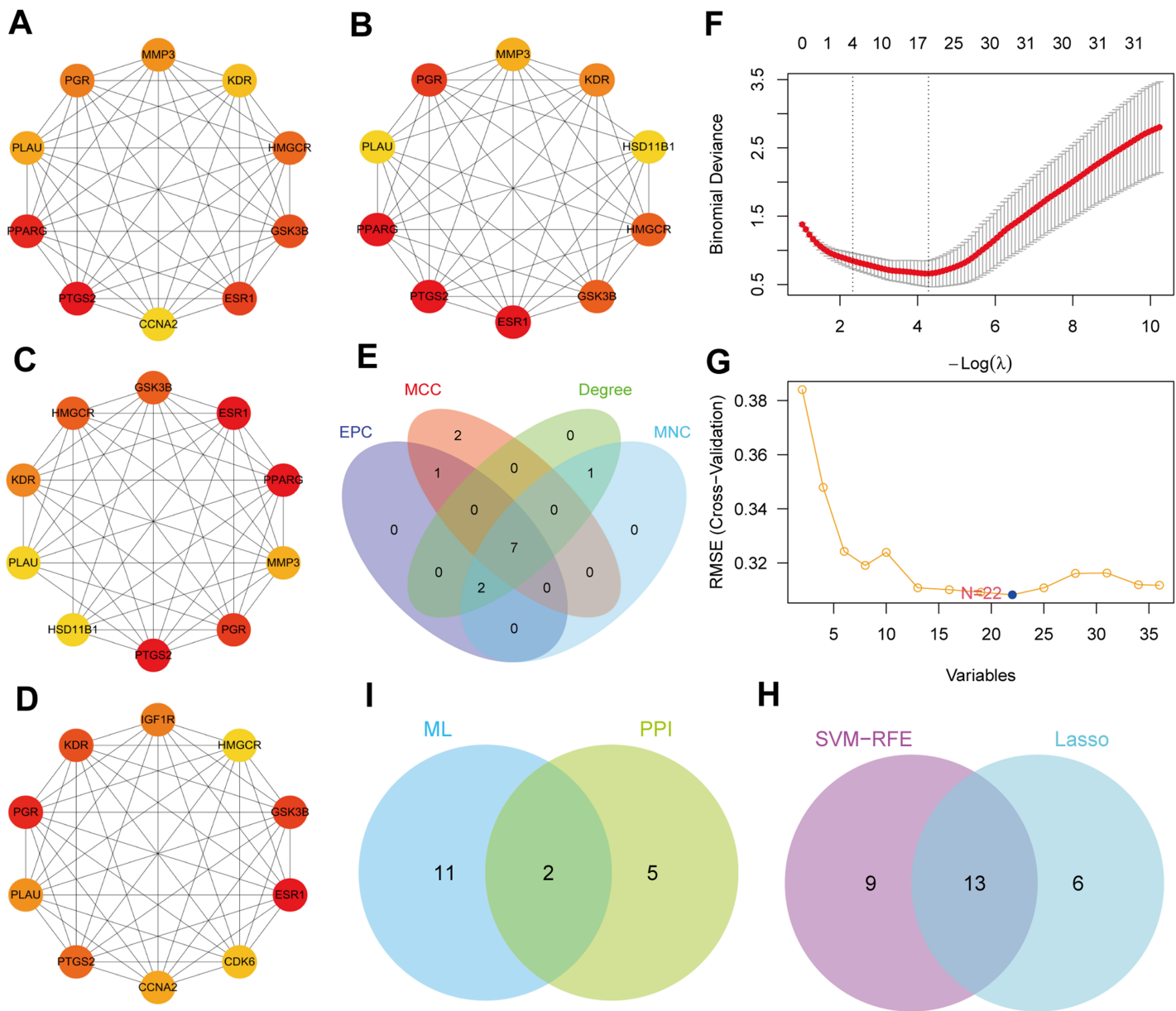


Figure 6

Screening of Key Target Genes.

(A-D) Top 10 genes from DMNC, MCC, Degree, and EPC algorithms, respectively, with node colors transitioning from red to yellow indicating decreasing scores; (E) Venn diagram of four algorithm results identifying 7 consensus core targets; (F) Lasso regression-based core target screening; (G) Support vector machine-based core target screening; (H) Intersection of two machine learning approaches yielding 13 core targets; (I) Final intersection identifying 2 core targets.

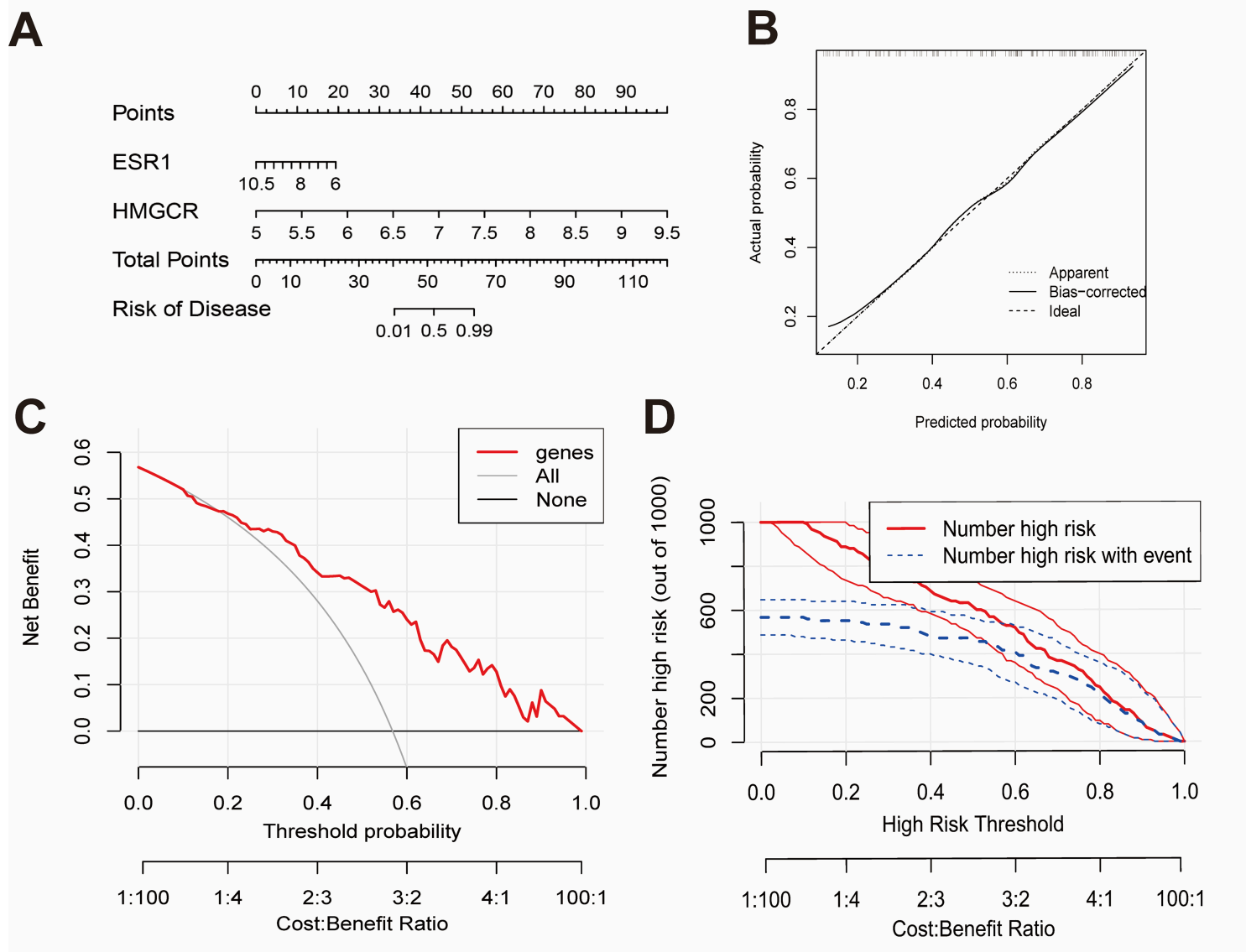


Figure 7

Risk Model Construction and Validation.

(A) Nomogram based on 2 core cross-target genes; (B) Calibration curve; (C) Decision curve; (D) Clinical impact curve.

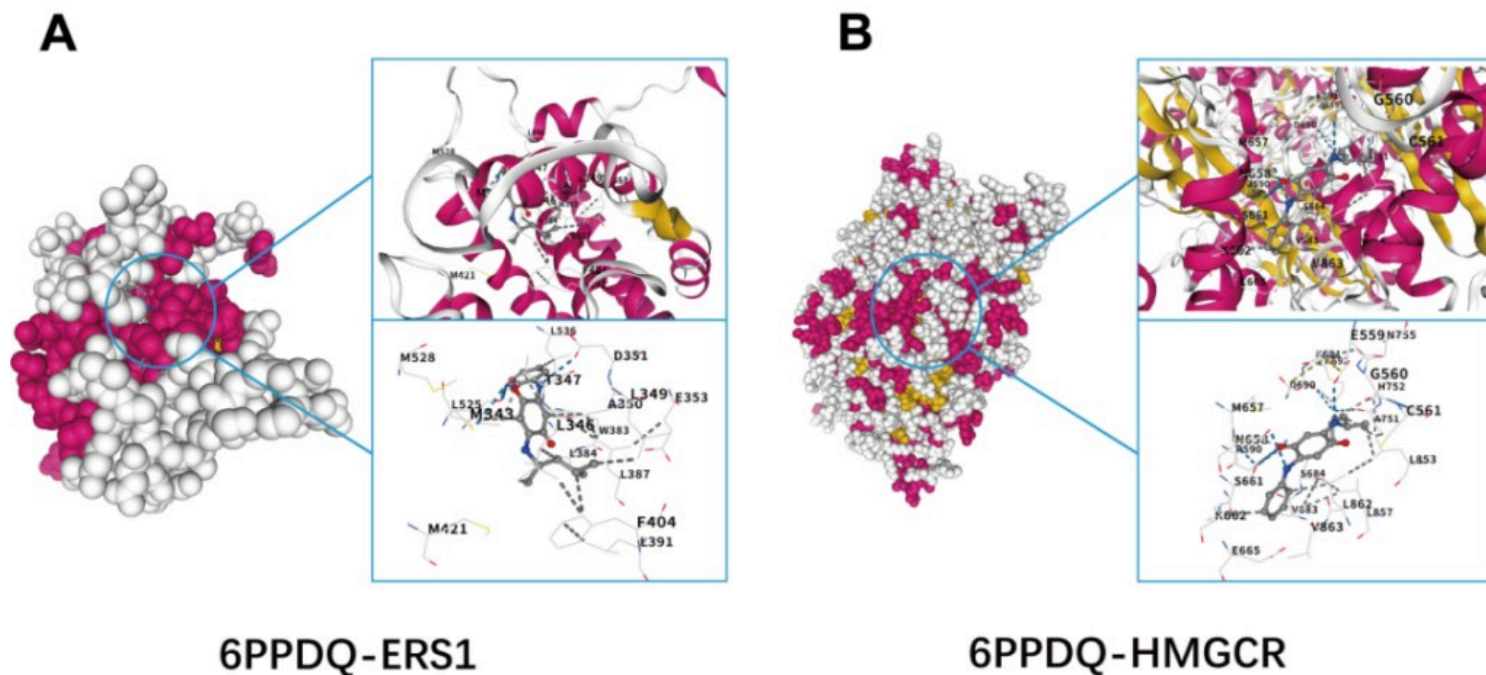


Figure 8

Molecular Docking Results.

(A) 6PPDQ-ERS1 molecular docking; (B) 6PPDQ-HMGCR molecular docking.

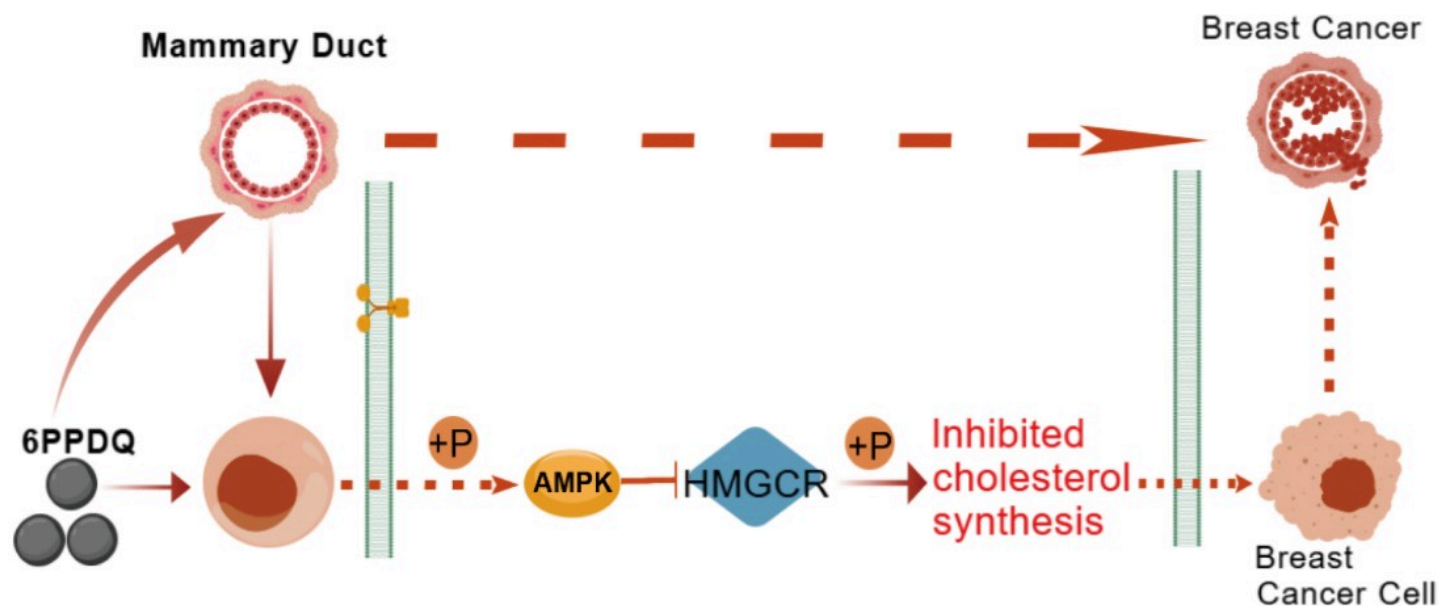


Figure 9

Proposed mechanism of 6PPDQ-mediated breast cancer through AMPK-HMGCR signaling.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1.docx](#)
- [SupplementaryFile1.pdf](#)
- [Supplementaryfigures..docx](#)
- [GA.jpg](#)