Current Cancer Research

https://ccr.cultechpub.com/ccr

Cultech Publishing

Article

Integrative Bioinformatics and Systems Biology Approach for the Identification of Key Genes Responses to Masoprocol Treatment in Breast Cancer

Mayank Roy Chowdhury*

Department of Biotechnology, National Institute of Technology, Tadepalligudem, Andhra Pradesh, India

*Corresponding author: Mayank Roy Chowdhury, mayank.sclr@nitandhra.ac.in

Abstract

This study presents a system-level investigation repositioning Masoprocol as a potential therapeutic candidate for breast cancer. Previously withdrawn due to limited mechanistic insights, Masoprocol was re-evaluated using integrative bioinformatics and systems biology to reveal its multi-target potential and pharmacological relevance. Absorption, distribution, metabolism, excretion, and toxicity (ADMET) profiling confirmed favourable drug-like characteristics, including Lipinski compliance, low predicted toxicity, and limited central nervous system penetration. Target prediction and enrichment analyses were performed, followed by construction of a protein-protein interaction network and a gene regulatory network. Masoprocol was found to interact with multiple breast cancer-associated genes, with TP53 emerging as a prominent regulatory hub across PPI and GRN layers. Transcriptomic profiling showed ESR1, MMP9, and AR were upregulated in tumors, whereas PTGS2 was downregulated. Elevated MMP9 and AR levels correlated with poor overall survival. Co-expression analysis linked Masoprocol's targets to oncogenic regulators involved in epithelial-mesenchymal transition and hormone signaling. Immune infiltration analysis suggested potential immunomodulatory effects. Gene ontology and Kyoto Encyclopedia of Genes and Genomes enrichment highlighted estrogen signaling and multiple cancer-associated pathways. Docking and molecular dynamics simulations with the Y220C TP53 mutant showed strong binding but moderate conformational stability, indicating a need for structural optimization. These findings support Masoprocol's repositioning as a polypharmacological breast cancer candidate and provide a roadmap for its future optimization and experimental validation.

Keywords

Masoprocol, Breast cancer, TP53, Mutation, Molecular docking, Protein-protein interactions, Pathway enrichment, Systems biology

Article History

Received: 10 November 2025 Revised: 17 December 2025

Accepted: 24 December 2025 Available Online: 31 December 2025

1. Introduction

Natural products have long served as a cornerstone in the development of therapeutics, particularly in oncology [1], where complex diseases such as cancer demand multi-targeted and systems-level treatment strategies [2]. Among natural sources, marine-derived actinomycetes have emerged as a prolific reservoir of bioactive compounds, exhibiting anticancer activities through mechanisms including apoptosis induction, cell-cycle arrest, and inhibition of metastasis [3]. Recent studies, including our own work on marine actinomycetes for Alzheimer's disease treatment [4,5], have highlighted their untapped therapeutic potential across multiple disease domains, including cancer.

Parallel to marine bioprospecting, traditional systems of medicine—particularly Ayurveda—have gained renewed interest due to their holistic, multi-component approach to disease modulation. Our previous research demonstrated the efficacy of Ayurvedic formulations such as *Bacopa monnieri*, *Hippophae rhamnoides*, and *Dioscorea bulbifera* in neurodegenerative disorders including Alzheimer's disease [6], and identified promising phytoconstituents relevant to Parkinson's disease and amyotrophic lateral sclerosis [7]. These findings underscore the relevance of Ayurvedic principles not only in neurological disorders but also in oncology. In a recent study, we further explored Ayurvedic strategies targeting breast cancer and identified plant-derived compounds capable of modulating oncogenic pathways through immune regulation, hormone signaling, and apoptotic mechanisms [8].

Beyond marine and Ayurvedic sources, indigenous medicinal systems worldwide have historically contributed to the discovery of anticancer agents. One such source is *Larrea tridentata*, commonly known as the creosote bush, which has been traditionally used by Native American communities for its broad therapeutic applications. Extracts of this plant have been employed in folk medicine for anti-inflammatory, antioxidant, and antimicrobial purposes, prompting scientific interest in its pharmacological potential [9,10]. Among its bioactive constituents, Masoprocol—also known as nordihydroguaiaretic acid—has received attention for its reported anticancer properties, particularly in breast cancer [11,12]. Mechanistically, Masoprocol has been shown to inhibit lipoxygenase enzymes, modulate growth factor signaling, and induce apoptosis in malignant cells [13]. However, despite these promising attributes, Masoprocol was withdrawn from clinical use due to toxicity concerns, most notably hepatotoxicity, highlighting the limitations of reductionist drug-development paradigms that fail to capture complex biological interactions [14].

To overcome these challenges, systems biology and integrative bioinformatics approaches provide a robust alternative by enabling multi-target and multi-pathway evaluations of bioactive compounds. These approaches facilitate a holistic understanding of drug-target-disease relationships, moving beyond single-gene analyses toward network-level dynamics [15,16]. In the present study, we applied a systems biology-driven framework to investigate the therapeutic potential of Masoprocol in breast cancer. The overall research design is summarized in Figure 1. Collectively, this integrative strategy aims to reposition Masoprocol as a potential multi-target therapeutic candidate for breast cancer while demonstrating the value of computational systems pharmacology in bridging ethnopharmacology, molecular docking, network biology, and translational oncology.

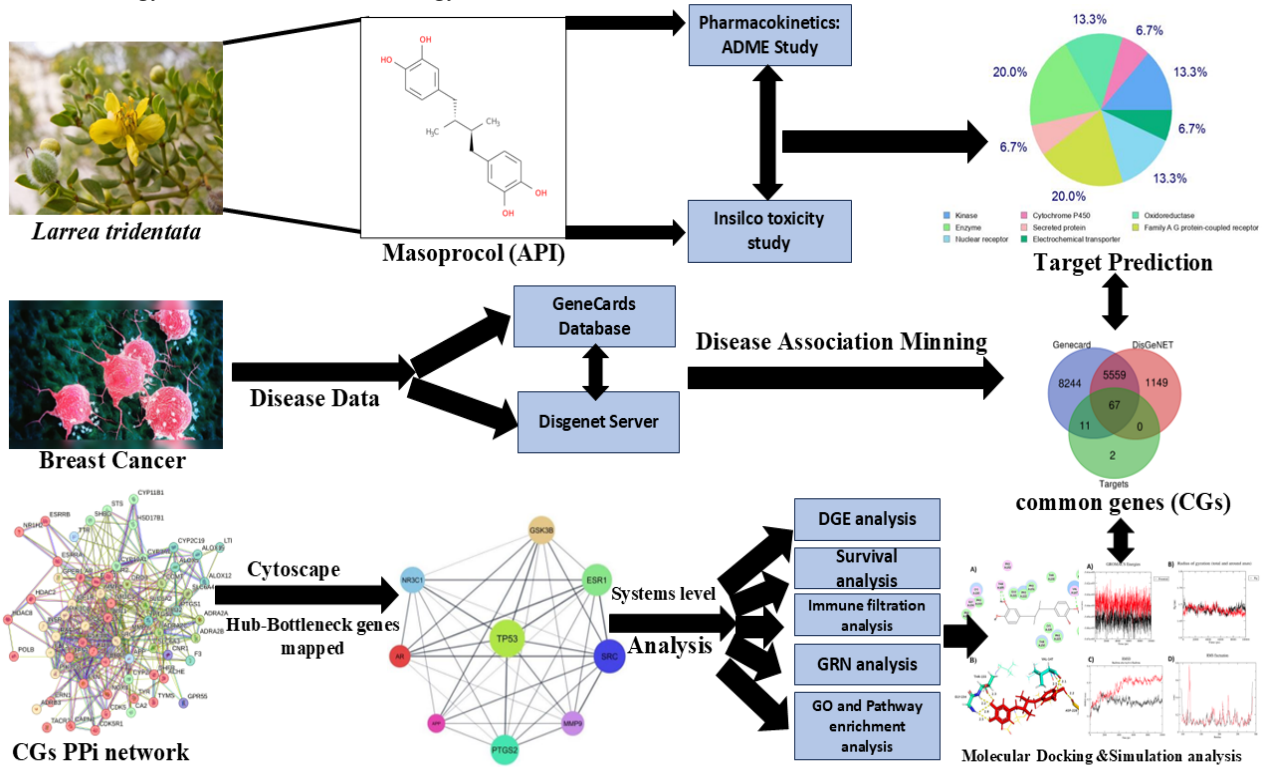


Figure 1. Overall workflow of the systems biology-driven investigation of Masoprocol in breast cancer. The figure summarizes the sequential analytical pipeline, including ADME-toxicity profiling, target prediction, disease association mining, network-based hub

identification, expression and survival analysis, immune infiltration assessment, pathway enrichment, molecular docking and dynamics simulations.

2. Materials and Methods

2.1 Pharmacokinetics and Toxicity Prediction

The pharmacokinetic properties of Masoprocol were systematically assessed using in silico tools to predict its absorption, distribution, metabolism, and excretion (ADME) profiles. The molecule's SMILES notation was entered into the SwissADME platform (http://www.swissadme.ch/, accessed May 2024) [17], a widely used web-based tool for evaluating pharmacokinetic and drug-likeness properties. SwissADME was used to compute critical parameters such as oral bioavailability, gastrointestinal (GI) solubility, blood-brain barrier (BBB) permeability, and compliance with Lipinski's Rule of Five [18]. SwissADME also provided a comprehensive evaluation of molecular descriptors, including molecular weight, hydrogen bond donors and acceptors, number of rotatable bonds, topological polar surface area (TPSA), and multiple logP values. Predictions of water solubility and GI absorption further contributed to understanding the suitability of Masoprocol for oral drug development. To maintain continuity with the subsequent analytical phase, the pharmacokinetic filtering ensured that only molecules with acceptable drug-likeness and absorption profiles proceeded to downstream target prediction and systems-level analyses.

Following the ADME analysis, the molecule underwent toxicity prediction using the ProTox-III platform (https://toxnew.charite.de/protox_III/) [19] accessed in May 2024 (Version 3.0), an advanced computational tool employing machine learning algorithms to estimate various toxicological endpoints. The SMILES string of Masoprocol was entered into ProTox-III to evaluate potential oral toxicity and organ-specific toxicities, including hepatotoxicity, neurotoxicity, and nephrotoxicity. Additional assessments included carcinogenicity, mutagenicity, cytotoxicity, and immunotoxicity. The results provided a detailed safety profile, with each toxicity endpoint categorized as active or inactive, accompanied by probability scores. Radar plots were generated to visually summarize the molecule's toxicological risks. This toxicity assessment served as a preparatory step for the next analytical phase, ensuring that only compounds exhibiting acceptable safety margins were subjected to further molecular network, expression, and survival analyses.

2.2 Target Prediction and Disease Association Mining

To predict the potential protein targets of Masoprocol, two well-established platforms were utilized: SwissTargetPrediction (http://www.swisstargetprediction.ch/) and the BindingDB database (https://www.bindingdb.org). SwissTargetPrediction (Version accessed May 2024) employs a probabilistic learning model that evaluates both 2D and 3D chemical similarities between the query compound and a curated database of known bioactive molecules to estimate the likelihood of interaction with specific human protein targets [20]. BindingDB (accessed May 2024), in contrast, uses a ligand-based similarity approach where the compound is matched against known ligands with experimentally validated target profiles, inferring interactions based on Tanimoto similarity scores derived from chemical fingerprints [21].

A refined set of high-confidence targets was generated by selecting only those with a similarity score ≥ 0.70 in BindingDB and a probability score ≥ 0.70 in SwissTargetPrediction, ensuring the inclusion of the most relevant predicted interactions.

To identify breast cancer-associated genes, curated disease-gene association platforms were queried, including DisGeNET (https://www.disgenet.org) [22] and GeneCards (https://www.genecards.org) [23]. The keyword "Breast Cancer" was used to extract disease-related genes, and each entry was subsequently validated against the UniProt database (https://www.uniprot.org) [24] to ensure that only Homo sapiens genes with experimentally supported associations were included. After curation and removal of duplicates, overlapping targets common to Masoprocol and breast cancer were identified.To visualize this overlap, a Venn diagram was constructed using the Bioinformatics & Evolutionary Genomics VenPlot tool (https://bioinformatics.psb.ugent.be/webtools/Venn/), effectively highlighting the common genes (CGs) shared between predicted targets and disease-associated genes. Importantly, the identification of these overlapping targets provided the foundational input for the subsequent network-based analyses, where the interconnectedness and regulatory importance of these genes were examined to determine key hubs driving Masoprocol's potential therapeutic effects.

2.3 Protein-protein Interaction Mapping and Clustering

A comprehensive protein-protein interaction (PPI) network was constructed based on the previously identified CGs. To accomplish this, we utilized the STRING v11.5 (accessed May 2024) database (https://string-db.org/) [25], which integrates both experimental data and computational predictions to identify potential protein interactions. We specifically focused on first-degree interactors, selecting only those with a medium confidence score (≥ 0.4) to ensure the inclusion of biologically relevant connections. The network was refined by removing duplicate edges and isolated nodes (singletons) to enhance clarity and biological accuracy. For clustering and identification of functionally

significant modules, the Molecular Complex Detection (MCODE) algorithm (version 1.5.1) was employed within the Cytoscape platform. The MCODE parameters were set as follows: maximum depth = 100, node score threshold = 0.2, and K-core = 2. This strategy enabled the detection of highly interconnected regions, offering valuable insights into key biological modules and potential functional hubs [26]. The final PPI network was visualized and analyzed using Cytoscape v3.10.0 (https://cytoscape.org/) [27].This network construction step directly informed the subsequent topological and regulatory analyses, where hub genes were prioritized for expression, survival, and pathway-level evaluation to understand their relevance in breast cancer progression.

2.4 Hub-Bottleneck Genes Identification

To gain deeper insights into potential driver genes, we employed hub-bottleneck analysis on the PPI network derived from the CGs. In PPI networks, nodes represent genes or proteins, while edges denote biological relationships such as physical interactions or co-expression. Network biology leverages topological metrics–most notably degree and betweenness centrality–to identify key nodes governing network structure and function.

The degree of a node reflects its number of direct interactions, with highly connected nodes termed hub genes. Betweenness centrality quantifies how frequently a node lies on the shortest paths connecting other nodes, identifying nodes with high centrality as bottlenecks essential for regulating information flow [28]. To identify hub genes, we calculated the average degree (A.D) and standard deviation (SD) across the network. Genes with a degree greater than (A.D + SD) were classified as hub proteins. In parallel, all nodes were ranked by betweenness centrality, and the top 5% were extracted as bottleneck proteins. Nodes meeting both criteria–high degree and high betweenness centrality–were designated hub-bottleneck genes.

All computations were performed using Cytoscape v3.10.0 (accessed May 2024), enabling the extraction and visualization of hub proteins, hub-bottleneck genes, and their first-degree neighbors. This integrative, network-centric approach allowed the identification of critical regulators within the PPI network, shedding light on their potential roles in cellular signaling, disease progression, and therapeutic intervention. The prioritized hub-bottleneck genes identified here served as key inputs for the subsequent expression and survival analyses performed using GEPIA and TIMER, enabling systematic evaluation of their clinical relevance in breast cancer.

2.5 Hub Gene Regulatory Network Analysis

Gene regulatory networks (GRNs) are fundamental to cellular function, as they coordinate gene expression through complex interactions among transcription factors (TFs), microRNAs (miRNAs), and their target genes. Understanding these regulatory dynamics is essential for pinpointing molecular contributors to disease progression and identifying potential therapeutic targets.

To explore the regulatory framework governing the hub drug-target genes, we employed NetworkAnalyst (https://www.networkanalyst.ca/, accessed May 2024) [29], a robust platform designed for visualizing and analyzing molecular interaction networks. For transcription factor-gene interactions, we utilized the JASPAR 2024 database (https://jaspar.elixir.no/) [30], a comprehensive repository of TF-binding profiles across multiple species. This enabled the identification of TFs that may serve as upstream regulators of the prioritized hub genes.

To further refine the regulatory network, we incorporated miRNA-gene interactions using data from miRTarBase v9.0 (accessed May 2024) [31], a curated database of experimentally validated miRNA targets. miRNAs, as small non-coding RNAs, post-transcriptionally modulate gene expression by binding to complementary sequences on target mRNAs, thereby influencing hub gene activity.

By integrating TF and miRNA interactions through NetworkAnalyst, we constructed a comprehensive regulatory network, enabling detailed topological assessment and identification of regulatory elements that may critically modulate hub gene function and contribute to disease-associated signalling pathways. The resulting GRN served as an essential framework for subsequent differential expression and survival analyses, allowing us to validate whether these regulatory interactions translate into clinically meaningful patterns in breast cancer patient cohorts.

2.6 Expression and Prognostic Significance of Hub Genes in Breast Cancer

To assess the expression profiles of the identified hub genes and their relevance to breast cancer prognosis, we conducted an in-depth analysis using RNA sequencing datasets from The Cancer Genome Atlas-Breast Invasive Carcinoma (TCGA-BRCA) and Genotype-Tissue Expression (GTEx) databases, accessed via the GEPIA 2.0 platform (http://gepia2.cancer-pku.cn/, accessed May 2024) [32]. GEPIA is a robust web-based tool that integrates TCGA and GTEx data to support gene expression comparisons, functional interpretation, correlation assessment, and survival outcome analysis.

For this study, we analyzed 1,085 TCGA breast cancer samples and 291 GTEx normal breast tissue samples, and performed differential expression analysis to compare the transcriptional levels of each hub gene between tumor and normal tissues. Statistical significance was set at $p < 0.05$. Because GEPIA integrates data from two independent sources (TCGA and GTEx), batch effects were addressed using the GEPIA normalization framework, which applies

log2(TPM + 1) transformation along with cross-platform normalization to mitigate discrepancies inherent to multi-source RNA-seq datasets. No additional batch-correction methods (such as ComBat) were applied.

To examine the prognostic potential of the identified hub genes, Kaplan-Meier survival analysis was performed. Patients were stratified into high- and low-expression groups based on the median expression value of each gene, allowing evaluation of their impact on overall survival. The expression and prognostic patterns identified in this section formed the basis for subsequent immune infiltration analysis using TIMER, enabling assessment of how these hub genes may influence tumor-immune interactions in the breast cancer microenvironment.

2.7 Co-expression and Functional Network Analysis of Hub Genes

To explore the functional interplay and interaction landscape of the identified hub drug-target genes, we utilized the GeneMANIA platform (https://genemania.org/, accessed May 2024) (Warde-Farley et al. 2010). This comprehensive tool integrates diverse biological datasets–including co-expression profiles, physical and genetic interactions, shared protein domains, and subcellular co-localization–to predict gene function and connectivity.

Among these datasets, co-expression analysis was particularly valuable for identifying genes with correlated expression patterns, suggesting potential co-regulation or involvement in shared biological process (BP). By leveraging these integrated data sources, we constructed a robust gene interaction network that revealed complex functional associations and potential cooperative roles of the hub genes across key cellular pathways.

The functional relationships identified through GeneMANIA provided the foundation for subsequent pathway enrichment analyses, enabling systematic evaluation of the BP and signaling cascades influenced by the prioritized hub genes.

2.8 Tumor Immune Infiltration Analysis

To explore the relationship between hub gene expression and immune cell infiltration in breast cancer, we utilized the Tumor Immune Estimation Resource (TIMER) database (https://cistrome.shinyapps.io/timer, accessed May 2024) [33]. TIMER is a comprehensive and user-friendly web platform designed to analyze and visualize tumor-immune interactions across multiple cancer types based on RNA-seq data. It employs a statistical deconvolution method to infer the abundance of six key tumor-infiltrating immune cell types: B cells, $CD4^+$ T cells, $CD8^+$ T cells, macrophages, neutrophils, and dendritic cells.

Using gene expression profiles from The Cancer Genome Atlas (TCGA), we assessed the correlation between the expression levels of the selected hub genes and the infiltration levels of these immune cell populations in breast cancer tissues. Spearman's correlation analysis was performed, and results were visualized using scatter plots generated within the TIMER interface. A significance threshold of $p < 0.05$ was applied to identify meaningful associations.

This analysis enabled us to investigate whether key hub genes may influence the tumor immune microenvironment, potentially contributing to cancer progression or therapy resistance. Understanding these correlations offers insights into how these genes might modulate immune surveillance and highlights their potential relevance as immunomodulatory targets or biomarkers in breast cancer.

The immune infiltration patterns observed here provided additional biological context for downstream pathway enrichment and molecular docking analyses, allowing us to connect immune-associated gene behaviour with mechanistic interactions at the protein level.

2.9 Gene Ontology and Pathway Enrichment Analysis

Gene Ontology (GO) enrichment analysis serves as a key computational strategy for understanding gene functions and biological roles. Using ShinyGO v0.77 (http://bioinformatics.sdstate.edu/go77/, accessed May 2024) [34] under the Homo sapiens setting, we performed GO enrichment analysis across three domains–BP, Cellular Component (CC), and Molecular Function (MF). A stringent significance threshold of $p < 0.05$ and false discovery rate (FDR) < 0.05 was applied, with terms ranked by fold enrichment. The top 10 significantly enriched terms in each category were selected for interpretation, including the number of associated genes.

To gain deeper insights into functional pathways, KEGG pathway enrichment analysis was performed using the Kyoto Encyclopedia of Genes and Genomes (KEGG, https://www.genome.jp/kegg/, accessed May 2024) [35], with visual support from the Pathview web tool (accessed May 2024). The top 10 enriched signaling pathways ($p < 0.05$) were identified, revealing key regulatory interactions and shared gene involvement across multiple pathways. Core genes were visualized in red to enhance intuitive understanding of their centrality within biological networks. Gene-pathway associations were then mapped using Cytoscape v3.10.0, enabling systems-level visualization of pathway crosstalk and functional interconnectivity. These enriched BP and signaling pathways formed the mechanistic basis for subsequent molecular docking and dynamics analyses, allowing us to connect computational pathway insights with structure-level interaction studies.

2.10 Molecular Docking Analysis

To elucidate the binding interactions of MB710 and Masoprocol with the Y220C mutant form of TP53, molecular docking studies were conducted using BIOVIA Discovery Studio v3.1 (accessed May 2024) following a validated and systematic protocol. MB710 is a small-molecule stabilizer specifically designed to bind the mutation-induced cavity in Y220C mutant TP53, thereby restoring its native conformation. It was selected as the reference compound due to its well-characterized binding affinity and structural compatibility with the mutant pocket, as demonstrated in previous crystallographic and functional studies [36].

The crystal structure of the Y220C mutant TP53 protein (PDB ID: 5O1I) was used for this analysis. Protein preparation included removal of crystallographic water molecules, addition of hydrogen atoms, and energy minimization to obtain an optimized and stable conformation. The binding site was defined based on the co-crystallized ligand MB710, ensuring physiologically relevant docking conditions that mimic native interaction scenarios.

Both MB710 and Masoprocol were prepared and energy-minimized using the CHARMm force field [37] to ensure standardized structural parameters. Docking simulations were performed using the CDOCKER algorithm [38], with both ligands docked into the same predefined binding pocket within the Y220C mutant protein. By employing MB710 as a benchmark–owing to its proven capacity to stabilize the mutation-induced cavity–this comparative docking strategy enabled a robust evaluation of Masoprocol's binding affinity, interaction profile, and its potential to restore structural stability to the destabilized TP53 mutant. The docking outcomes served as a critical foundation for subsequent molecular dynamics (MD) simulations, which were performed to assess the stability and conformational behavior of the ligand-protein complexes under physiologically relevant conditions.

2.11 Molecular Dynamic Simulations

To further elucidate the stability and dynamic behaviour of the protein-ligand complexes, MD simulations were conducted using a systematic and validated protocol. Protein structures were retrieved from the RCSB PDB database (accessed May 2024) and prepared using the Prepare Protein tool in BIOVIA Discovery Studio v3.1, where missing atoms, residues, and loops were reconstructed to generate a complete, energy-minimized model.

Simulations were performed using GROMACS v2022.4 (accessed May 2024), employing the GROMOS54a7 force field, which is well-suited for biomolecular systems and accurately captures protein dynamics [39]. Ligand topologies were generated with the PRODRG server (accessed May 2024), ensuring compatibility with GROMOS parameters for seamless integration into the simulation environment [40].

Each system was solvated in a dodecahedron box (1.0 nm) using the SPC216 water model, providing a realistic aqueous environment. System neutralization was achieved through the addition of chloride ions–7 for the wild-type TP53 system, 4 for the mutant TP53 system, and 5 for the mutant TP53 -ligand complex–to counterbalance net charges and prevent electrostatic artifacts.

Before initiating the main simulation, energy minimization using the steepest descent algorithm for 50,000 steps was performed to remove steric clashes and stabilize the starting structures. Two equilibration phases followed: (1) NVT ensemble (100 ps) to stabilize temperature. (2) NPT ensemble (100 ps) to equilibrate pressure and system density.

Subsequently, a 10-nanosecond production MD simulation was executed for each system. These trajectories enabled real-time assessment of conformational changes, ligand stability within the binding site, and dynamic interaction patterns.

The MD simulation outputs provided a critical basis for downstream analyses–including RMSD, RMSF, hydrogen bonding, and radius of gyration (Rg)–to evaluate whether Masoprocol maintains stable binding within the Y220C mutant TP53 pocket, thereby supporting its potential to restore functional TP53 conformation.

3. Results

3.1 Pharmacokinetics and Toxicity Prediction

The pharmacokinetic analysis of Masoprocol revealed a favorable profile, highlighting its potential as an orally bioavailable therapeutic agent. The compound fully complied with Lipinski's rule of five, meeting the criteria for molecular weight, hydrogen bond donors and acceptors, lipophilicity, and GI solubility. These attributes suggest that Masoprocol is likely to possess good oral absorption and systemic distribution, supporting its classification as a drug-like molecule. Moreover, the prediction for BBB permeability indicated that Masoprocol is unlikely to penetrate the BBB, potentially minimizing central nervous system (CNS)-related side effects–an advantage depending on the intended therapeutic application. Collectively, these pharmacokinetic features support the compound's suitability for further drug development, as illustrated in Figure 2.

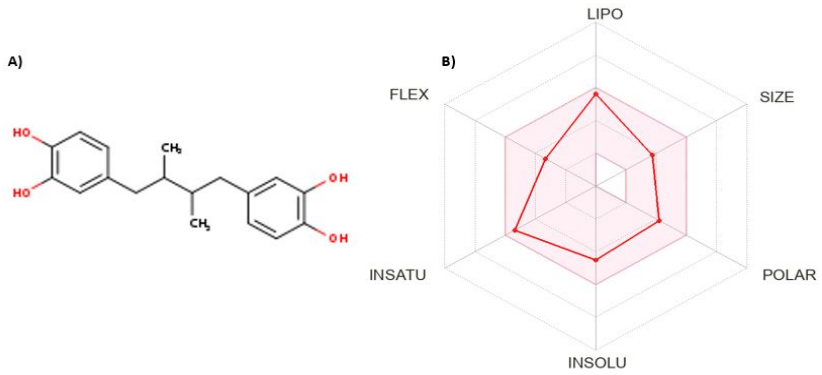


Figure 2. ADME analysis of Masoprocol. A) 2D structure of Masoprocol displaying the chemical representation of the compound. B) Radar chart for ADME prediction showing Insolubility (INSOLU), Polarity (POLAR), Size (SIZE), Lipophilicity (LIPO), Flexibility (FLEX), and Unsaturation (INSATU) as analyzed using the SwissADME tool.

In terms of safety, Masoprocol underwent an extensive in silico toxicity evaluation using the ProTox-III platform. Notably, the compound was predicted to be inactive across all examined toxicological endpoints, including organ-specific toxicities (hepatotoxicity, neurotoxicity, nephrotoxicity) and broader risks such as carcinogenicity, mutagenicity, cytotoxicity, and immunotoxicity. These predictions strongly suggest that Masoprocol possesses a favorable safety profile, which is promising for its further therapeutic development. A comprehensive summary of the pharmacokinetic and toxicity data is presented in Table 1, consolidating the key ADME and toxicity findings. Additionally, a radar chart generated via ProTox-III, shown in Figure 3, visually illustrates the compound's non-toxic profile across multiple toxicity endpoints, reinforcing its potential as a safe drug candidate. ProTox-III predictions indicated that Masoprocol is largely inactive across major toxicity endpoints; however, the radar chart (Figure 3B) revealed a few modest positive toxicity probabilities, particularly for neurotoxicity, hepatotoxicity, and selected receptor-mediated effects. These signals, although below high-risk thresholds, represent potential toxicity domains that may require experimental confirmation. Overall, the pattern supports a predominantly low predicted toxicity profile.

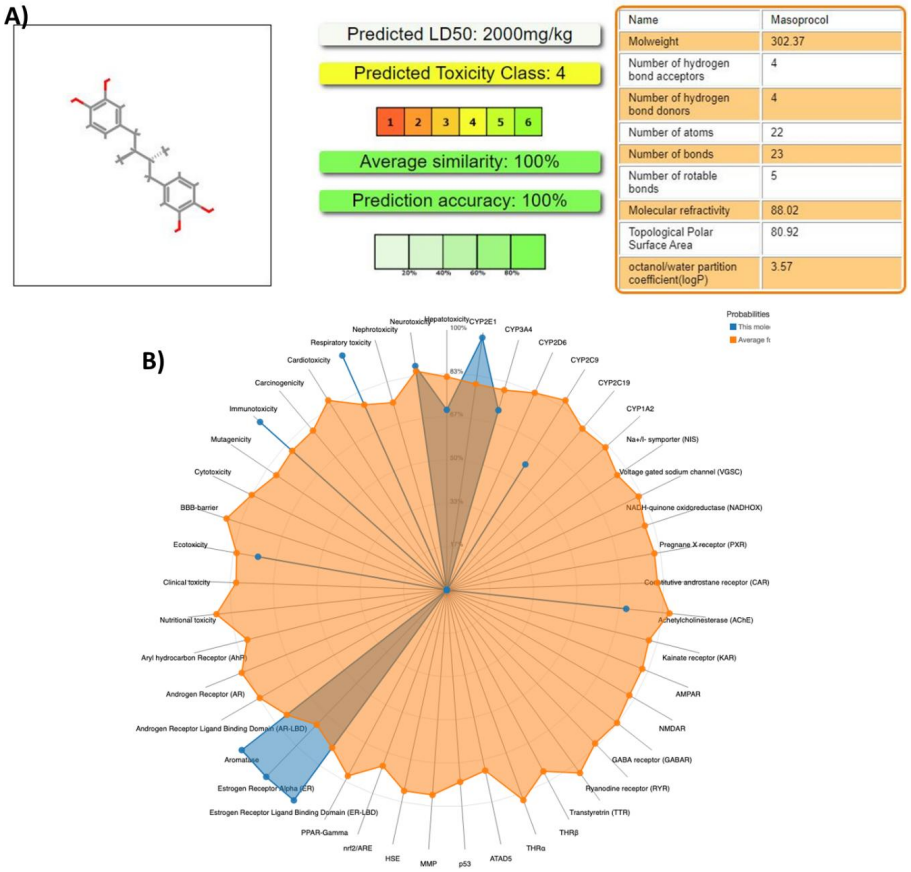


Figure 3. Toxicity prediction of Masoprocol. A) ProTox-III summary showing a predicted LD_{50} of 2000 mg/kg and Toxicity Class 4, indicating low acute toxicity with mild adverse effects only at higher doses. B) Toxicity radar plot illustrating predicted probabilities for multiple toxicity endpoints. Peaks extending outward represent positive (higher) toxicity probabilities, while shorter radii indicate lower risk. Masoprocol shows generally low toxicity across most endpoints, with only a few modest positive signals (e.g., neurotoxicity/hepatotoxicity probability and selected receptor interactions), suggesting an overall favorable safety profile that still warrants experimental verification.

Table 1. Summary of pharmacokinetic properties and toxicity predictions for Masoprocol.

Category	Parameter	Value
General Information	Molecule	Masoprocol
	Formula	C18H22O4
	Molecular Weight (MW)	302.36
	Canonical SMILES	CC(C(Cc1ccc(c(c1)O)O)C)Cc1ccc(c(c1)O)O
Molecular Descriptors	Heavy Atoms	22
	Aromatic Heavy Atoms	12
	Fraction Csp3	0.33
	Rotatable Bonds	5
	H-bond Acceptors	4
	H-bond Donors	4
	TPSA	80.92
	Molar Refractivity	88.02
Physicochemical Properties (LogP)	iLOGP	2.38
	XLOGP3	4.31
	WLOGP	3.57
	MLOGP	2.74
	Silicos-IT LogP	3.45
	Consensus LogP	3.29
Solubility - ESOL Model	ESOL LogS	-4.5
	ESOL Solubility (mg/mL)	0.00948
	ESOL Solubility (mol/L)	3.14E-05
	ESOL Class	Moderately soluble
Solubility - Ali Model	Ali LogS	-5.72
	Ali Solubility (mg/mL)	5.72E-04
	Ali Solubility (mol/L)	1.89E-06
	Ali Class	Moderately soluble
Solubility - Silicos-IT Model	Silicos-IT LogSw	-4.28
	Silicos-IT Solubility (mg/mL)	1.57E-02
	Silicos-IT Solubility (mol/L)	5.20E-05
	Silicos-IT Class	Moderately soluble
ADME / Pharmacokinetics	GI Absorption	High
	BBB Permeant	No
	P-gp Substrate	Yes
	log Kp (cm/s)	-5.08
CYP450 Interactions	CYP1A2	No
	CYP2C19	No
	CYP2C9	Yes
	CYP2D6	Yes
	CYP3A4	Yes
Drug-Likeness Rules	Lipinski Violations	0
	Ghose Violations	0
	Veber Violations	0
	Egan Violations	0
	Muegge Violations	0
	Lead-likeness Violations	1
	Bioavailability Score	0.55
Structural Alerts	PAINS Alerts	1
	Brenk Alerts	1
	Synthetic Accessibility	2.68
Toxicity Predictions (ProTox-III)	Hepatotoxicity	Inactive (0.76)
	Neurotoxicity	Inactive (0.84)
	Nephrotoxicity	Inactive (0.51)
	Respiratory Toxicity	Inactive (0.63)
	Cardiotoxicity	Inactive (0.85)
	Carcinogenicity	Inactive (0.61)
	Immunotoxicity	Inactive (0.96)
	Mutagenicity	Inactive (0.90)
	Cytotoxicity	Inactive (0.97)
	Ecotoxicity	Inactive (0.60)
	Clinical Toxicity	Inactive (0.57)
	Nutritional Toxicity	Inactive (0.80)

3.2 Target Prediction and Disease Association Mining

A crucial step in elucidating the mechanism of action involves target fishing, wherein the bioactive compound Masoprocol is assessed for its potential interactions with specific proteins based on its structural features. This approach aims to identify proteins through which Masoprocol may exert protective effects against pathological processes. To achieve this, the molecular structure of Masoprocol was analyzed using multiple online target prediction platforms. Following the removal of duplicates, a total of 80 unique protein targets were identified, as presented in supplementary table 1 and categorized into various protein classes in Figure 4A. To further refine these predictions, genetic information related to breast cancer-associated targets was systematically retrieved from multiple reputable databases, including DisGeNET and GeneCards. The search yielded 13,881 breast cancer-related genes from GeneCards and 6,776 genes from DisGeNET. By intersecting these disease-associated genes with the 80 predicted targets of Masoprocol, we identified 67 CGs, as illustrated in Figure 4B and 4C, and detailed in supplementary table 2. This curated set of 67 breast cancer-associated targets forms the basis for the downstream network and pathway analyses, providing critical insights into the potential therapeutic mechanism of Masoprocol.

3.3 Protein-protein Interaction Mapping and Clustering

The 67 CGs were subjected to PPI analysis using the STRING v11.5 database, resulting in a comprehensive interaction network, as illustrated in Figure 4D. The constructed network comprises 67 nodes and 385 edges, representing a dense and intricate web of molecular interactions. Topological properties such as degree and betweenness centrality were computed using Cytoscape v3.10.0, with detailed metrics presented in supplementary table 3. Quantitative analysis of the network revealed a network diameter of 4, radius of 3, and a characteristic path length of 2.056, indicating relatively short paths between protein nodes. Additional metrics include a network density of 0.174, centralization of 0.398, average number of neighbors of 11.493, and network heterogeneity of 0.699, collectively providing insights into the structural complexity and interaction diversity within the network. The biological significance and robustness of the interaction map were confirmed by a PPI enrichment p-value < 1.0e-16, suggesting that the observed interactions are significantly more than expected by chance. To further dissect the functional architecture, we performed sub-cluster analysis using the MCODE plugin in Cytoscape. Clustering parameters included a degree cutoff of 2, K-core value of 2, and node score cutoff of 0.2. The resulting highly interconnected subnetworks and functional modules are summarized in Table 2, offering deeper insight into potential biological roles of the CGs.

Table 2. Top 5 gene clusters from PPI network.

Cluster ID	Sub Cluster	Nodes	Edges	Genes
Cluster 1		12	58	BRAF PTGS2 MMP9 ESR2 NR3C1 APP TP53 IGF1R SRC ESR1 MMP2 AR
Cluster 2		9	16	PIK3CA ABL1 NOX4 PDGFRB GPER1 KDR CYP19A1 GSK3B RAF1
Cluster 3		9	16	ALOX15 INSR FYN CYP2C19 PIK3CB ALOX5 PTGS1 ALOX12 FLT3
Cluster 4		5	7	CYP3A4 SLC6A4 DRD2 ADRA2B SLC6A2
Cluster 5		3	3	CDK5 CAPN1 CDK5R1

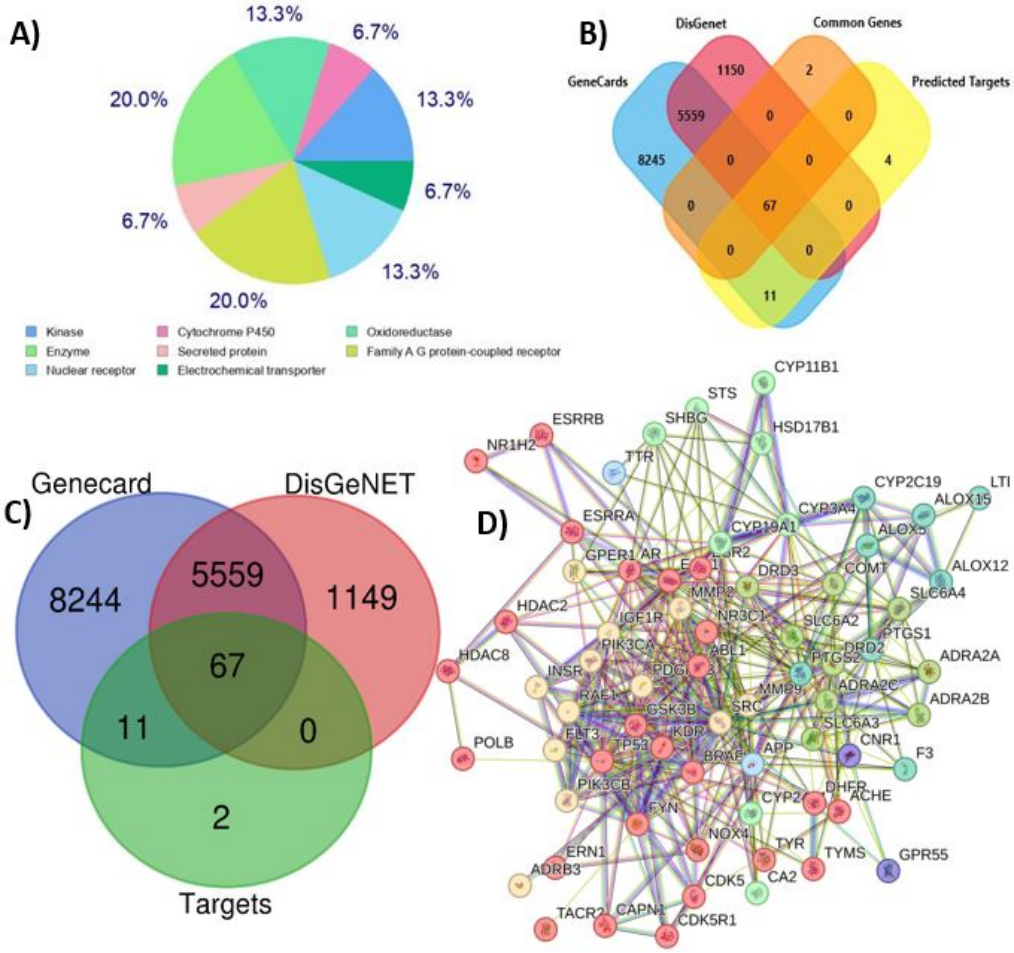


Figure 4. Characterization of predicted targets and breast cancer-associated genes. A) Distribution of protein classes among the predicted targets of Masoprocol. B) Gene counts obtained from different curated disease-gene association databases for breast cancer-associated genes. C) Overlap of CGs between predicted targets and breast cancer-associated genes, as shown by the Venn diagram. D) PPI network representing interactions among the 67 common breast cancer-associated genes identified.

3.4 Hub-bottleneck Genes

A distinct subset of genes exhibiting both hub-like connectivity and bottleneck properties within the PPI network is summarized in Table 3. Prominent hub-bottleneck genes include TP53, SRC, ESR1, PTGS2, GSK3B, NR3C1, MMP9, AR, and APP. These genes surpass the hub classification threshold, defined as degree ≥ 20 (mean + SD), and fall within the top 5% of nodes based on betweenness centrality, underscoring their dual role as key structural and regulatory nodes in the network. Interestingly, all identified hub-bottleneck genes are clustered within Gene Cluster 1, with the exception of *GSK3B*, which is located in Cluster 2, suggesting a potential cross-cluster regulatory influence. This observation emphasizes the centrality and functional importance of Cluster 1 within the broader interaction network. The strategic positioning of these genes within the PPI network highlights their role in maintaining network stability, coordinating signal transduction, and potentially driving disease pathology. Their central role is further illustrated in Figure 5A, which highlights the core positioning of hub-bottleneck genes within the network, while Figure 5B represents the degree distribution of these hub genes, with TP53 exhibiting the highest degree (37), followed by SRC (33), ESR1 (32), PTGS2 (31), GSK3B (25), NR3C1 (23), MMP9 (22), AR (21), and APP (20)—underscoring their pivotal roles in maintaining the structural integrity and connectivity of the network. Collectively, these genes represent critical regulatory hubs and may serve as candidate targets for therapeutic intervention in breast cancer. The summary of degree and betweenness centrality values for each hub-bottleneck gene has been provided in Table 3, further highlighting their topological importance within the network.

3.5 Hub Gene Regulatory Network Analysis

To decipher the transcriptional and post-transcriptional regulatory landscape of hub drug target genes influenced by Masoprocol interactions, we employed a network-based approach that integrates both microRNA (miRNA) and transcription factor (TF) interactions. The miRNA-hub gene interaction network is depicted in Figure 5C, while Figure 5D illustrates TF-mediated regulatory interactions. Our analysis revealed that TP53 is the most extensively regulated gene post-transcriptionally, targeted by 138 miRNAs, followed by NR3C1 (103 miRNAs), AR (100 miRNAs), APP (85 miRNAs), ESR1 (76 miRNAs), and GSK3B (74 miRNAs). In contrast, SRC (8 miRNAs) and PTGS2 (20 miRNAs) showed fewer miRNA regulators, indicating a more specific or tightly controlled regulatory mechanism. Transcriptionally, ESR1 was modulated by 22 TFs, followed by NR3C1 (20 TFs), TP53 (16 TFs), APP (16 TFs),

GSK3B (15 TFs), AR (10 TFs), SRC (7 TFs), MMP9 (7 TFs), and PTGS2 (6 TFs). These findings highlight TP53 as a central regulatory hub under extensive control at both levels, suggesting its crucial role in maintaining cellular homeostasis and its frequent involvement in disease. Similarly, the complex regulation of NR3C1, AR, and APP implies their participation in multiple signaling pathways and potential vulnerability in pathological conditions. In contrast, the limited regulation observed for SRC and PTGS2 may reflect context-specific expression patterns. The integration of TF and miRNA data offers a holistic view of the gene control network, revealing key upstream regulators that could serve as therapeutic targets. Full details of regulatory interactions are summarized in Table 3.

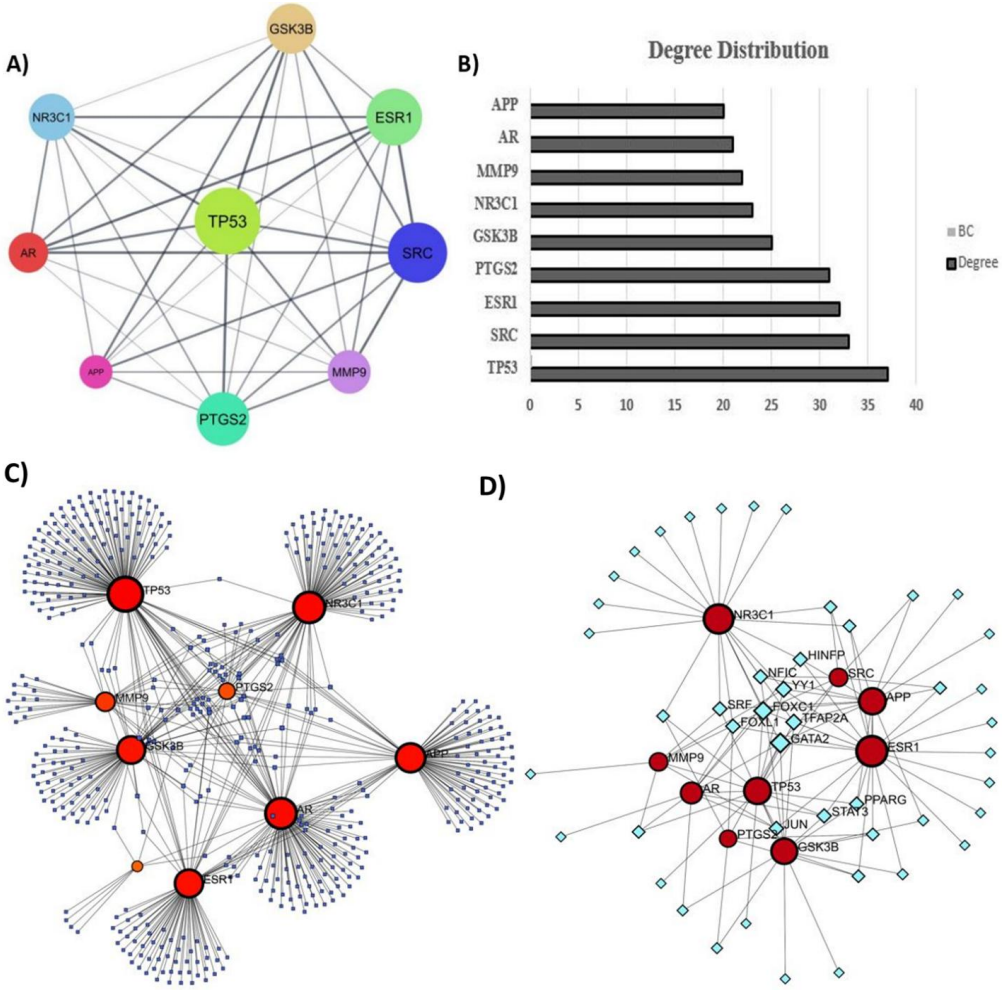


Figure 5. Hub-Bottleneck genes network. A) Core Gene Network with node color and size reflecting degree, and B) Bar chart showing degree analysis of core genes. C) miRNA-Hub GRN. This network illustrates the post-transcriptional regulation of hub drug target genes by miRNAs. Nodes represent genes (red) and miRNAs (blue), while edges denote validated regulatory interactions derived from mirTarBase. TP53, NR3C1, AR, and APP show extensive miRNA regulation, indicating their roles as major control points within the network. D) Transcription Factor-Hub Gene RegulatoryThis network shows transcriptional regulatory interactions between TFs and hub drug target genes. Genes are depicted in red, and TFs in green. Data were obtained from the JASPAR database. ESR1, NR3C1, and TP53 exhibit the highest number of TF interactions, suggesting their involvement in multiple transcriptional regulatory pathways.

Table 3. Summary of hub-bottleneck genes with network topology and regulatory interactions.

R	Hub-Bottleneck	Degree	Betweenness Centrality	miRNA	TF	Gene Cluster
1	TP53	37	0.165592	138	16	Cluster 1
2	SRC	33	0.126146	8	7	Cluster 1
3	ESR1	32	0.101879	76	22	Cluster 1
4	PTGS2	31	0.108869	20	6	Cluster 1
5	GSK3B	25	0.034487	74	15	Cluster 2
6	NR3C1	23	0.049432	103	20	Cluster 1
7	MMP9	22	0.029065	29	7	Cluster 1
8	AR	21	0.019026	100	10	Cluster 1
9	APP	20	0.048808	85	16	Cluster 1

3.6 Expression and Prognostic Value of Hub Genes

Transcriptomic analysis revealed that ESR1, MMP9, and AR were significantly upregulated in breast cancer tissues compared to adjacent normal tissues ($P < 0.05$; Figure 6A-D), while PTGS2 was downregulated. Interestingly, TP53 expression was also elevated in cancer tissues, but this increase did not reach statistical significance, suggesting variability across tumor contexts. Kaplan-Meier survival analysis was conducted based on median expression levels of each gene, dividing patients into high- and low-expression cohorts. The survival analysis revealed that higher expression of ESR1, MMP9, and AR was significantly associated with poorer overall survival outcomes, with MMP9 exhibiting the strongest negative prognostic value (HR = 1.76, log-rank P = 0.02). Similarly, ESR1 and AR showed negative trends with overall survival, suggesting that overexpression of these genes may promote more aggressive tumor behavior and confer a worse prognosis in breast cancer. In contrast, PTGS2 downregulation was associated with better survival outcomes, consistent with its role as an inhibitor of apoptosis and potential tumor suppressor. The results align with existing studies, where MMP9 overexpression has been linked to increased metastatic potential, ESR1 overexpression is often observed in aggressive hormone receptor-positive breast cancers, and AR dysregulation has been implicated in poor prognosis, particularly in triple-negative breast cancer. This highlights the importance of gene overexpression as a prognostic marker and supports the view that higher expression levels of MMP9, ESR1, and AR may indeed be indicative of more aggressive tumor phenotypes and reduced survival chances in breast cancer patients. These findings underscore their potential involvement in breast cancer pathophysiology and highlight the need for further functional validation to elucidate their mechanistic roles in disease progression (Figure 6E-H). Taken together, these results emphasize the complexity of gene regulation in breast cancer, where overexpression of key hub genes like MMP9, ESR1, and AR may serve as poor prognostic indicators. The downregulation of PTGS2, however, warrants further investigation as a potential therapeutic target or biomarker of favorable prognosis.

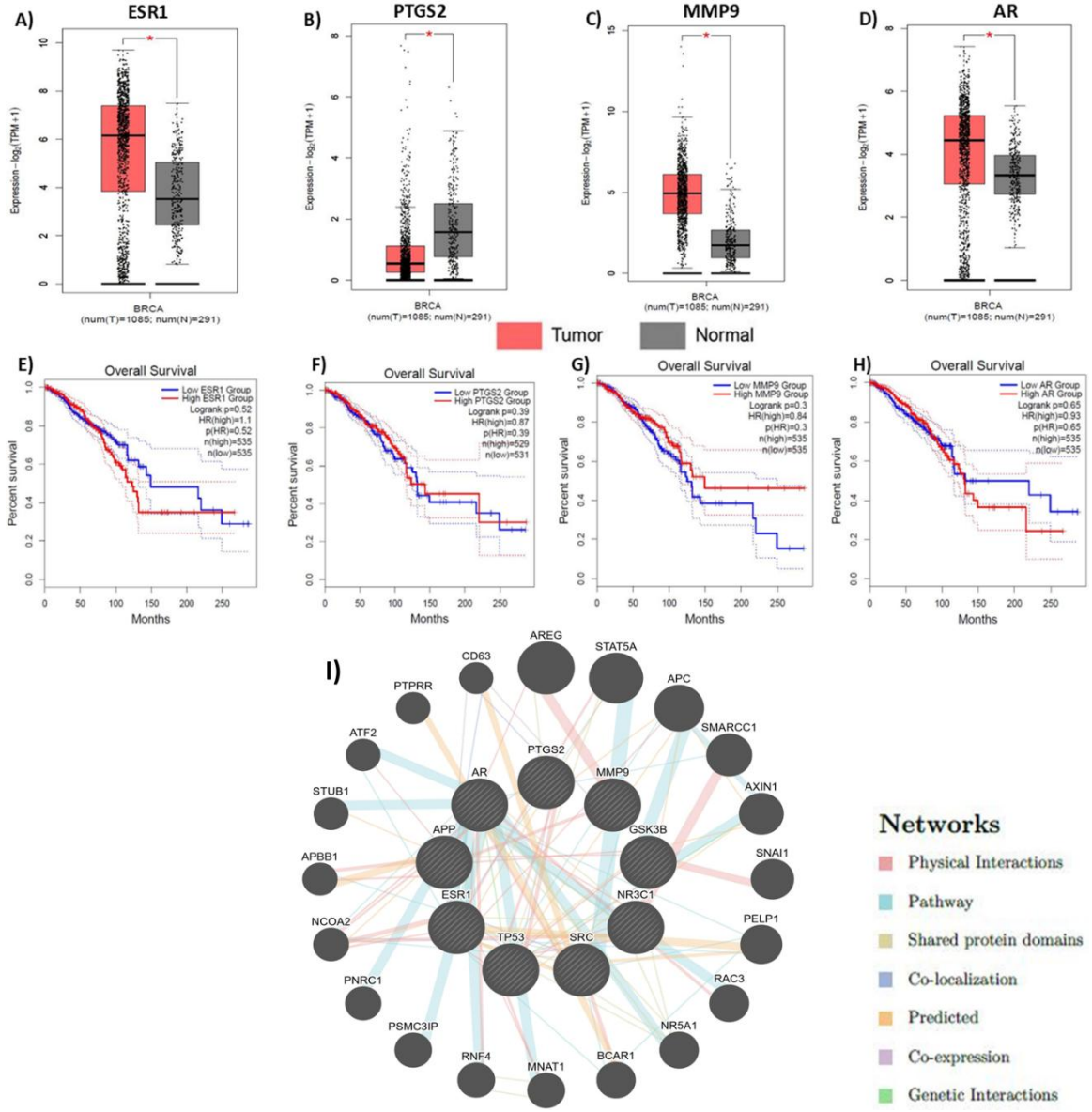


Figure 6. Differential expression, prognostic relevance, and functional connectivity of hub genes in breast cancer. A–D) Box plots illustrating the differential expression of the hub genes ESR1, PTGS2, MMP9, and AR in breast cancer tissues (red) compared with

normal breast tissues (gray), based on integrated TCGA–GTEx RNA-seq data analyzed using the GEPIA platform. Statistical significance between tumor and normal groups is indicated ($p < 0.05$). E–H) Kaplan–Meier overall survival curves stratifying breast cancer patients into high- and low-expression groups for each hub gene, demonstrating the prognostic impact of ESR1, MMP9, and AR overexpression and PTGS2 downregulation on patient survival outcomes. I) GeneMANIA-derived functional interaction network of core hub genes, integrating multiple evidence layers including physical interactions, shared pathways, common protein domains, co-localization, predicted associations, co-expression, and genetic interactions. This network highlights the systems-level functional interconnectivity of hub genes and their cooperative roles in breast cancer-associated regulatory processes.

3.7 Co-expression Network Analysis of Hub Genes in Breast Cancer

To unravel the regulatory interplay among hub drug target genes in the context of breast cancer, we utilized the GeneMANIA platform to construct a detailed co-expression network (Figure 6I). This network centered around nine core hub genes—TP53, SRC, ESR1, PTGS2, GSK3B, NR3C1, MMP9, AR, and APP—previously identified for their high degree and betweenness centrality in the protein-protein interaction (PPI) network. These core genes were found to be intricately linked with multiple co-expressed partners, suggesting a complex regulatory framework potentially disrupted in breast cancer. Among the interaction types, physical interactions formed the largest proportion (41.51%), followed by shared pathways (20.63%), common protein domains (20.63%), co-localization (10.62%), predicted associations (9.28%), co-expression (4.28%), and genetic interactions (1.81%). Although co-expression accounted for a smaller percentage, it played a crucial role in highlighting genes with similar expression profiles across datasets—an important indicator of functional synergy or coordinated regulation. In the context of breast cancer, co-expression analysis revealed that several key oncogenes and tumor suppressors, such as AREG, STAT5A, SNAI1, PELP1, APC, and NCOA2, are co-expressed with the core hub genes. These genes are known to influence critical cancer-related pathways, including hormone signaling, cell proliferation, epithelial-mesenchymal transition (EMT), and immune modulation. For instance, SNAI1 and PELP1 are involved in metastatic progression and estrogen receptor signaling, respectively, while APC and STAT5A play roles in WNT and JAK-STAT pathways, both of which are dysregulated in breast cancer. The observed co-expression patterns not only reinforce the functional relevance of the hub genes but also suggest the presence of shared transcriptional control mechanisms that may be therapeutically targetable. Overall, the network underscores how these genes may operate within a coordinated regulatory landscape that drives breast cancer pathophysiology, supporting their consideration as potential biomarkers or combinatorial therapeutic targets.

3.8 Immune Infiltration Analysis in Breast Cancer

To assess the potential immunomodulatory roles of the hub genes in the breast cancer (BRCA) microenvironment, we analyzed their correlation with immune cell infiltration using the TIMER database. The analysis revealed distinct interaction profiles between specific hub genes and various immune cell types, underscoring their relevance in shaping the tumor immune landscape. TP53 demonstrated a positive correlation with B cell infiltration, suggesting a potential role in promoting B cell-mediated immune responses. However, it showed no significant associations with other immune cell types. SRC exhibited strong positive correlations with $CD4^+$ T cells, neutrophils, and dendritic cells, implying its involvement in modulating adaptive and innate immune responses. This pattern points toward its potential function in immune cell recruitment or activation within the tumor microenvironment. ESR1 (Estrogen Receptor 1) showed a negative correlation with multiple immune cell types, indicating a possible immune-suppressive role. This suggests that high ESR1 expression may contribute to an immunologically "cold" tumor microenvironment, which could have implications for therapeutic resistance. PTGS2 (COX-2) was strongly associated with $CD8^+$ T cells, neutrophils, and dendritic cells, highlighting its link with an immune-infiltrated and potentially inflammatory tumor phenotype. Given its role in prostaglandin synthesis and inflammation, this association supports its relevance in immune modulation. GSK3B showed a moderate correlation with immune cell infiltration and a weak positive association with tumor purity, suggesting a context-dependent influence on immune interactions and tumor biology. NR3C1 was positively correlated with $CD8^+$ T cells, macrophages, and neutrophils, reflecting its role in fostering an immune-rich tumor microenvironment. Its expression appears to be linked with both adaptive and innate immune activity in breast cancer. MMP9 showed a strong positive association with dendritic cells and neutrophils, supporting its known role in matrix remodeling and facilitating immune cell trafficking and activation within tumors. AR (Androgen Receptor) exhibited weak to moderate correlation with $CD8^+$ T cells but showed no significant association with other immune cell types or tumor purity, indicating a limited but potentially focused immune interaction profile. APP (Amyloid Beta Precursor Protein) displayed moderate to strong correlations with macrophages and dendritic cells, suggesting a putative role in immune regulation and microglial/macrophage-like activity within the BRCA tumor microenvironment. These findings highlight the diverse immunological roles of the hub genes and their potential influence on breast cancer progression, immune evasion, or response to immunotherapy. Figure 7 illustrates the correlation between gene expression and immune cell infiltration levels in BRCA. Table 4 provides a summary of the correlation trends of each hub gene with tumor purity and six key immune cell types (↑ indicates positive correlation, ↓ indicates negative correlation).

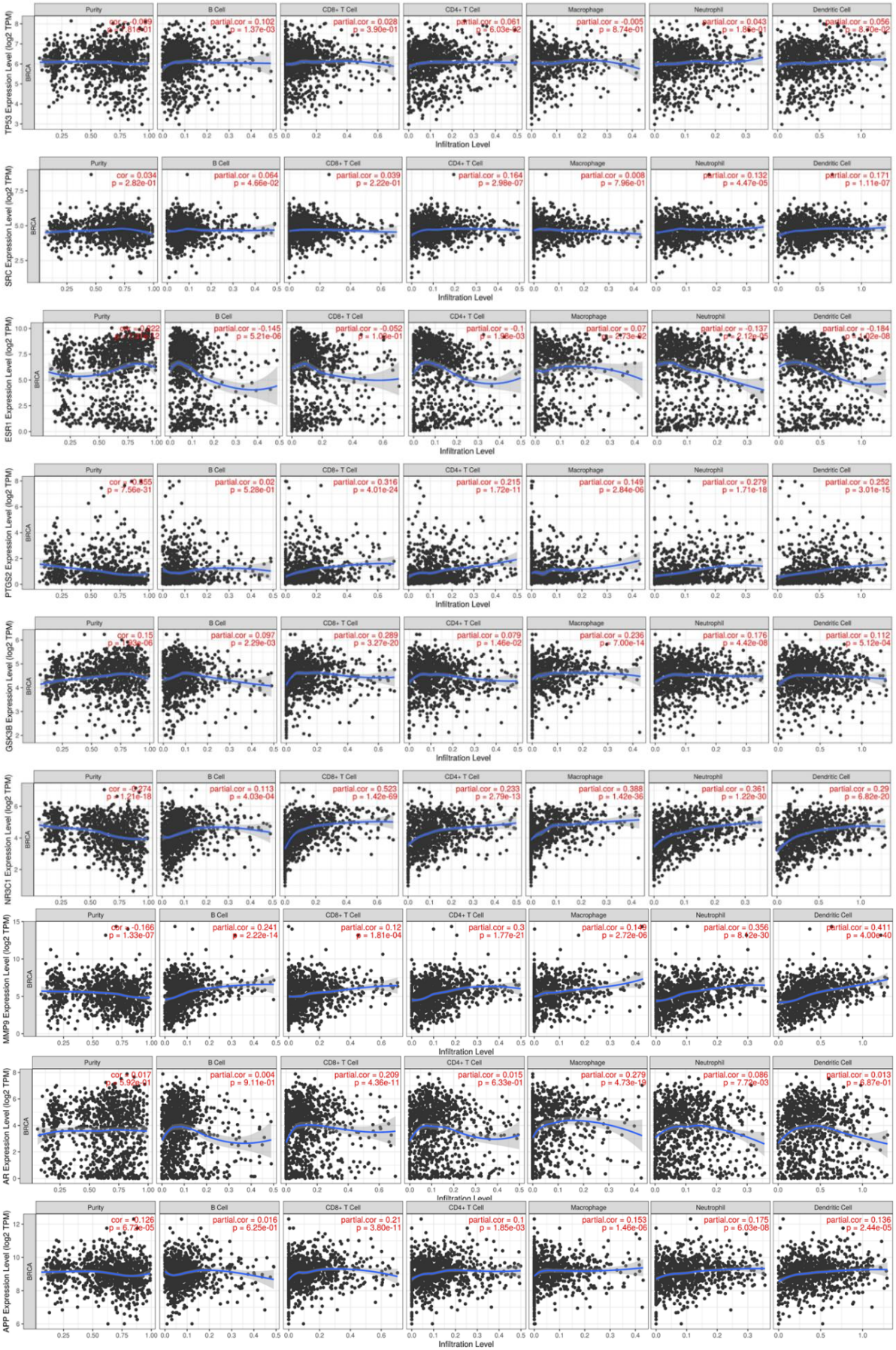


Figure 7. A) Correlation of Hub Genes with Immune Cell Infiltration in BRCA. This figure presents the correlation between the expression of six hub genes (ESR1, PTGS2, GSK3B, TP53, SRC, and NR3C1) and immune cell infiltration in breast cancer, as analyzed via the TIMER database. Notably, PTGS2, SRC, and NR3C1 showed strong positive associations with $CD8^+$ T cells, neutrophils, and dendritic cells, suggesting a role in promoting immune cell recruitment. GSK3B and TP53 exhibited moderate correlations with various immune cells, indicating their involvement in shaping the tumor microenvironment. In contrast, ESR1 demonstrated predominantly negative correlations, particularly with dendritic cells and neutrophils, pointing towards an immune-

suppressive function. B) Correlation of Additional Hub Genes (MMP9, AR, APP) with Immune Infiltration in BRCA. This continuation figure depicts the immune cell infiltration profiles of MMP9, AR, and APP in breast cancer. MMP9 showed robust positive correlations with dendritic cells and neutrophils, highlighting its role in immune activation and extracellular matrix remodeling. APP exhibited strong associations with macrophages and dendritic cells, indicating potential involvement in immune modulation. AR displayed only weak to moderate correlations, with its highest link observed with $CD8^+$ T cells, suggesting a relatively limited role in immune cell engagement within the BRCA microenvironment.

Table 4. Comparative correlation of hub genes with tumor purity and immune cell infiltration in breast cancer.

Gene	Tumor Purity	$CD8^+$ T Cell	$CD4^+$ T Cell	Macrophage	Neutrophil	Dendritic Cell
ESR1	0.2223 (↑)	-0.0523 (↓)	-0.0996 (↓)	0.0704 (↑)	-0.1374 (↓)	-0.1844 (↓)
PTGS2	-0.3546 (↓)	0.3163 (↑)	0.2147 (↑)	0.1487 (↑)	0.2792 (↑)	0.2521 (↑)
GSK3B	0.1503 (↑)	0.2889 (↑)	0.0787 (↑)	0.2358 (↑)	0.1763 (↑)	0.1125 (↑)
TP53	0.0987 (↑)	0.1124 (↑)	0.2518 (↑)	-0.0452 (↓)	0.0917 (↑)	0.1035 (↑)
SRC	-0.2745 (↓)	0.3871 (↑)	0.2153 (↑)	0.0765 (↑)	0.3126 (↑)	0.2759 (↑)
NR3C1	-0.1127 (↓)	0.2985 (↑)	0.1981 (↑)	0.3527 (↑)	0.2912 (↑)	0.2449 (↑)
MMP9	-0.2186 (↓)	0.1872 (↑)	0.1324 (↑)	0.1279 (↑)	0.2893 (↑)	0.3105 (↑)
AR	0.0715 (↑)	0.1396 (↑)	0.0873 (↑)	-0.0268 (↓)	0.0563 (↑)	0.0932 (↑)

Note: Upward arrows (↑) indicate positive correlation and downward arrows (↓) indicate negative correlation.

3.9 Gene Ontology and Pathway enrichment analysis

To elucidate the functional implications of the identified hub genes in breast cancer, we performed comprehensive GO and KEGG pathway enrichment analyses. GO enrichment, conducted via the ShinyGO tool under "Homo sapiens" settings with $p < 0.05$ and FDR < 0.05, revealed the top 10 enriched terms across BP, CC, and MF, presented in Figure 8. Notably, key hub genes such as TP53, PTGS2, GSK3B, MMP9, AR, and SRC were significantly enriched in the regulation of apoptotic signaling pathways–a hallmark of cancer progression–while genes like ESR1 and APP were implicated in transcriptional control and subcellular localization to the growth cone lamellipodium, respectively. These findings underscore the multifunctionality of hub genes and their diverse roles in tumorigenesis, immune response modulation, and transcriptional regulation in breast cancer.

KEGG pathway enrichment analysis identified several significantly enriched pathways (FDR < 0.05), as summarized in Table 5. These included cancer-related pathways such as bladder Cancer, endocrine Resistance, prostate Cancer, estrogen Signaling, breast cancer, and IL-17 Signaling, all showing high fold enrichment values. Using Pathview, these pathways were visualized in Figure 9, with hub-bottleneck genes highlighted to illustrate their involvement across multiple oncogenic processes. To further contextualize these findings, a bioactive-target-pathway network was generated in Figure 10, mapping Masoprocol to its key gene targets and the enriched pathways. This integrative network demonstrates Masoprocol's potential to modulate multiple cancer-associated signaling mechanisms through its interaction with core regulatory genes.

Table 5. Top significantly enriched KEGG pathways (FDR < 0.05) for the hub-bottleneck genes, ranked by fold enrichment.

Enrichment FDR	nGenes	Pathway Genes	Fold Enrichment	Pathways
1.0E-05	3	41	185.3	Bladder cancer
2.6E-05	3	70	108.6	Prolactin signaling pathway
1.4E-06	4	95	106.6	Endocrine resistance
1.4E-06	4	97	104.4	Prostate cancer
1.3E-03	2	59	85.9	VEGF signaling pathway
2.6E-06	4	121	83.7	Thyroid hormone signaling pathway
5.6E-05	3	93	81.7	IL-17 signaling pathway
1.5E-04	3	138	55.1	Estrogen signaling pathway
1.2E-05	4	194	52.2	Kaposi sarcoma-associated herpesvirus infection
1.7E-04	3	147	51.7	Breast cancer
1.2E-05	4	202	50.2	Proteoglycans in cancer
2.0E-04	3	161	47.2	MicroRNAs in cancer
3.4E-04	3	197	38.6	Chemical carcinogenesis
1.4E-06	6	530	28.7	Pathways in cancer

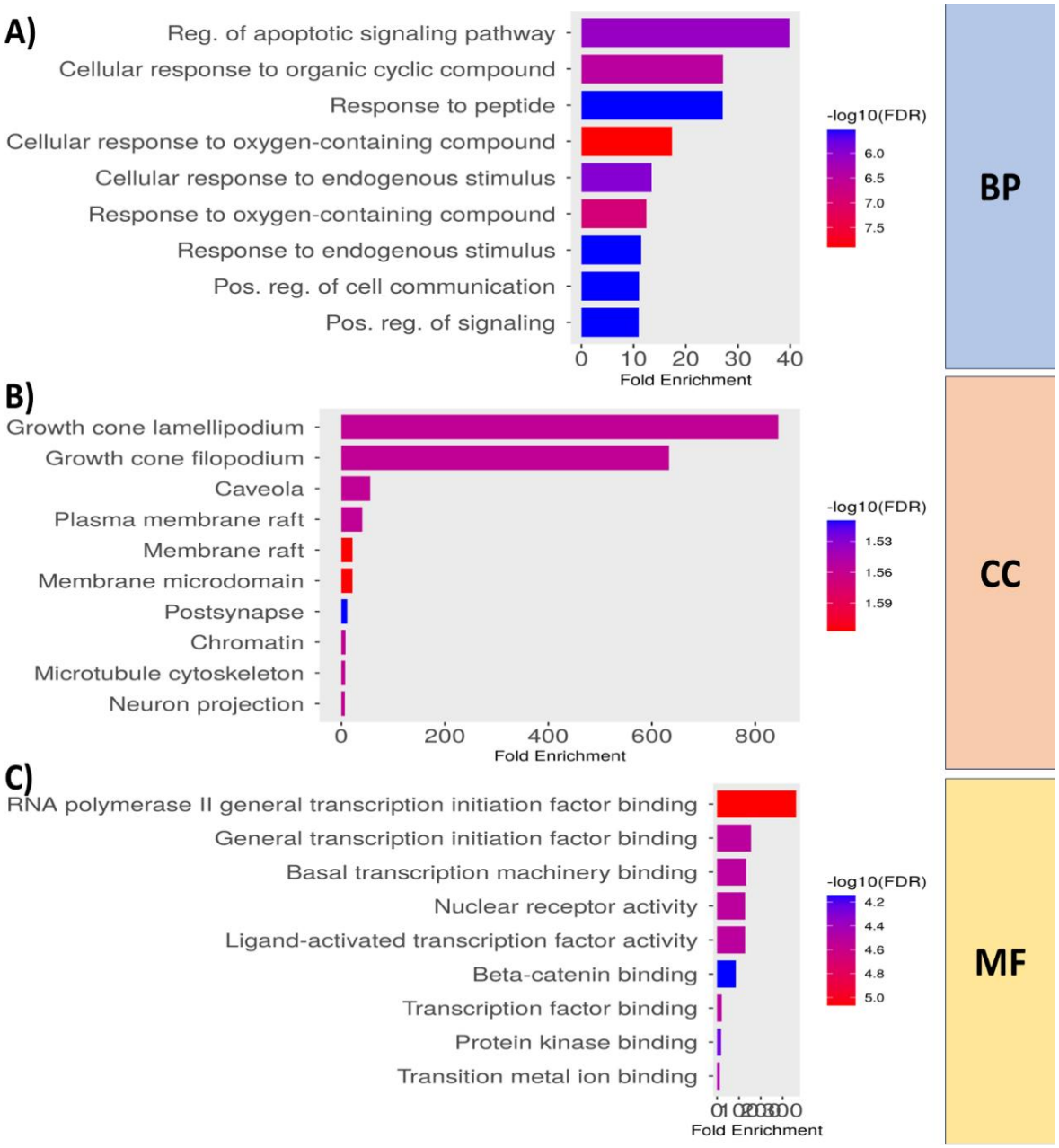


Figure 8. GO Enrichment Analysis of Hub Genes. GO enrichment analysis revealed key biological insights across three categories: (A) BP–the most significantly enriched term was regulation of apoptotic signaling pathway (FDR: 7.89×10^{-7}, Fold Enrichment: 39.88), involving six genes: PTGS2, GSK3B, MMP9, TP53, AR, and SRC; (B) CC–the top enriched term was growth cone lamellipodium (FDR: 0.02771, Fold Enrichment: 844.30), represented solely by the APP gene; and (C) MF–RNA polymerase II general transcription initiation factor binding was the most enriched term (FDR: 8.57×10^{-6}, Fold Enrichment: 361.84), with contributions from ESR1, TP53, and AR. These results highlight critical MF and subcellular localizations relevant to breast cancer pathology. Exact Fold Enrichment and FDR values are provided since the axis scaling is software-generated and cannot be manually adjusted.

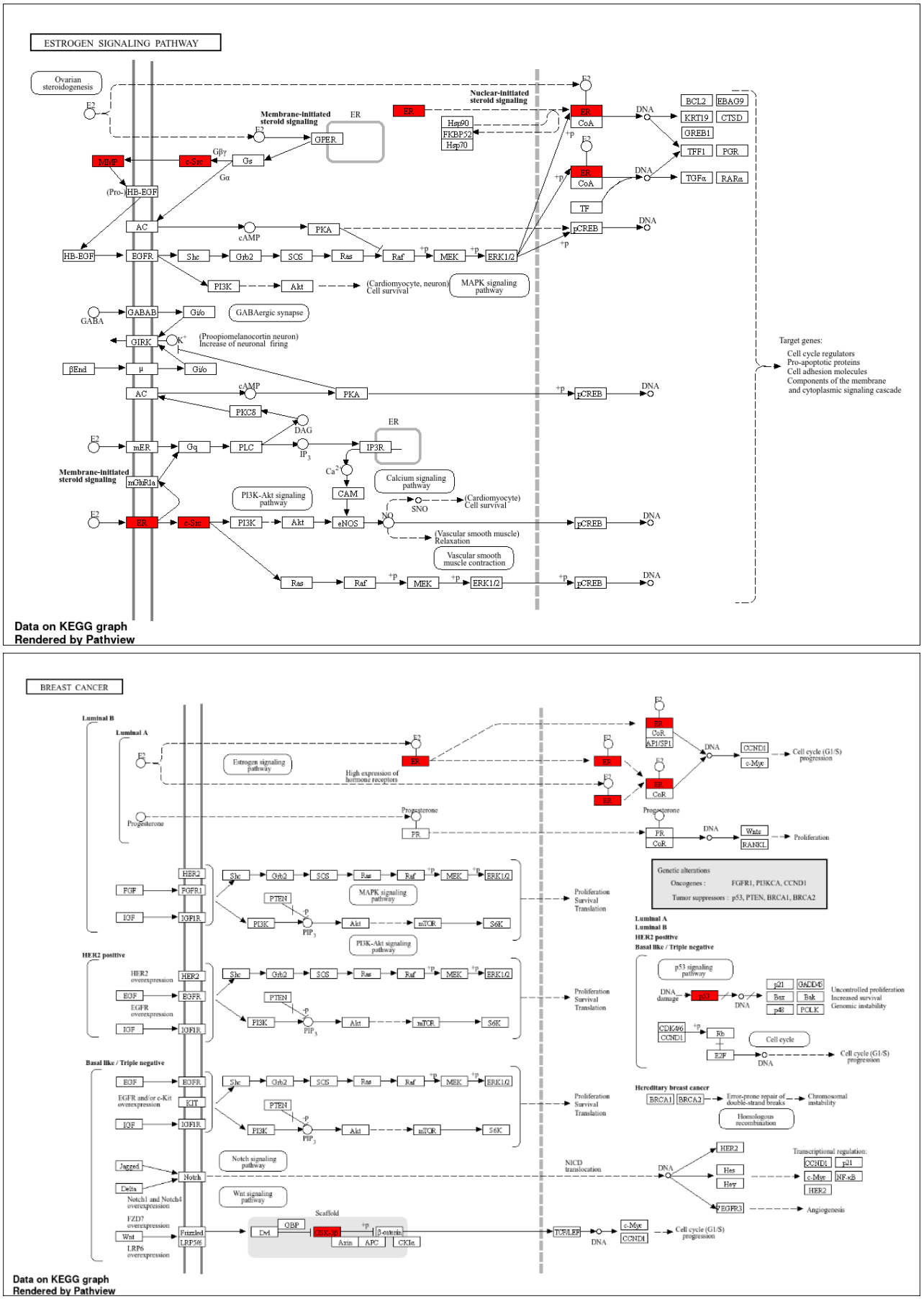


Figure 9. KEGG Pathway Mapping of Hub-Bottleneck Genes in Breast Cancer. Visualization of the top enriched KEGG pathways associated with the hub-bottleneck genes. Pathways relevant to breast cancer biology—such as the Estrogen Signaling Pathway and the Breast Cancer Pathway—are highlighted, with core genes mapped in red. This figure illustrates how the prioritized genes participate across multiple interconnected oncogenic pathways.

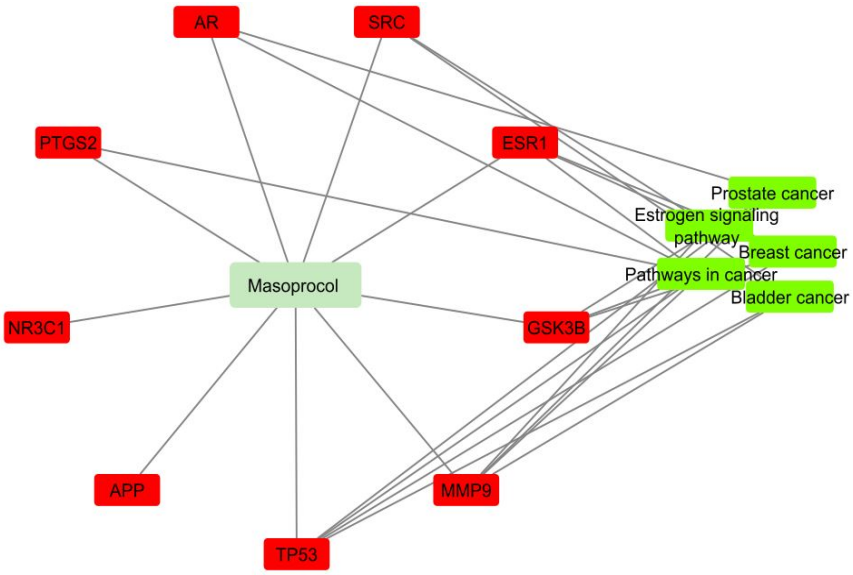
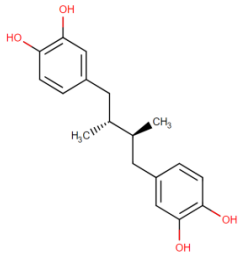
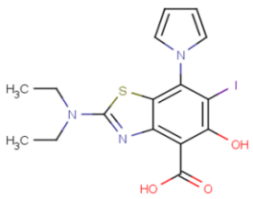


Figure 10. Bioactive-Target-Pathway Network of Masoprocol in Breast Cancer. Network representation showing Masoprocol (light green), its hub-bottleneck gene targets (red nodes), and associated breast cancer-related pathways (dark green). Edges indicate functional associations, illustrating how Masoprocol may modulate multiple oncogenic pathways through key gene targets.

3.10 Molecular Docking Analysis

Molecular docking studies were performed to evaluate the binding interactions of Masoprocol and MB710 with the Y220C mutant TP53 protein (PDB ID: 5O1I), offering valuable insights into their binding affinities and structural complementarity. As detailed in Table 6, Masoprocol achieved a notably higher docking score (136.156) compared to the reference ligand MB710 (121), indicating stronger binding potential within the mutation-induced cavity. Masoprocol established a network of stabilizing hydrogen bonds with key residues such as ASP228 (1:1.9 Å), GLY154 (2:2.2 Å, 2.5 Å), THR155 (2:2.1 Å), and VAL147 (2:2.1 Å, 2.8 Å). These interactions suggest a robust and energetically favorable conformation within the hydrophobic pocket created by the Y220C mutation. In contrast, MB710 formed fewer hydrogen bonds, notably with THR150 (1:2.5 Å), but still retained a relevant binding pose consistent with its reported stabilizing role in mutant TP53. Figure 11 further illustrates these interactions through both 2D and 3D binding maps. The 2D map highlights specific hydrogen bonds, hydrophobic contacts, and π-π stacking interactions, while the 3D map visualizes Masoprocol's spatial orientation within the binding pocket. Key interacting residues are labeled, with hydrogen bonds shown as yellow dotted lines, hydrophobic contacts shaded in gray, and π-π interactions depicted as blue stacks, offering a comprehensive view of the ligand-target interface. Together, this analysis underscores Masoprocol's superior interaction profile relative to MB710 and supports its potential as a promising small-molecule stabilizer of the Y220C mutant TP53, relevant in the context of cancer therapeutics.

Table 6. Comparative docking of Masoprocol Vs. reference compound (MB710).

Sr. No.	Compound Name	Structure	Docking Score PDB ID: 5O1I	H-bond
1	Masoprocol		136.156	ASP228-1:1.9 Å GLY154-2:2.2 Å,2.5 Å THR155-2:2.1 Å Val147-2:2.1 Å, 2.8Å
2	MB710		121	THR150-1:2.5 Å

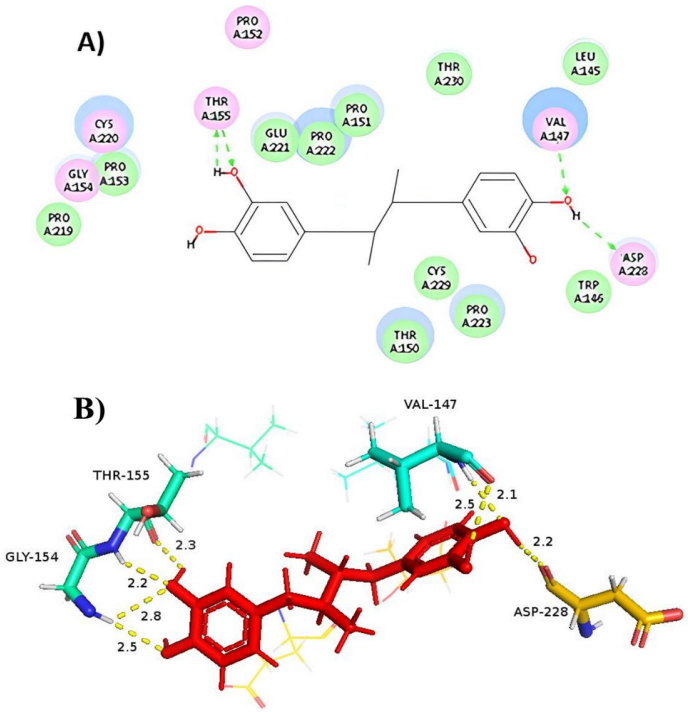


Figure 11. 2D and 3D interaction representation of Masoprocol with the target protein (PDB ID: 5O1I). A) 2D interaction diagram showing key hydrogen bonds, hydrophobic contacts, and π-π interactions, with color-coded ligand atoms (green: C, red: O, blue: N). B) 3D interaction visualization depicting Masoprocol's binding orientation within the pocket and highlighting key interacting residues and bond features.

3.11 Molecular Dynamic Simulations

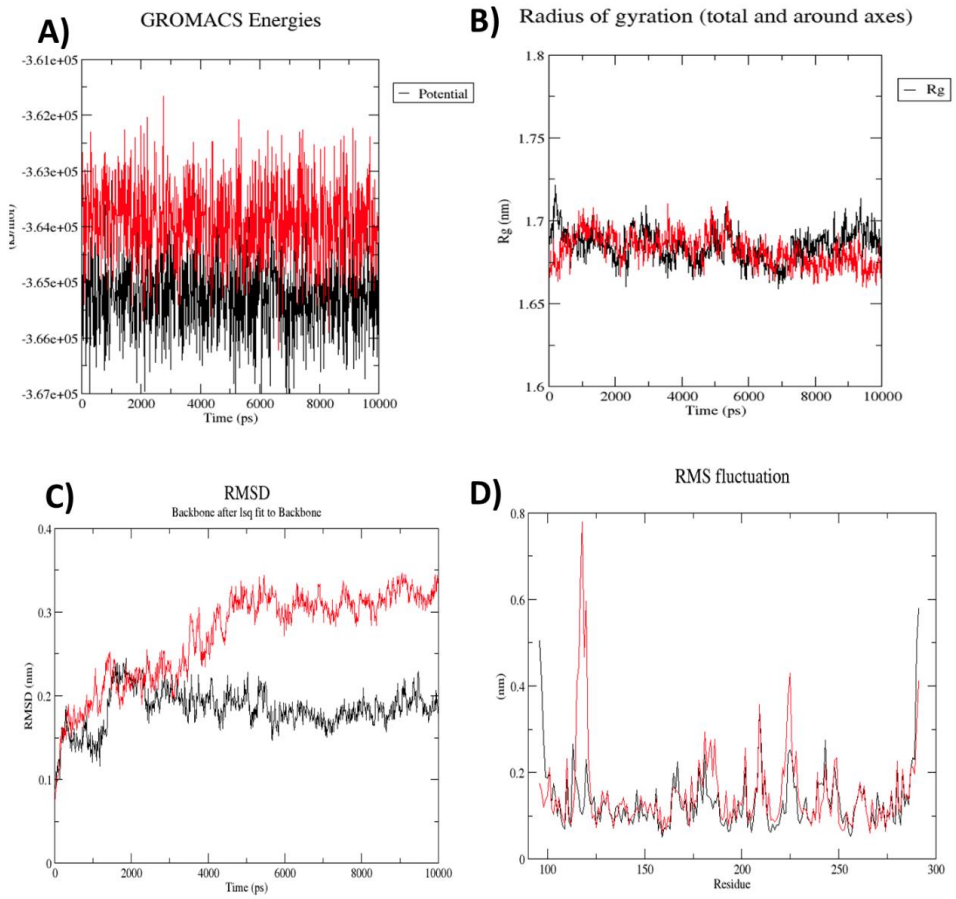


Figure 12. Comparative MD analysis of mutant p53 bound to Masoprocol (red) and MB710 (black). A) Potential energy profiles showing system stability and convergence. B) Rg indicating differences in protein compactness over time. C) RMSD plots depicting overall structural deviations of the complexes during the simulation. D) RMSF analysis showing residue-level flexibility, with notable fluctuations near the mutation site in the Masoprocol complex.

To assess the dynamic stability and structural effects of small molecule binding on the Y220C mutant TP53 protein, MD simulations were performed for the complexes with Masoprocol and MB710. These simulations provided time-resolved data on the conformational behavior of the protein-ligand complexes, revealing how each compound influenced the stability of the mutant protein. The simulation trajectories were analyzed using several key parameters: Potential Energy, Rg, Root Mean Square Deviation (RMSD), and Root Mean Square Fluctuation (RMSF), as illustrated in Figure 12. The potential energy profiles (Figure 12A) indicated that both complexes achieved stable convergence over the course of the simulation, suggesting proper system equilibration and energetically favorable conformations for both Masoprocol and MB710. The RMSD plots (Figure 12C) demonstrated that the mutant TP53 protein complexed with Masoprocol stabilized around 4500 picoseconds, though with slightly greater fluctuations than the MB710 complex, which showed a more consistent and stable RMSD profile throughout the simulation. This suggests that MB710 may bind more tightly or induce fewer structural perturbations than Masoprocol. The Rg curves (Figure 12B) were nearly identical for both complexes, reflecting comparable levels of overall protein compactness and suggesting that neither ligand caused significant unfolding or structural collapse. RMSF analysis (Figure 12D) highlighted the flexibility of individual residues across the protein. Notably, the region spanning residues 222 to 225, corresponding to the destabilized S7-S8 loop caused by the Y220C mutation, exhibited higher fluctuations in the Masoprocol complex compared to MB710. This region is critical, as the mutation exposes hydrophobic residues to the solvent, disrupting local structure. The comparatively lower fluctuations in the MB710 complex imply better stabilization of the loop. Overall, these simulations underscore the potential of both compounds in stabilizing the Y220C mutant TP53 protein. However, MB710 demonstrated a slightly superior stabilizing effect

4. Discussions & Perspectives

The present study provides a systems-level perspective on Masoprocol in the context of breast cancer by integrating target prediction, network topology, pathway enrichment, and structure-based analyses. Rather than proposing Masoprocol as a definitive therapeutic agent, the findings position it as a network-intersecting compound whose predicted target space overlaps with key regulatory axes of breast cancer biology, warranting further investigation. A central insight emerging from the network analysis is the convergence of Masoprocol-associated targets on a limited set of hub-bottleneck genes—most notably TP53, ESR1, SRC, PTGS2, AR, NR3C1, GSK3B, MMP9, and APP. These genes are well-recognized contributors to hormone signaling, inflammation, extracellular matrix remodeling, and transcriptional control in breast cancer [41-44]. Their prioritization here reflects topological importance within the inferred interaction network, rather than direct evidence of Masoprocol-mediated regulation. Importantly, similar hub-centric architectures have been reported in prior network-pharmacology studies, where compounds intersecting multiple regulatory nodes were suggested to exhibit broader biological influence than single-pathway modulators, particularly in hormone-responsive and inflammation [45]. The co-expression patterns observed between Masoprocol-associated targets and regulators such as SNAI1, STAT5A, APC, and PELP1 further suggest that the inferred target set aligns with transcriptional programs linked to epithelial-mesenchymal transition, endocrine signaling, and invasive remodeling. These associations should be interpreted as contextual reinforcement rather than functional confirmation, as co-expression does not imply causality. Nevertheless, their consistency with established oncogenic transcriptional modules supports the biological plausibility of the network-level findings [46,47]. Immune infiltration correlations provided an additional layer of contextual insight. The observed associations between certain hub genes and immune cell populations—such as TP53 with B cells, SRC with antigen-presenting cells, and PTGS2 with cytotoxic T cells—are consistent with existing evidence that genomic alterations and inflammatory signaling shape the breast cancer immune microenvironment. However, no direct immunomodulatory activity of Masoprocol can be inferred from these correlations alone, and these results should be viewed as hypothesis-generating indicators of potential immune relevance rather than mechanistic evidence [48,49]. Among the identified hubs, TP53 was selected for structure-based evaluation due to its high network centrality and clinical relevance, not because it represents the sole or primary target of Masoprocol. TP53 mutations—particularly the destabilizing Y220C variant—representing one of the most clinically impactful alterations influencing tumor proliferation, immune evasion, therapeutic resistance, and metastatic potential [50]. Docking results demonstrated that Masoprocol can occupy the Y220C mutation-induced cavity with favorable interaction geometry, while MD simulations indicated localized instability near the mutation site. These findings suggest structural feasibility but incomplete stabilization, reinforcing the interpretation that Masoprocol itself is unlikely to function as an optimized mutant-p53 stabilizer without further chemical refinement. Importantly, these simulations do not demonstrate functional rescue of TP53 activity and are limited to short timescales. Pathway enrichment analyses further indicated that the hub genes paarticipate in interconnected signaling circuits related to estrogen signaling, endocrine resistance, inflammatory pathways, and extracellular matrix organization. Rather than implying pathway inhibition, these enrichments highlight that Masoprocol's predicted target landscape overlaps with pathways commonly dysregulated in breast cancer, supporting its relevance as a system-intersecting compound rather than a pathway-specific agent [51-53]. From a clinical perspective, the survival associations observed for ESR1, MMP9, AR, and PTGS2 are consistent with existing literature and lend external validity to the prioritization of these genes within the network. However, these prognostic trends should not be interpreted as evidence that Masoprocol modulates patient outcomes, but rather that its predicted targets reside within clinically meaningful regions of the disease network [54-56]. Overall, this study is inherently constrained by its computational design. Target prediction algorithms, network metrics, enrichment analyses, and docking simulations are probabilistic and do not substitute for experimental validation. The

restriction of MD simulations to TP53 and the absence of functional assays further limit mechanistic inference. Consequently, the conclusions of this work should be viewed as a systems-pharmacology framework for hypothesis generation, not confirmation of therapeutic efficacy. Future work should therefore focus on targeted experimental validation, including in vitro assays assessing apoptosis, hormone responsiveness, and inflammatory signaling, as well as extended structural optimization and validation across additional hub-bottleneck targets. Additional incorporation of PROTAC technologies, quantitative proteomics, and unbiased chemoproteomic pulldown methods may uncover novel Masoprocol targets and deepen mechanistic resolution [57] Within these boundaries, the present study contributes a structured and conservative systems-level rationale for considering Masoprocol as a candidate for further exploration in breast cancer research.

5. Conclusion

This study presents a comprehensive systems-level investigation that repositions Masoprocol as a promising lead compound for breast cancer therapy. Previously withdrawn due to limited mechanistic insights from reductionist approaches, our integrative bioinformatics and systems biology analysis reveals its multi-target potential and pharmacological relevance. ADMET profiling confirmed its favorable drug-like properties, including Lipinski compliance, low predicted toxicity, and limited CNS penetration—supporting its suitability as an orally bioavailable and systemically safe agent. Target prediction and enrichment analyses identified Masoprocol's interaction with multiple breast cancer-associated genes, with TP53 emerging as a key regulatory hub with high centrality in PPI and GRN networks. Expression and survival analyses underscored the clinical importance of hub genes like ESR1, PTGS2, MMP9, and AR—particularly MMP9, whose overexpression correlated with poorer survival, indicating its role as a potential adverse prognostic marker. Co-expression analysis revealed transcriptional alignment of Masoprocol's targets with oncogenic regulators such as SNAI1, STAT5A, APC, and PELP1, suggesting its potential to influence key cancer-related pathways including EMT and hormone signaling. Immune infiltration patterns further pointed to its possible immunomodulatory effects. Finally, docking and MD simulations against the Y220C mutant of TP53 demonstrated strong binding but moderate conformational stability—highlighting the need for structural optimization. Altogether, this work supports Masoprocol's repositioning as a polypharmacological candidate for breast cancer, providing a roadmap for its future optimization and experimental validation in preclinical models.

Data Availability

Data available in the article's supplementary material.

Conflict of Interest

The authors declare no conflict of interest.

Generative AI Statement

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

References

[1] Huang M, Lu JJ, Ding J. Natural products in cancer therapy: past, present and future. Natural Products and Bioprospecting, 2021, 11, 5-13. DOI: 10.1007/s13659-020-00293-7

[2] Azer K, Leaf I. Systems biology platform for efficient development and translation of multitargeted therapeutics. Frontiers in Systems Biology, 2023, 3, 1229532. DOI: 10.3389/fsysb.2023.1229532

[3] Khalifa SAM, Elias N, Farag MA, Chen L, Saeed A, Hegazy MF, et al. Marine natural products: A source of novel anticancer drugs. Marine Drugs, 2019, 17(9), 491. DOI: 10.3390/md17090491

[4] Ghareeb A, Fouda A, Kishk RM, El Kazzaz WM. Unlocking the therapeutic potential of bioactive exopolysaccharide produced by marine Actinobacterium *Streptomyces vinaceusdrappus* AMG31: A novel approach to drug development. International Journal of Biological Macromolecules, 2024, 276(Pt 2), 133861. DOI: 10.1016/j.ijbiomac.2024.133861

[5] Chowdhury MR, Kizhakedathil MPJ, Kumar V, Saktheeswaran M, Kutty Mathesh K, Deepa VS. Exploring the therapeutic potential of marine actinomycetes: A systems biology-based approach for Alzheimer's disease treatment. Discover Molecules, 2024, 1(4). DOI: /10.1007/s44345-024-00004-6

[6] Chowdhury MR, Karamveer K, Tiwary BK, Nampoothiri NK, Erva RR, Deepa VS. Integrated systems pharmacology, molecular docking, and MD simulations investigation elucidating the therapeutic mechanisms of BHD in Alzheimer's disease treatment. Metabolic Brain Disease, 2025, 40(1), 8. DOI: 10.1007/s11011-024-01460-2

[7] Chowdhury MR, Reddy RVS, Nampoothiri NK, Erva RR, Vijaykumar SD. Exploring bioactive natural products for treating neurodegenerative diseases: A computational network medicine approach targeting the estrogen signaling pathway in amyotrophic lateral sclerosis and Parkinson's disease. Metabolic Brain Disease, 2025, 40(4), 169. DOI: 10.1007/s11011-025-01585-y

[8] Chowdhury MR, Tiwari A, Karamveer K, Dubey GP, Tiwary BK, Vijaykumar SD. Clinical application and pharmacological mechanism of polyherbal phytoformulations in breast cancer and depression treatment: review and network pharmacological analysis. Proceedings of the Indian National Science Academy, 2023, 89(3), 560-583. DOI: 10.1007/s43538-023-00193-7

[9] Luo J, Chuang T, Cheung J, Quan J, Tsai J, Sullivan C, et al. Masoprocol (nordihydroguaiaretic acid): A new antihyperglycemic agent isolated from the creosote bush (*Larrea tridentata*). European Journal of Pharmacology, 1998, 346(1), 77-79. DOI: 10.1016/s0014-2999(98)00139-3

[10] Morales-Ubaldo AL, Rivero-Perez N, Valladares-Carranza B, Madariaga-Navarrete A, Higuera-Piedrahita RI, Delgadillo-Ruiz L, et al. Phytochemical compounds and pharmacological properties of *Larrea tridentata*. Molecules, 2022, 27(17), 5393. DOI: 10.3390/molecules27175393

[11] Lambert JD, Dorr RT, Timmermann BN. Nordihydroguaiaretic acid: A review of its numerous and varied biological activities. Pharmaceutical Biology, 2004, 42(2), 149-158. DOI: 10.1080/13880200490512016

[12] Hwu JR, Hsu CI, Hsu MH, Liang YC, Huang RC, Lee YC. Glycosylated nordihydroguaiaretic acids as anti-cancer agents. Bioorganic & Medicinal Chemistry Letters, 2011, 21(1), 380-382. DOI: 10.1016/j.bmcl.2010.10.137

[13] Hernández-Damián J, Andérica-Romero AC, Pedraza-Chaverri J. Paradoxical cellular effects and biological role of the multifaceted compound nordihydroguaiaretic acid. Archiv der Pharmazie, 2014, 347(10), 685-697. DOI: 10.1002/ardp.201400159

[14] Choulis, N.H. Dermatological drugs, topical agents, and cosmetics. Side Effects of Drugs Annual. 2012, 34, 257-269. DOI: 10.1016/B978-0-444-59499-0.00014-3

[15] Romano JD, Tatonetti NP. Informatics and computational methods in natural product drug discovery: A review and perspectives. Frontiers in Genetics, 2019, 10, 368. DOI: 10.3389/fgene.2019.00368

[16] Khan A, Rehman Z, Hashmi HF, Khan AA, Junaid M, Sayaf AM, et al. An integrated systems biology and network-based approach to identify novel biomarkers in breast cancer cell lines using gene expression data. Interdisciplinary Sciences, 2020, 12(2), 155-168. DOI: 10.1007/s12539-020-00360-0

[17] Daina A, Michielin O, Zoete V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness, and medicinal chemistry friendliness of small molecules. Scientific Reports, 2017, 7, 42717. DOI: 10.1038/srep42717

[18] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced Drug Delivery Reviews, 2001, 46, 3-26. DOI: 10.1016/s0169-409x(00)00129-0

[19] Banerjee P, Kemmler E, Dunkel M, Preissner R. ProTox 3.0: A webserver for the prediction of toxicity of chemicals. Nucleic Acids Research, 2024, 52(W1), W513-W520. DOI: 10.1093/nar/gkae303

[20] Daina A, Michielin O, Zoete V. SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. Nucleic Acids Research, 2019, 47(W1), W357-W364. doi: 10.1093/nar/gkz38

[21] Liu T, Lin Y, Wen X, Jorissen RN, Gilson MK. BindingDB: A web-accessible database of experimentally determined protein-ligand binding affinities. Nucleic Acids Research, 2007, 35(Database issue), 198-201. DOI: 10.1093/nar/gkl999

[22] Piñero J, Queralt-Rosinach N, Bravo À, Deu-Pons J, Bauer-Mehren A, Baron M, et al. DisGeNET: A discovery platform for the dynamical exploration of human diseases and their genes. Database (Oxford), 2015, 2015, bav028. DOI: 10.1093/database/bav028

[23] Safran M, Dalah I, Alexander J, Rosen N, Iny Stein T, Shmoish M, et al. GeneCards Version 3: The human gene integrator. Database (Oxford), 2010, 2010, baq020. DOI: 10.1093/database/baq020

[24] Coudert E, Gehant S, de Castro E, Pozzato M, Baratin D, Neto T, et al. Annotation of biologically relevant ligands in UniProtKB using ChEBI. Bioinformatics, 2023, 39(1), btac793. DOI: 10.1093/bioinformatics/btac793

[25] Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: Customizable protein-protein networks and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Research, 2021, 49(D1), D605-D612. DOI: 10.1093/nar/gkaa1074

[26] Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics, 2001, 2(1), 2. DOI: 10.1186/1471-2105-4-2

[27] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Research, 2003, 13(11), 2498-2504. DOI: 10.1101/gr.1239303

[28] Barthélemy M. Betweenness centrality in large complex networks. European Physical Journal B, 2004, 38(2), 163-168. DOI: 10.1140/epjb/e2004-00111-4

[29] Zhou G, Soufan O, Ewald J, Hancock REW, Basu N, Xia J. NetworkAnalyst 3.0: A visual analytics platform for comprehensive gene expression profiling and meta-analysis. Nucleic Acids Research, 2019, 47(W1), W234-W241. DOI: 10.1093/nar/gkz240

[30] Fornes O, Castro-Mondragon JA, Khan A, van der Lee R, Zhang X, Richmond PA, et al. JASPAR 2020: Update of the open-access database of transcription factor binding profiles. Nucleic Acids Research, 2020, 48(D1), D87-D92. DOI: 10.1093/nar/gkz1001

[31] Huang HY, Lin YC, Li J, Huang KY, Shrestha S, Hong HC, et al. miRTarBase 2020: Updates to the experimentally validated microRNA-target interaction database. Nucleic Acids Research, 2020, 48(D1), D148-D154. DOI: 10.1093/nar/gkz896

[32] Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: Biological network integration for gene prioritization and predicting gene function. Nucleic Acids Research, 2010, 38(Suppl_2), W214-W220. DOI: 10.1093/nar/gkq537

[33] Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, et al. TIMER: A web server for comprehensive analysis of tumor-infiltrating immune cells. Cancer Research, 2017, 77(21), e108-e110. DOI: 10.1158/0008-5472.CAN-17-0307

[34] Ge SX, Jung D, Yao R. ShinyGO: A graphical gene-set enrichment tool for animals and plants. Bioinformatics, 2020, 36(8), 2628-2629. DOI: 10.1093/bioinformatics/btz931

[35] Kanehisa M, Sato Y. KEGG Mapper for inferring cellular functions from protein sequences. Protein Science, 2020, 29(1), 28-35. DOI: 10.1002/pro.3711

[36] Bauer MR, Krämer A, Settanni G, Jones RN, Ni X, Khan Tareque R, et al. Targeting cavity-creating p53 cancer mutations with small-molecule stabilizers: the Y220X paradigm. ACS Chemical Biology, 2020, 15(3), 657-668. DOI: 10.1021/acschembio.9b00748

[37] Brooks BR, Brooks CL 3rd, Mackerell AD Jr, Nilsson L, Petrella RJ, Roux B, et al. CHARMM: The biomolecular simulation program. Journal of Computational Chemistry, 2009, 30(10), 1545-1614. DOI: 10.1002/jcc.21287

[38] Wu G, Robertson DH, Brooks CL 3rd, Vieth M. Detailed analysis of grid-based molecular docking: A case study of CDOCKER—a CHARMm-based MD docking algorithm. Journal of Computational Chemistry, 2003, 24(13), 1549-1562. DOI: 10.1002/jcc.10306

[39] Schmid N, Eichenberger AP, Choutko A, Riniker S, Winger M, Mark AE, et al. Definition and testing of the GROMOS force field versions 54A7 and 54B7. European Biophysics Journal, 2011, 40(7), 843-856. DOI: 10.1007/s00249-011-0700-9

[40] van Aalten DM, Bywater R, Findlay JB, Hendlich M, Hooft RW, Vriend G. PRODRG: A program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. Journal of Computer-Aided Molecular Design, 1996, 10(3), 255-262. DOI: 10.1007/BF00355047

[41] Joseph C, Alsaleem M, Orah N, Narasimha PL, Miligy IM, Kurozumi S, et al. Elevated MMP9 expression in breast cancer is a predictor of shorter patient survival. Breast Cancer Research and Treatment, 2020, 182(2), 267-282. DOI: 10.1007/s10549-020-05670-x

[42] Saatci O, Huynh-Dam KT, Sahin O. Endocrine resistance in breast cancer: from molecular mechanisms to therapeutic strategies. Journal of Molecular Medicine, 2021, 99(12), 1691-1710. DOI: 10.1007/s00109-021-02136-5

[43] Herzog SK, Fuqua SAW. ESR1 mutations and therapeutic resistance in metastatic breast cancer: Progress and remaining challenges. British Journal of Cancer, 2022, 126(2), 174-186. DOI: 10.1038/s41416-021-01564-x

[44] Markosyan N, Li J, Sun YH, Richman LP, Lin JH, Yan F, et al. Tumor cell-intrinsic EPHA2 suppresses antitumor immunity by regulating PTGS2 (COX-2). Journal of Clinical Investigation, 2019, 129(9), 3594-3609. DOI: 10.1172/JCI127755

[45] Chowdhury MR, Vijaykumar SD. A systems biology approach reveals dual neurotherapeutic mechanisms of *Dioscorea bulbifera* in Alzheimer's disease via estrogen signaling and cholinergic modulation. Inflammopharmacology, 2025, 33, 5483-5508. DOI: 10.1007/s10787-025-01872-1

[46] Daniel AR, Gaviglio AL, Knutson TP, Ostrander JH, D'Assoro AB, Ravindranathan P, et al. Progesterone receptor-B enhances estrogen responsiveness of breast cancer cells via scaffolding PELP1- and estrogen receptor-containing transcription complexes. Oncogene, 2015, 34(4), 506-515. DOI: 10.1038/onc.2013.579

[47] Ogara MF, Rodríguez-Seguí SA, Marini M, Nacht AS, Stortz M, Levi V, et al. The glucocorticoid receptor interferes with progesterone receptor-dependent genomic regulation in breast cancer cells. Nucleic Acids Research, 2019, 47(20), 10645-10661. DOI: 10.1093/nar/gkz857

[48] Gong Y, Nagarathinam R, Arisi MF, Gerratana L, Winn JS, Slifker M, Pei J. Genetic variants and tumor immune microenvironment: Clues for targeted therapies in inflammatory breast cancer (IBC). International Journal of Molecular Sciences, 2021, 22(16), 8924. DOI: 10.3390/ijms22168924

[49] Harris MA, Savas P, Virassamy B, O'Malley MMR, Kay J, Mueller SN, et al. Towards targeting the breast cancer immune microenvironment. Nature Reviews Cancer, 2024, 24, 554-577. DOI: 10.1038/s41568-024-00714-6

[50] Chowdhury MR, Tiwari A, Dubey GP. In silico investigation of Y220C mutant p53 for lead design. bioRxiv, 2019, 820761. DOI:10.1101/820761

[51] Chakraborty B, Byemerwa J, Krebs T, Lim F, Chang CY, McDonnell DP. Estrogen receptor signaling in the immune system. Endocrine Reviews, 2023, 44(1), 117-141. DOI: 10.1210/endrev/bnac017

[52] Clusan L, Ferrière F, Flouriot G, Pakdel F. A basic review on estrogen receptor signaling pathways in breast cancer. International Journal of Molecular Sciences, 2023, 24(7), 6834. DOI: 10.3390/ijms24076834

[53] Belachew EB, Sewasew DT. Molecular mechanisms of endocrine resistance in estrogen-receptor-positive breast cancer. Frontiers in Endocrinology, 2021, 12, 599586. DOI: 10.3389/fendo.2021.599586

[54] Chauhan P, Yadav R, Kaushal V, Ranga S, Yadav C, Kaur M. Overall survival of breast cancer patients in association with ESR1 polymorphism. Human Genetics, 2023, 142, 201217. DOI: 10.1016/j.humgen.2023.201217

[55] Jiang H, Li H. Prognostic values of tumoral MMP2 and MMP9 overexpression in breast cancer: A systematic review and meta-analysis. BMC Cancer, 2021, 21, 149. DOI: 10.1186/s12885-021-07860-2

[56] Hackbart H, Cui X, Lee JS. Androgen receptor in breast cancer and its clinical implication. *Transl.* Breast Cancer Research, 2023, 4, 30. DOI: 10.21037/tbcr-23-44

[57] Li X, Song Y. Proteolysis-targeting chimera (PROTAC) for targeted protein degradation and cancer therapy. Journal of Hematology & Oncology, 2020, 13, 50. DOI: 10.1186/s13045-020-00885-3