

Research Article

Computational Prediction of Candidate Targets and Pathways of Wild Ginseng-Derived Adventitious Root Constituents Against Triple-Negative Breast Cancer: An *In Silico* Study

Mi-Kyeong Jang^{1,2}, Su-Yun Lyu^{2,3,*}¹Department of Chemical Engineering, Suncheon National University, 57922 Suncheon, Republic of Korea²Smart Bio Research Institute, Suncheon National University, 57922 Suncheon, Republic of Korea³Department of Pharmacy, Suncheon National University, 57922 Suncheon, Republic of Korea*Correspondence: suyun@scnu.ac.kr (Su-Yun Lyu)

Academic Editor: Mehmet Ozaslan

Submitted: 24 July 2026 Revised: 7 September 2026 Accepted: 14 September 2026 Published: 20 October 2026

Abstract

Background: Triple-negative breast cancer (TNBC) is an aggressive breast cancer subtype that lacks validated molecular targets and effective targeted therapies. Wild ginseng-derived adventitious root (WGAR) extract was previously shown experimentally to reduce MDA-MB-231 TNBC cell viability (half-maximal inhibitory concentration 79 $\mu\text{g/mL}$), induce caspase-9/caspase-3-mediated apoptosis, and suppress epithelial-mesenchymal transition (EMT) markers and interleukin (IL)-6 secretion, but that study examined only two docking targets and left the constituent-level targets undefined. The present work is a purely computational (*in silico*) follow-up study predicting, at the systems level, the candidate targets and pathways underlying this validated phenotype; no new experimental data were generated. **Methods:** The 15 WGAR constituents previously identified in that study were subjected to *in silico* target prediction, and the predicted targets were intersected with TNBC-associated disease genes. Shared targets were analyzed by protein-protein interaction network construction and hub (centrality) analysis with Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Reactome enrichment. Four aglycones were docked against six hub targets, protein-ligand interactions were profiled computationally, and absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties were predicted with a machine learning-based predictor. **Results:** Disease-gene intersection yielded 128 predicted candidate targets forming a network of 128 nodes and 1623 edges centered on signal transducer and activator of transcription 3 (STAT3) and epidermal growth factor receptor (EGFR), with an inflammatory nuclear factor κB -IL-6 module and hypoxia-inducible factor-1 signaling. Enrichment analysis returned 182, 951, and 389 KEGG, GO, and Reactome terms, including cancer, hypoxia, and immune checkpoint pathways. In docking, the aglycones did not separate from the glycosides in predicted binding energy but did so in predicted drug-likeness; the predicted energies at EGFR are not interpreted, since native-ligand re-docking did not reproduce the crystallographic pose at that receptor. The predicted IL-6 module and glycogen synthase kinase-3 β hub corresponded to inflammatory and EMT endpoints measured in the previous study, whereas predicted EGFR and STAT3 centrality remains experimentally untested. **Conclusions:** This *in silico* analysis nominates STAT3, EGFR, and associated inflammatory and hypoxic pathways as candidate targets of the WGAR aglycones, directionally consistent with the previously validated phenotype. All findings are computational predictions, not experimental observations; they are hypothesis-generating and prioritize the target-level experimental validation to follow.

Keywords: triple negative breast neoplasms; network pharmacology; molecular docking simulation; Panax; plant roots; ginsenosides; computer simulation; ErbB receptors

1. Introduction

Breast cancer is the most frequently diagnosed malignancy and the leading cause of cancer-related death among women worldwide, with more than 2.26 million new cases and approximately 685,000 deaths reported in 2020 [1]. Triple-negative breast cancer (TNBC), defined by the absence of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) expression, is one of the most heterogeneous and aggressive breast cancer subtypes and carries a high risk of death upon recurrence [2,3]. Because it lacks the receptor targets exploited by endocrine and HER2-directed therapies, cytotoxic chemotherapy remains the mainstay of

systemic treatment, and TNBC is associated with a greater propensity for early metastasis and a poorer prognosis than other subtypes [4]. Although immune checkpoint inhibitors (ICI) and poly(ADP-ribose) polymerase (PARP) inhibitors have recently expanded the therapeutic options for selected patients, durable benefit remains limited and no broadly effective targeted therapy is available [3]. The persistent lack of validated molecular targets therefore constitutes an unmet clinical need that continues to drive the search for novel agents and actionable targets [4].

Natural products provide a rich source of structurally diverse anticancer agents, and *Panax ginseng* C.A. Meyer, one of the most widely used medicinal herbs



in East Asia, has attracted particular attention in oncology. Its principal bioactive constituents, the ginsenosides, are dammarane-type triterpenoid saponins that exert diverse anticancer effects by inhibiting tumor-cell proliferation, promoting apoptosis, and suppressing invasion and migration [5]. Recent studies have specifically implicated ginsenosides in TNBC: ginsenoside Rh2 suppresses epithelial-mesenchymal transition by disrupting the hypoxia-inducible factor-1 α (HIF-1 α)/Yes-associated protein (YAP) nuclear-translocation axis [6], ginsenoside Rg3 enhances atezolizumab immune-checkpoint therapy by targeting fucosyltransferase 8 (FUT8)-mediated programmed death-ligand 1 (PD-L1) N-glycosylation [7], and ginsenosides synergize with programmed cell death protein 1 (PD-1) blockade through gut-microbiota and metabolic regulation [8]. However, wild-grown ginseng is scarce and its field cultivation is constrained by long growth cycles and continuous-cropping obstacles, limiting the sustainable supply of these constituents. Adventitious root culture offers a scalable, controllable alternative, and elicitation with methyl jasmonate markedly enhances saponin accumulation in *Panax* adventitious root cultures [9]. Wild ginseng-derived adventitious roots (WGAR) thus represent a renewable source of ginseng saponins whose anticancer potential against TNBC remains largely uncharacterized at the systems level.

We previously reported, in a cell-based experimental study, that WGAR, prepared from cultured adventitious roots of a 100-year-old wild ginseng, exerts pro-apoptotic and anti-metastatic effects on MDA-MB-231 TNBC cells: it reduced cell viability with a half-maximal inhibitory concentration (IC₅₀) of 79 μ g/mL, activated the intrinsic apoptotic pathway through sequential caspase-9 and caspase-3 activation followed by PARP cleavage, and suppressed epithelial-mesenchymal transition markers in association with glycogen synthase kinase-3 β (GSK-3 β)/ β -catenin signaling [10]. In that study, 17 constituents were tentatively identified by liquid chromatography-tandem mass spectrometry (LC-MS/MS), and molecular docking was limited to two targets, PARP-1 and β -catenin. As a multi-component extract acting on multiple targets, however, WGAR is likely to engage a broader network of molecular targets than the two examined previously. The identity of these targets, together with the pathways that connect them and the individual constituents responsible, has remained largely undefined. This gap motivated the present *in silico* systems-level analysis of the WGAR constituents, which was designed to generate mechanistic hypotheses extending beyond the experimentally established phenotype, without itself producing new experimental data.

Elucidating the mechanism of such a multi-component, multi-target extract is poorly served by the conventional one-drug, one-target paradigm, which has repeatedly failed to capture the polypharmacological action of complex natural-product mixtures [11,12]. Network

pharmacology addresses this limitation by integrating compound-target prediction, protein-protein interaction analysis, and functional enrichment to model the action of herbal constituents at a systems level, and it has become a widely adopted framework for revealing the mechanisms of traditional herbal medicines [11,13]. Within this framework, compound-target relationships are commonly inferred from structure-based similarity models such as SwissTargetPrediction [14], and the resulting predictions are complemented with structural evidence from molecular docking, which estimates the binding modes and affinities of individual constituents for their putative targets [15]. In parallel, *in silico* profiling of absorption, distribution, metabolism, excretion, and toxicity (ADMET) provides an early assessment of drug-likeness and safety liabilities for candidate constituents [16]. Integrating target prediction, interaction networks, enrichment analysis, docking, and ADMET profiling allows candidate constituents, priority targets, and their associated pathways to be prioritized for subsequent experimental validation [11,12].

In the present study, we performed an entirely computational (*in silico*) network-pharmacology and molecular-docking analysis of the 15 structurally defined constituents of WGAR, building on the validated cell-based findings of that earlier work, to predict their putative multi-target action against TNBC [10]. Predicted compound targets were intersected with TNBC-associated disease genes to define a candidate target set, which was then analyzed by protein-protein interaction networking and hub-centrality analysis, followed by Kyoto Encyclopedia of Genes and Genomes (KEGG), Gene Ontology (GO), and Reactome enrichment to identify the biological processes and pathways engaged. Representative aglycone constituents were docked against the resulting hub targets to assess predicted binding feasibility at the structural level, and the drug-likeness and ADMET profiles of all constituents were evaluated. By assembling a compound-target-pathway view of WGAR, this work aims to generate mechanistic hypotheses and prioritize constituents and targets for future experimental validation. All results presented below are computational predictions rather than experimental measurements, and none of them constitutes direct pharmacological proof.

2. Materials and Methods

2.1 Preparation of Compound Structures

The wild ginseng-derived adventitious root material analyzed in the previous study was derived from a specimen of *P. ginseng* estimated to be over 100 years old, authenticated by the Korea Sanwoncho Wild Ginseng Association on 5 June 2021, from which adventitious roots were established by callus culture and cultivated by MIRAE FARM Agricultural Corporation Co., Ltd. (Gokseong-gun, Jeollanam-do, Republic of Korea) [10]. No plant material was used in the present study, which is entirely computational. The overall *in silico* workflow is summarized

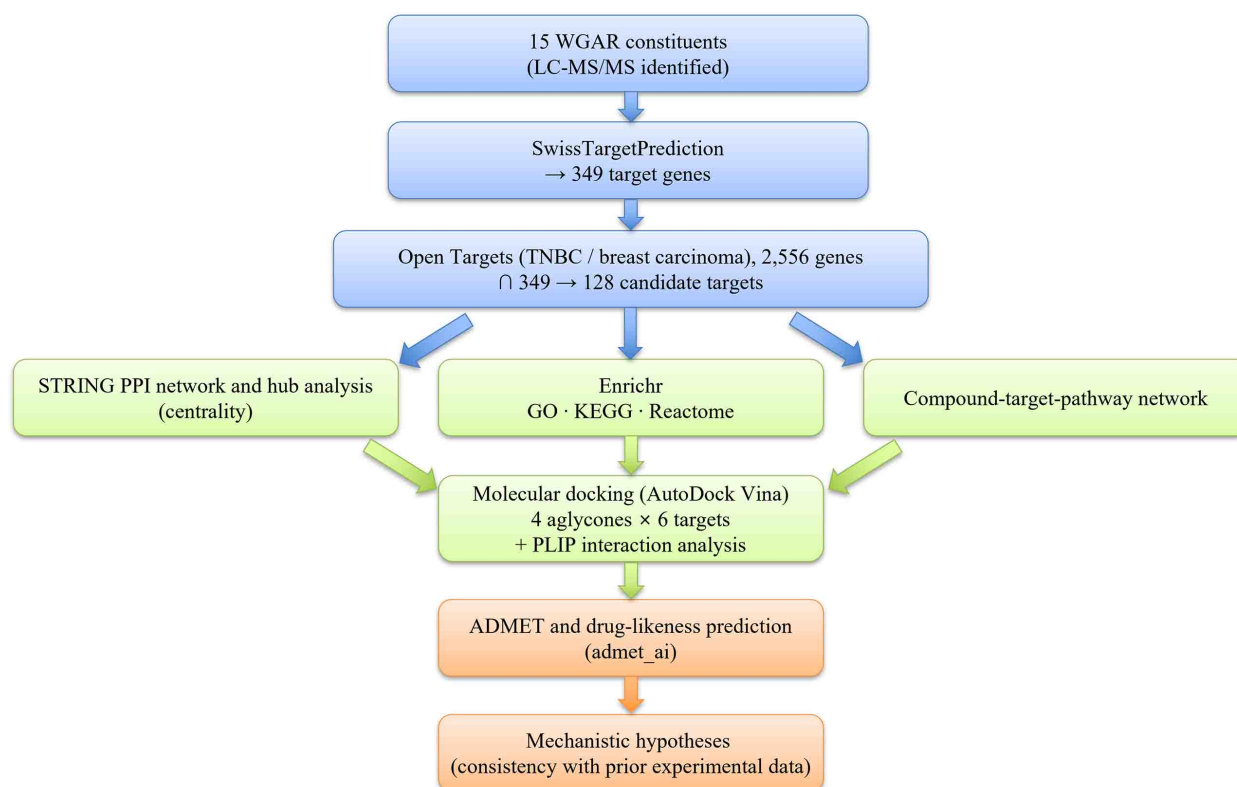


Fig. 1. Overall *in silico* workflow of the study. The 15 constituents of wild ginseng-derived adventitious root (WGAR), previously identified by liquid chromatography-tandem mass spectrometry (LC-MS/MS), were subjected to protein target prediction (SwissTargetPrediction), yielding 349 target genes. These were intersected with 2556 triple-negative breast cancer (TNBC) and breast carcinoma-associated genes from the Open Targets Platform to define 128 candidate targets. The candidate targets were analyzed in parallel by protein-protein interaction (PPI) network construction and hub (centrality) analysis (STRING), functional enrichment analysis [Gene Ontology (GO) biological processes, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, and Reactome pathways; Enrichr], and compound-target-pathway network construction. Four aglycone constituents were then docked against six selected targets (AutoDock Vina), and protein-ligand interactions were profiled with the Protein-Ligand Interaction Profiler (PLIP). Drug-likeness and absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties of the constituents were predicted *in silico* (admet_ai). The integrated results were interpreted as mechanistic hypotheses in the context of prior experimental data.

in Fig. 1. From the 17 constituents tentatively identified by LC-MS/MS in the WGAR extract in our previous study [10], the 15 constituents with publicly available chemical structures were selected for analysis (Table 1; two structures that could not be retrieved, poecillastroside D and an octahydronaphthalene-2-carboxylic acid derivative, were excluded). The canonical and isomeric Simplified Molecular-Input Line-Entry System (SMILES) strings and PubChem compound identifiers (CIDs) of each constituent were obtained from the PubChem database (National Center for Biotechnology Information, Bethesda, MD, USA; accessed 4 June 2026; all SMILES in **Supplementary Table 1**) [17], and structural validity and molecular formulae were verified with RDKit (version 2026.03.2; open source, The RDKit Development Team, <https://www.rdkit.org>). Because 20(S)- and 20(R)-ginsenoside Rg3 share an identical two-dimensional structure, they were treated as a single input for target prediction.

2.2 Compound Target Prediction

Potential protein targets of each constituent were predicted with SwissTargetPrediction (Swiss Institute of Bioinformatics, Lausanne, Switzerland; accessed 1 June 2026; *Homo sapiens*), which is based on two- and three-dimensional structural similarity between the query compound and known bioactive molecules [14]. Owing to the 200-character limit on the input SMILES, flat SMILES with stereochemistry removed were used for the two-dimensional similarity-based target prediction (stereochemical structures were used for docking; Section 2.7). All probability-weighted predicted targets (up to 100 per constituent) were retained and mapped to official gene symbols, yielding 349 unique gene targets.

Table 1. Chemical classification and physicochemical properties of the 15 constituents of wild ginseng-derived adventitious root (WGAR) with publicly available chemical structures, tentatively identified by liquid chromatography-tandem mass spectrometry (LC-MS/MS) in a previous study.

No.	Compound	Class	Molecular formula	MW (g/mol)	PubChem CID
83	Pseudoginsenoside F11	Triterpenoid glycoside	C ₄₂ H ₇₂ O ₁₄	801.0	21633072
84	Myristic acid	Fatty acid	C ₁₄ H ₂₈ O ₂	228.4	11005
86	Licoricesaponin G2	Triterpenoid glycoside	C ₄₂ H ₆₂ O ₁₇	838.9	14891565
87	Majonoside R2	Triterpenoid glycoside	C ₄₁ H ₇₀ O ₁₄	787.0	9962384
90	Soyasapogenol C	Triterpenoid (aglycone)	C ₃₀ H ₄₈ O ₂	440.7	3083637
91	Glochidone	Triterpenoid (aglycone)	C ₃₀ H ₄₆ O	422.7	44559199
92	Methyl palmitate	Fatty acid ester	C ₁₇ H ₃₄ O ₂	270.5	8181
93	Ginsenoside Rf	Triterpenoid glycoside	C ₄₂ H ₇₂ O ₁₄	801.0	441922
94	Ginsenoside Rh1	Triterpenoid glycoside	C ₃₆ H ₆₂ O ₉	638.9	12855920
95	Ursolic acid	Triterpenoid (aglycone)	C ₃₀ H ₄₈ O ₃	456.7	64945
98	Austinoneol	Triterpenoid (aglycone)	C ₂₄ H ₃₀ O ₆	414.5	139584857
99	Octadecane	Alkane	C ₁₈ H ₃₈	254.5	11635
101	Methyl stearate	Fatty acid ester	C ₁₉ H ₃₈ O ₂	298.5	8201
102	20(S)-Ginsenoside Rg3	Triterpenoid glycoside	C ₄₂ H ₇₂ O ₁₃	785.0	9918693
103	20(R)-Ginsenoside Rg3	Triterpenoid glycoside	C ₄₂ H ₇₂ O ₁₃	785.0	46887680

2.3 Disease Target Collection and Candidate Target Identification

Genes associated with TNBC (Experimental Factor Ontology term EFO_0005537) and breast carcinoma (EFO_0000305) were collected from the Open Targets Platform (EMBL-EBI/Open Targets, Hinxton, Cambridge, UK; GraphQL API v4; accessed 2 June 2026) without an association-score cut-off (1495 and 1499 genes, respectively; 2556 in the union) [18]. The intersection of the predicted compound targets (349) with the disease-associated genes (2556) yielded 128 candidate targets.

2.4 Protein-Protein Interaction Network and Hub Analysis

The 128 candidate targets were submitted to the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (version 12.0; STRING Consortium: SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland; Novo Nordisk Foundation Center for Protein Research, Copenhagen, Denmark; European Molecular Biology Laboratory, Heidelberg, Germany; <https://string-db.org>; taxonomy ID 9606; minimum confidence score 0.4; no additional interactors; accessed 4 June 2026) [19]. The symbol VDR is returned by STRING's identifier resolution as CYP27B1, with which it shares database aliases; this entry was therefore removed, giving a protein-protein interaction (PPI) network of 128 nodes and 1623 edges. VDR and MBOAT4 had no interaction partners at the applied confidence threshold and were retained as isolated nodes. The network was reconstructed in NetworkX (version 3.6.1; open source, BSD license; originally developed at Los Alamos National Laboratory, Los Alamos, NM, USA; <https://networkx.org>), and the degree and betweenness centrality of each node were computed; the highest-degree nodes were defined as hub genes. All probability-

weighted target predictions and all disease-gene associations were retained without a threshold, since the SwissTargetPrediction probability and the Open Targets association score are similarity- and evidence-derived estimates rather than confidences of engagement, and are systematically low for chemical scaffolds and disease subtypes that are sparsely represented in the underlying annotation; specificity was instead imposed downstream by the disease-gene intersection, the hub analysis, and the docking step. To assess how this choice affected the hub ranking, the target-to-network workflow was repeated under five alternative criteria: restriction of the disease-gene set to TNBC alone; an Open Targets association-score threshold of 0.1; SwissTargetPrediction probability thresholds of 0.1 and of 0.5; and the combination of TNBC-only disease genes with a probability threshold of 0.1. For each scenario the candidate set was re-derived, the PPI network was rebuilt in STRING under identical settings, degree centrality was recomputed, and the KEGG pathway enrichment was recomputed in Enrichr with the same library (Supplementary Table 2).

2.5 Functional and Pathway Enrichment Analysis

GO biological-process, KEGG, and Reactome pathway enrichment analyses of the 128 candidate targets were performed with Enrichr (Ma'ayan Laboratory, Icahn School of Medicine at Mount Sinai, New York, NY, USA; accessed 4 June 2026) [20], using the GO_Biological_Process_2021, KEGG_2021_Human, and Reactome_2022 libraries, respectively [21,22,23]. Terms with a Benjamini-Hochberg-adjusted $p < 0.05$ were considered significant and ranked by adjusted p -value.

2.6 Compound-Target-Pathway Network Construction

An integrated network comprising the 15 constituents, the 25 highest-degree hub targets, and the top 10 KEGG pathways as nodes was constructed and visualized in NetworkX (version 3.6.1). Compound-target edges were assigned from the SwissTargetPrediction results, and target-pathway edges from the gene composition of each KEGG pathway.

2.7 Molecular Docking

Molecular docking was performed with AutoDock Vina (version 1.2.5; Scripps Research, La Jolla, CA, USA) [24,25]. The receptors were selected from the hub targets of the predicted network, so the docking was not designed to test the target predictions independently; its purpose was to rank the constituents at each selected receptor, a question the network analysis cannot address. Receptor crystal structures were obtained from the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB; Rutgers, The State University of New Jersey, Piscataway, NJ, USA and University of California San Diego, La Jolla, CA, USA; <https://www.rcsb.org>; accessed 4 June 2026) [26]: epidermal growth factor receptor (EGFR, PDB 1M17), B-cell lymphoma 2 (BCL2, 6O0K), heat shock protein 90 alpha family class A member 1 (HSP90AA1, 1YET), prostaglandin-endoperoxide synthase 2 (PTGS2, 3LN1), matrix metalloproteinase (MMP) 9 (1GKC), and signal transducer and activator of transcription 3 (STAT3, 1BG1). Receptors were prepared by removing water and non-protein molecules (the catalytic Zn^{2+} of MMP9 was retained) and assigning Gasteiger partial charges with Open Babel (version 3.1.1; open source, The Open Babel Development Team, <http://openbabel.org>), followed by conversion to Protein Data Bank, Partial Charge, and Atom Type (PDBQT) format [27]. Ligand three-dimensional (3D) conformers were generated with the Experimental-Torsion-Knowledge Distance Geometry version 3 (ETKDGv3) method implemented in RDKit, energy-minimized with the Merck Molecular Force Field 94 (MMFF94), and converted to PDBQT with Open Babel. The grid box (18 to 26 Å per side) was centered on the co-crystallized ligand of each receptor. PDB 1BG1 contains no co-crystallized small-molecule ligand, its only non-solvent heterogroup being phosphotyrosine 705, which the entry annotates as a modified residue of the protein chain (MOD-RES) rather than as a bound ligand; for STAT3 the box was therefore centered manually on the alpha-carbon ($\text{C}\alpha$) positions of the phosphotyrosine-binding residues of the Src homology 2 (SH2) domain (Lys591, Arg609, Ser611, Ser613, and Thr620), which lie within a 12.2 Å span and form a pocket that is unoccupied in this structure. Docking parameters were set to an exhaustiveness of 8, num_modes of 3, and a random seed of 42, and the lowest-binding-energy (kcal/mol) pose was reported. Because glycosidic (saponin) constituents of high molecular weight and with

many rotatable bonds yield docking scores of low reliability (consistent with the drug-likeness results; Section 2.9), binding-pose interpretation focused on the four aglycone triterpenoids (glochidone, ursolic acid, soyasapogenol C, and austinoneol), although all 15 constituents were docked at all six receptors for comparison. The docking protocol was validated by re-docking the co-crystallized ligand of each receptor under the same settings and computing the heavy-atom root-mean-square deviation (RMSD) of the top-scoring pose from the crystallographic pose without superposition. The crystallographic pose was reproduced to within 2 Å for HSP90AA1 (geldanamycin, 0.37 Å), PTGS2 (celecoxib, 0.47 Å), BCL2 (0.61 Å) and MMP9 (1.84 Å), but not for EGFR (erlotinib, 6.9 Å), where the same outcome was obtained across box edges of 18 to 26 Å, exhaustiveness values of 8 and 32, and receptor preparations with and without polar hydrogens; STAT3 could not be tested because 1BG1 contains no co-crystallized ligand. Because the re-docked scores of these five experimentally confirmed binders span 5.4 kcal/mol, binding energies were compared only between compounds at the same receptor and not between receptors.

2.8 Protein-Ligand Interaction Analysis and Visualization

Protein-ligand interactions of the docked complexes were analyzed with the Protein-Ligand Interaction Profiler (PLIP, version 3.0.0; open source, Biotechnology Center (BIOTEC), Technische Universität Dresden, Dresden, Germany; <https://plip-tool.biotech.tu-dresden.de>) to detect hydrogen bonds, hydrophobic contacts, π -stacking, salt bridges, and metal coordination [28]. Three-dimensional binding poses were rendered in PyMOL (The PyMOL Molecular Graphics System, open-source version 2.5.0; Schrödinger, LLC, New York, NY, USA), and two-dimensional interaction diagrams were generated from the PLIP output with RDKit and Matplotlib (version 3.10.8; open source, The Matplotlib Development Team, <https://matplotlib.org>).

2.9 ADMET and Drug-Likeness Prediction

The ADMET properties and drug-likeness of the 15 constituents were predicted with the graph neural network-based *admet_ai* (version 2.0.1; open source, Stanford University, Stanford, CA, USA and Greenstone Biosciences, Palo Alto, CA, USA; https://github.com/swansonk14/admet_ai) [16], yielding 104 indices covering physicochemical properties, drug-likeness (Lipinski's rule, Quantitative Estimate of Drug-likeness [QED]), cytochrome P450 (CYP) enzyme metabolism, and toxicity. All in silico analyses were performed in Python (version 3.12.3; Python Software Foundation, Beaverton, OR, USA).

3. Results

3.1 WGAR Constituents and Predicted Targets

Of the 17 constituents tentatively identified by LC-MS/MS in the WGAR extract in our previous experimental study, the 15 with publicly available chemical structures were selected for the present *in silico* analysis (Table 1). These comprised four aglycone triterpenoids (soyasapogenol C, glochidone, ursolic acid, and austinoneol), seven triterpenoid glycosides including the ginsenosides Rf, Rh1, and 20(S)- and 20(R)-Rg3, and four fatty acid or alkane-type compounds. Their molecular weights ranged from 228.4 g/mol (myristic acid) to 838.9 g/mol (licoricesaponin G2). Because 20(S)- and 20(R)-ginsenoside Rg3 share an identical flat structure, they were treated as a single input for target prediction. *In silico* target prediction with SwissTargetPrediction, a structure-based similarity model, using the flat SMILES of each constituent (Supplementary Table 1), yielded 349 unique predicted gene targets across the 15 constituents (Supplementary Table 3).

3.2 Candidate Targets Against TNBC

In silico intersection of the 349 predicted compound targets with 2556 TNBC and breast carcinoma-associated genes retrieved from the Open Targets Platform (Supplementary Table 4) yielded 128 predicted candidate targets (Fig. 2, Supplementary Table 5). These predicted targets included regulators of apoptosis (STAT3, BCL2, MCL1, BCL2L1), epithelial-mesenchymal transition (EMT) and extracellular matrix degradation (MMP9, MMP2), inflammation and survival signaling (NFKB1, RELA, HIF1A, JUN, IL6, MTOR), and breast cancer signaling (EGFR, ESR1, AR, PIK3CA).

3.3 PPI Network and Hub Target Analysis

The 128 predicted candidate targets were assembled *in silico* into a PPI network of 128 nodes and 1623 edges in the STRING database, which infers protein-protein associations from curated and predicted evidence (Fig. 3, Supplementary Table 6). Degree and betweenness centrality were computed for each node (Supplementary Table 7). In this predicted network, the highest-degree targets were STAT3 (degree 79), EGFR (78), TNF (76), IL6 (76), BCL2 (73), HIF1A (68), and NFKB1 (65), whereas betweenness centrality was highest for EGFR (0.063), followed by TNF (0.050), STAT3 (0.047), and IL6 (0.040). Topologically, the predicted network was organized around STAT3 and EGFR, with closely connected inflammation- and survival-related nodes including TNF, IL6, NFKB1, and IL1B; this organization is a property of the predicted interaction network and was not tested experimentally in the present study. STAT3, EGFR, TNF, and IL6 are recurrently high-degree proteins in cancer interaction networks, so their centrality here prioritizes them for testing rather than identifying them as pharmacological targets of WGAR. Sensitivity analysis showed that this organization was robust to

the definition of the disease-gene set but not to the stringency of the target prediction, whereas the enriched pathways were robust to both (Supplementary Table 2). Restricting the disease genes to TNBC alone reduced the candidate set to 105 targets and the network to 1289 edges, yet all 15 baseline hub targets were retained, 14 of them among the 15 highest-degree nodes of the reduced network, with EGFR second (degree 68) and STAT3 third (68); an Open Targets association-score threshold of 0.1 gave 82 targets and 1016 edges, retained 14 of the 15 baseline hub targets, and placed EGFR first (60) and STAT3 third (58). Applying a SwissTargetPrediction probability threshold of 0.1 or 0.5 reduced the candidate set to 80 and 43 targets and retained 10 and 6 of the baseline hub targets, respectively. STAT3 remained among the two highest-degree nodes at both thresholds (degree 48 and 23), whereas EGFR was not retained at either, having been predicted for only two constituents, at probabilities of 0.061 and 0.035.

3.4 Enriched Pathways and Biological Processes

In silico enrichment analysis of the 128 predicted candidate targets returned 182 significant KEGG pathways, 951 GO biological-process terms, and 389 Reactome pathways (adjusted $p < 0.05$; Fig. 4, Supplementary Tables 8–10). In KEGG, the most significant pathway was “Pathways in cancer” (adjusted $p = 4.85 \times 10^{-36}$, 45 genes), followed by the HIF-1 signaling pathway ($p = 3.76 \times 10^{-22}$, 20 genes), Proteoglycans in cancer ($p = 1.29 \times 10^{-20}$, 23 genes), and the PD-L1 expression and PD-1 checkpoint pathway in cancer ($p = 2.95 \times 10^{-19}$, 17 genes). The top GO biological-process terms were positive regulation of intracellular signal transduction ($p = 3.26 \times 10^{-23}$, 36 genes) and cellular response to cytokine stimulus ($p = 1.32 \times 10^{-20}$, 32 genes), with positive regulation of nuclear factor κ B (NF- κ B) transcription factor activity also enriched ($p = 1.13 \times 10^{-16}$, 19 genes). In Reactome, Signal Transduction ($p = 7.61 \times 10^{-28}$, 70 genes), Signaling by Interleukins ($p = 7.00 \times 10^{-21}$, 31 genes), and Signaling by Receptor Tyrosine Kinases ($p = 1.18 \times 10^{-12}$, 24 genes) were among the most enriched terms in the predicted target set.

3.5 Compound-Target-Pathway Network

An integrated *in silico* network was constructed from the 15 constituents, the 25 highest-degree predicted hub targets, and the 10 most significant KEGG pathways, comprising 133 predicted compound-target edges and 115 target-pathway edges (Fig. 5). Among the constituents, austinoneol and licoricesaponin G2 connected to the largest number of predicted hub targets (14 of the 25 each), followed by pseudoginsenoside F11, majonoside R2, soyasapogenol C and ginsenoside Rf (10 each) and by glochidone, ursolic acid, ginsenoside Rh1 and the two ginsenoside Rg3 stereoisomers (9 each), whereas the fatty acid and alkane constituents connected to the fewest (4 to 6 each).

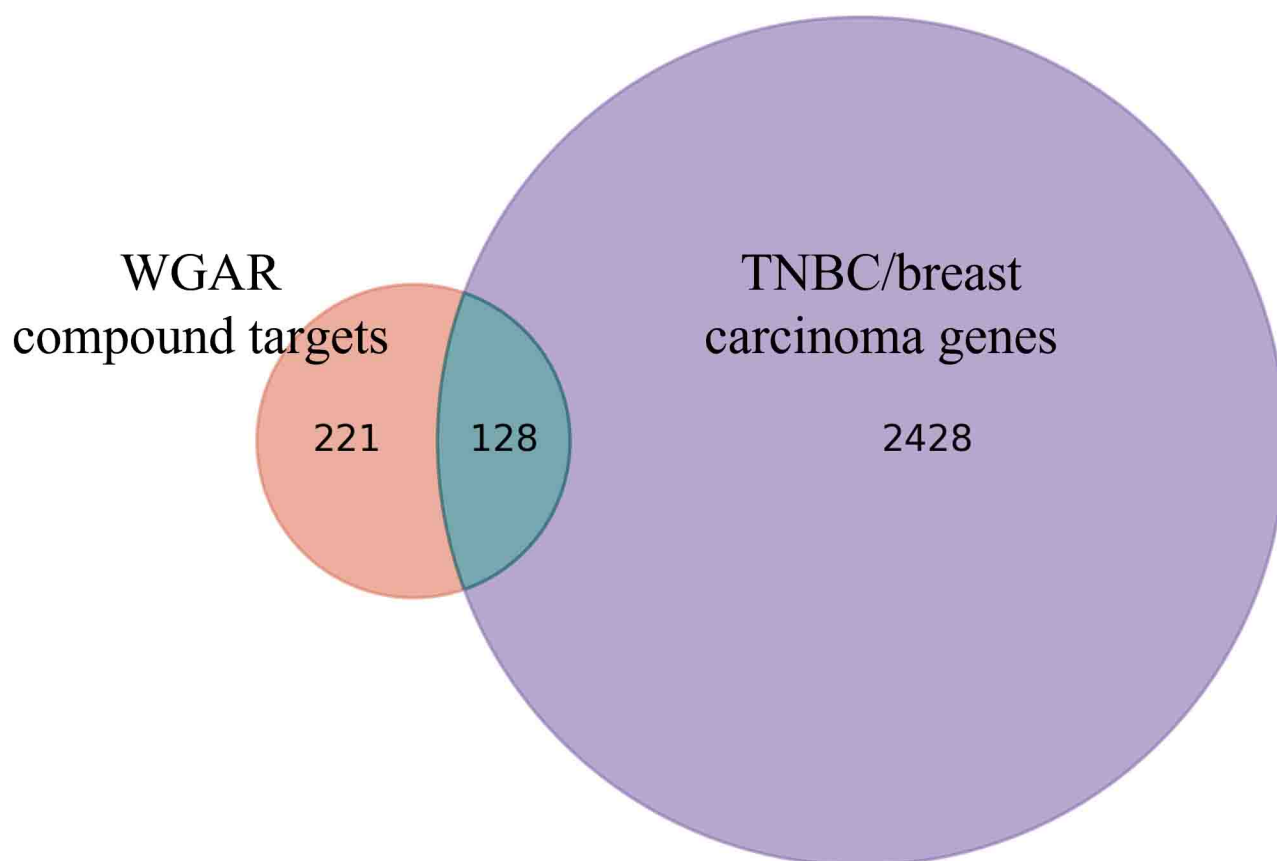


Fig. 2. Venn diagram of the intersection between wild ginseng-derived adventitious root (WGAR) compound-predicted targets and triple-negative breast cancer (TNBC)/breast carcinoma-associated genes. Of the 349 unique targets predicted for the 15 WGAR constituents by SwissTargetPrediction and the 2556 disease-associated genes retrieved from the Open Targets Platform, 128 were shared and taken forward as candidate targets.

3.6 Molecular Docking and Protein-Ligand Interactions

The four aglycone triterpenoids were docked *in silico* into six hub targets selected from the predicted PPI network described above (Fig. 6, Table 2). For STAT3 the search box was defined manually from the unoccupied SH2 phosphotyrosine-binding pocket, since the structure contains no co-crystallized ligand, so the docking protocol could not be validated by native-ligand re-docking for this receptor. All four aglycones gave consistent values at BCL2 (−7.31 to −8.23 kcal/mol) and HSP90AA1 (−7.25 to −8.32 kcal/mol), and the least favorable values in absolute terms were obtained at STAT3 (−4.90 to −6.06 kcal/mol); the numerically lowest values were obtained at EGFR (−7.21 to −9.58 kcal/mol), but they are not interpreted further because native-ligand re-docking did not reproduce the crystallographic pose at this receptor (Section 2.7). As set out in Section 2.7, these energies are not interpretable as relative affinities between receptors, and the ordering of receptors reported here is descriptive. The fatty acid and alkane constituents reached only −4.9 to −5.2 kcal/mol in predicted affinity against EGFR (Table 2). The seven triterpenoid glycosides, docked under the same protocol for

comparison, gave more favorable predicted binding energies than the aglycones at EGFR (−8.44 to −9.92 kcal/mol), BCL2 (−8.00 to −8.96 kcal/mol) and MMP9 (−6.29 to −8.20 kcal/mol), and less favorable ones at HSP90AA1 (−7.32 to −7.98 kcal/mol) and STAT3 (−3.25 to −5.62 kcal/mol), with all seven values positive at PTGS2 (Table 2). In the docked poses, PLIP analysis identified predicted contacts between glochidone and hydrophobic residues of the BCL2 homology 3 (BH3) groove (Phe104, Tyr108, and Leu137, among others; Fig. 7B); seven hydrophobic contacts were also predicted in the EGFR ATP-binding pocket (Fig. 7A), but they are listed in **Supplementary Table 11** without residue-level interpretation here, since the protocol validation does not support that pose. In the HSP90AA1 pose, glochidone was predicted to form one hydrogen bond with Ser50 (3.88 Å) and additional hydrophobic contacts (Fig. 7C), and ursolic acid was predicted to interact with Lys591 (hydrophobic) and Glu594 (hydrogen bond, 4.01 Å) in the STAT3 SH2 domain (Fig. 7D). The contacts detected in the four complexes for which PLIP identified defined ligand-protein interactions are listed in **Supplementary Table 11**, and the corresponding two-dimensional interaction diagrams are shown

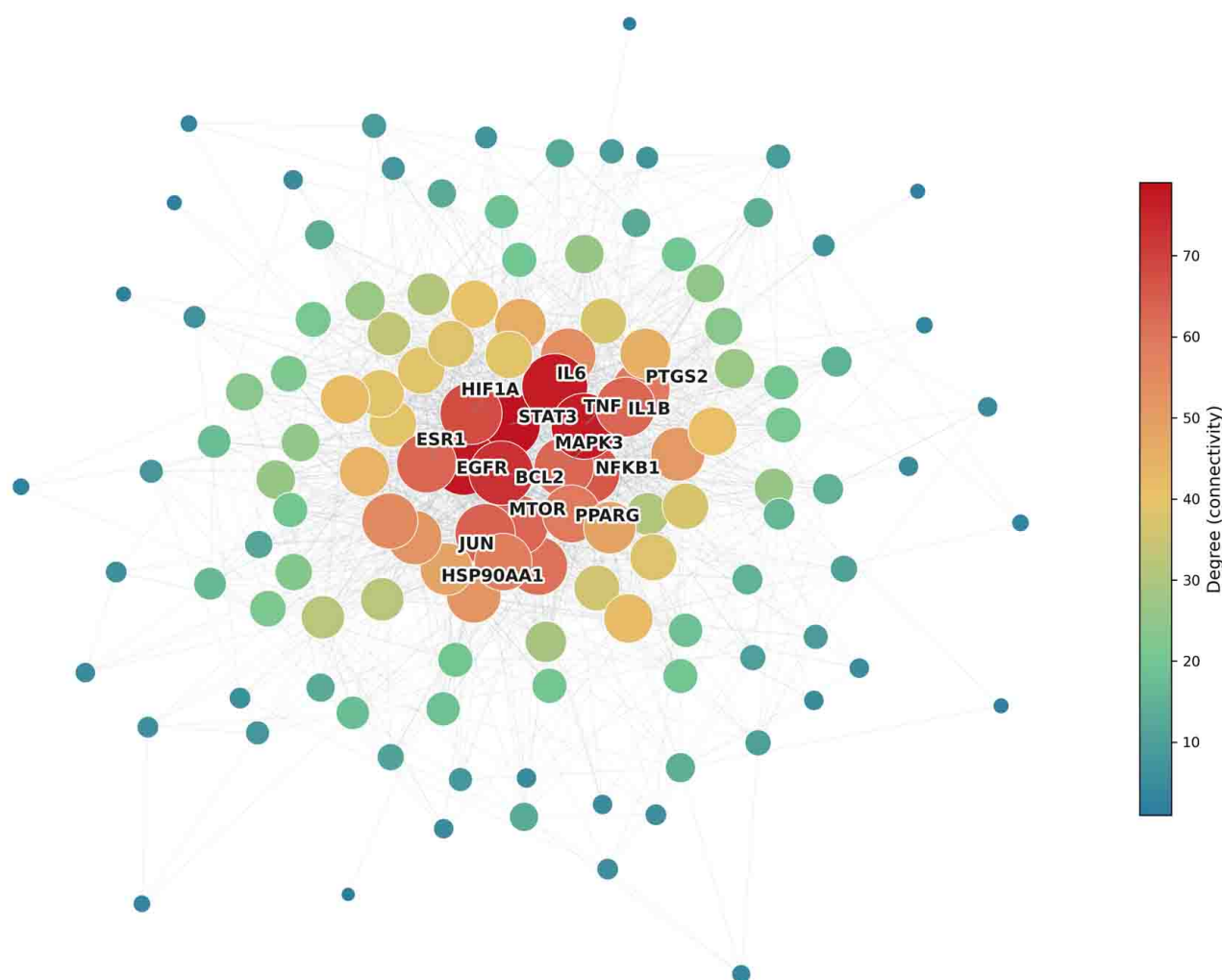


Fig. 3. *In silico* protein-protein interaction (PPI) network of the 128 predicted candidate targets of wild ginseng-derived adventitious root (WGAR) against triple-negative breast cancer (TNBC). The network was constructed computationally in the STRING database (128 nodes, 1623 edges). Node size and color are scaled to degree centrality, and the 15 highest-degree hub targets are labeled. Hub assignment reflects predicted network topology and was not verified experimentally.

in **Supplementary Fig. 1**. Because the docking protocol did not reproduce the crystallographic pose of erlotinib in the same EGFR structure (Section 2.7), the residue-level contacts predicted for the EGFR complexes are reported as predictions that the protocol validation does not support, whereas the corresponding contacts for BCL2 and HSP90AA1 are supported by re-docking to within 2 Å.

For PTGS2, most aglycones yielded positive predicted binding energies (ursolic acid, +7.32 kcal/mol), and glochidone-MMP9, although showing an intermediate predicted binding energy (−7.18 kcal/mol), showed no specific predicted ligand-protein interaction by PLIP analysis; the only coordination detected was that of the catalytic zinc of MMP9 by its own histidine residues, which is an intrinsic feature of the receptor rather than a ligand contact and is therefore not tabulated in **Supplementary Table 11**.

3.7 Predicted ADMET and Drug-Likeness Profiles

In the machine learning-based *in silico* profile, the aglycone triterpenoids satisfied on average 3.2 of the four Lipinski rules and showed a predicted QED of approximately 0.41, whereas the triterpenoid glycosides satisfied on average only 1.1 rules and showed a predicted QED of approximately 0.14 (Fig. 8, Table 3, **Supplementary Table 12**). The main predicted drug-likeness liability of the aglycones was a high computed logP (soyasapogenol C, 6.9; ursolic acid, 7.1; glochidone, 8.0), above the conventional limit of 5 for three of the four, whereas that of the glycosides was their high molecular weight (639 to 839 g/mol) and large number of hydrogen bond donors (7 to 10); predicted Caco-2 permeability was low for every constituent, the aglycones included.

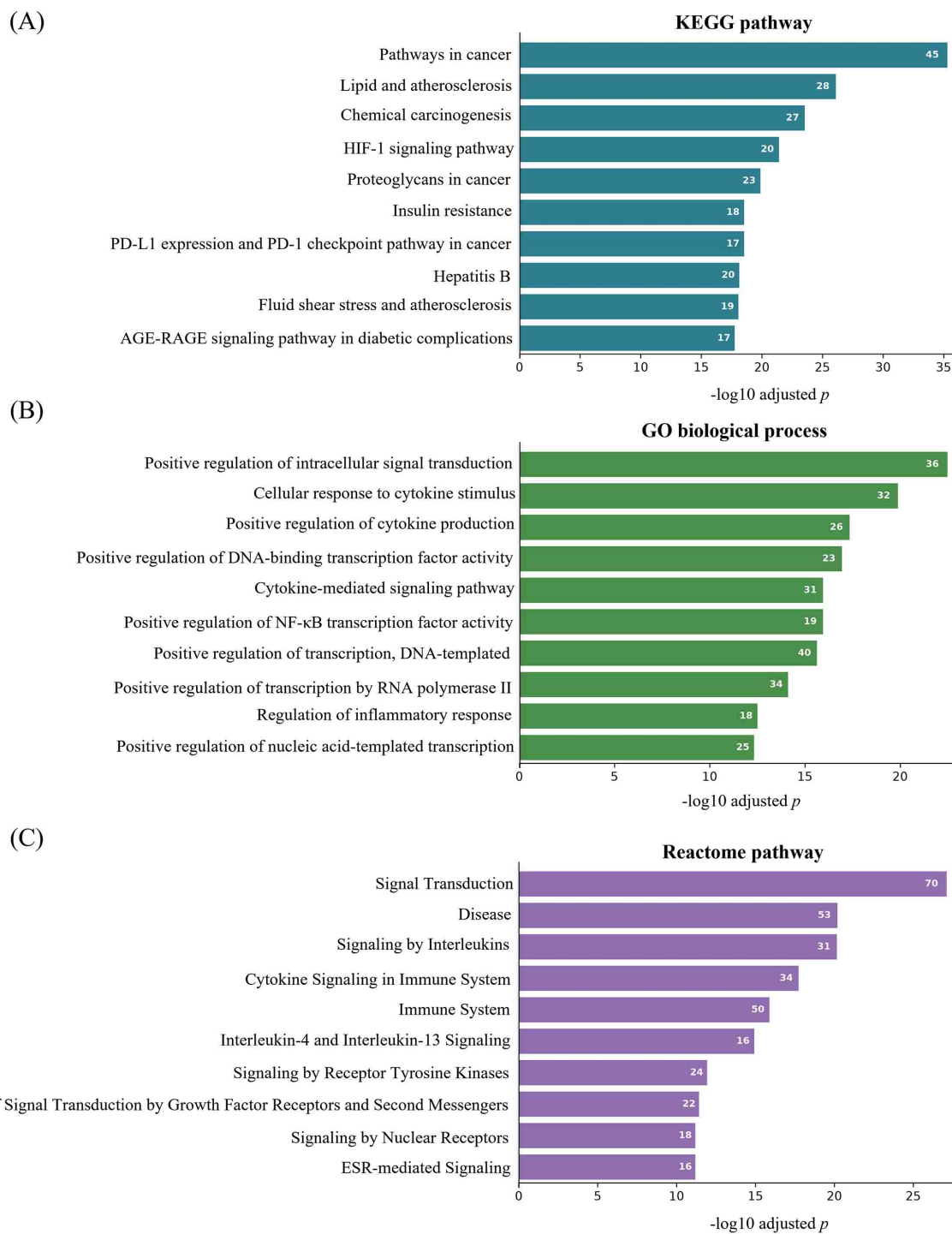


Fig. 4. *In silico* functional and pathway enrichment analysis of the 128 predicted candidate targets. The top 10 enriched terms are shown for (A) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, (B) Gene Ontology (GO) biological processes, and (C) Reactome pathways, ranked by $-\log_{10}$ adjusted p -value (Benjamini-Hochberg). The number at the end of each bar is the count of predicted candidate targets annotated to that term. All terms shown were significant at adjusted $p < 0.05$. Enrichment was computed from database annotations of the predicted target set and does not constitute experimental pathway validation.

4. Discussion

Before the findings are interpreted, the scope of this study should be stated explicitly. All results presented above are computational predictions generated from pub-

lic databases and structure-based modelling; no cells were treated, no proteins were assayed, and no new experimental measurement was performed in the present work. The experimental basis for interpreting them comes entirely from

Compound (15)

Hub target (top 25)

KEGG pathway (top 10)

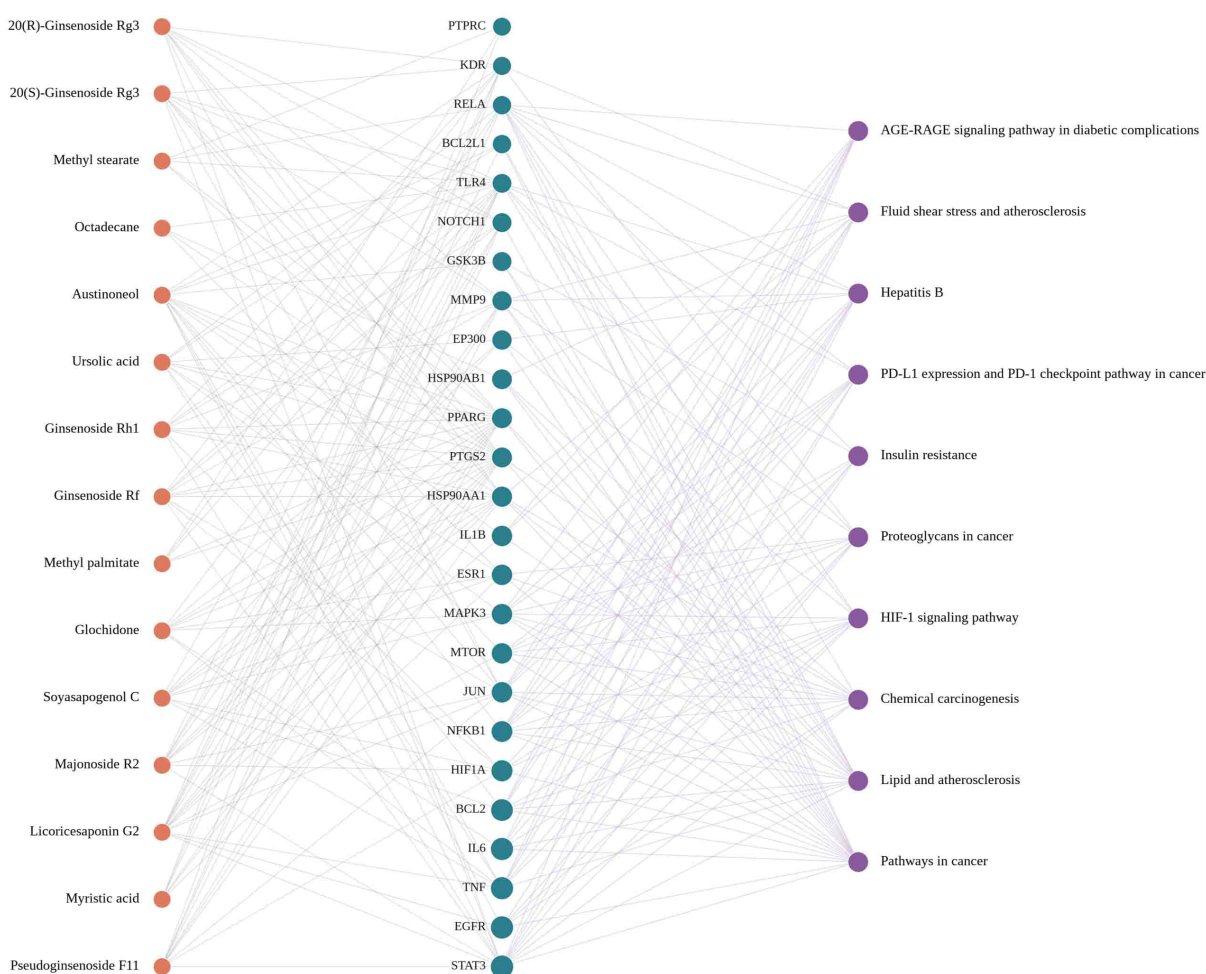


Fig. 5. *In silico* compound-target-pathway network integrating the 15 wild ginseng-derived adventitious root (WGAR) constituents (left), the 25 highest-degree predicted hub targets (center), and the 10 most significant Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (right). Edges connect each constituent to its predicted targets and each target to the pathways in which it participates. Hub node size is scaled to degree centrality in the predicted protein-protein interaction (PPI) network.

our earlier cell-based study of the same extract, in which WGAR reduced MDA-MB-231 viability (IC₅₀ 79 µg/mL), activated the caspase-9/caspase-3 cascade with subsequent PARP cleavage, lowered N-cadherin, β-catenin and Slug, reduced inhibitory Ser9 phosphorylation of GSK-3β, and suppressed secreted IL-6 and u-PA [10]. Accordingly, the interpretation below distinguishes throughout between predictions that correspond to those measured endpoints and predictions that, while biologically plausible, have not yet been examined experimentally in this model.

This study integrated *in silico* network pharmacology and molecular docking to place the previously validated anticancer activity of WGAR on a predicted systems-level map, and the analysis converged on a compact set of proliferation-, survival-, and inflammation-associated predicted hub targets, with EGFR and STAT3 the most central

in the network topology. This computational profile agrees in direction with our previous experimental finding that WGAR induces apoptosis and suppresses EMT in MDA-MB-231 TNBC cells, although directional agreement with a phenotype does not itself validate the individual predicted targets [10]. The correspondence is specific rather than general: the earlier study measured experimentally that WGAR downregulates the secreted inflammatory mediators interleukin (IL)-6 and urokinase-type plasminogen activator (u-PA) and acts through GSK-3β/β-catenin signaling, and each of these reappears in the present predicted network as the IL-6 inflammatory module, the GSK3B hub, and the β-catenin-linked pathways in cancer term. These three correspondences are the parts of the present analysis that rest on measured endpoints; the centrality of EGFR and STAT3, by contrast, was not examined in that study and remains a



Fig. 6. Heatmap of predicted molecular docking binding affinities (kcal/mol) between the four aglycone triterpenoids and six hub targets. Docking was performed *in silico* with AutoDock Vina. Values are computational estimates of binding and were not determined experimentally. Cell values are the lowest-energy Vina scores; more negative values (green) indicate stronger predicted binding and positive values (red) indicate no favorable binding. Targets: EGFR, epidermal growth factor receptor; BCL2, B-cell lymphoma 2; HSP90AA1, heat shock protein 90 alpha family class A member 1; STAT3, signal transducer and activator of transcription 3; MMP9, matrix metalloproteinase 9; PTGS2, prostaglandin-endoperoxide synthase 2.

prediction. Whereas that work experimentally established the phenotype and examined only two docking targets, the present analysis maps a broader predicted target-pathway network and nominates candidate constituent-level interactions that may underlie those effects. Ginseng triterpenoids are increasingly recognized as multi-target anti-cancer agents whose polypharmacology suits a heterogeneous disease such as TNBC [5,29].

The predicted centrality of STAT3 and EGFR is biologically plausible in the light of the TNBC literature, although this plausibility is contextual support rather than evidence that the WGAR constituents engage these proteins. The two nodes are not equally well supported by the prediction step itself. STAT3 was predicted as a target of eight of the 15 constituents, with a maximum probability of 0.90, and remained among the highest-degree nodes under every filtering criterion tested, whereas EGFR was predicted for two constituents only, at probabilities below 0.07, and was lost from the candidate set once a probability threshold was applied (Supplementary Table 2). The predicted role of EGFR therefore rests on its position in the interaction network rather than on the strength of the constituent-level prediction, and is the more provisional of the two. STAT3 is constitutively activated in TNBC and drives survival, proliferation, and invasion, functioning immediately downstream of RTK signaling [30]. EGFR is frequently overexpressed in basal-like and triple-negative tumors even without gene amplification and contributes to their aggressive

phenotype, yet single-agent EGFR inhibition has shown limited clinical efficacy in TNBC [31]. The STAT3-EGFR relationship is bidirectional in this setting, as STAT3 activation mediates resistance to EGFR tyrosine kinase inhibitors in TNBC, which both explains the limited benefit of targeting either node alone and supports the network-level rationale of the present approach [32]. Whether the WGAR aglycones in fact engage either node remains untested, and testing it - for example by measuring EGFR and STAT3 phosphorylation in WGAR-treated MDA-MB-231 cells - is the most direct experimental follow-up implied by this analysis. The predicted network was further organized around an inflammatory NF- κ B-IL-6 module comprising tumor necrosis factor (TNF), IL-6, NFKB1, and IL1B, consistent with the role of chronic inflammation and NF- κ B signaling in TNBC metastasis [33]. This module is the one part of the predicted topology with direct experimental support in the same model, since IL-6 secretion was suppressed to 0.13-fold on the proteome array and to 0.28-fold by independent ELISA at the 20% inhibitory concentration in WGAR-treated cells [10].

The enriched pathways reinforced these predicted target-level relationships and added a hypoxic and immune dimension to the predicted profile. Beyond the broad pathways in cancer term, HIF-1 signaling and proteoglycans in cancer were among the most significant KEGG pathways in the predicted target set, in line with the strong hypoxic drive of TNBC, where HIF-1 α sustains a positive-

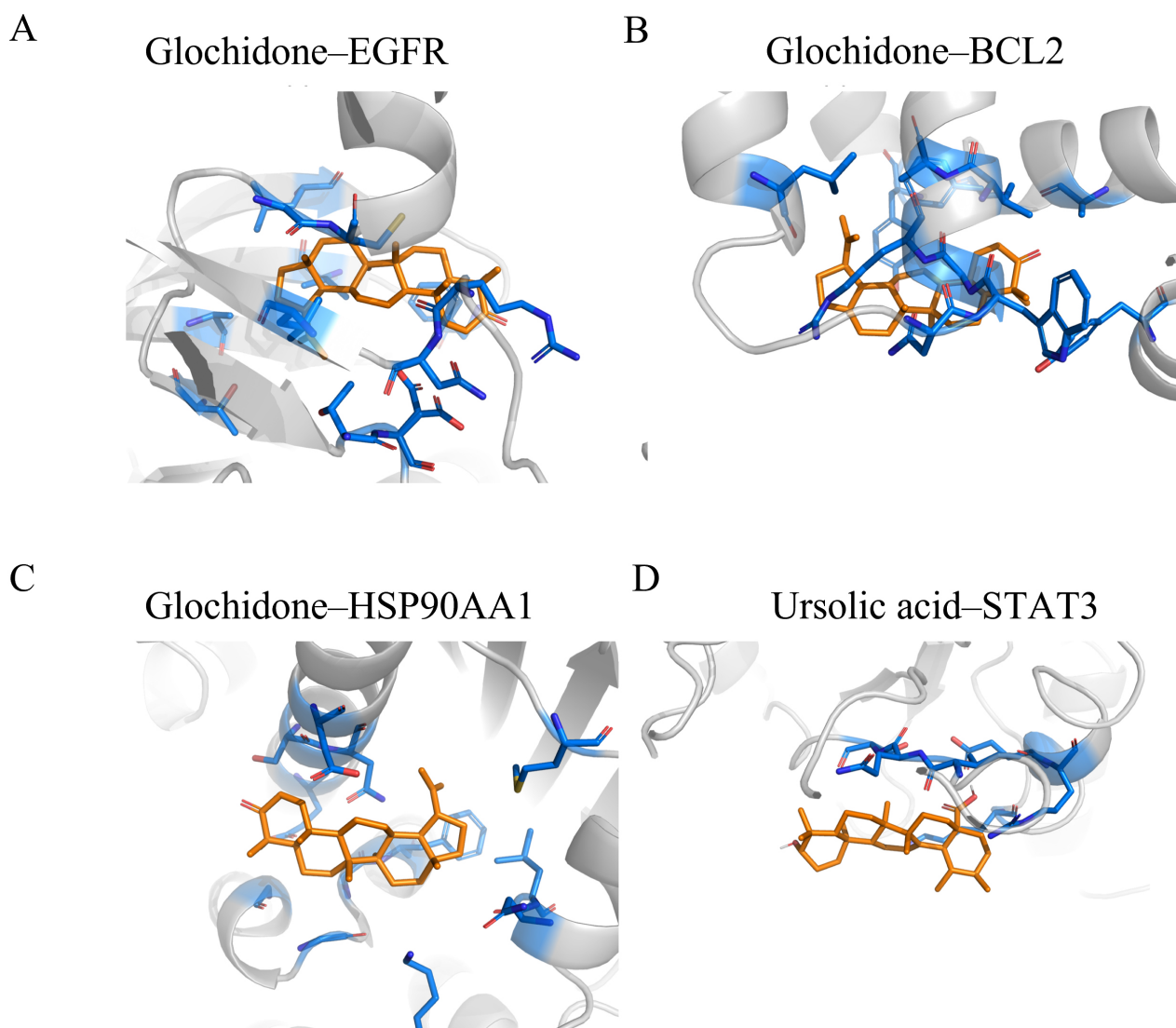


Fig. 7. Predicted binding poses of the aglycone triterpenoids in the predicted hub targets. (A) Glochidone in the ATP-binding pocket of epidermal growth factor receptor (EGFR); (B) glochidone in the BCL2 homology 3 (BH3) groove of B-cell lymphoma 2 (BCL2, -8.23 kcal/mol); (C) glochidone in the N-terminal pocket of heat shock protein 90 alpha family class A member 1 (HSP90AA1, -8.32 kcal/mol); (D) ursolic acid in the Src homology 2 (SH2) domain of signal transducer and activator of transcription 3 (STAT3, -6.06 kcal/mol). Ligands are shown as orange sticks and interacting protein residues as blue sticks; the protein backbone is shown as a gray cartoon. Binding affinities were calculated *in silico* with AutoDock Vina, and the poses shown are computational predictions rather than experimentally determined structures. Native-ligand re-docking reproduced the crystallographic pose to within $2\text{ \AA}$ for BCL2, HSP90AA1 and MMP9 but not for EGFR (Section 2.7), so the pose shown in panel A is a prediction that the protocol validation does not support.

feedback circuit that promotes angiogenesis, invasion, and malignant progression [34,35]. The PD-L1 expression and PD-1 checkpoint pathway in cancer was also significantly enriched (adjusted $p = 2.95 \times 10^{-19}$, 17 genes), which is notable because TNBC is the breast cancer subtype for which immune checkpoint inhibitors have entered clinical practice, albeit with benefit confined to defined patient subgroups [36,37]. Because several WGAR constituents were predicted to target genes annotated to this pathway,

the extract may intersect with immune-checkpoint biology; this remains a hypothesis generated by pathway annotation alone, but it is directionally consistent with reports that ginsenoside Rg3 targets FUT8-mediated PD-L1 N-glycosylation to enhance atezolizumab therapy and that ginsenosides synergize with PD-1 blockade [7,8].

The docking predictions provided a structure-based complement to the topological analysis, both being computational and therefore not mutually independent forms of

Table 2. Predicted molecular docking binding affinities (kcal/mol) of the four aglycone triterpenoids, seven triterpenoid glycosides, and four fatty acid or alkane constituents against six hub target proteins, calculated *in silico* with AutoDock Vina (version 1.2.5).

Compound	Class	EGFR	BCL2	HSP90AA1	STAT3	MMP9	PTGS2
Glochidone	aglycone	-9.583	-8.227	-8.324	-5.846	-7.179	1.564
Ursolic acid	aglycone	-8.999	-7.945	-7.843	-6.063	-6.984	7.317
Soyasapogenol C	aglycone	-8.329	-8.049	-7.247	-5.783	-6.341	7.223
Austinoneol	aglycone	-7.212	-7.313	-7.392	-4.898	-5.691	-3.886
Licoricesaponin G2	glycoside	-9.919	-8.314	-7.339	-4.061	-7.501	12.701
Majonoside R2	glycoside	-9.463	-8.019	-7.637	-5.165	-8.200	12.936
20(R)-Ginsenoside Rg3	glycoside	-9.223	-8.949	-7.514	-3.254	-7.908	6.949
20(S)-Ginsenoside Rg3	glycoside	-8.998	-8.312	-7.978	-5.623	-7.031	6.346
Pseudoginsenoside F11	glycoside	-8.709	-8.001	-7.771	-5.172	-7.718	13.396
Ginsenoside Rh1	glycoside	-8.547	-8.212	-7.323	-5.010	-7.681	2.419
Ginsenoside Rf	glycoside	-8.435	-8.964	-7.485	-5.303	-6.286	7.858
Myristic acid	fatty/alkane	-5.210	-	-	-4.121	-	-
Methyl palmitate	fatty/alkane	-4.943	-	-	-3.657	-	-
Octadecane	fatty/alkane	-5.088	-	-	-3.88	-	-
Methyl stearate	fatty/alkane	-5.142	-	-	-	-	-

The Protein Data Bank (PDB) crystal structures used were EGFR (epidermal growth factor receptor, 1M17), BCL2 (B-cell lymphoma 2, 6O0K), HSP90AA1 (heat shock protein 90 alpha family class A member 1, 1YET), STAT3 (signal transducer and activator of transcription 3, 1BG1), MMP9 (matrix metalloproteinase 9, 1GKC), and PTGS2 (prostaglandin-endoperoxide synthase 2, 3LN1). More negative values indicate stronger predicted binding. The triterpenoid glycosides were docked under the same protocol for comparison; because the re-docked scores of five experimentally confirmed native ligands span 5.4 kcal/mol (Section 2.7), values are comparable between compounds at the same receptor but not between receptors.

Blank cells (-) indicate combinations that were not docked. Binding-pose interpretation focused on the four aglycone triterpenoids, since the glycoside constituents (compounds 83, 86, 87, 93, 94, 102, and 103) show greater conformational flexibility and less favorable predicted drug-likeness (Section 3.7); all constituents were nevertheless docked under the same protocol, and their predicted energies are reported here.

evidence. Because the docking targets were themselves selected from the network hubs, the agreement between high centrality and favorable docking is not an independent validation but rather a check for internal consistency, and that check is itself limited, since it compares predicted energies between different receptors, which Section 2.7 shows is not interpretable as a comparison of affinity. The docking did, however, discriminate between constituents at a fixed target: the fatty acid and alkane constituents of the same extract reached only -4.94 to -5.21 kcal/mol at EGFR against -7.21 kcal/mol or better for the aglycones, and because the receptor is identical in both cases this separation cannot arise from the way the receptor was selected. They were also predicted to bind the anti-apoptotic protein BCL2 and the molecular chaperone HSP90AA1 with favorable affinities. The predicted energies obtained at STAT3 were weak in absolute terms (-4.90 to -6.06 kcal/mol) and came from a manually defined search box that could not be validated by native-ligand re-docking, so the standing of STAT3 in this analysis rests on its network centrality and target-prediction support rather than on the docking result. Both are mechanistically relevant to TNBC: MCL-1, a member of the BCL2 family, is a clinically targetable apoptotic vulner-

ability in breast cancer [38], and HSP90 stabilizes multiple oncogenic clients and supports tumor growth, migration, and invasion [39]. Predicted engagement of BCL2 and HSP90 by the aglycones would be consistent with the pro-apoptotic phenotype previously observed for WGAR, which proceeded through the intrinsic caspase-9/caspase-3 cascade [10]; neither protein, however, was measured in that study, so this correspondence links a prediction to a phenotype rather than to a matching target-level measurement.

The four aglycones did not outperform the triterpenoid glycosides in predicted binding energy, which was more favorable for the glycosides at EGFR, BCL2 and MMP9 and less favorable at HSP90AA1, STAT3 and PTGS2; the two classes separated in predicted drug-likeness alone. The aglycones were predicted to bind the hub targets favorably and satisfied most Lipinski criteria, whereas the glycosides combined high molecular weight, an excess of hydrogen bond donors, and very low predicted intestinal absorption (Section 3.7). That the glycosides, at 639 to 839 g/mol against 415 to 457 g/mol for the aglycones, scored more favorably at the three receptors with the more open binding sites and were alone in giving uniformly positive val-

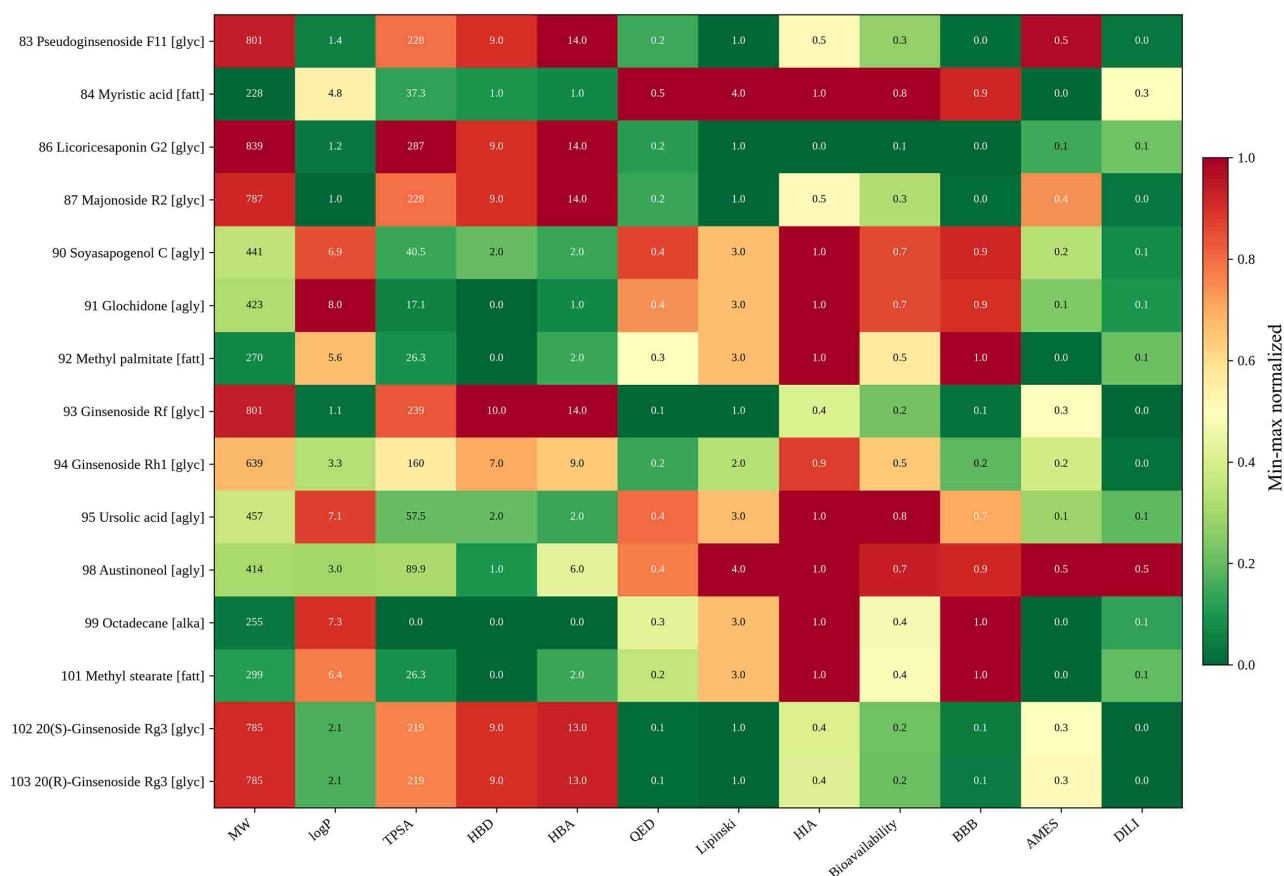


Fig. 8. Heatmap of predicted ADMET and drug-likeness properties of the 15 WGAR constituents. Twelve representative indices predicted with `admet_ai` are shown: molecular weight (MW), lipophilicity (logP), topological polar surface area (TPSA), hydrogen bond donors (HBD), hydrogen bond acceptors (HBA), quantitative estimate of drug-likeness (QED), number of Lipinski rules satisfied, human intestinal absorption (HIA), bioavailability, blood-brain barrier (BBB) permeability, Ames mutagenicity (AMES), and drug-induced liver injury (DILI). Cell color is min-max normalized per column; raw values are labeled. The class of each constituent is indicated in brackets: agly, aglycone triterpenoid; glyc, triterpenoid glycoside; fatt, fatty acid; alka, alkane. All values are machine learning predictions and were not measured experimentally.

ues in the enclosed PTGS2 channel indicates that the predicted binding energies track molecular size and contact area rather than affinity, which is why they are not compared across receptors. The drug-likeness advantage of the aglycones is therefore carried by molecular weight and hydrogen-bond donor count, not by lipophilicity, which is itself a liability of this class (Section 3.7). The poor predicted drug-likeness of the glycosides reflects their unmodified form and does not exclude a contribution *in vivo*, because the parent glycosides are extensively deglycosylated in the gastrointestinal tract, largely by gut microbiota, to less-glycosylated forms such as compound K and 20(S)-protopanaxadiol, which are more potent than the parent glycosides and are the species that reach the systemic circulation [40,41]. On this view the glycosides act as precursors that deliver the aglycone forms bearing the target-level activity, so restricting the docking analysis to the aglycones targets the species most likely to engage the proteins directly, rather than dismissing the glycosides as inactive.

PTGS2 and MMP9 were retained as inflammation- and invasion-associated hubs to test whether the constituents would engage them, and they served as informative negative cases: the predicted ranking is a computational estimate rather than a measure of activity, as shown by the discordance between binding energy and profiled interactions for these two targets, where a formally intermediate score for the glochidone-MMP9 complex corresponded to no specific contact and most PTGS2 scores were unfavorable.

5. Limitations

Several considerations qualify these predictions, none of which was verified experimentally in the present work. Network pharmacology assembles static, database-derived interactions and cannot capture tissue context, expression level, or signaling dynamics, so its hub assignments are hypotheses about relative importance rather than validated dependencies [42]. The high-degree hubs recovered here

Table 3. Key absorption, distribution, metabolism, excretion, and toxicity (ADMET) and drug-likeness indices of the 15 constituents predicted in silico with the machine learning-based predictor *admet_ai* (version 2.0.1).

No.	Compound	MW	logP	Lipinski	QED	HIA	BBB	hERG
83	Pseudoginsenoside F11	801.0	1.36	1	0.17	0.51	0.04	0.59
84	Myristic acid	228.4	4.77	4	0.49	1.00	0.92	0.22
86	Licoricesaponin G2	838.9	1.22	1	0.16	0.00	0.03	0.09
87	Majonoside R2	787.0	0.97	1	0.16	0.51	0.04	0.59
90	Soyasapogenol C	440.7	6.92	3	0.44	1.00	0.91	0.89
91	Glochidone	422.7	8.01	3	0.39	1.00	0.90	0.72
92	Methyl palmitate	270.5	5.64	3	0.30	1.00	1.00	0.34
93	Ginsenoside Rf	801.0	1.12	1	0.11	0.41	0.05	0.64
94	Ginsenoside Rh1	638.9	3.30	2	0.16	0.88	0.21	0.71
95	Ursolic acid	456.7	7.09	3	0.41	1.00	0.71	0.65
98	Austinoneol	414.5	3.05	4	0.40	1.00	0.92	0.31
99	Octadecane	254.5	7.27	3	0.27	1.00	1.00	0.28
101	Methyl stearate	298.5	6.42	3	0.24	1.00	1.00	0.36
102	20(S)-Ginsenoside Rg3	785.0	2.15	1	0.12	0.41	0.07	0.69
103	20(R)-Ginsenoside Rg3	785.0	2.15	1	0.12	0.42	0.07	0.68

The table includes Lipinski rules satisfied, quantitative estimate of drug-likeness (QED), lipophilicity (logP), human intestinal absorption (HIA), blood-brain barrier (BBB) permeability, and human ether-à-go-go-related gene (hERG) inhibition. Ames mutagenicity and drug-induced liver injury are reported for all 15 constituents in Fig. 8, and the full predicted profile is given in **Supplementary Table 12**.

are also pleiotropic proteins that recur across many disease networks, and the degree distribution of interaction databases is itself shaped by how unevenly individual proteins have been studied, so high centrality may reflect research attention as much as biological importance and does not establish target specificity [43]. Target prediction and docking are hypothesis-generating steps that carry a risk of false positives, and the constituents were tentatively identified without authentic-standard confirmation. Two features of the prediction step further limit its resolution. All probability-weighted predictions were retained, so the candidate set favors recall over precision, and the sensitivity analysis shows that the hub ranking of EGFR in particular depends on this choice (**Supplementary Table 2**). The 200-character input limit of SwissTargetPrediction also required stereochemistry to be removed for target prediction, so 20(S)- and 20(R)-ginsenoside Rg3 collapsed to a single query and the 15 constituents yielded 14 distinct prediction queries; stereochemistry-dependent differences in target profile therefore cannot be resolved for the ginsenosides, although isomeric structures were used for docking. These results therefore define priorities for experimental testing rather than established mechanisms.

6. Conclusions

This *in silico* study provides a systems-level hypothesis for the experimentally established anti-TNBC activity of WGAR. Integrating *in silico* target prediction, protein-protein interaction analysis, functional enrichment, molecular docking, and machine learning-based ADMET profil-

ing, the analysis predicts that the aglycone triterpenoids of WGAR may act on a network centered on STAT3 and EGFR together with an inflammatory NF- κ B-IL-6 module and a hypoxic HIF-1 axis, and that these predicted targets and pathways are consistent in direction with the proapoptotic and anti-EMT effects previously observed experimentally in TNBC cells. Of the two central nodes, STAT3 was the more robust to the filtering criteria applied, whereas the predicted centrality of EGFR was threshold-dependent.

These conclusions rest on computational prediction without new experimental validation, and the constituents were tentatively identified without authentic-standard confirmation; the target and docking analyses are therefore hypothesis-generating and may include false positives. The results are best used to prioritize subsequent experimental work, in particular the evaluation of EGFR and STAT3 engagement and downstream pathway activity by the aglycone constituents, which would test the mechanistic hypotheses proposed here.

Abbreviations

3D, three-dimensional; ADMET, absorption, distribution, metabolism, excretion, and toxicity; BCL2, B-cell lymphoma 2; BH3, BCL2 homology 3; CID, compound identifier; C α , alpha-carbon; CYP, cytochrome P450; EGFR, epidermal growth factor receptor; EMT, epithelial-mesenchymal transition; ER, estrogen receptor; ETKDGv3, Experimental-Torsion-Knowledge Distance Geometry version 3; FUT8, fucosyltransferase 8; GO, Gene Ontology; GSK-3 β , glycogen synthase kinase-3 β ;

HER2, human epidermal growth factor receptor 2; HIF-1 α , hypoxia-inducible factor-1 α ; HSP90AA1, heat shock protein 90 alpha family class A member 1; IC50, half-maximal inhibitory concentration; ICI, immune checkpoint inhibitors; IL, interleukin; KEGG, Kyoto Encyclopedia of Genes and Genomes; LC-MS/MS, liquid chromatography-tandem mass spectrometry; MMFF94, Merck Molecular Force Field 94; MMP, matrix metalloproteinase; NF- κ B, nuclear factor κ B; PARP, poly(ADP-ribose) polymerase; PD-1, programmed cell death protein 1; PDB, Protein Data Bank; PDBQT, Protein Data Bank, Partial Charge, and Atom Type; PD-L1, programmed death-ligand 1; PLIP, Protein-Ligand Interaction Profiler; PPI, protein-protein interaction; PR, progesterone receptor; PTGS2, prostaglandin-endoperoxide synthase 2; QED, quantitative estimate of drug-likeness; RCSB, Research Collaboratory for Structural Bioinformatics; RTK, receptor tyrosine kinases; SH2, Src homology 2; SMILES, Simplified Molecular-Input Line-Entry System; STAT3, signal transducer and activator of transcription 3; STRING, Search Tool for the Retrieval of Interacting Genes/Proteins; TNBC, triple-negative breast cancer; TNF, tumor necrosis factor; u-PA, urokinase-type plasminogen activator; WGAR, wild ginseng-derived adventitious roots; YAP, Yes-associated protein.

Availability of Data and Materials

All data generated and analyzed during this study are included in this article. The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Author Contributions

Conceptualization, SYL; methodology, MKJ; investigation, MKJ; formal analysis, MKJ; writing—original draft preparation, MKJ; writing—review and editing, SYL; supervision, SYL; funding acquisition, SYL. Both authors contributed to editorial changes in the manuscript. Both authors have read and agreed to the published version of the manuscript. Both authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

Not applicable.

Acknowledgment

The authors thank MIRAE FARM Agricultural Corporation Co., Ltd. (Gokseong-gun, Jeollanam-do, Republic of Korea) for kindly providing the wild ginseng-derived adventitious root material, commercially known as “Nunkkot Sansam” (Snow Flower Wild Ginseng), used in this study.

Funding

This research was supported by the ANCHOR program through the Jeollanamdo ANCHOR center, funded by the Ministry of Education (MOE) and the Jeollanamdo, Republic of Korea (2026-ANCHOR-14-003).

Conflicts of Interest

The authors declare no conflicts of interest. MIRAE FARM Agricultural Corporation Co., Ltd., which provided the wild ginseng-derived adventitious root material, had no role in the study design, data collection, analysis, interpretation of results, or manuscript preparation, and the authors have no financial or commercial relationship with the company. Given her role as the Guest Editor, Su-yun Lyu had no involvement in the peer-review of this article and has no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to Mehmet Ozaslan.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at <https://doi.org/10.31083/IJP56325>.

References

- [1] Torres Quintas S, Canha-Borges A, Oliveira MJ, Sarmiento B, Castro F. Special Issue: Nanotherapeutics in Women's Health Emerging Nanotechnologies for Triple-Negative Breast Cancer Treatment. *Small* (Weinheim an Der Bergstrasse, Germany). 2024; 20: e2300666. <https://doi.org/10.1002/sml.202300666>
- [2] Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. *The New England Journal of Medicine*. 2010; 363: 1938–1948. <https://doi.org/10.1056/NEJMra1001389>
- [3] Masci D, Naro C, Puxeddu M, Urbani A, Sette C, La Regina G, et al. Recent Advances in Drug Discovery for Triple-Negative Breast Cancer Treatment. *Molecules* (Basel, Switzerland). 2023; 28: 7513. <https://doi.org/10.3390/molecules28227513>
- [4] Yao H, He G, Yan S, Chen C, Song L, Rosol TJ, et al. Triple-negative breast cancer: is there a treatment on the horizon? *Oncotarget*. 2017; 8: 1913–1924. <https://doi.org/10.18632/oncotarget.12284>
- [5] Yang Y, Nan Y, Du Y, Liu W, Ning N, Chen G, et al. Ginsenosides in cancer: Proliferation, metastasis, and drug resistance. *Biomedicine & Pharmacotherapy = Biomedecine & Pharmacotherapie*. 2024; 177: 117049. <https://doi.org/10.1016/j.biopha.2024.117049>
- [6] Liu H, Ye H, Han G. Ginsenoside Rh2 inhibits epithelial-mesenchymal transition in triple-negative breast cancer cell via disrupting the HIF-1 α /YAP nuclear translocation axis. *International Immunopharmacology*. 2026; 184: 116917. <https://doi.org/10.1016/j.intimp.2026.116917>
- [7] Tian M, Li X, Yang J, Bai X, Deng R, Ye W, et al. Ginsenoside Rg3 improves atezolizumab immune checkpoint therapy in triple-negative breast cancer by targeting FUT8-mediated PD-L1 N-glycosylation. *Phytomedicine: International Journal of Phytotherapy and Phytomedicine*. 2026; 156: 158272. <https://doi.org/10.1016/j.phymed.2026.158272>
- [8] Luo N, Li X, Hu Y, Guo L, Zhang K. Ginsenoside Suppresses Triple-Negative Breast Cancer Growth and Synergizes With Programmed Cell Death Protein 1 (PD-1) Immunotherapy

- Through Gut Microbiota and Metabolic Regulation. *Phytotherapy Research*. PTR. 2026; 40: 3190–3211. <https://doi.org/10.1002/ptr.70278>
- [9] Liu K, Li P, Li W. Methyl Jasmonate Enhances Saponin Accumulation in Cultured *Panax notoginseng* Adventitious Roots. *Plants* (Basel, Switzerland). 2025; 14: 3462. <https://doi.org/10.3390/plants14223462>
 - [10] Hong CE, Le D, Lee M, Lyu SY. Pro-Apoptotic and Anti-EMT Activity of Wild Ginseng Adventitious Root Extract in MDA-MB-231 TNBC Cells: Association with GSK-3 β / β -Catenin Signaling. *Pharmaceuticals* (Basel, Switzerland). 2026; 19: 216. <https://doi.org/10.3390/ph19020216>
 - [11] Zhao L, Zhang H, Li N, Chen J, Xu H, Wang Y, et al. Network pharmacology, a promising approach to reveal the pharmacology mechanism of Chinese medicine formula. *Journal of Ethnopharmacology*. 2023; 309: 116306. <https://doi.org/10.1016/j.jep.2023.116306>
 - [12] Habtemariam S. Natural Products as a Pipeline for Next-Generation Neurodegenerative Drugs: From Single-Target Failure to Multi-Target Opportunity in Alzheimer's and Parkinson's Disease. *Molecules* (Basel, Switzerland). 2026; 31: 1489. <https://doi.org/10.3390/molecules31091489>
 - [13] Wu Y, Huang J, Guo J, Lian W, Suo M. Drug repurposing in traditional Chinese medicine: from empirical wisdom to modern therapeutic strategies. *Frontiers in Pharmacology*. 2025; 16: 1631727. <https://doi.org/10.3389/fphar.2025.1631727>
 - [14] 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: W357–W364. <https://doi.org/10.1093/nar/gkz382>
 - [15] Kittelson KS, Martins ACF, Santos RP, Celante G, da Silva Gomes R. Reproducibility, validation, and failure modes across classical and AI-driven molecular docking. *Journal of Computer-aided Molecular Design*. 2026; 40: 137. <https://doi.org/10.1007/s10822-026-00849-8>
 - [16] Swanson K, Walther P, Leitz J, Mukherjee S, Wu JC, Shivaraine RV, et al. ADMET-AI: a machine learning ADMET platform for evaluation of large-scale chemical libraries. *Bioinformatics* (Oxford, England). 2024; 40: btae416. <https://doi.org/10.1093/bioinformatics/btae416>
 - [17] Kim S. Exploring Chemical Information in PubChem. *Current Protocols*. 2021; 1: e217. <https://doi.org/10.1002/cpz1.217>
 - [18] Ochoa D, Hercules A, Carmona M, Suveges D, Baker J, Malangone C, et al. The next-generation Open Targets Platform: reimaged, redesigned, rebuilt. *Nucleic Acids Research*. 2023; 51: D1353–D1359. <https://doi.org/10.1093/nar/gkac1046>
 - [19] Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Research*. 2023; 51: D638–D646. <https://doi.org/10.1093/nar/gkac1000>
 - [20] Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*. 2013; 14: 128. <https://doi.org/10.1186/1471-2105-14-128>
 - [21] Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics*. 2000; 25: 25–29. <https://doi.org/10.1038/75556>
 - [22] Kanehisa M. Toward understanding the origin and evolution of cellular organisms. *Protein Science: a Publication of the Protein Society*. 2019; 28: 1947–1951. <https://doi.org/10.1002/pro.3715>
 - [23] Gillespie M, Jassal B, Stephan R, Milacic M, Rothfels K, Senff-Ribeiro A, et al. The reactome pathway knowledgebase 2022. *Nucleic Acids Research*. 2022; 50: D687–D692. <https://doi.org/10.1093/nar/gkab1028>
 - [24] Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*. 2010; 31: 455–461. <https://doi.org/10.1002/jcc.21334>
 - [25] Eberhardt J, Santos-Martins D, Tillack AF, Forli S. AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *Journal of Chemical Information and Modeling*. 2021; 61: 3891–3898. <https://doi.org/10.1021/acs.jcim.1c00203>
 - [26] Goodsell DS, Zardecki C, Di Costanzo L, Duarte JM, Hudson BP, Persikova I, et al. RCSB Protein Data Bank: Enabling biomedical research and drug discovery. *Protein Science: a Publication of the Protein Society*. 2020; 29: 52–65. <https://doi.org/10.1002/pro.3730>
 - [27] O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: An open chemical toolbox. *Journal of Cheminformatics*. 2011; 3: 33. <https://doi.org/10.1186/1758-2946-3-33>
 - [28] Adasme MF, Linnemann KL, Bolz SN, Kaiser F, Salentin S, Haupt VJ, et al. PLIP 2021: expanding the scope of the protein-ligand interaction profiler to DNA and RNA. *Nucleic Acids Research*. 2021; 49: W530–W534. <https://doi.org/10.1093/nar/gkz382>
 - [29] Cui D, Zhang C, Zhang L, Zheng J, Wang J, He L, et al. Natural anti-cancer products: insights from herbal medicine. *Chinese Medicine*. 2025; 20: 82. <https://doi.org/10.1186/s13020-025-01124-y>
 - [30] Qin JJ, Yan L, Zhang J, Zhang WD. STAT3 as a potential therapeutic target in triple negative breast cancer: a systematic review. *Journal of Experimental & Clinical Cancer Research: CR*. 2019; 38: 195. <https://doi.org/10.1186/s13046-019-1206-z>
 - [31] Shapira I, Lee A, Vora R, Budman DR. P53 mutations in triple negative breast cancer upregulate endosomal recycling of epidermal growth factor receptor (EGFR) increasing its oncogenic potency. *Critical Reviews in Oncology/hematology*. 2013; 88: 284–292. <https://doi.org/10.1016/j.critrevonc.2013.05.003>
 - [32] Jin LL, Lu HJ, Shao JK, Wang Y, Lu SP, Huang BF, et al. Relevance and mechanism of STAT3/miR-221-3p/Fascin-1 axis in EGFR TKI resistance of triple-negative breast cancer. *Molecular and Cellular Biochemistry*. 2024; 479: 3037–3047. <https://doi.org/10.1007/s11010-023-04907-9>
 - [33] Lim YX, Lin H, Chu T, Lim YP. WBP2 promotes BTBR mRNA stability to drive migration and invasion in triple-negative breast cancer via NF- κ B activation. *Molecular Oncology*. 2022; 16: 422–446. <https://doi.org/10.1002/1878-0261.13048>
 - [34] Liu L, Bai J, Hu L, Jiang D. Hypoxia-mediated activation of hypoxia-inducible factor-1 α in triple-negative breast cancer: A review. *Medicine*. 2023; 102: e35493. <https://doi.org/10.1097/MD.00000000000035493>
 - [35] Son S, Jo S, Lim H, Kang MG, Lee JH, Lee CJ, et al. CCN3/IMP3/HIF1 α positive feedback loop enhances malignant progression of triple-negative breast cancer. *Cellular Signaling*. 2025; 136: 112179. <https://doi.org/10.1016/j.cellsig.2025.112179>
 - [36] Corti C, Koca B, Rahman T, Mittendorf EA, Tolaney SM. Recent Advances in Immune Checkpoint Inhibitors for Triple-Negative Breast Cancer. *ImmunoTargets and Therapy*. 2025; 14: 339–357. <https://doi.org/10.2147/ITT.S495751>
 - [37] Lin S, Fu B, Khan M. Identifying subgroups deriving the most benefit from PD-1 checkpoint inhibition plus chemotherapy in advanced metastatic triple-negative breast cancer: a systematic review and meta-analysis. *World Journal of Surgical Oncology*. 2024; 22: 346. <https://doi.org/10.1186/s12957-024-03424-3>
 - [38] Winder ML, Campbell KJ. MCL-1 is a clinically targetable vulnerability in breast cancer. *Cell Cycle* (Georgetown, Tex.). 2022; 21: 1439–1455. <https://doi.org/10.1080/15384101.2022.>

2054096

- [39] Suwannalert P, Panpinyaporn P, Wantanachaisaeng P, Teepaibul T, Khumsri W, Koomsang T, et al. A Potential Combination of Targeting HSP90 and mTOR in Breast Cancer Cell Growth, Migration, and Invasion Through Inhibiting AKT Phosphorylation and F-actin Organization. *Anticancer Research*. 2024; 44: 2555–2565. <https://doi.org/10.21873/anticancer.17061>
- [40] Morshed MN, Akter R, Karim MR, Iqbal S, Kang SC, Yang DC. Bioconversion, Pharmacokinetics, and Therapeutic Mechanisms of Ginsenoside Compound K and Its Analogues for Treating Metabolic Diseases. *Current Issues in Molecular Biology*. 2024; 46: 2320–2342. <https://doi.org/10.3390/cimb46030148>
- [41] Fan W, Fan L, Wang Z, Mei Y, Liu L, Li L, et al. Rare ginsenosides: A unique perspective of ginseng research. *Journal of Advanced Research*. 2024; 66: 303–328. <https://doi.org/10.1016/j.jare.2024.01.003>
- [42] Nogales C, Mamdouh ZM, List M, Kiel C, Casas AI, Schmidt HHHW. Network pharmacology: curing causal mechanisms instead of treating symptoms. *Trends in Pharmacological Sciences*. 2022; 43: 136–150. <https://doi.org/10.1016/j.tips.2021.11.004>
- [43] Blumenthal DB, Lucchetta M, Kleist L, Fekete SP, List M, Schaefer MH. Emergence of power law distributions in protein-protein interaction networks through study bias. *elife*. 2024; 13: e99951. <https://doi.org/10.7554/eLife.99951>