

Deciphering the Intricacies of Breast Cancer Signaling Network and the Potential of Soy-derived Isoflavones on Cancer Therapeutics

Pharmacognosy Magazine
22(1) 53–66, 2026

© The Author(s) 2025

Article reuse guidelines:

in.sagepub.com/journals-permissions-india

DOI: 10.1177/09731296251322233

journals.sagepub.com/home/phm



Amir Mahgoub Awadelkareem¹, Abd Elmoneim O. Elkhalfi¹ , Mohd Adnan²,
Mohammed Kuddus³, Mohammed Idreesh Khan⁴, Manojkumar Sachidanandan⁵,
Fauzia Ashfaq⁶, Mirza Sarwar Baig⁷ and Syed Amir Ashraf¹

Abstract

Background: Breast cancer remains a significant global health challenge despite the emergence of various drug molecules. However, the adverse side effects of several drugs and chemotherapy necessitate the exploration of novel therapeutic strategies. Identifying effective therapeutic proteins specific to breast cancer is complex, and finding potential natural, non-cytotoxic inhibitors presents an even more significant challenge in this field.

Objectives: In this study, we aimed to identify various proteins responsible for the development of breast cancer, as well as explore the potential therapeutic application of various isoflavones as complementary agents for breast cancer management.

Materials and Methods: Analysis of The Cancer Genome Atlas (TCGA) RNA-Seq and protein expression data at the Human Protein Atlas was performed for the identification of proteins. Furthermore, selected proteins were used for molecular docking and dynamics against various isoflavone derivatives. In addition, pharmacokinetic activity was performed for the isoflavone derivatives.

Results: Molecular docking exhibited the most potent binding energy of -9.6 kcal/mol for the CRMP2-genistin complex, closely followed by the HER2-daidzin complex with a binding energy of -9.4 kcal/mol. Subsequent molecular dynamics simulations showed dynamic behavior, structural integrity, stability, and interaction stability of HER2 protein with ligand daidzin. According to ADMET data, most soy isoflavones satisfy the Lipinski, Pfizer, Ghose, and GoldenTriangle criteria, indicating drug-like properties. Immunotoxicity projections indicate daidzein has the least adverse effects, while *in silico*, cytotoxicity assays indicate minimal overall risk. Glycitin and daidzin have the lowest levels of cytotoxicity. According to the comprehensive ADMET profiles, soy-derived isoflavones can safely complement current breast cancer therapeutics.

Conclusion: Computational analysis revealed that these ligands had inhibitory potential against BC-related HER2 and CRMP2 proteins. These isoflavones could be used to develop nutraceuticals to ensure safe and effective breast cancer management.

Keywords

ADMET, breast cancer, dynamic simulation, intra-PPI network, isoflavones, molecular docking, molecular interactions, nutraceutical

Received 19 October 2024; accepted 31 January 2025

¹Department of Clinical Nutrition, College of Applied Medical Sciences, University of Ha'il, Ha'il, Saudi Arabia

²Department of Biology, College of Science, University of Ha'il, Ha'il, Saudi Arabia

³Department of Biochemistry, College of Medicine, University of Ha'il, Ha'il, Saudi Arabia

⁴Department of Basic Health Sciences, College of Applied Medical Sciences, Qassim University, Buraydah, Saudi Arabia

⁵Department of Oral and Maxillofacial Surgery and Diagnostic Sciences, College of Dentistry, University of Ha'il, Ha'il, Saudi Arabia

⁶Department of Clinical Nutrition, College of Nursing and Health Sciences, Jazan University, Jazan, Saudi Arabia

⁷Center for Virology, School of Interdisciplinary Science and Technology, Jamia Hamdard, New Delhi, India

Corresponding authors:

Syed Amir Ashraf, Department of Clinical Nutrition, College of Applied Medical Sciences, University of Ha'il, P.O. Box 2440, Ha'il 55476, Saudi Arabia.

E-mail: amirashrafy2007@gmail.com



Introduction

Breast cancer or carcinoma (BC) mostly arises in the epithelial cells that line the ducts (85%) and are known as ductal carcinomas, while sometimes in lobules (15%) in the glandular tissue known as lobular carcinomas of the breast (Feng et al., 2018; Makki, 2015). BC can be non-invasive, which does not go beyond the milk ducts or lobules in the breast, or invasive, which spreads into surrounding breast tissue and then spreads to the nearby lymph nodes or other distant organs in the body. The hallmarks of advanced metastatic BC are stromal invasion and metastasis to regional lymph nodes or distant organs, which become incurable because of metastasis (Makki, 2015; Nwabo Kamdje et al., 2014; Testa et al., 2020). However, BC treatment can be highly effective, especially when the disease is identified early. BC remains a global health concern, exacerbated by issues such as drug resistance and adverse effects from conventional therapies. Studies also emphasize the role of various proteins in BC signaling. For instance, the human epidermal growth factor receptor-2 (HER2), also known as the ERBB2 pathway, is crucial for tumor progression and has been extensively targeted in therapies, although resistance remains a significant challenge (Angus et al., 2023). Several cancer therapeutic studies are focusing on the HER/EGFR/ErbB family of receptor tyrosine kinases (RTKs) and have exhibited promising outcomes, offering an alternative therapy with efficacy comparable to chemotherapy, especially in mutated forms of EGFR (Kjær et al., 2020; Schoeberl et al., 2017). Comprising four homologous RTKs—EGFR/ErbB1/HER1, ErbB2/HER2, ErbB3/HER3, and ErbB4/HER4—the EGFR/ErbB family kinases play a pivotal role in intercellular signaling, cell proliferation, differentiation, and migration (Ding et al., 2020; Green et al., 2016; Kunte et al., 2020; Loibl & Gianni, 2017). These single-pass transmembrane proteins translate extracellular signals, such as ligands or growth factors, into activating specific cell signaling cascades. BC is categorized into molecular subtypes by the presence or absence of estrogen, progesterone hormone receptors, and HER2 (Kunte et al., 2020; Testa et al., 2020; Webster et al., 2009). It has been well-established that estrogen participates in the pathophysiology of BC (Al-Shami et al., 2023). Estrogen receptor 1 (ESR1) or ER alpha (Era) promotes the proliferation of cancer tissues, while estrogen receptor 2 (ESR2) or ER beta (Erb) can protect against the mitogenic effect of estrogen in breast tissue. The expression status of Era and Erb may highly influence the development, treatment, and prognosis of BC (Paterni et al., 2014; Zhou & Liu, 2020). The top protein for target prediction of BC in our study is HER2, a member of the family of RTKs whose overexpression has been reported in numerous cancers, including BC (Galogre et al., 2023; Swain et al., 2023). Research indicates that about 20–30% of BC exhibit an overexpression of HER2, resulting in heightened activation of HER2 and the initiation of multiple downstream pathways that lead to uncontrolled cancer cell proliferation.

HER2 binds tightly to other ligand-bound epidermal growth factor (EGF) receptor family members to form a heterodimer, stabilizing ligand binding and enhancing kinase-mediated activation of downstream signaling pathways, such as those involving mitogen-activated protein kinase (MAPK) and phosphatidylinositol-3 kinase (PI3K) (Hsu & Hung, 2016; Mayer & Arteaga, 2016). Drugs targeting these specific proteins or their interactions could offer novel treatment strategies for BC patients. Molecular docking and dynamic simulation studies provide valuable insights into the potential therapeutic interactions between distinct inhibitor drugs and crucial protein targets. In particular, a range of inhibitory drugs, including afatinib, pertuzumab, trastuzumab, ado-trastuzumab emtansine, and ceritinib, demonstrate substantial interactions with these protein targets. Notably, each of these drugs exhibits different inhibitory potentials, reflecting their unique mechanisms of action, and is accompanied by well-known side effects, emphasizing the importance of re-evaluating treatment strategies for BC.

Isoflavones, particularly those derived from soy, have garnered attention for their potential therapeutic effects. In this work, we use an *in silico* approach to predict the binding interactions between the different isoflavones and targeted proteins involved in the development of BC in order to investigate the anti-cancer potential of the isoflavones derived from soy. Furthermore, molecular dynamics simulation (MDS) analysis was carried out to verify the protein and ligand complexes' stability. Additionally, ADMET and computational pharmacokinetic study of these isoflavones was conducted to understand their physicochemical properties.

Materials and Methods

Analysis of The Cancer Genome Atlas (TCGA) RNA-Seq and Protein Expression Data at the Human Protein Atlas (HPA)

For this metadata analysis, we used data from TCGA, which offered transcriptomics information for 1,075 patients. There were 12 men and 1,063 women in this sample. It categorizes them into different groups, including “cancer tissue enriched,” “cancer group enriched,” “cancer tissue enhanced,” “expressed in all,” “mixed,” and “undetected.” Analysis of RNA expression data for five putative BC-associated proteins was obtained from TCGA and is accessible on the Protein Atlas website (<https://www.proteinatlas.org>).

Intra-species Protein–Protein Interaction (PPI) Network Construction

The PPI was conducted in the EMBL-EBI IntAct database (<https://www.ebi.ac.uk/intact>) (Del Toro et al., 2022). The five shortlisted gene transcript Ensembl IDs searched in the IntAct database were ENSG00000141736 (HER2),

ENSG00000092964 (CRMP2), ENSG00000185499 (CA15.3), ENSG00000091831 (ESR1), and ENSG00000140009 (ESR2). A network representation depicts the proteins' interaction partners, with proteins shown as nodes and the kind and confidence (measured by molecular interaction (MI) score) as edges. The edge color corresponds to the type of interaction, distinguishing between direct exchanges and physical linkages. Furthermore, the thickness of the edges represents the degree of confidence in the interactions, with thicker lines representing higher confidence scores (MI > 0.6) compared to medium confidence MI scores between 0.45 and 0.6.

Prediction and Selection of Diagnostic Marker Protein Target for BC

This modeling study selected five proteins based on cancer specificity, tau specificity score (RNA), and tissue specificity to predict diagnostic marker proteins associated with BC. A critical protein is HER2, well known for its overexpression in specific subtypes of BC, and it is upregulated in blood from such BC patients (Magis et al., 2020).

Retrieval of Protein Structures

After TCGA RNA-Seq data, HPA, PPI, and pathway analysis, we identified HER, CRMP2, CA15.3, ESR1 and ESR2 as promising drug targets for BC. The crystal structure of the human HER2 (PDB ID- 7PCD), CRMP2 (PDB ID- 5LXX), CA15.3 (PDB ID- 1Y8X), MUC1 (PDB ID- 5T6P), ESR1 (PDB ID- 6CHZ), and ESR2 (PDB ID- 5TOA) was obtained from the RCSB-Protein Data Bank (Rutgers University, NJ, USA).

Retrieval of Molecules or Compounds

Soybean-derived six prominently known isoflavonoid compounds, daidzein (CID: 5281708), daidzin (CID: 107971), genistein (CID: 5280961), genistin (CID: 5281377), glycitein (CID: 5317750), and glycitin (CID: 187808), were retrieved from NCBI-PubChem. Meanwhile, known inhibitors and drugs of HER2 were downloaded from EMBL-CHEBI and DrugBank and used as positive controls for molecular docking.

Molecular Docking

This computational method employed AutoDock Vina (version 4.2) to predict the molecular-level interactions between two entities: anti-cancer drugs or inhibitors and target proteins (Trott & Olson, 2009). We performed molecular docking to determine the best binding configuration and energy (kcal/mol) between these compounds and target proteins. First, we assigned Gasteiger-Marsili partial charges to the ligand molecules in PDB format and converted them to

PDBQT format. Next, we processed the target protein by removing water molecules and adding polar hydrogen atoms. We then introduced partial atomic Kollman charges and distributed the charge deficit. Finally, we converted the PDB file to PDBQT format (Ashraf et al., 2021; Elkhailifa et al., 2023).

MDS Trajectory Analysis

MDS was performed on a Linux operating system (LOS)-based machine using the GROMACS 2019.2 program and the GROMOS 43a1 force field (Abraham et al., 2015; Páll et al., 2015). A 50 ns MDS production run was conducted for individual systems (unbound and docked proteins) at a constant temperature of 300 K and pressure of 1 bar. Following the completion of MDSs, individual atomic coordinates within the overall MD trajectory for both the protein and the docked complex were subject to thorough analysis. Various GROMACS tools, namely, *g_rmsd*, *g_rmsf*, *g_gyr*, *g_sasa*, and *g_hbond*, were employed for trajectory analyses. The resulting trajectory data were visualized using *xmgrace* 5.0.5 software, and the generated plots for root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), solvent-accessible surface area (SASA), and the number of hydrogen bonds, presented in .xvg files, were further illustrated using Microsoft 365 Excel Worksheet.

Assessment of ADMET and Computational Pharmacokinetics

The SwissADME (release 2017) (Daina & Zoete, 2016; Daina et al., 2017) and ADMETlab 3.0 (Fu et al., 2024) are freely accessible online resources that assess small compounds' pharmacokinetics, drug-likeness, and medicinal chemistry using comprehensive experimental datasets. The SMILES representation of the compounds was inputted, and 2D structure files were produced using SwissADME and ADMETlab 3.0 (Ghose et al., 1999; Lipinski et al., 2001; Martin, 2005; Muegge et al., 2001; Veber et al., 2002).

Results

Evaluation of RNA-Seq Information from TCGA

The RNA tissue category classification is based on analyzing mRNA expression levels of CDK4 and CDK6 in 17 different cancer tissues, using data from TCGA. This data is presented as the median fragments per kilobase of exon per million (FPKM) reads and provides valuable insights into mRNA expression across these cancer tissues. We used the HPA database to confirm the protein expression, which indicates that HER2, CRMP2, CA15.3, and ESR1 are strongly associated with BC (Figure 1A–B). This information makes understanding the gene expression patterns and possible prognostic value within different cancer types easier.

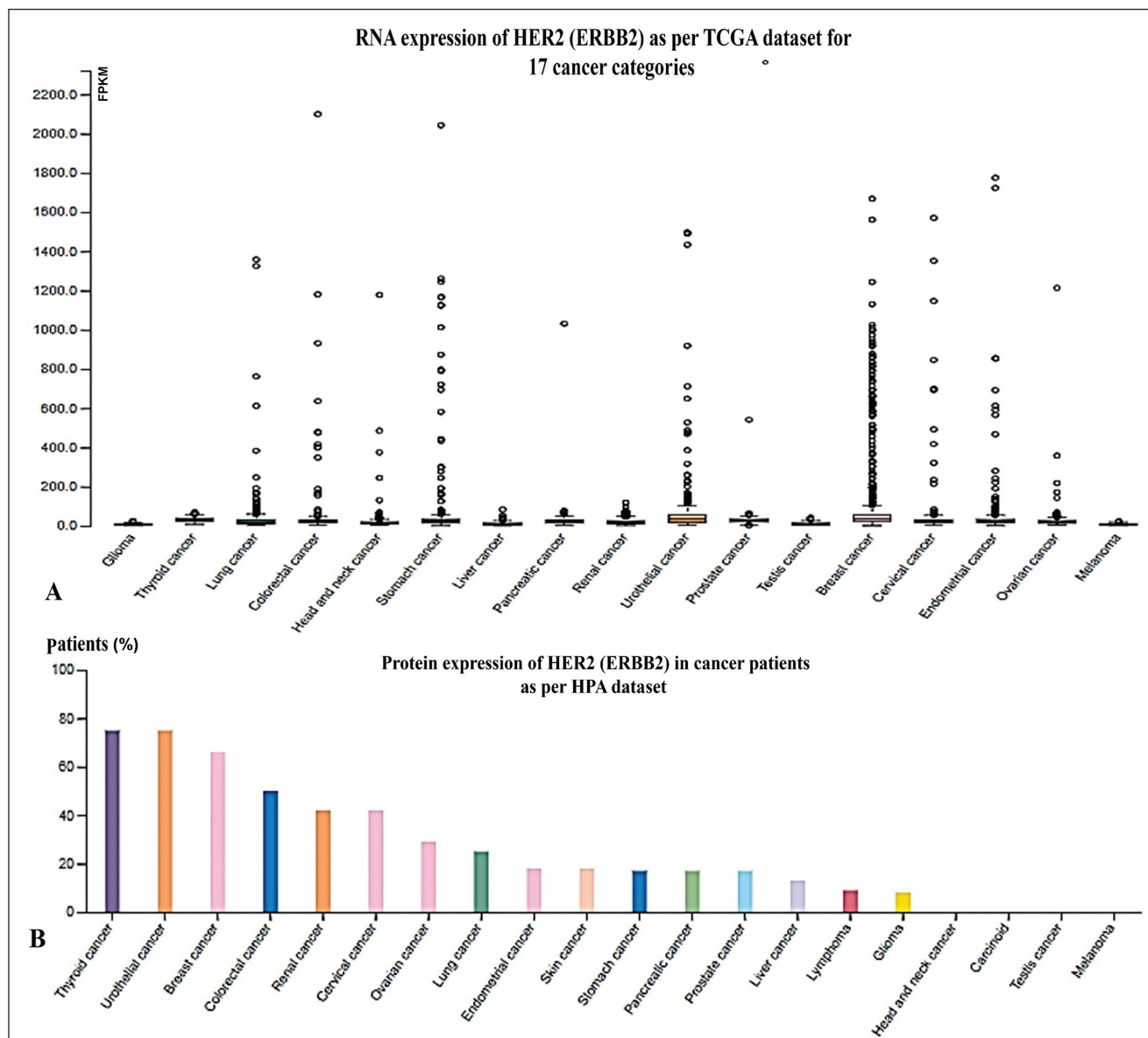


Figure 1. (A) A Color-coded Box Plot Diagram Depicting the RNA-Seq Data of the *HER2* Gene Across 17 Cancer Types. The Graph Represents the Median Number of Fragments per Kilobase of Exon per Million (FPKM) Reads. (B) The *HER2* Gene and Protein Expression Has Been Implicated in Developing Several Cancers, Exclusively in Humans' Thyroid, Urethral, Breast, and Colorectal Cancers.

Source: Human Protein Atlas (HPA) as Adapted from The Cancer Genome Atlas (TCGA) Datasets.

Most cancer tissues showed moderate to strong cytoplasmic immunoreactivity for CA15.3 or MUC1, especially for lung, breast, and ovarian cancer types. Nuclear positivity was observed in a few cases. TCGA RNA expression data of *ESR2* categorizes it as undetected in any cancer category. In contrast, protein expression of *ESR2* showed weak to moderate membranous positivity in cervical and colorectal cancer only. TCGA data (cancer tissue) suggest low cancer and low tissue specificity for *HER2* gene transcript, located intracellularly, and has 11 different gene transcripts. Cell line expression analysis and group-enriched cell line

specificity at HPA show that it is BC-specific, with a tau specificity score of 0.44. The tau specificity score ranging from 0 to 1 serves as a quantitative, graded scalar assessment of the specificity of gene expression. It is also known as a gene characterization score (Lüleci & Yılmaz, 2022; Fredholm et al., 2017). The other top four predicted proteins as diagnostic markers of BC in our model were CRMP2, CA15.3, *ESR1*, and *ESR2*, showing strong involvement in BC (Table S1 in the supplementary material). Table S1 in the supplementary material presents a comprehensive list of cancer-related human proteins that serve as Food and Drug

Administration (FDA)-approved drug targets for various cancer types.

Analysis of Intra-species PPI Network

The intra-PPI data of five shortlisted gene transcripts were filtered to focus on their effective interactor proteins. The IntAct circular layout visual representation of the PPI network of HER2, a gene with the Ensembl ID ENSG00000141736, showed the interactions between HER2 and 860 other human proteins. Based on stringent criteria, this result was further enriched with a high-confidence MI score, robust experimental evidence, their interaction types, and associated interactions. Twenty-four proteins were selected that were interacting with HER2 by colocalization, phosphorylation, and dephosphorylation. ERBB2 engages in colocalization with STAT3, supported by a confidence MI score of 0.78. Similarly, LRIG1 colocalizes with ERBB2, indicated by interaction and an MI score of 0.75. ERBB2 and PTPN12 participate in dephosphorylation interactions and have a confidence MI score of 0.59. Additionally, ERBB2 colocalizes with EEA1, supported by an MI score of 0.57. The analysis was further refined, and the network was recreated using HPA data. It identified 57, 49, 7, 6, and 4 human proteins interacting with ESR1, HER2, DPYSL2, ESR2, and

MUC1, respectively (Figure 2A–E). The edges and nodes in this network are color-coded, with green edges representing direct interactions between proteins and black edges indicating physical associations. The PPI network implies that ERBB2 and ESR1 proteins are strongly regulatory within the cellular context.

The PPI results indicate a network of protein interactions with associated interaction types, confidence levels, and MI scores. Notably, HER2 directly interacts with ERBB3 with a high confidence level, supported by an MI score of 0.97. ESR1 demonstrates a highly interactive network with NCOA1, showing a high-confidence physical association with an MI score of 0.84, contrary RELA and BRCA1 exhibit direct interactions with high confidence, featuring MI scores of 0.83 and 0.81, respectively.

Molecular Docking

By applying AutoDock Vina, molecular binding energies (kcal/mol) for various proteins against these isoflavones (daidzein, daidzin, genistein, genistin, glycitein, and glycitin) were presented in Table 1. ESR2's binding energies range from -6.6 to -9.1 kcal/mol. Meanwhile, ESR1 demonstrates binding energies varying from -6.9 to -8.5 kcal/mol. CRMP2 exhibits binding energies between -8.1 and -9.6 kcal/mol.

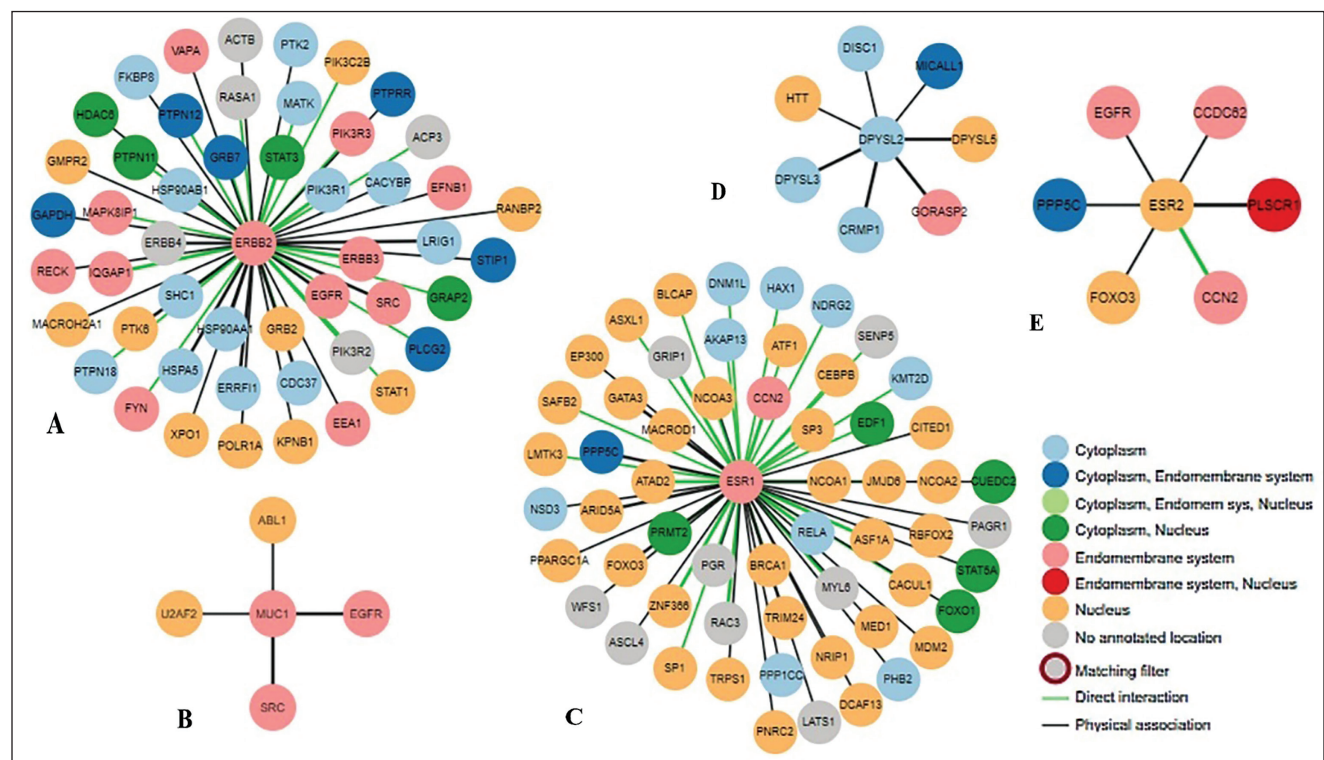


Figure 2. The Intra-Protein-Protein Interaction (PPI) Network Demonstrates (A) HER2, (B) MUC1, (C) ESR1, (D) DPYSL2, and (E) ESR2 and Their Respective Interacting Partners (Derived from Data in the Human Protein Atlas). The Edges' and Nodes' Colors Signify Proteins' Interaction Type and Subcellular Localization. The Edges' Colors Distinguish Direct Interactions (Green) and Physical Associations (Black) Among the Chosen Protein and Interacting Partners.

Table 1. Selected Breast Cancer Biomarkers and Their Molecular Docking with Soybean-derived Isoflavones and Their Respective Metabolites—Daidzein, Daidzin, Genistein, Genistin, Glycitein, and Glycitin. The Binding Energies are Measured in kcal/mol for Various Proteins and are Presented.

PDB ID	Protein Name	AutoDock Vina Binding Energy (kcal/mol)						Average Binding Energy of Protein
		Daidzein	Daidzin	Genistein	Genistin	Glycitein	Glycitin	
6CHZ	ESR1	-6.9	-8.5	-6.9	-8.2	-6.9	-7.9	-7.55
5TOA	ESR2	-9.1	-7.5	-6.6	-7.9	-6.6	-7.6	-7.55
5LXX	CRMP2	-8.1	-9.3	-8.7	-9.6	-8.1	-9.3	-8.85
1Y8X	CA15.3	-6.8	-7.8	-6.8	-7.5	-6.2	-7.1	-7.03
5T6P	MUC1	-7.5	-7.3	-7.7	-7.7	-7.4	-8.3	-7.65
7PCD	HER2	-6.6	-9.4	-7.5	-7.9	-6.5	-7.4	-7.55
Average Binding Energy of Ligand		-7.5	-8.3	-7.4	-8.1	-6.9	-7.9	

CA15.3 shows binding energies spanning from -6.2 to -7.8 kcal/mol. MUC1 displays binding energies ranging from -7.3 to -8.3 kcal/mol. Finally, for HER2, the binding energies range from -6.5 to -9.4 kcal/mol. Daidzein exhibits the highest affinity of -9.1 kcal/mol with Erβ, whereas it has an exceptionally low affinity to HER2 (-6.6 kcal/mol). Daidzin exhibits the highest binding energy of -9.4 kcal/mol with HER2. Genistein binds to CRMP2 with the highest affinity of -9.6 kcal/mol. The average binding energy of these six compounds was highest (-8.85 kcal/mol) for CRMP2 while lowest (-7.03 kcal/mol) for cancer antigen 15-3. This comparative analysis reveals variations in binding energies across different proteins and isoflavones.

Binding Affinity and Interaction Profile of Isoflavones with HER2

The molecular docking of isoflavones with HER2 is visually represented using PyMol in 3D. This visualization illustrates the interactions between isoflavones and the active site residues of the HER2 protein (Figure 3). Further, the 2D analysis delves into specific interactions for each isoflavone. Daidzein and daidzin, for instance, form two hydrogen bonds each. However, daidzin exhibits stronger interactions due to the formation of 12 van der Waals (VDW), 3 pi-sigma, and 1 pi-alkyl interaction. Genistein establishes four hydrogen bonds, five VDW interactions, three pi-alkyl, one pi-sigma, and one unfavorable acceptor-acceptor bond with HER2 binding cavity residues. Conversely, genistin forms 3 hydrogen bonds, 11 VDW interactions, 2 pi-alkyl, 1 pi-sigma, and 1 carbon-hydrogen bond. Glycitein and glycitin also display distinct interaction patterns with HER2 binding cavity residues, encompassing hydrogen bonds, VDW interactions, pi-alkyl, pi-sigma, and other specific bonds. This comprehensive analysis provides a nuanced understanding of how each isoflavone interacts with the HER2 protein at the molecular level, offering valuable insights for further research in the context of HER2-related conditions.

Elucidation of MDS of Daidzin with HER2

The average binding energy of the ligand daidzin is notably negative across all six proteins studied, with an exceptionally high value of -8.3 kcal/mol. Notably, for the HER2 protein, the binding energy reaches an exceptionally high level of -9.4 kcal/mol. Examination of the RMSD graph for the unbound HER2 protein backbone reveals a deviation ranging from 0.2 to 0.35 nm throughout the 50 ns simulation. In contrast, the ligand-bound HER2 protein displays a deviation between 0.1 and 0.4 nm over the same simulation time. Interestingly, the RMSD plots for the unbound and ligand-bound complexes exhibit numerous overlapped regions, suggesting stability throughout the simulation. After the 50 ns simulation, the unbound and ligand-bound complexes demonstrate equivalent RMSD values, affirming their stability and suitability for further analysis (Figure 4A). Considering that the HER2 protein's binding domain encompasses 281 residues (amino acids 710–991), the elevated RMSF values observed for specific Cα-backbone residues (ranging from 0.3 to 0.45 nm) signify a heightened level of flexibility, particularly in a distinct region characterized by increased flexibility and accessibility. Conversely, many residues exhibit low RMSF values (ranging from 0.1 to 0.25 nm), indicating constrained movement and a region of reduced flexibility within the HER2 protein structure (Figure 4B). These variations in RMSF values suggest a potential correlation between flexibility and the diverse binding affinities of the residues. The Rg graph illustrates an overall decrease in Rg values from 1.98 to 1.88 nm for the unbound HER2 protein structure, suggesting a compactly packed state and stable folding of the HER2 protein (Figure 4C). The calculation of local SASA per residue revealed a range of 0 to 2.0 nm², indicating varying exposure of HER2 protein residues to the solvent. Higher SASA values suggest increased surface accessibility for solvent or ligand interactions, potentially leading to more conformational changes (Figure 4D). Conversely, the overall global SASA of the unbound HER2 protein per residue decreased after 50 ns of MDS, ranging from 150 to 125 nm².

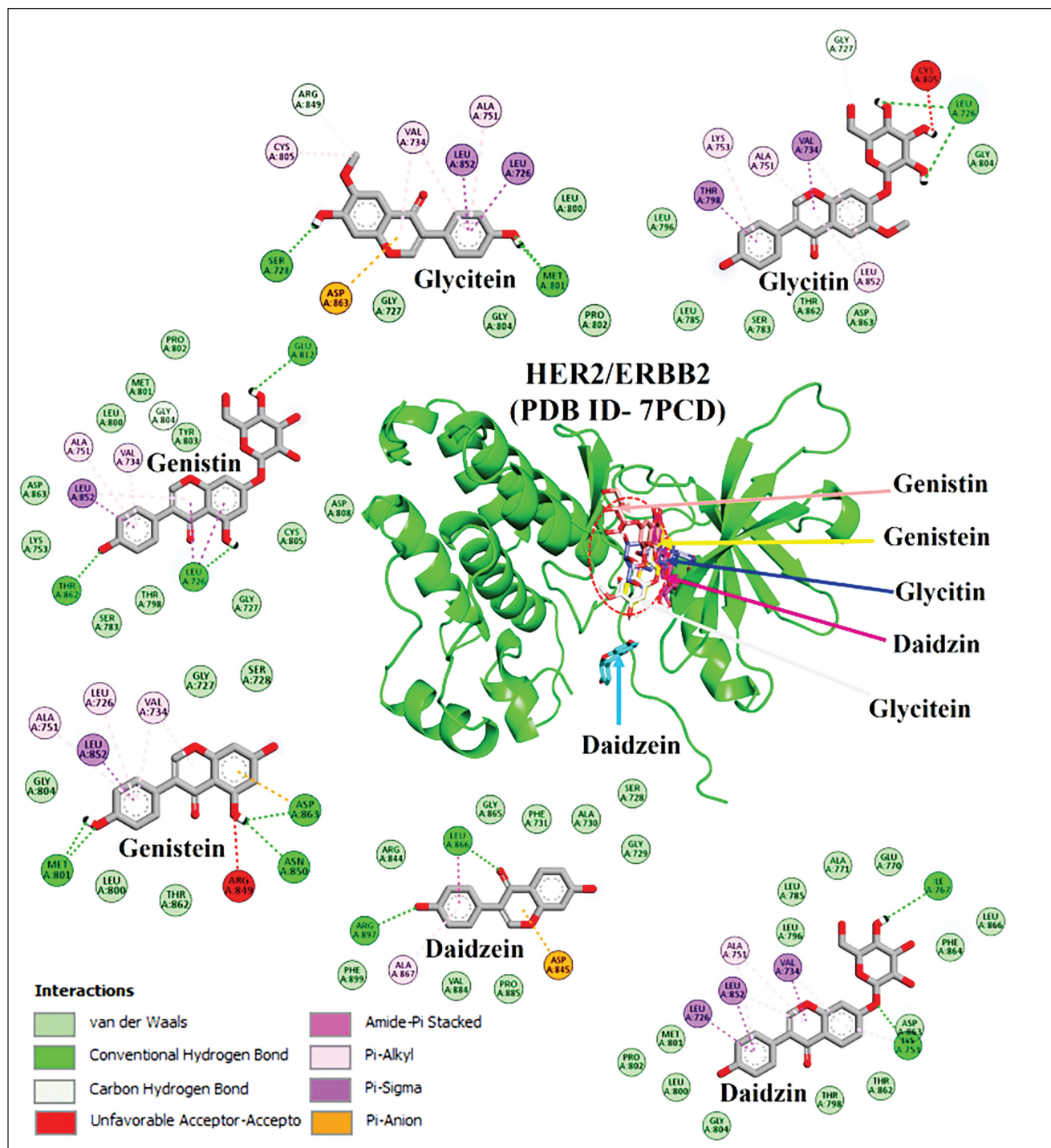


Figure 3. The Molecular Docking of Isoflavones with HER2 (PDB ID-7PCD). The PyMol 3D Visualization Highlights the Interactions Between Isoflavones and Active Site Residues. The Remaining Five Molecules are Docked Within a Single Binding Cavity Apart from Daidzein. The 2D Discovery Studio Illustration Provides Insights into Eight Types of Interactions: van der Waals (VDW) (Depicted in Green), Conventional H-bond (Dark Green), Carbon-Hydrogen Bond (Light Green), Unfavorable Acceptor-Acceptor (Red), Amide-Pi Stacked (Dark Pink), Pi-Alkyl (Pink), Pi-Sigma (Purple), and Pi-Anion (Orange). This 2D Analysis Offers a Detailed Perspective on the Molecular-level Interactions Occurring Within the Binding Pocket Residues of HER2. Daidzein (LEU866, ARG897) and Daidzin (LYS753, ILE767) Each Form Two Hydrogen Bonds, with Daidzin Displaying Stronger Interactions Than Daidzein Through The Formation of 12 VDW, 3 Pi-Sigma, and 1 Pi-Alkyl Interaction.

Throughout the 50 ns MDs, hydrogen bonds between the atoms of unbounded and ligand-bounded HER2 proteins were quantified individually. Figure 4F illustrates the fluctuation in the number of hydrogen bonds among unbounded HER2 backbone residues, ranging from 152 to 208. In contrast, the number of hydrogen bonds between the ligand (daidzin) and HER2 protein varies from 1 to 6 per residue, with an average of 3 hydrogen bonds forming per residue (Figure 4E–G). This analysis provides insights into the dynamic interactions and stability of the protein and its complex with the ligand during the simulation.

ADMET Analysis

ADMET analysis assessed the drug-like properties of six isoflavones derived from soy: daidzein, daidzin, genistein, genistin, glycitein, and glycitin. The evaluation of these compounds revealed that they get efficiently absorbed, distributed, metabolized, and excreted (ADMET) within the body. ADMET analysis assesses different molecular properties, including the count of hydrogen bond acceptors (nHA), the count of hydrogen bond donors (nHD), the count of rotatable bonds (nRot), the count of rings (nRing), the maximum count

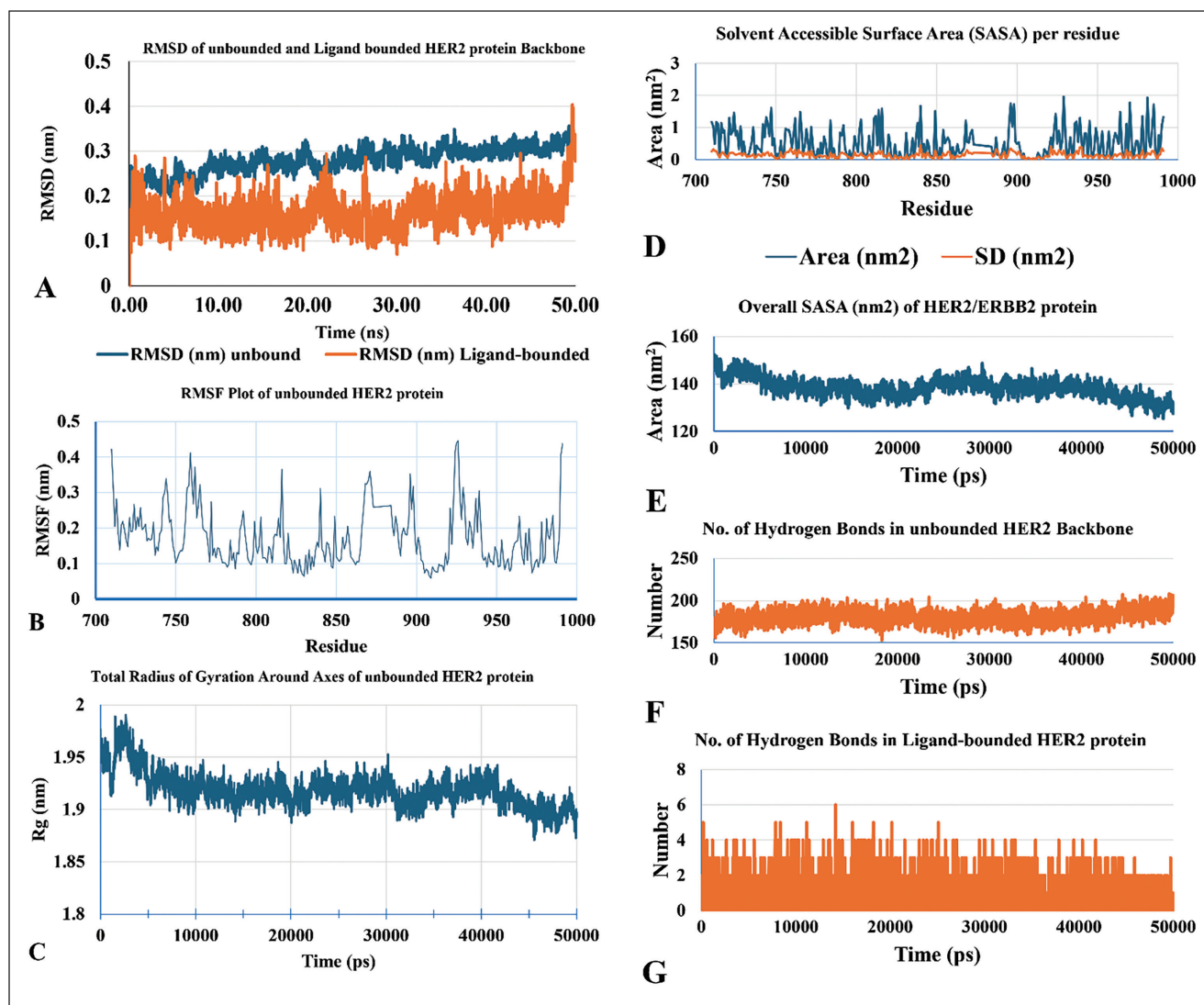


Figure 4. This Figure Illustrates the Plots Derived from the 50 ns (50,000 ps) Molecular Dynamics Simulations of the Unbounded and Ligand-bound (Daidzin) HER2 Protein Backbone. (A) Root Mean Square Deviation (RMSD), (B) Root Mean Square Fluctuation (RMSF), and (C) Radius of Gyration (Rg). (D) Local Solvent-accessible Surface Area (SASA) per Residue, (E) Overall Global SASA, (F) the Number of Hydrogen Bonds in HER2 Backbone Residues, and (G) the Number of Hydrogen Bonds Between the Ligand (Daidzin) and HER2 Residues. These Representations Show the Protein's Structural Stability, Flexibility, and Compactness During the Simulation Period, Showing the Dynamic Behavior of HER2 in Both Unbounded and Ligand-bound States.

of atoms in the most extensive ring (MaxRing), the count of rigid bonds (nRig), the count of heteroatoms (nHet), and the formal charge (Fchar). Most of the examined isoflavones complied with the Lipinski, Pfizer, Ghose, and GoldenTriangle criteria, suggesting positive ADMET characteristics (Table S2 in the supplementary material). The Boiled-Egg plot depicts the blood–brain barrier (BBB) penetration capacity of daidzein (1) and glycitein (2); meanwhile, the human intestinal absorption (HIA) tendency of genistein (3) was higher than daidzein (4), glycitin (5), and genistin (6) (Table S2 in the supplementary material). The comprehensive analysis of absorption attributes for the compounds reveals distinct predicted values across various parameters. Daidzin and glycitin (glycosides) have significantly higher HIA values than their aglycone counterparts (daidzein and genistein). Daidzin and glycitin again outperform daidzein and genistein across all bioavailability levels, supporting the potential benefit of glycosylation (Table S3 in the supplementary material).

Discussion

The findings of our study provide a comprehensive analysis of the mRNA and protein expression levels of four key genes—*ESR1*, *CA15.3*, *HER2*, and *CRMP2*, shedding light on their significant relationships with BC (Chen et al., 2003; Ding et al., 2020; Galogre et al., 2023; Schoeberl et al., 2017; Testa et al., 2020). Their distinct expression patterns provide valuable insights into the potential diagnostic significance of these proteins against BC. *CRMP2* exhibited moderate to strong cytoplasmic positivity, predominantly in tumor stroma and a subset of tumor cells in glioma and testis cancer (Barnard et al., 2015). Conversely, colorectal, cervical, and endometrial cancers were negative for *CRMP2* expression. *CA15.3* showed moderate to strong cytoplasmic immunoreactivity, especially in lung, breast, and ovarian cancers (Goodwin et al., 2021; Li et al., 2022). RNA expression analysis of *ESR1*, as per TCGA cancer enhanced dataset, showed strong nuclear positivity in breast, endometrial, and ovarian cancers. Other cancer categories were negative for *ESR1* RNA and protein expression. *ESR2*, based on TCGA RNA expression data, was not detected in any cancer category. However, protein expression showed weak to moderate membranous positivity in cervical and colorectal cancer. *HER2* receptors command attention, particularly in BC. Approximately 20–30% of BC exhibit *HER2* overexpression, fueling uncontrolled cell proliferation (Galogre et al., 2023; Kunte et al., 2020; Schoeberl et al., 2017). As a therapeutic target, trastuzumab emerges as an effective intervention, highlighting its diagnostic significance in BC. With a tau specificity score of 0.44, *HER2* demonstrates low tissue specificity, featuring cytoplasmic and membranous expression in various tissues and interactions with 49 proteins (Figure 1A). Glycoprotein mucin 1 (*MUC1*) or cancer antigen 15-3 (*CA15-3*) navigates a complex landscape, a transmembrane

glycoprotein overexpressed in BC cells. *CA15-3*, an associated glycoprotein, aids in monitoring disease progression and treatment response. Expressing primarily on the apical surface of epithelial cells, especially in the breast, and uterus, it interacts with four proteins, including *EGFR*, *SRC*, *U2AF2*, and *ABL1* (Table S1 in the supplementary material). *ESR1*, also known as *Era*, emerges as a pivotal biomarker in BC, dictating treatment strategies and prognostication (Fredholm et al., 2017). Its prevalence or absence significantly influences therapeutic responses in hormone receptor-positive BC. *ER1* proves particularly prominent in BC, ovarian, and prostate cancers, earning the distinction as a prognostic and diagnostic marker, with a tau specificity score of 0.67 (Table S1 in the supplementary material).

The investigation into the intra-species PPI network landscape, specifically its association with BC progression, delved into a meticulous analysis refined with data from the HPA (Digre & Lindskog, 2023; Thul & Lindskog, 2018). The outcome was a discerning identification of interacting proteins associated with key biomarkers, namely, *ESR1*, *HER2*, *DPYSL2*, *ESR2*, and *MUC1*. This network, vividly depicted in Figure 2A–E, unfolds a complex web of 57, 49, 7, 6, and 4 human proteins intricately connected with *ESR1*, *ERBB2*, *DPYSL2*, *ESR2*, and *MUC1*. The PPI network's visual representation underscores the robust regulatory role of *ERBB2* and *ESR1* proteins within the cellular context, as evidenced by their high number of interacting partners and network topology (Figure 2A and C). This points toward their potential importance in BC development and progression. The subcellular localization of the PPI network reveals that *ERBB2*, *ESR1*, and *MUC1* express in the endomembrane system, ensuring the inclusion of interactions specific to *Homo sapiens* in the datasets. A notable finding within this intricate network is the direct interaction between *HER2* and *ERBB3*, supported by a high confidence level and an MI score of 0.97. These findings contribute significantly to our understanding of the molecular landscape associated with BC progression, opening avenues for further exploration and targeted therapeutic interventions. This overexpression, in turn, triggers the activation of the *PI3K/AKT* signaling pathway, a crucial pathway in cell proliferation and survival, known to be dysregulated in many types of cancers (Mayer & Arteaga, 2016; Testa et al., 2020). The network pinpoints a specific cascade where *ERBB2* upregulates *ERBB4*, activating *PIK3CA*, a key enzyme in the *PI3K/AKT* pathway (Figure 2).

In the current study, the therapeutic potential of six soybean-derived bioactive isoflavones and their metabolites (daidzein, daidzin, genistein, genistin, glycitein, and glycitin) was docked with six promising protein targets in BC: *ESR1*, *ESR2*, *CRMP2*, *CA15.3*, *MUC1*, and *HER2* (Cameron et al., 2017; Shimada et al., 2014). Interestingly, daidzin exhibited the most favorable binding energy and interactions compared to other isoflavones across all studied protein targets (Table 1 and Figure 3). The observed prominence of daidzin in binding energy and interactions compared to daidzein, genistein,

genistin, glycitein, and glycitin underscores its potential as a promising candidate for further investigation in BC therapy. On average, these findings contribute valuable information indicating the differential affinities of soybean-derived compounds toward specific protein targets and encouraging further exploration of daidzin as a potential therapeutic agent for BC treatment (Elkhalifa et al., 2023). The study lays the groundwork for future research to understand the molecular mechanisms underlying the observed interactions and potential translational applications in BC therapeutics. Several research studies suggest that they can modulate BC signaling pathways by targeting specific human proteins like HER2, CRMP2, CA15.3, ESR1, and ESR2. Soy isoflavone genistein has demonstrated anti-cancer properties, including suppressing HER2 and ER α expression. HER2 is a RTK overexpressed in specific BC (HER2+). Genistein exhibits a variety of biological functions that contribute to its BC preventive effects. It competes with estrogens for binding to estrogen receptors (ERs) and regulates estrogen-dependent gene expression (Markiewicz et al., 1993). Additionally, genistein acts as a protein tyrosine kinase inhibitor (Shao et al., 1998), suppresses angiogenesis (Fotsis et al., 1995), inhibits DNA topoisomerase II activity (Markovits et al., 1989), induces apoptosis of BC cells (Katdare et al., 2002; Pagliacci et al., 1994), and downregulates HER2 (Katdare et al., 2002; Li et al., 1999) and ER α expression (Chen et al., 2003; Mai et al., 2007). Therefore, the supplementation of genistein suppresses HER2-overexpressing BC cells (Bhat et al., 2021).

The MDS results of the daidzin and HER2 complex provide a comprehensive understanding of the structural dynamics and stability of the binding interaction over a 50 ns simulation period. The RMSD analysis of the unbound HER2 protein backbone reveals a deviation ranging from 0.2 to 0.35 nm, indicating a stable and compactly packed state throughout the simulation. Interestingly, the ligand-bound HER2 protein exhibits a slightly higher deviation (0.1–0.4 nm), suggesting that daidzin induces subtle structural changes, a common phenomenon in ligand binding (Figure 4A). While slight deviations exist, the overlap between unbound and bound RMSD plots suggests overall stability. Specific regions of the HER2 protein exhibit higher flexibility (RMSF values 0.3–0.45 nm), particularly in the binding domain (residues 710–991), indicating potential interaction points. Other regions remain rigid (RMSF 0.1–0.25 nm), suggesting key structural elements (Figure 4B). This flexibility variation might correlate with binding affinities of different residues. The Rg analysis shows a decrease in Rg values from 1.98 to 1.88 nm for the unbound HER2 protein, indicating a compactly packed and stably folded structure (Figure 4C). The analysis of local SASA per residue reveals varying exposure of HER2 protein residues to the solvent, with higher SASA values suggesting increased surface accessibility for solvent or ligand interactions (Figure 4D). The overall global SASA of the unbound HER2 protein per residue exhibits a decrease, indicating a stable protein structure with reduced solvent exposure (Figure 4E).

The diminished SASA values imply a lower content of exposed surface amino acids or higher hydrophobicity of the HER2 protein. Previous studies on ERBB kinases have shown hydrophobic sub-regions, including the C-spine, R-spine, hydrophobic core, and α C- β 4 regions. Notably, active conformations minimize SASA compared to inactive forms, and the transition from inactive to active states involves an increase in hydrophobicity, settling into a hydrophobically favorable region indicative of a well-formed C-spine in active conformations (Shi et al., 2011; Telesco & Radhakrishnan, 2009). The evaluation of hydrogen bonds within the unbound HER2 protein backbone demonstrates fluctuations in the number of hydrogen bonds formed among residues (Figure 4F). In contrast, the ligand-bound HER2 complex exhibits varying hydrogen bonds between the ligand (daidzin) and HER2 protein, ranging from 1 to 6 per residue, with an average of 3 hydrogen bonds forming per residue (Figure 4G). Highlighting the dynamic interaction between daidzin and HER2 provides compelling evidence for the stable interaction between daidzin and HER2, identifying potential binding hotspots and dynamic hydrogen bonding patterns. BCs that are positive for HER2 demonstrate a particularly aggressive clinical profile, exhibiting limited responsiveness to standard chemotherapy and a heightened propensity for recurrence and metastasis.

Significantly, some research has highlighted the enduring impacts of trastuzumab, whether administered independently or in conjunction with chemotherapy, on patients diagnosed with HER2-positive BC. Across these studies, a consistent finding emerges—trastuzumab demonstrates a substantial influence on disease-free and overall survival (Cameron et al., 2017; Testa et al., 2020; Tolaney et al., 2019). Two novel drugs designed to target HER2 have been incorporated into the clinical treatment landscape for HER2-positive BCs: (a) trastuzumab emtansine, an antibody-drug conjugate pairing trastuzumab with the cytotoxic agent emtansine—a microtubule inhibitor, and (b) pyrotinib, an irreversible pan-ERB RTK inhibitor specifically targeting HER1, HER2, and HER4. Notably, adjuvant trastuzumab emtansine demonstrated a 60% lower risk of recurrence of invasive BC or mortality compared to trastuzumab alone in patients with HER2-positive early BC who exhibited residual invasive disease post neoadjuvant therapy completion (von Minckwitz et al., 2019).

Most of the isoflavones studied were found to meet Lipinski, Pfizer, Ghose, and GoldenTriangle criteria, indicating drug-like qualities, according to ADMET data. Genistein has the highest HIA potential in the Boiled-Egg plot; however, daidzein, glycitein, and genistein have different BBB penetration capacities. Interestingly, all isoflavones are unlikely PGP substrates, indicating negligible medication absorption interference. The PGP pessimistic predictions for daidzin, genistin, and glycitin show they will not inhibit medication absorption. These detailed ADMET profiles suggest that soy-derived isoflavones like daidzin, genistin, and glycitin can supplement the existing BC drugs without any risks.

Toxicity predictions suggest minimal overall risk, with glycin and daidzin showing the lowest cytotoxicity. Daidzein might also have the least potential for immune system effects.

Conclusion

Our study explored BC treatment and identified potential therapeutic proteins associated with this prevalent and challenging disease. Through various databases, RNA-Seq data analysis, and PPI networks, essential proteins such as HER2, CRMP2, CA15.3, and ESR1 were identified for BC progression. Furthermore, molecular docking analyses revealed strong binding affinities, with the CRMP2–genistin complex exhibiting the highest binding energy, closely followed by the HER2–daidzin complex. Further, a comprehensive 50 ns MDS revealed the dynamic behavior, solvent exposure, structural integrity, and stability of the HER2 protein in a complex with daidzin. This study presents promising initial results for daidzin as a potential BC therapeutic candidate. However, further research is necessary to validate its efficacy, elucidate its mechanisms of action, and assess its safety and specificity *in vivo*.

Abbreviations

ADMET: Absorption, distribution, metabolism, excretion, and toxicity; CA15.3: Carcinoma-associated mucin; CRMP2: Collapsin response mediator protein 2; DPYSL2: Dihydropyrimidinase like 2; EGF: Epidermal growth factors; ESR1: Estrogen receptor 1; ESR2: Estrogen receptor 2; FPKM: Fragments per kilobase of exon per million; HER2: Human epidermal growth factor receptor-2; HIA: Human intestinal absorption; HPA: Human Protein Atlas; LOS: Linux operating system; MAPK: Mitogen-activated protein kinase; MDS: Molecular dynamics simulation; MUC1: Glycoprotein mucin 1; PIEK: Phosphatidylinositol-3 kinase; PPI: Protein–protein interaction; Rg: Radius of gyration; RMSD: Root mean square deviation; RMSF: Root mean square fluctuation; RTKs: Receptor tyrosine kinases; TCGA: The Cancer Genome Atlas.

Acknowledgments

This research has been funded by Scientific Research Deanship at University of Ha'il - Saudi Arabia through project number RG-21 121.

Authors' Contributions

Conceptualization, SAA (Syed Amir Ashraf) and AEOE (Abd Elmoneim O Elkhalfifa); methodology, MA (Mohd Adnan), FA (Fauzia Ashfaq), and MSB (Mirza Sarwar Baig); validation, AMA (Amir Mahgoub Awadelkareem), MS (Manojkumar Sachidanandan), and MIK (Mohammad Idrees Khan); formal analysis, MSB and SAA; investigation, SAA, MSB, FA, and MA; data curation, MK (Mohammed Kuddus) FSA (Fahad Saad Alhodieb), MIK, AMA, and MS; writing—original draft preparation, AEOE, AMA, MSB, FA, MA, and MK; writing—review and editing, SAA, FA, MA, MS, FSA, and AMA; software, MIK, MK, FSA, MSB, and AEOE; visualization, MS, MK, FSA, FA, MA, MIK, and AMA; supervision, AEOE and SAA; project administration, AEOE and SAA. All

authors have revised the manuscript and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical Approval and Informed Consent

No humans or animals were used in the present research.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research has been funded by the Scientific Research Deanship at the University of Ha'il, Ha'il, Saudi Arabia, through project number RG-21121.

Supplementary Material

Supplementary material for this article is available online.

ORCID iDs

Abd Elmoneim O. Elkhalfifa  <https://orcid.org/0000-0001-6726-4842>
Syed Amir Ashraf  <https://orcid.org/0000-0002-5333-9505>

References

- Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1–2, 19–25.
- Al-Shami, K., Awadi, S., Khamees, A., Alsheikh, A. M., Al-Sharif, S., Ala' Bereshy, R., Al-Eitan, S. F., Banikhaled, S. H., Al-Qudimat, A. R., Al-Zoubi, R. M., & Al Zoubi, M. S. (2023). Estrogens and the risk of breast cancer: A narrative review of literature. *Heliyon*, 9(9), e20224. <https://doi.org/10.1016/j.heliyon.2023.e20224>
- Angus, L., Smid, M., Wilting, S. M., Bos, M. K., Steeghs, N., Konings, I. R. H. M., Tjan-Heijnen, V. C. G., van Riel, J. M. G. H., van de Wouw, A. J., Cpct, Consortium, Cuppen, E., Lolkema, M. P., Jager, A., Sleijfer, S., & Martens J. W. M. (2023). Genomic alterations associated with estrogen receptor pathway activity in metastatic breast cancer have a differential impact on downstream ER signaling. *Cancers*, 15(17), 4416. <https://doi.org/10.3390/cancers15174416>
- Ashraf, S. A., Elkhalfifa, A. E. O., Mehmood, K., Adnan, M., Khan, M. A., Eltoum, N. E., Krishnan, A., & Baig, M. S. (2021). Multi-targeted molecular docking, pharmacokinetics, and drug-likeness evaluation of okra-derived ligand abscisic acid targeting signaling proteins involved in the development of diabetes. *Molecules (Basel, Switzerland)*, 26(19), 5957. <https://doi.org/10.3390/molecules26195957>
- Barnard, M. E., Boeke, C. E., & Tamimi, R. M. (2015). Established breast cancer risk factors and risk of intrinsic tumor subtypes. *Biochimica et Biophysica Acta*, 1856(1), 73–85. <https://doi.org/10.1016/j.bbcan.2015.06.002>

- Bhatt, D., Verma, A. K., Bharti, P. S., Goyal, Y., Alsahli, M. A., Almatroudi, A., Rahmani, A. H., Almatroodi, S., Joshi, P. C., Alam, M. M., Ahmad, I., Zaman, G. S., & Dev, K. (2021). BCL-2 (-938C>A), BAX (-248G>A), and HER2 Ile655Val polymorphisms and breast cancer risk in Indian population. *Journal of Oncology*, 2021, 8865624. <https://doi.org/10.1155/2021/8865624>
- Cameron, D., Piccart-Gebhart, M. J., Gelber, R. D., Procter, M., Goldhirsch, A., de Azambuja, E., Castro, G. Jr, Untch, M., Smith, I., Gianni, L., Baselga, J., Al-Sakaff, N., Lauer, S., McFadden, E., Leyland-Jones, B., Bell, R., Dowsett, M., Jackisch, C., & Herceptin Adjuvant (HERA) Trial Study Team (2017). 11 years' follow-up of trastuzumab after adjuvant chemotherapy in HER2-positive early breast cancer: final analysis of the HERceptin Adjuvant (HERA) trial. *Lancet (London, England)*, 389(10075), 1195–1205. [https://doi.org/10.1016/S0140-6736\(16\)32616-2](https://doi.org/10.1016/S0140-6736(16)32616-2)
- Chen, W. F., Huang, M. H., Tzang, C. H., Yang, M., & Wong, M. S. (2003). Inhibitory actions of genistein in human breast cancer (MCF-7) cells. *Biochimica et Biophysica Acta*, 1638(2), 187–196. [https://doi.org/10.1016/s0925-4439\(03\)00082-6](https://doi.org/10.1016/s0925-4439(03)00082-6)
- Daina, A., & Zoete, V. (2016). A BOILED-Egg to predict gastrointestinal absorption and brain penetration of small molecules. *ChemMedChem*, 11(11), 1117–1121. <https://doi.org/10.1002/cmdc.201600182>
- Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, 7, 42717. <https://doi.org/10.1038/srep42717>
- Del Toro, N., Shrivastava, A., Ragueneau, E., Meldal, B., Combe, C., Barrera, E., Peretto, L., How, K., Ratan, P., Shirodkar, G., Lu, O., Mészáros, B., Watkins, X., Pundir, S., Licata, L., Iannuccelli, M., Pellegrini, M., Martin, M. J., Panni, S., Duesbury, M., ... Hermjakob, H. (2022). The IntAct database: Efficient access to fine-grained molecular interaction data. *Nucleic Acids Research*, 50(D1), D648–D653. <https://doi.org/10.1093/nar/gkab1006>
- Digre, A., & Lindskog, C. (2021). The Human Protein Atlas—Spatial localization of the human proteome in health and disease. *Protein Science: A Publication of the Protein Society*, 30(1), 218–233. <https://doi.org/10.1002/pro.3987>
- Digre, A., & Lindskog, C. (2023). The human protein atlas—Integrated omics for single cell mapping of the human proteome. *Protein Science: A Publication of the Protein Society*, 32(2), e4562. <https://doi.org/10.1002/pro.4562>
- Ding, L., Cao, J., Lin, W., Chen, H., Xiong, X., Ao, H., Yu, M., Lin, J., & Cui, Q. (2020). The roles of cyclin-dependent kinases in cell-cycle progression and therapeutic strategies in human breast cancer. *International Journal of Molecular Sciences*, 21(6), 1960. <https://doi.org/10.3390/ijms21061960>
- Elkhalifa, A. E. O., Al-Shammari, E., Kuddus, M., Adnan, M., Sachidanandan, M., Awadelkareem, A. M., Qattan, M. Y., Khan, M. I., Abduljabbar, S. I., Sarwar Baig, M., & Ashraf, S. A. (2023). Structure-based multi-targeted molecular docking and dynamic simulation of soybean-derived isoflavone genistin as a potential breast cancer signaling proteins inhibitor. *Life (Basel, Switzerland)*, 13(8), 1739. <https://doi.org/10.3390/life13081739>
- Feng, Y., Spezia, M., Huang, S., Yuan, C., Zeng, Z., Zhang, L., Ji, X., Liu, W., Huang, B., Luo, W., Liu, B., Lei, Y., Du, S., Vuppapalapati, A., Luu, H. H., Haydon, R. C., He, T. C., & Ren, G. (2018). Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. *Genes & Diseases*, 5(2), 77–106. <https://doi.org/10.1016/j.gendis.2018.05.001>
- Fotsis, T., Pepper, M., Adlercreutz, H., Hase, T., Montesano, R., & Schweigerer, L. (1995). Genistein, a dietary ingested isoflavonoid, inhibits cell proliferation and in vitro angiogenesis. *The Journal of Nutrition*, 125(3 Suppl.), 790S–797S. https://doi.org/10.1093/jn/125.suppl_3.790S
- Fredholm, H., Magnusson, K., Lindström, L. S., Tobin, N. P., Lindman, H., Bergh, J., Holmberg, L., Pontén, F., Frisell, J., & Fredriksson, I. (2017). Breast cancer in young women and prognosis: How important are proliferation markers? *European Journal of Cancer (Oxford, England: 1990)*, 84, 278–289. <https://doi.org/10.1016/j.ejca.2017.07.044>
- Fu, L., Shi, S., Yi, J., Wang, N., He, Y., Wu, Z., Peng, J., Deng, Y., Wang, W., Wu, C., Lyu, A., Zeng, X., Zhao, W., Hou, T., & Cao, D. (2024). ADMETlab 3.0: An updated comprehensive online ADMET prediction platform enhanced with broader coverage, improved performance, API functionality and decision support. *Nucleic Acids Research*, 52(W1), W422–W431. <https://doi.org/10.1093/nar/gkac236>
- Galogre, M., Rodin, D., Pyatnitskiy, M., Mackelprang, M., & Koman, I. (2023). A review of HER2 overexpression and somatic mutations in cancers. *Critical Reviews in Oncology/Hematology*, 186, 103997. <https://doi.org/10.1016/j.critrevonc.2023.103997>
- Ghose, A. K., Viswanadhan, V. N., & Wendoloski, J. J. (1999). A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *Journal of Combinatorial Chemistry*, 1(1), 55–68. <https://doi.org/10.1021/cc9800071>
- Goodwin, P. J., Dowling, R. J. O., Ennis, M., Chen, B. E., Parulekar, W. R., Shepherd, L. E., Gelmon, K. A., Whelan, T. J., Ligibel, J. A., Hershman, D. L., Mayer, I. A., Hobday, T. J., Rastogi, P., Rabaglio-Poretti, M., Lemieux, J., Thompson, A. M., Rea, D. W., & Stambolic, V. (2021). Cancer antigen 15-3/mucin 1 levels in CCTG MA.32: A breast cancer randomized trial of metformin vs placebo. *JNCI Cancer Spectrum*, 5(5), pkab066. <https://doi.org/10.1093/jncics/pkab066>
- Green, A. R., Soria, D., Stephen, J., Powe, D. G., Nolan, C. C., Kunkler, I., Thomas, J., Kerr, G. R., Jack, W., Cameron, D., Piper, T., Ball, G. R., Garibaldi, J. M., Rakha, E. A., Bartlett, J. M., & Ellis, I. O. (2016). Nottingham Prognostic Index Plus: Validation of a clinical decision making tool in breast cancer in an independent series. *The Journal of Pathology. Clinical Research*, 2(1), 32–40. <https://doi.org/10.1002/cjp.2.32>
- Hsu, J. L., & Hung, M. C. (2016). The role of HER2, EGFR, and other receptor tyrosine kinases in breast cancer. *Cancer Metastasis Reviews*, 35(4), 575–588. <https://doi.org/10.1007/s10555-016-9649-6>
- Katdare, M., Osborne, M., & Telang, N. T. (2002). Soy isoflavone genistein modulates cell cycle progression and induces apoptosis in HER-2/neu oncogene expressing human breast epithelial cells. *International Journal of Oncology*, 21(4), 809–815.
- Kjær, I. M., Olsen, D. A., Brandslund, I., Bechmann, T., Jakobsen, E. H., Bogh, S. B., & Madsen, J. S. (2020). Prognostic impact of serum levels of EGFR and EGFR ligands in early-stage breast cancer. *Scientific Reports*, 10(1), 16558. <https://doi.org/10.1038/s41598-020-72944-1>

- Kunte, S., Abraham, J., & Montero, A. J. (2020). Novel HER2-targeted therapies for HER2-positive metastatic breast cancer. *Cancer*, 126(19), 4278–4288. <https://doi.org/10.1002/cncr.33102>
- Li, Q., Chu, Y., Li, S., Yu, L., Deng, H., Liao, C., Liao, X., Yang, C., Qi, M., Cheng, J., Chen, G., & Huang, L. (2022). The oncoprotein MUC1 facilitates breast cancer progression by promoting Pink1-dependent mitophagy via ATAD3A destabilization. *Cell Death & Disease*, 13(10), 899. <https://doi.org/10.1038/s41419-022-05345-z>
- Li, Y., Bhuiyan, M., & Sarkar, F. H. (1999). Induction of apoptosis and inhibition of c-erbB-2 in MDA-MB-435 cells by genistein. *International Journal of Oncology*, 15(3):525–533. <https://doi.org/10.3892/ijo.15.3.525>
- Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, 46(1–3), 3–26. [https://doi.org/10.1016/s0169-409x\(00\)00129-0](https://doi.org/10.1016/s0169-409x(00)00129-0)
- Loibl, S., & Gianni, L. (2017). HER2-positive breast cancer. *Lancet (London, England)*, 389(10087), 2415–2429. [https://doi.org/10.1016/S0140-6736\(16\)32417-5](https://doi.org/10.1016/S0140-6736(16)32417-5)
- Lülecı, H. B., & Yılmaz, A. (2022). Robust and rigorous identification of tissue-specific genes by statistically extending tau score. *BioData Mining*, 15(1), 31. <https://doi.org/10.1186/s13040-022-00315-9>
- Magis, A. T., Rappaport, N., Conomos, M. P., Omenn, G. S., Lovejoy, J. C., Hood, L., & Price, N. D. (2020). Untargeted longitudinal analysis of a wellness cohort identifies markers of metastatic cancer years prior to diagnosis. *Scientific Reports*, 10(1), 16275. <https://doi.org/10.1038/s41598-020-73451-z>
- Mai, Z., Blackburn, G. L., & Zhou, J. R. (2007). Genistein sensitizes inhibitory effect of tamoxifen on the growth of estrogen receptor-positive and HER2-overexpressing human breast cancer cells. *Molecular Carcinogenesis*, 46(7), 534–542. <https://doi.org/10.1002/mc.20300>
- Makki, J. (2015). Diversity of breast carcinoma: Histological subtypes and clinical relevance. *Clinical Medicine Insights. Pathology*, 8, 23–31. <https://doi.org/10.4137/CPATH.S31563>
- Markiewicz, L., Garey, J., Adlercreutz, H., & Gurdip, E. (1993). In vitro bioassays of non-steroidal phytoestrogens. *The Journal of Steroid Biochemistry and Molecular Biology*, 45(5), 399–405. [https://doi.org/10.1016/0960-0760\(93\)90009-1](https://doi.org/10.1016/0960-0760(93)90009-1)
- Markovits, J., Linassier, C., Fossé, P., Couprie, J., Pierre, J., Jacquemin-Sablon, A., Saucier, J. M., Le Pecq, J. B., & Larsen, A. K. (1989). Inhibitory effects of the tyrosine kinase inhibitor genistein on mammalian DNA topoisomerase II. *Cancer Research*, 49(18), 5111–5117.
- Martin, Y. C. (2005). A bioavailability score. *Journal of Medicinal Chemistry*, 48(9), 3164–3170. <https://doi.org/10.1021/jm0492002>
- Mayer, I. A., & Arteaga, C. L. (2016). The PI3K/AKT pathway as a target for cancer treatment. *Annual Review of Medicine*, 67, 11–28. <https://doi.org/10.1146/annurev-med-062913-051343>
- Muegge, I., Heald, S. L., & Brittelli, D. (2001). Simple selection criteria for drug-like chemical matter. *Journal of Medicinal Chemistry*, 44(12), 1841–1846. <https://doi.org/10.1021/jm015507e>
- Nwabo Kamdje, A. H., Seke Etet, P. F., Vecchio, L., Muller, J. M., Krampera, M., & Lukong, K. E. (2014). Signaling pathways in breast cancer: Therapeutic targeting of the microenvironment. *Cellular Signalling*, 26(12), 2843–2856. <https://doi.org/10.1016/j.cellsig.2014.07.034>
- Pagliacci, M. C., Smacchia, M., Migliorati, G., Grignani, F., Riccardi, C., & Nicoletti, I. (1994). Growth-inhibitory effects of the natural phyto-oestrogen genistein in MCF-7 human breast cancer cells. *European Journal of Cancer (Oxford, England: 1990)*, 30A(11), 1675–1682. [https://doi.org/10.1016/0959-8049\(94\)00262-4](https://doi.org/10.1016/0959-8049(94)00262-4)
- Páll, S., Abraham, M. J., Kutzner, C., Hess, B., & Lindahl, E. (2015). *Tackling exascale software challenges in molecular dynamics simulations with GROMACS*. Solving Software Challenges for Exascale: International Conference on Exascale Applications and Software, EASC 2014, Stockholm, Sweden, April 2–3, 2014, Revised Selected Papers 2 (pp. 3–27). Springer International Publishing.
- Paterni, I., Granchi, C., Katzenellenbogen, J. A., & Minutolo, F. (2014). Estrogen receptors alpha (ER α) and beta (ER β): Subtype-selective ligands and clinical potential. *Steroids*, 90, 13–29. <https://doi.org/10.1016/j.steroids.2014.06.012>
- Schoeberl, B., Kudla, A., Masson, K., Kalra, A., Curley, M., Finn, G., Pace, E., Harms, B., Kim, J., Kearns, J., Fulgham, A., Burenkova, O., Grantcharova, V., Yasar, D., Paragas, V., Fitzgerald, J., Wainszelbaum, M., West, K., Mathews, S., Nering, R., ... MacBeath, G. (2017). Systems biology driving drug development: From design to the clinical testing of the anti-ErbB3 antibody seribantumab (MM-121). *NPJ Systems Biology and Applications*, 3, 16034. <https://doi.org/10.1038/npsba.2016.34>
- Shao, Z. M., Wu, J., Shen, Z. Z., & Barsky, S. H. (1998). Genistein exerts multiple suppressive effects on human breast carcinoma cells. *Cancer Research*, 58(21), 4851–4857.
- Shi, M., Liu, D., Duan, H., Qian, L., Wang, L., Niu, L., Zhang, H., Yong, Z., Gong, Z., Song, L., Yu, M., Hu, M., Xia, Q., Shen, B., & Guo, N. (2011). The β 2-adrenergic receptor and Her2 comprise a positive feedback loop in human breast cancer cells. *Breast Cancer Research and Treatment*, 125(2), 351–362. <https://doi.org/10.1007/s10549-010-0822-2>
- Shimada, K., Ishikawa, T., Nakamura, F., Shimizu, D., Chishima, T., Ichikawa, Y., Sasaki, T., Endo, I., Nagashima, Y., & Goshima, Y. (2014). Collapsin response mediator protein 2 is involved in regulating breast cancer progression. *Breast Cancer (Tokyo, Japan)*, 21(6), 715–723. <https://doi.org/10.1007/s12282-013-0447-5>
- Swain, S. M., Shastry, M., & Hamilton, E. (2023). Targeting HER2-positive breast cancer: Advances and future directions. *Nature Reviews. Drug Discovery*, 22(2), 101–126. <https://doi.org/10.1038/s41573-022-00579-0>
- Telesco, S. E., & Radhakrishnan, R. (2009). Atomistic insights into regulatory mechanisms of the HER2 tyrosine kinase domain: A molecular dynamics study. *Biophysical Journal*, 96(6), 2321–2334. <https://doi.org/10.1016/j.bpj.2008.12.3912>
- Testa, U., Castelli, G., & Pelosi, E. (2020). Breast cancer: A molecularly heterogeneous disease needing subtype-specific treatments. *Medical Sciences (Basel, Switzerland)*, 8(1), 18. <https://doi.org/10.3390/medsci8010018>
- Thul, P. J., & Lindskog, C. (2018). The Human Protein Atlas: A spatial map of the human proteome. *Protein Science: A Publication of the Protein Society*, 27(1), 233–244. <https://doi.org/10.1002/pro.3307>
- Tolaney, S. M., Guo, H., Pernas, S., Barry, W. T., Dillon, D. A., Ritterhouse, L., Schneider, B. P., Shen, F., Fuhrman, K., Baltay, M., Dang, C. T., Yardley, D. A., Moy, B., Marcom, P. K., Albain, E. R., et al. (2018). A phase I study of the HER2 tyrosine kinase inhibitor, neratinib, in combination with trastuzumab and epirubicin in patients with HER2-positive breast cancer. *Journal of Clinical Oncology*, 36(15), 1595–1603. <https://doi.org/10.1200/JCO.2017.728599>

- K. S., Rugo, H. S., Ellis, M. J., Shapira, I., Wolff, A. C., Carey, L. A., ... Winer, E. P. (2019). Seven-year follow-up analysis of adjuvant paclitaxel and trastuzumab trial for node-negative, human epidermal growth factor receptor 2-positive breast cancer. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 37(22), 1868–1875. <https://doi.org/10.1200/JCO.19.00066>
- Trott, O., & Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31(2), 455–461. <https://doi.org/10.1002/jcc.21334>
- Veber, D. F., Johnson, S. R., Cheng, H. Y., Smith, B. R., Ward, K. W., & Kopple, K. D. (2002). Molecular properties that influence the oral bioavailability of drug candidates. *Journal of Medicinal Chemistry*, 45(12), 2615–2623. <https://doi.org/10.1021/jm020017n>
- von Minckwitz, G., Huang, C.S., Mano, M.S., Loibl, S., Mamounas, E. P., Untch, M., Wolmark, N., Rastogi, P., Schneeweiss, A., Redondo, A., Fischer, HH, Jacot, W, Conlin, AK, Arce-Salinas, C, Wapnir, IL, Jackisch, C, DiGiovanna, M. P., Fasching, P. A., Crown, J. P., Wülfing, P., ... Geyer, C. E. Jr. (2019). Trastuzumab emtansine for residual invasive HER2-positive breast cancer. *The New England Journal of Medicine*, 380(7), 617–628. <https://doi.org/10.1056/NEJMoa1814017>
- Webster, R. J., Giles, K. M., Price, K. J., Zhang, P. M., Mattick, J. S., & Leedman, P. J. (2009). Regulation of epidermal growth factor receptor signaling in human cancer cells by microRNA-7. *The Journal of Biological Chemistry*, 284(9), 5731–5741. <https://doi.org/10.1074/jbc.M804280200>
- Zhou, Y., & Liu, X. (2020). The role of estrogen receptor beta in breast cancer. *Biomarker Research*, 8, 39. <https://doi.org/10.1186/s40364-020-00223-2>