

RESEARCH

Open Access



# Uncovering the therapeutic potential of green pea waste in breast cancer: a multi-target approach utilizing LC-MS/MS metabolomics, molecular networking, and network pharmacology

Asmaa M. Khalil<sup>1\*</sup>, Omar M. Sabry<sup>1,2</sup>, Hesham I. El-Askary<sup>1</sup>, Soheir M. El Zalabani<sup>1</sup>, Basma M. Eltanany<sup>3</sup>, Laura Pont<sup>4,5</sup>, Fernando Benavente<sup>4</sup>, Ahmed F. Mohamed<sup>6,7</sup> and Nesrin M. Fayed<sup>1</sup>

## Abstract

**Background** *Pisum sativum* (PS) is a universal legume plant utilized for both human and animal consumption, particularly its seeds, known as green peas. The processing of PS in food industries and households produces a significant amount of waste that needs to be valorized.

**Methods** In this study, the metabolite profiles of the 70% ethanolic extracts of PS wastes, namely peels (PSP) and a combination of leaves and stems (PSLS), were investigated by liquid chromatography-electrospray ionization-quadrupole time-of-flight tandem mass spectrometry (LC-ESI-QTOF-MS/MS) followed by molecular networking.

**Results** Different classes of metabolites were identified, being flavonoids and their derivatives, along with phenolic acids, the most abundant categories. Additionally, a comprehensive network pharmacology strategy was applied to elucidate potentially active metabolites, key targets, and the pathways involved in cytotoxic activity against breast cancer. This cytotoxic activity was investigated in MCF-7 and MCF-10a cell lines. Results revealed that PSLS extract exhibited a potent cytotoxic activity with a good selectivity index ( $IC_{50} = 17.67$  and selectivity index of 3.51), compared to the reference drug doxorubicin ( $IC_{50} = 2.69$   $\mu\text{g/mL}$  and selectivity index of 5.28). Whereas PSP extract appeared to be less potent and selective ( $IC_{50} = 32.92$   $\mu\text{g/mL}$  and selectivity index of 1.62). A similar performance was also observed for several polyphenolics isolated from the PSLS extract, including methyl cis *p*-coumarate, trans *p*-coumaric acid, and liquiritigenin/ 7-methyl liquiritigenin mixture. Methyl cis *p*-coumarate showed the most potent cytotoxic activity against MCF-7 cell line and the highest selectivity ( $IC_{50} = 1.18$   $\mu\text{g/mL}$  (6.91  $\mu\text{M}$ ) and selectivity index of 27.42). The network pharmacology study revealed that the isolated compounds could interact with several breast cancer-associated protein targets including carbonic anhydrases 1, 2, 4, 9, and 12, as well as aldo-keto reductase family 1 member B1, adenosine A3 receptor, protein tyrosine phosphatase non-receptor type 1, and estrogen receptor 2.

\*Correspondence:

Asmaa M. Khalil  
asmaa.khalil@pharma.cu.edu.eg; asmaa.khalil@cu.edu.eg

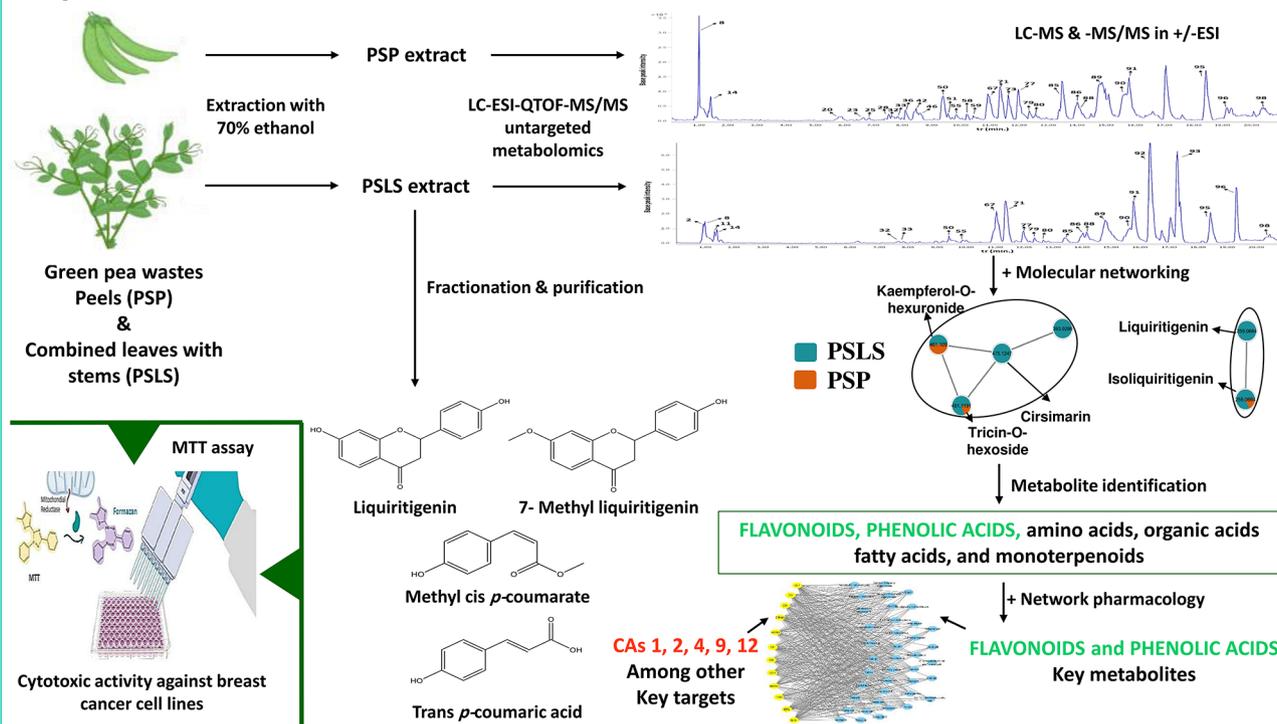
Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

**Conclusion** The uncovered therapeutic potential of PSLs and its metabolite constituents pave the way for an efficient and mindful PS waste valorization, calling for further in-vitro and in-vivo research.

### Graphical Abstract



**Keywords** Breast cancer, Green pea, LC-MS/MS metabolomics, Molecular networking, Network pharmacology, Waste valorization

### Background

*Pisum sativum* (PS) is a well-known legume plant cultivated worldwide as a source of food for both humans and animals. With its remarkable adaptability, PS has a vast geographical distribution, making it a staple crop in many regions of the world [1]. Its small, spherical seeds, commonly known as green peas, are valued as low-cost source of several nutrients, including vitamins, minerals, proteins, and complex carbohydrates [2]. Moreover, PS seeds have been used traditionally for various medicinal purposes, such as treating acne, diabetes, hemorrhoids, and intestinal inflammations [1, 3]. Recent research has confirmed the antioxidant, antimicrobial, antihyperglycemic, antihyperlipidemic, antihypertensive, cardioprotective, and cytotoxic activities of PS seeds. These beneficial effects are attributed to the presence of several bioactive secondary metabolites, including flavonoids, phenolic acids, cinnamyl phenols, lectins, proteins, peptides, polysaccharides, and saponins. PS seeds also constitute a well-known source of various types of nutrients including proteins, soluble and insoluble fibers, complex

carbohydrates, folate, vitamin B, as well as minerals such as calcium, potassium, and iron [4].

The primary wastes generated from PS processing include peels, leaves, and stems [5, 6]. PS seeds are located inside pods and peels are typically discarded as waste. Approximately, 22 million tons of PS seeds are produced annually around the world. PS peels constitute 55% of the total volume of PS pods and are an abundant source of phenolic and bioactive natural dietary antioxidants, along with dietary fibers, protein, and calcium [4]. They are commonly used as animal feed or serve as starting material for the production of fuel or compost [7]. Additionally, PS peel flour is reported to be incorporated in biscuit and cake preparation to impart a natural green pistachio color without adding synthetic colors [4]. Nowadays, new strategies for the effective valorization of food waste are emerging. These strategies involve the extraction of bioactive metabolites and the development of value-added products, such as enzymes, prebiotics, and bioactive functional phytonutrients, from food waste. These bioactive or value-added products can be further incorporated into pharmaceutical, cosmetic, or

food industries, thereby generating additional income and ensuring regional food sustainability [8].

Several biological studies have explored extracts derived from various waste organs of PS. PS peels have been shown to possess antioxidant, cardioprotective, antidiabetic, cytotoxic, and antibacterial properties [9–11], while leaves and roots have exhibited antioxidant activity [12–14]. Various phytochemical classes, including different types of flavonoids such as flavonols, flavones, isoflavones, chalcones, and anthocyanins, along with phenolic acids and pterocarpanes, have been isolated from aerial parts, leaves, shoots, flowers, and roots [15].

Liquid chromatography-mass spectrometry (LC-MS) analysis of PS seeds has revealed the presence of phenolic acids, flavonoids and amino acids [14, 16], whereas only a limited number of phenolic acids, flavonoids, and terpenoids have been identified in PS peels and aerial parts [3, 9, 17]. Consequently, the comprehensive metabolite profiles of common PS wastes, namely peels (PSP) and a combination of leaves and stems (PSLS), as well as their potential anticancer effects, have not yet been fully investigated, despite other biological effects have been reported on these wastes [9–13]. Our study will focus on breast cancer, the most prevalent cancer type among women, second only to lung cancer as the leading cause of death worldwide [18, 19]. Cancer therapies often cause undesirable effects on normal cells, necessitating novel anticancer agents with high selectivity for cancer treatment [19].

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) untargeted metabolomics is a large-scale analysis focused on comprehensive metabolite profiling and chemotaxonomic studies. It significantly contributes to obtaining complete metabolite profiles of natural products, detecting qualitative and quantitative differences between metabolites, and developing hypotheses to explain these differences. This may eventually lead to the identification of bioactive metabolites and the discovery of lead compounds from natural sources [20–24]. Interpreting the massive datasets generated by LC-MS/MS metabolomics requires the use of databases and literature searches. Additionally, it can be aided by molecular networking (MN) [21]. MN is a graph-based workflow that helps organize massive MS data by finding spectral similarities between structurally related precursor ions with different MS/MS fragmentation patterns. MN compares the MS/MS spectra of different ions in a sample and organizes them based on similarities. Many free web-based platforms, such as MetaboAnalyst and the Global Natural Products Social Molecular Networking (GNPS-MN) platform can be utilized for MN. This approach holds promise for metabolite identification, dereplication, and comprehensive analysis of natural products [20–23].

Network pharmacology presents a novel strategy for drug discovery. It facilitates the identification of new bioactive agents from natural products by studying the intricate relationships among drugs, targets, and diseases, as well as the associated pathways and underlying mechanisms of action. This approach is particularly beneficial for complex diseases, including cancer, neurological disorders, and cardiovascular diseases, which involve various pathways, genes, and functional proteins. Consequently, optimized combinations of multiple drugs or drugs having multiple targets, such as plant extracts, could be disclosed to synergistically target different pathways, leading to improved curative effects [25].

In this study, the detailed metabolite profiles of the 70% ethanolic extracts of PSP and PSLS were investigated using liquid chromatography-electrospray ionization quadrupole time-of-flight-tandem mass spectrometry (LC-ESI-QTOF-MS/MS, simplified as LC-MS/MS) untargeted metabolomics followed by MN. The potentially active metabolites, key targets, and the pathways involved in cytotoxic activity against breast cancer were determined through network pharmacology. The efficacy and safety of both extracts and several polyphenolics isolated from the PSLS extract were further investigated in human breast adenocarcinoma (MCF-7) and human mammary epithelial (MCF-10a) cell lines. Our findings will contribute to the valorization of PS waste through the development of effective natural anticancer agents with high safety margins and low incidence of undesirable effects.

## Materials and methods

### Cell lines, chemicals, and reagents

Solvents used in LC-MS/MS analysis, namely acetonitrile, methanol, formic acid ( $\geq 95.0\%$ ), and water, were of LC-MS grade and supplied by Merck (Darmstadt, Germany). Deuterated methanol ( $\text{CD}_3\text{OD}$ , Sigma-Aldrich, St. Louis, USA) was used as solvent for NMR analysis. All other solvents and reagents, used for the extraction and fractionation of PS wastes, including ethanol, distilled water, *n*-hexane, dichloromethane, ethyl acetate, *n*-butanol, and methanol, were of analytical grade and purchased from Piochem (6th of October city, Giza, Egypt). MCF-7 and MCF-10a cell lines were supplied by the American Type Culture Collection (Manassas, USA) and maintained frozen in liquid nitrogen at  $-180^\circ\text{C}$ . The 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) assay kit and the reference anticancer drug, doxorubicin, were provided by Sigma-Aldrich (St. Louis, USA).

### Plant material

PS wastes (ripe green pods without seeds (PSP) and the combined leaves with stems (PSLS)) were collected from

the National Center of Vegetable Research, Dokki, Giza, Egypt, in February 2021. Authentication of the plant material was done by Professor Abdelhaleem Abdelmogally, head of the Taxonomy Department, Agricultural Museum, Dokki, Giza, Egypt. A voucher specimen was deposited in the Herbarium of the Department of Pharmacognosy, Faculty of Pharmacy, Cairo University (No. 30.7.2022).

### Extraction

PSP and PSLs samples were individually air-dried and powdered. Five hundred grams of each powdered sample were extracted separately with 1 L of 70% ethanol using cold maceration (25°C) with the aid of ultra-sonication (Elma Terrasonic TS-540, Bedford, UK). The extraction process was repeated 7 times, then each collective extract was evaporated to dryness at 60°C under vacuum using a rotatory evaporator (Rotavapor® R-100, Büchi, Switzerland), to yield the corresponding solid extracts. The recorded weights of the solid extracts of PSP and PSLs were 50 and 60 g, respectively (i.e. 10 and 12% mass per mass (m/m) of the original waste samples).

### LC-MS/MS untargeted metabolomics

Samples were prepared by individually dissolving 10 mg of each of the solid extracts, previously prepared (Sect. 2.3), in 1 mL of methanol with the aid of ultra-sonication. LC-MS/MS experiments were performed in a 1260 Infinity liquid chromatograph coupled with an orthogonal ESI interface to a 6546 LC/QTOF mass spectrometer (Agilent Technologies, Waldbronn, Germany). A Zorbax SB-C18 column (150 mm total length ( $L_T$ ) × 2.1 mm internal diameter (ID), 5 µm particle size, 90 Å pore diameter, Agilent Technologies) was used for the separations. The chromatographic, MS, and MS/MS conditions were as described in a previous work [26]. Briefly, a gradient elution using mobile phase solvents, (A) water and (B) acetonitrile (both with 0.1% (v/v) of formic acid). The optimized elution was achieved through a gradient of 5% (v/v) of solvent B for 1 min, then raising the percentage of solvent B from 5 to 95% (v/v) in 15 min, followed by cleaning and re-equilibration with 95% (v/v) of solvent B for 2 min, then from 95 to 5% (v/v) of solvent B in 2 min, and finally 5% (v/v) of solvent B for 5 min. Injection volume was 5 µL and flow rate was 350 µL/minute. Experiments were done in triplicate at room temperature. Data acquisition and processing were performed with the ChemStation LC3D software (Agilent Technologies). The obtained raw data files were converted using MS Convert 3.0 open-source software ([www.proteowizard.org](http://www.proteowizard.org)) to mzXML files, which were then imported to MZmine 2.53 open-source software (<https://github.com/mzmine/mzmine2/releases/tag/v2.53>) for peak picking, deconvolution, deisotoping, alignment, and formula

prediction. Detected metabolites were characterized by their accurate molecular monoisotopic mass, retention time ( $t_r$ ), and MS/MS spectra. Metabolite identification was done by comparison with reference literature and online databases (e.g., KEGG, <https://www.genome.jp/kegg/compound/> and PubChem, <https://pubchem.ncbi.nlm.nih.gov/>). Identities were further confirmed by MN as will be discussed in the upcoming sections. Therefore, annotated metabolites were identified at a high confidence level (Level 2a: probable structure, MS, MS/MS, and library/bibliography search) [27]. Examples of MS/MS fragmentation patterns of selected metabolites from each class are represented by Figures S1-S23.

### Molecular networking

The converted mzXML files were uploaded, using WinSCP cross-platform, to the GNPS server (<https://gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp>). The parameters for network generation were set as mentioned by Hamed et al. [28] except that the minimum pairs cosine and the minimum matched fragment ions were set to 0.65 and 3, respectively. The created MN and parameters for the negative and positive ESI modes can be accessed through the following links: <https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=f56500ffb179472bb8b7d3b3098925e8> and <https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=aa58b89c813a4dc38e1e72c37acf670d>, respectively. The output network was merged to Feature Based Molecular Network (FBMN) and imported to the open-source software platform, Cytoscape 3.9.1 software (<https://cytoscape.org/download.html>) for network visualization and interpretation.

### Network pharmacology study of potential therapeutic activity against breast cancer

#### Construction of compound-target-disease network

The respective Simplified Molecular Input Line Entry Specification (SMILES) formulas of metabolites identified through LC/MS-MS untargeted metabolomics were generated and used for screening of compound-target interactions. Protein targets for the identified metabolites were extracted from two databases, SwissTargetPrediction (<http://www.swisstargetprediction.ch/>) and Similarity Ensemble Approach (SEA) (<https://sea.bkmlab.org/>). The extracted protein targets were combined, dereplicated, and code-unified using UniProt Retrieve/ID mapping (<https://www.uniprot.org/uploadlists/>). Disease targets were determined using GeneCards (<https://www.genecards.org/>) and DisGeNet (<https://www.disgenet.org/home/>) databases. “Breast cancer” and “breast adenocarcinoma” were used as search terms. Then the identified genes were assembled and dereplicated. A Venn’s diagram was generated using Venny (<https://bioinformatics.csic.es/tools/venny/>) to show the common tar

gets between compound-target and disease-target pairs. Cytoscape 3.9.1 software was used for visualizing and interpreting the compound-target-disease network.

#### **Protein-protein interaction (PPI) network**

The aforementioned common targets were imported to the STRING database (<https://string-db.org/>) to reveal the PPI. Confidence limit was set to  $\geq 0.7$ .

#### **Enrichment of gene ontology (GO) terms and Kyoto encyclopedia of genes and genome (KEGG) pathways**

Functional analysis and enrichment of the gene targets were performed utilizing the online tool ShinyGO 0.77 and its integrated tools (<http://bioinformatics.sdstate.edu/go77/>).

#### **Isolation of phenolic components from PSLs**

Fifty grams of the solid extract of PSLs were fractionated using solvents with increasing polarities, namely *n*-hexane, dichloromethane, ethyl acetate, and *n*-butanol (10 times  $\times$  300 mL each). The ethyl acetate fraction was evaporated under vacuum at 60°C resulting in 3 g of the dried fraction, which was then suspended in 30 mL of distilled water and subjected to column chromatography on a diaion® HP-20 column (45 g, 35 cm  $L_T \times$  2.5 cm ID, Merck, Darmstadt, Germany). Elution started with 500 mL of distilled water. Then, the percentage of methanol was increased by 25% increments in each step. The resulting five sub-fractions, namely water, 25%, 50%, 75%, and 100% methanol (500 mL each) were analyzed by thin layer chromatography (TLC). Then they were evaporated under vacuum at 60°C, weighed, and kept in a desiccator at room temperature for further analysis. The 75% methanol sub-fraction (100 mg) was dissolved in a 9:1 mixture of methanol and water and subjected to column chromatography on a Sephadex® LH-20 (8 g, 28 cm  $L_T \times$  1.5 cm ID, Merck, Darmstadt, Germany) using isocratic elution with 90% methanol. Fractions of 2 mL were collected and analyzed by TLC. Fractions 12 and 13 were pooled and evaporated under vacuum at 50°C resulting in 10 mg of component 1. Similarly, fractionation by column chromatography of the 50% sub-fraction (137 mg) on the Sephadex® LH-20 column using isocratic elution with 50% methanol resulted in the isolation of components 2 and 3 (5 and 9 mg, respectively). Schematic representation of the steps of extraction, fractionation, and isolation to obtain these components is represented in Figure S24. The structures of the isolated phenolic components were elucidated by nuclear magnetic resonance (NMR), using a U28-E04 Bruker Ascend 400 MHz NMR spectrometer (Karlsruhe, Germany). The NMR spectra were compared to those found in reference literature for confirmation of their identities [29, 30].

#### **MTT cytotoxicity assay**

The MTT assay is a colorimetric method used for measurement of cell proliferation, activity, and viability. It relies on the ability of mitochondrial dehydrogenase enzymes within living cells to transform the color of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) from yellow to purple. This color change serves as an indicator of cell viability [31]. To assess the in-vitro cytotoxic activity against MCF-7 and MCF-10a cell lines of PSP and PSLs ethanolic extracts, as well as the isolated phenolic components from PSLs extract and the reference drug doxorubicin, the MTT assay was conducted as described in a previous study [31]. The presence of viable cells was visualized through the development of purple color. The optical density (OD) at 595 nm was measured using a ROBONIK P2000 ELISA reader (Robonik India Pvt Ltd, Morivali, India), and the %viability was calculated at different concentration levels using the following equation:

$$\% \text{Viability} = \frac{\text{OD of sample}}{\text{OD of control (no drug)}} \times 100$$

The half-maximal inhibitory concentration ( $IC_{50}$ ) values were obtained for each sample after plotting the %viability against the concentration.

## **Results**

#### **LC-MS/MS untargeted metabolomics**

The LC-MS/MS untargeted metabolomic study revealed the presence of several classes of secondary metabolites including flavonoids and their derivatives, phenolic acids, amino acids and their derivatives, organic acids, fatty acids and their derivatives, monoterpenoids, and other miscellaneous compounds. A total of 98 metabolites were identified in PSP and PSLs extracts, including 39 flavonoids and their derivatives, 9 phenolic acids, 14 amino acids and their derivatives, 6 organic acids, 18 fatty acids and their derivatives, 4 monoterpenoids, and 8 other miscellaneous compounds. Metabolite identities,  $t_r$ , mass to charge ratios ( $m/z$ ) of the detected molecular ions and fragment ions, and the predicted molecular formulas are summarized in Table 1. The base peak chromatograms of PSP and PSLs extracts in the positive and negative ionization modes are depicted in Figs. 1 and 2, respectively.

#### **MN study**

MN was used to further confirm the metabolite identification in PS samples. In this approach, molecules are grouped into clusters based on common fragmentation patterns. Molecules with similar fragmentation patterns are interconnected, while those with dissimilar patterns are separated. Molecules that do not form groups are represented as single nodes. The color of network nodes

**Table 1** Metabolites identified in *Pisum sativum* L. (green pea, PS) waste extracts by LC-MS/MS untargeted metabolomics in negative and positive ESI modes

Peak	t <sub>r</sub> (min.)	Mo- lecular ion (M-H) <sup>-</sup> (m/z)	MS/MS fragments (m/z)	Negative mode	Molecu- lar ion (M+H) <sup>+</sup> (m/z)	MS/MS fragments (m/z)	Positive mode	Molecular Formula	Identification	Extract PSP	Ref.
<b>Flavonoids and their derivatives</b>											
Flavonols											
33	7.83	461.1089	285					C <sub>21</sub> H <sub>18</sub> O <sub>12</sub>	Kaempferol-O-hexuronide	+	[32]
37	8.24	593.1504	447, 285, 283, <b>253*</b>					C <sub>27</sub> H <sub>30</sub> O <sub>15</sub>	Kaempferol-O-neohesperidoside	+	[33]
42	8.62	577.1559	431, <b>285*</b> , 283, 255		579.1730	287		C <sub>27</sub> H <sub>30</sub> O <sub>14</sub>	kaempferol-3,7-O-di-deoxyhexoside	+	
43	8.63	933.2301	787, 625, 463, 413, 301, <b>300*</b> , 255					C <sub>42</sub> H <sub>46</sub> O <sub>24</sub>	Quercetin-O-coumaroyl sophorotrioside	+	(+)
44	8.64	963.2385	787, 623, 463, 343, 301, <b>300*</b> , 255					C <sub>43</sub> H <sub>48</sub> O <sub>25</sub>	Quercetin-O-feruloyl sophorotrioside	+	(+)
45	8.88	947.2451	786, 771, 609, 542, 448, 383, <b>285*</b> , 283, 255					C <sub>43</sub> H <sub>48</sub> O <sub>24</sub>	Kaempferol-O-feruloyl sophorotrioside	+	(+)
46	8.92				433.1138	<b>287*</b> , 271, 153, 99		C <sub>21</sub> H <sub>20</sub> O <sub>10</sub>	Kaempferol-O-deoxyhexoside	+	
47	8.94	917.2339	867, 771, 609, 447, 429, 312, <b>285*</b> , 198					C <sub>42</sub> H <sub>46</sub> O <sub>23</sub>	Kaempferol-O-coumaroyl sophorotrioside	+	(+)
48	9.09	447.1293	365, <b>285*</b> , 270, 161, 123					C <sub>21</sub> H <sub>20</sub> O <sub>11</sub>	Kaempferol-O-hexoside	+	(+)
63	10.60	269.0464	<b>269*</b> , 225, 211, 185, 59		271.0601	<b>192*</b> , 176, 98		C <sub>15</sub> H <sub>10</sub> O <sub>5</sub>	Galangin	+	[34]
Flavones											
34	7.91	563.1404	563, 473, 443, <b>383*</b> , 353, 282, 238, 208, 165, 143, 89					C <sub>26</sub> H <sub>28</sub> O <sub>14</sub>	Apigenin-6-C-pentoside-8-C-hexoside	+	- 35-37
35	8.04	491.1201	329, <b>314*</b> , 299, 271, 189, 123					C <sub>23</sub> H <sub>24</sub> O <sub>12</sub>	Tricin-O-hexoside	+	
36	8.20	577.1559	577, 487, 473, 439, 413, 395, 383, 365, <b>353*</b> , 282, 221, 170, 91		579.1708	443, 409, <b>379*</b> , 309, 294, 271, 119		C <sub>27</sub> H <sub>30</sub> O <sub>14</sub>	Apigenin-6-C-deoxyhexoside-8-C-hexoside	+	
54	9.81	475.1247	475, 343, <b>239*</b> , 187, 125		477.1413	<b>315*</b> , 300, 272, 127, 85		C <sub>23</sub> H <sub>24</sub> O <sub>11</sub>	Cirsimaritin	+	
56	9.94	447.1293	447, <b>285*</b> , 270, 137					C <sub>21</sub> H <sub>20</sub> O <sub>11</sub>	Luteolin-O-hexoside	+	(+)
60	10.53	269.0464	<b>269*</b> , 240, 223, 195, 169, 135, 121, 109, 93		271.0600	<b>271*</b> , 215, 161, 137, 121, 105, 88		C <sub>15</sub> H <sub>10</sub> O <sub>5</sub>	7,3,4'-Trihydroxyflavone	+	
62	10.59	299.0563	284, 271, <b>255*</b> , 227, 211, 185, 175, 149, 135, 123, 108, 81, 59					C <sub>16</sub> H <sub>12</sub> O <sub>6</sub>	Chrysoeriol	+	(+)
73	11.44	537.0823	451, 417, 385, <b>375*</b> , 309, 257, 203					C <sub>30</sub> H <sub>18</sub> O <sub>10</sub>	3,8'-Biapigenin	+	
75	11.64	313.0719	313, 298, 283, 270, 255, 239, <b>226*</b> , 211, 183, 167		315.0873	<b>315*</b> , 300, 272, 255, 227, 167		C <sub>17</sub> H <sub>14</sub> O <sub>6</sub>	Cirsimaritin	+	(+)
79	12.31	297.0775	<b>282*</b> , 267, 254, 239, 151, 93, 73					C <sub>17</sub> H <sub>14</sub> O <sub>5</sub>	3,7-Dimethoxy-3-hydroxyflavone	+	(+)
88	14.54	269.0456	<b>269*</b> , 254, 225, 220, 195, 179, 105					C <sub>15</sub> H <sub>10</sub> O <sub>5</sub>	Baicalin	+	(+)
Flavanones											
28	7.56	323.1248	292, 282, 229, 203, <b>189*</b> , 175, 161, 146, 138, 97, 71, 59					C <sub>20</sub> H <sub>20</sub> O <sub>4</sub>	6-Prenyl pinocembrin	+	[38]

**Table 1** (continued)

Peak	t <sub>r</sub> (min.)	Molecular ion (M-H) <sup>-</sup> (m/z)	MS/MS fragments (m/z) Negative mode	Molecular ion (M+H) <sup>+</sup> (m/z)	MS/MS fragments (m/z) Positive mode	Molecular Formula	Identification	Extract		Ref.
								PSLS	PSP	
<b>Flavonoids and their derivatives</b>										
39	8.45	549.1619	549, 417, <b>255*</b> , 135, 119, 89, 59			C <sub>26</sub> H <sub>30</sub> O <sub>13</sub>	Liquiritigenin-O-pentosyl hexoside	+	-	[39]
41	8.58	417.1194	255, 149, <b>135*</b> , 119			C <sub>21</sub> H <sub>22</sub> O <sub>9</sub>	Liquiritigenin-O-hexoside	+	-	
58	10.44	255.0665	255, 135, <b>119*</b> , 91	257.0815	257, 239, 168, 147, <b>137</b> , 93	C <sub>15</sub> H <sub>12</sub> O <sub>4</sub>	Liquiritigenin	+	-	
69	11.18	271.0977	255, 226, 184, <b>165*</b> , 150, 137, 123, 108, 93, 80			C <sub>15</sub> H <sub>12</sub> O <sub>5</sub>	Naringenin	+	(+)	
74	11.53			287.0924	269, 255, 241, 227, 209, <b>153</b> , 123, 110	C <sub>16</sub> H <sub>14</sub> O <sub>5</sub>	Isosakuranetin	(+)	+	
82	12.94	269.0822	<b>254*</b> , 221, 210, 135, 121, 119, 98	271.0979	271, 241, 210, 177, <b>137*</b> , 109, 79	C <sub>16</sub> H <sub>14</sub> O <sub>4</sub>	7-methyl liquiritigenin	+	(+)	
<b>Isoflavones</b>										
57	10.21	253.0505	<b>253*</b> , 224, 197, 179, 135, 117, 91	255.0652	<b>255*</b> , 237, 199, 150, 137, 131, 118	C <sub>15</sub> H <sub>10</sub> O <sub>4</sub>	Daidzein	+	+	[40]
59	10.50	283.0613	<b>268*</b> , 239, 211, 148, 120			C <sub>16</sub> H <sub>12</sub> O <sub>5</sub>	5-methyl genistein	+	(+)	
64	10.88			285.0765	285, <b>270*</b> , 253, 225, 183, 169, 157, 134	C <sub>16</sub> H <sub>12</sub> O <sub>5</sub>	Calycosin	(+)	+	
66	11.00			315.0874	<b>315*</b> , 300, 282, 255, 240, 199, 167, 139, 119	C <sub>17</sub> H <sub>14</sub> O <sub>6</sub>	Dipteryxin	+	(+)	
72	11.39	269.0464	<b>269*</b> , 254, 225, 201, 169, 133, 107			C <sub>15</sub> H <sub>10</sub> O <sub>5</sub>	Genistein	+	(+)	[40]
78	12.16	267.0664	<b>267</b> , <b>252*</b> , 223, 208, 195, 132	269.0816	<b>269*</b> , 253, 226, 213, 198, 185, 155, 136	C <sub>16</sub> H <sub>12</sub> O <sub>4</sub>	7-methyl daidzein	+	(+)	
81	12.60			299.0922	<b>299*</b> , 284, 243, 211	C <sub>17</sub> H <sub>14</sub> O <sub>5</sub>	5,7-Dimethyl genistein	+	(+)	
84	13.08			283.0973	<b>283*</b> , 268, 240, 212, 197, 181, 161, 148, 121, 91	C <sub>17</sub> H <sub>14</sub> O <sub>4</sub>	7,4'-Dimethoxy isoflavone	+	(+)	
85	13.36	283.0611	<b>268*</b> , 239, 211, 195, 183, 168, 148, 120	285.0765	<b>285*</b> , 270, 257, 242, 229, 167	C <sub>16</sub> H <sub>12</sub> O <sub>5</sub>	Prunetin	+	(+)	
<b>Chalcones</b>										
65	10.89	269.0822	<b>269*</b> , 253, 241, 225, 212, 193, 181, 161, 148, 136, 118, 108, 92	271.0979	271, 241, 177, 161, <b>137</b> , 109	C <sub>16</sub> H <sub>14</sub> O <sub>4</sub>	4,4'-dihydroxy-2'-methoxychalcone	+	(+)	[41]
77	11.98	255.0663	255, 213, 149, 135, <b>119*</b> , 91	257.0808	<b>257*</b> , 165, 137, 90, 68	C <sub>15</sub> H <sub>12</sub> O <sub>4</sub>	Isoliquiritigenin	+	(+)	
<b>Phenolic acids</b>										
23	6.66	165.0558	164, 150, 121, 107, <b>93*</b> , 80, 52			C <sub>9</sub> H <sub>10</sub> O <sub>3</sub>	Hydroxyphenyl propanoic acid	+	-	[42]
25	6.75	153.0194	153, 109, <b>108*</b> , 96, 81, 55			C <sub>7</sub> H <sub>6</sub> O <sub>4</sub>	Genistic acid	(+)	+	
32	7.64	153.0194	153, 109, <b>108*</b> , 96, 81, 55			C <sub>7</sub> H <sub>6</sub> O <sub>4</sub>	Protocatechuic acid	(+)	+	
40	8.54	163.0401	163, <b>119*</b> , 95	165.0551	<b>147*</b> , 119, 91, 65	C <sub>9</sub> H <sub>8</sub> O <sub>3</sub>	p-Coumaric acid	+	(+)	
53	9.75	181.0509	181, 109, <b>108*</b> , 93, 69			C <sub>9</sub> H <sub>10</sub> O <sub>4</sub>	Veratric acid	+	(+)	

**Table 1** (continued)

Peak	t <sub>r</sub> (min.)	Molecular ion (M-H) <sup>-</sup> (m/z)	MS/MS fragments (m/z) Negative mode	Molecular ion (M+H) <sup>+</sup> (m/z)	MS/MS fragments (m/z) Positive mode	Molecular Formula	Identification	Extract		Ref.
								PSLS	PSP	
<b>Flavonoids and their derivatives</b>										
55	9.89	137.0243	<b>93*</b> , 65			C <sub>7</sub> H <sub>6</sub> O <sub>3</sub>	Salicylic acid	(+)	+	[42]
61	10.57	151.0402	<b>107*</b> , 92, 65			C <sub>6</sub> H <sub>6</sub> O <sub>3</sub>	p-Hydroxyphenyl acetic acid	+	+	
80	12.57	193.0870	<b>193*</b> , 177, 121, 89, 74			C <sub>10</sub> H <sub>10</sub> O <sub>4</sub>	Ferulic acid	+	+	
83	12.97			181, 1221	181, 163, 145, <b>135*</b> , 121, 107, 91, 79	C <sub>9</sub> H <sub>8</sub> O <sub>4</sub>	Caffeic acid	+	(+)	
<b>Amino acids and their derivatives</b>										
1	0.93	118.0510	118, 100, <b>74*</b> , 56	120.0652	<b>74, 56*</b>	C <sub>4</sub> H <sub>9</sub> NO <sub>3</sub>	Threonine	(+)	+	[43]
2	0.95	132.0300	115, 100, <b>88*</b> , 79, 71, 59			C <sub>4</sub> H <sub>7</sub> NO <sub>4</sub>	Aspartic acid	+	(+)	
4	0.97			175.0584	175, 130, 112, <b>70*</b>	C <sub>6</sub> H <sub>14</sub> N <sub>4</sub> O <sub>2</sub>	Arginine	+	(+)	
7	1.02	114.0563	<b>114*</b> , 83, 70, 56, 50	116.0709	116, <b>70*</b>	C <sub>5</sub> H <sub>9</sub> NO <sub>2</sub>	Proline	(+)	+	
8	1.04	118.0510	118, <b>99*</b> , 74, 72, 60	120.0657	<b>74, 56*</b>	C <sub>4</sub> H <sub>9</sub> NO <sub>3</sub>	Homoserine	+	(+)	
9	1.05	290.0881	212, 200, 170, <b>128*</b> , 101, 84, 70			C <sub>11</sub> H <sub>17</sub> NO <sub>8</sub>	N-Fructosyl pyroglutamate	(+)	+	
10	1.06	160.0616	160, 142, 118, 98, <b>74*</b> , 58			C <sub>6</sub> H <sub>11</sub> NO <sub>4</sub>	O-acetyl homoserine	(+)	+	
11	1.30	116.0716	99, 92, 78, <b>75*</b>	118.0868	<b>118*</b> , 72, 55, 58	C <sub>5</sub> H <sub>11</sub> NO <sub>2</sub>	Valine	+	+	
15	1.60	130.0874	<b>130*</b> , 117, 83, 69, 58	132.1022	<b>86*</b> , 69, 58	C <sub>6</sub> H <sub>13</sub> NO <sub>2</sub>	Leucine	+	+	
16	1.67	180.0666	172, 163, 157, 120, 115, <b>93*</b> , 72	182.0817	165, 147, 136, <b>119*</b> , 91	C <sub>9</sub> H <sub>11</sub> NO <sub>3</sub>	Tyrosine	+	+	
17	1.77			132.1025	<b>86*</b> , 73, 69	C <sub>6</sub> H <sub>13</sub> NO <sub>2</sub>	Isoleucine	+	+	
18	2.57	164.0717	147, <b>103*</b> , 77, 72	205.0975	188, 170, 160, <b>146*</b> , 132, 118, 89, 56	C <sub>9</sub> H <sub>11</sub> NO <sub>2</sub>	Phenylalanine	+	+	
21	6.25	203.0830	186, 159, 142, <b>116*</b> , 72			C <sub>11</sub> H <sub>12</sub> N <sub>2</sub> O <sub>2</sub>	Tryptophan	+	+	
31	7.62	172.0978	<b>130*</b> , 82			C <sub>8</sub> H <sub>15</sub> NO <sub>3</sub>	N-acetyl leucine	(+)	+	
<b>Organic acids</b>										
3	0.96	195.0511	195, 177, 159, 141, 129, 99, 87, <b>75*</b> , 59			C <sub>6</sub> H <sub>12</sub> O <sub>7</sub>	Gluconic acid	(+)	+	[44]
5	0.99	135.0295	135, 87, <b>75*</b> , 59			C <sub>4</sub> H <sub>8</sub> O <sub>5</sub>	Threonic acid	+	+	
12	1.41	191.0195	129, <b>111*</b> , 87, 67, 57			C <sub>6</sub> H <sub>8</sub> O <sub>7</sub>	Citric acid	+	(+)	
13	1.48	117.0195	<b>73</b>			C <sub>4</sub> H <sub>6</sub> O <sub>4</sub>	Succinic acid	+	(+)	
14	1.57	161.0456	99, 59, <b>57*</b>			C <sub>6</sub> H <sub>10</sub> O <sub>5</sub>	3-Hydroxy-3-methylglutaric acid	(+)	+	
22	6.28	175.0613	<b>115*</b> , 85, 69, 59			C <sub>7</sub> H <sub>12</sub> O <sub>5</sub>	2-Isopropylmalic acid	(+)	+	
<b>Fatty acids and their derivatives</b>										

**Table 1** (continued)

Peak	t <sub>r</sub> (min.)	Molecular ion (M-H) <sup>-</sup> (m/z)	MS/MS fragments (m/z) Negative mode	Molecular ion (M+H) <sup>+</sup> (m/z)	MS/MS fragments (m/z) Positive mode	Molecular Formula	Identification	Extract		Ref.
								PSLS	PSP	
<b>Flavonoids and their derivatives</b>										
29	7.60	255.1241	255, 237, 240, 211, <b>182*</b> , 167, 149, 138, 97, 57			C <sub>13</sub> H <sub>20</sub> O <sub>5</sub>	Tricladic acid A	+	(+)	[44, 45]
30	7.61	433.2085	271, <b>234*</b> , 207, 159, 95, 75			C <sub>30</sub> H <sub>34</sub> O <sub>10</sub>	Peroxydisebacic acid	+	(+)	
51	9.40	187.0978	187, 169, 143, <b>125*</b> , 97, 57			C <sub>9</sub> H <sub>16</sub> O <sub>4</sub>	Azelaic acid	+	+	
67	11.06	327.2175	327, 309, 291, 283, 269, 251, 229, 211, <b>183*</b> , 171, 125, 97, 85, 71, 57			C <sub>18</sub> H <sub>32</sub> O <sub>5</sub>	9,12,13-trihydroxyoctadeca-10,15-dienoic acid	(+)	+	
68	11.14	215.1287	<b>197*</b> , 171, 153, 134, 102			C <sub>11</sub> H <sub>20</sub> O <sub>4</sub>	Undecanedioic acid	+	+	
71	11.35	329.2335	311, 293, 229, <b>211*</b> , 183, 171, 155, 139, 127, 99			C <sub>18</sub> H <sub>34</sub> O <sub>5</sub>	5,8,11-trihydroxyoctadec-9-enoic acid	(+)	+	
86	14.04	313.2384	313, 295, 277, 269, 195, <b>183*</b> , 129, 99, 58			C <sub>18</sub> H <sub>34</sub> O <sub>4</sub>	Octadecanedioic acid	(+)	+	[44, 45]
87	14.45			274.2752	<b>274*</b> , 256, 106, 88, 70, 57	C <sub>16</sub> H <sub>33</sub> NO <sub>2</sub>	Lauryldiethanolamine	+	+	
89	14.98	275.2017	275, 257, 231, 218, 167, <b>106*</b> , 101, 71, 59			C <sub>18</sub> H <sub>36</sub> O <sub>2</sub>	Stearidonic acid	+	(+)	
90	15.77	295.2278	295, <b>277*</b> , 251, 195, 171			C <sub>18</sub> H <sub>32</sub> O <sub>3</sub>	9-hydroxy-10,12-octadecadienoic acid	-	+	
91	16.36	297.2438	297, 279, 253, <b>185*</b>			C <sub>18</sub> H <sub>34</sub> O <sub>3</sub>	12-Hydroxyoctadec-9-enoic acid	-	+	
92	16.59			281.2483	245, 133, 109, 95, 83, <b>69*</b> , 55	C <sub>18</sub> H <sub>32</sub> O <sub>2</sub>	Chaulmoogric Acid	-	+	
93	17.29	299.2594	<b>299*</b> , 281, 255, 253, 141			C <sub>18</sub> H <sub>36</sub> O <sub>3</sub>	12-Hydroxyoctadecanoic acid	-	+	
94	17.49			283.2644	135, 121, 107, 97, 83, 69, <b>57*</b>	C <sub>18</sub> H <sub>34</sub> O <sub>2</sub>	Petroselinic acid	-	+	
95	18.44	271.2279	271, <b>225*</b> , 227			C <sub>16</sub> H <sub>32</sub> O <sub>3</sub>	Hydroxy palmitic acid	+	+	
96	19.31	279.2331	<b>279*</b> , 261, 235, 96			C <sub>18</sub> H <sub>32</sub> O <sub>2</sub>	Linoleic acid	(+)	+	
97	19.68	313.2752	313, <b>267*</b> , 251, 171, 57			C <sub>19</sub> H <sub>38</sub> O <sub>3</sub>	Hydroxy nonadecanoic acid	+	(+)	
98	20.48	281.2490	<b>281*</b> , 237, 242, 98			C <sub>18</sub> H <sub>34</sub> O <sub>2</sub>	Oleic acid	(+)	+	
<b>Monoterpenoids</b>										
27	7.34	431.1922	333, 136, <b>89*</b> , 59			C <sub>20</sub> H <sub>32</sub> O <sub>10</sub>	Sacranoside A	+	-	[46]
38	8.39	401.1823	401, 239, 221, 195, 177, 123, 101, 89, 85, <b>59*</b>			C <sub>19</sub> H <sub>30</sub> O <sub>9</sub>	Cannabisside D	+	(+)	[47]
50	9.38	239.1291	238, 220, 195, 179, <b>158*</b> , 125, 123, 109, 78, 57			C <sub>13</sub> H <sub>20</sub> O <sub>4</sub>	Aglycone of cannabisside D	+	(+)	[47]
52	9.62			197.1178	197, 161, <b>133*</b> , 93, 79	C <sub>11</sub> H <sub>16</sub> O <sub>3</sub>	Loliolide	+	(+)	[48]
<b>Miscellaneous compounds</b>										
6	1.00	179.0563	179, 161, 95, 85, <b>75*</b> , 59			C <sub>6</sub> H <sub>12</sub> O <sub>6</sub>	Glucose	(+)	+	[49]
19	2.60	218.1034	146, 99, <b>88*</b> , 71			C <sub>9</sub> H <sub>17</sub> NO <sub>5</sub>	Pantothenic acid	+	+	
20	5.87	217.0720	217, 171, 155, 143, 115, 85, <b>72*</b>			C <sub>9</sub> H <sub>14</sub> O <sub>6</sub>	Triacetin	+	-	[50]
24	6.72	109.0294	109, <b>108*</b>			C <sub>6</sub> H <sub>6</sub> O <sub>2</sub>	Catechol	+	+	[51]
26	7.04			188.0719	188, 170, 146, 127, <b>118*</b> , 85	C <sub>11</sub> H <sub>9</sub> NO <sub>2</sub>	Indoleacrylic acid	+	+	[49]
49	9.12	361.0570	<b>297*</b> , 285, 240, 227, 197, 179, 161, 148, 136, 122, 109			C <sub>20</sub> H <sub>26</sub> O <sub>6</sub>	Secosolaricresinol	+	(+)	[52]

**Table 1** (continued)

Peak	$t_r$ (min.)	Mo- lecular ion (M-H) <sup>-</sup> (m/z)	MS/MS fragments (m/z)		Molecu- lar ion (M + H) <sup>+</sup> (m/z)	MS/MS fragments (m/z) Positive mode	Molecular Formula	Identification	Extract		Ref.
			Negative mode						PSLS	PSP	
<b>Flavonoids and their derivatives</b>											
70	11.25	297.0395	297*, 269, 241, 211, 197, 181, 149, 93, 71		299.0922	299*, 284, 243, 211, 151	C <sub>17</sub> H <sub>14</sub> O <sub>5</sub>	Pterocarpin	+	+	[53]
76	11.96	941.5119	941*, 795, 733, 633, 615, 457, 205, 101, 59		943.5306	383, 271, 247, 203, 163, 141, 71*	C <sub>48</sub> H <sub>78</sub> O <sub>18</sub>	Soyasaponin Bb	+	+	[54]

+: Present, (+): Present in lower abundance, -: Absent, \*: Base peak ("The most intense peak"), PSLS: *Pisum sativum* leaves and stems extract, PSP: *Pisum sativum* peels extract

corresponds to the sample type and is labeled with the precursor ion  $m/z$  values. Nodes are displayed as pie charts to indicate the relative abundance of detected molecular ions in the investigated samples [20–23]. The MN study results are represented in Fig. 3.

#### Isolation of phenolic components from PSLS

In order to get a complementary insight on the PSLS extract composition, different phenolic components were isolated from its ethyl acetate fraction. NMR analysis of the isolated components led to the identification of component 1 as a mixture of two closely related flavanones (liquiritigenin and 7-methyl liquiritigenin). Components 2 and 3 were identified as trans *p*-coumaric acid, and methyl cis *p*-coumarate, respectively [29, 30]. The structures of the isolated components are represented by Figure S25. Figures S26–S34 represent the NMR spectra of the isolated components.

#### Network pharmacology study of potential therapeutic activity against breast cancer

Network pharmacology functions as a comprehensive in-silico approach that constructs a “protein–compound/disease–gene” network. This aids in revealing the underlying mechanisms and synergistic actions of plant extracts, shifting the paradigm from “one-target, one-drug” to a “network-target, multiple-component-therapeutics” [55]. The identified metabolites were investigated through network pharmacology to discern the key metabolites associated to breast cancer protein targets and genes, as well as to elucidate their potential mechanisms of action.

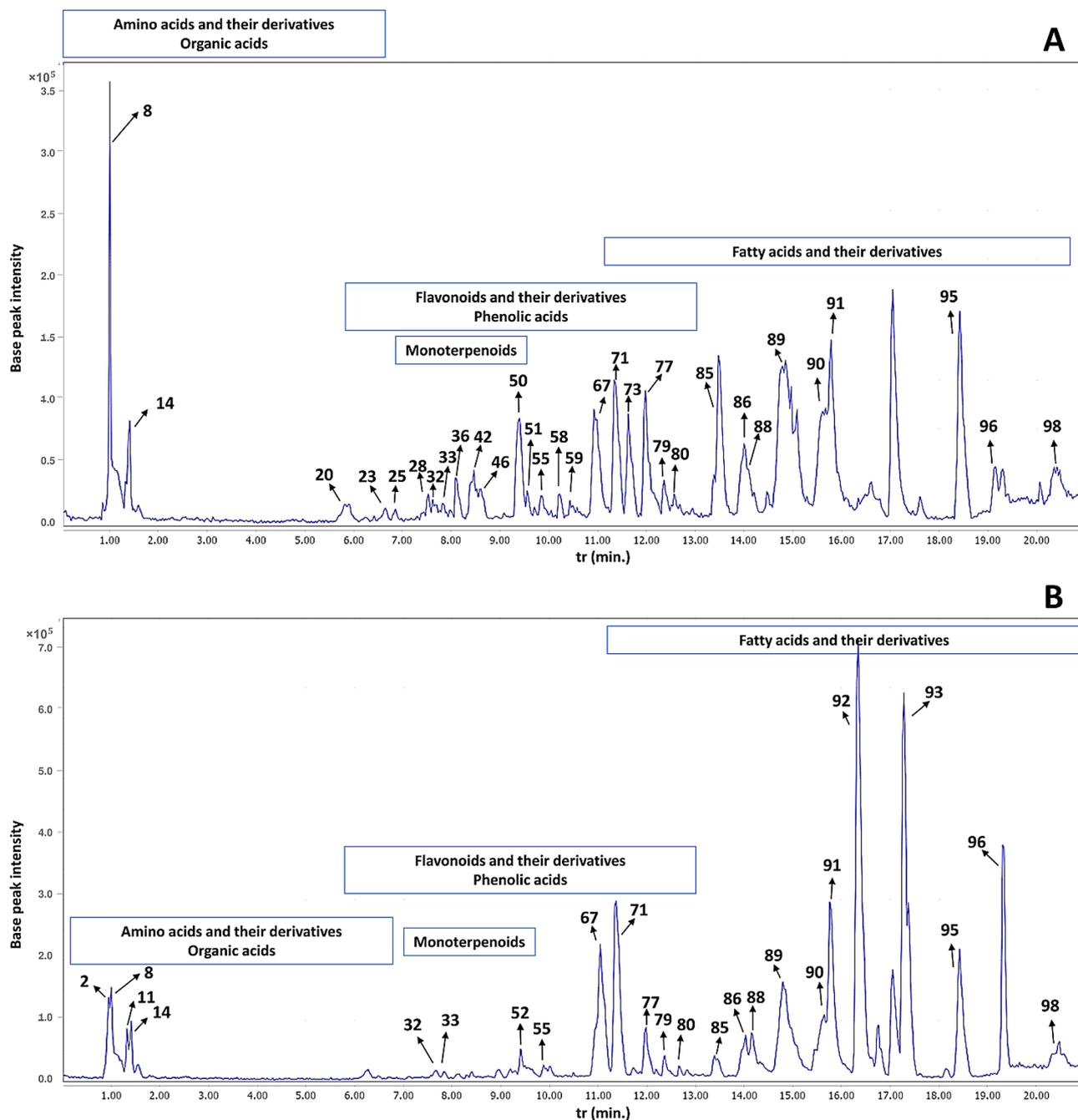
#### Compound-target-disease network

Potential protein targets for all the identified metabolites in both PS waste extracts were retrieved from the SwissTargetPrediction and SEA databases. The potential protein targets linked to breast cancer were also determined. Venn’s diagram of Fig. 4 showed 925 common targets between the two domains. The common targets with the highest degree of involvement are summarized in Table S1.

The analysis of the constructed network (Fig. 5) revealed that flavonoids and their derivatives along with phenolic acids were the key metabolites, as they exhibited the highest degree of involvement, as shown in Table 2.

#### PPI network

A PPI network is a subcellular web of nodes (proteins) connected by edges (interactions between these proteins). This approach studies the interactions between the common protein targets identified from the constructed compound-target-disease network, thus aids in gaining a global view of the investigated biological or pathological



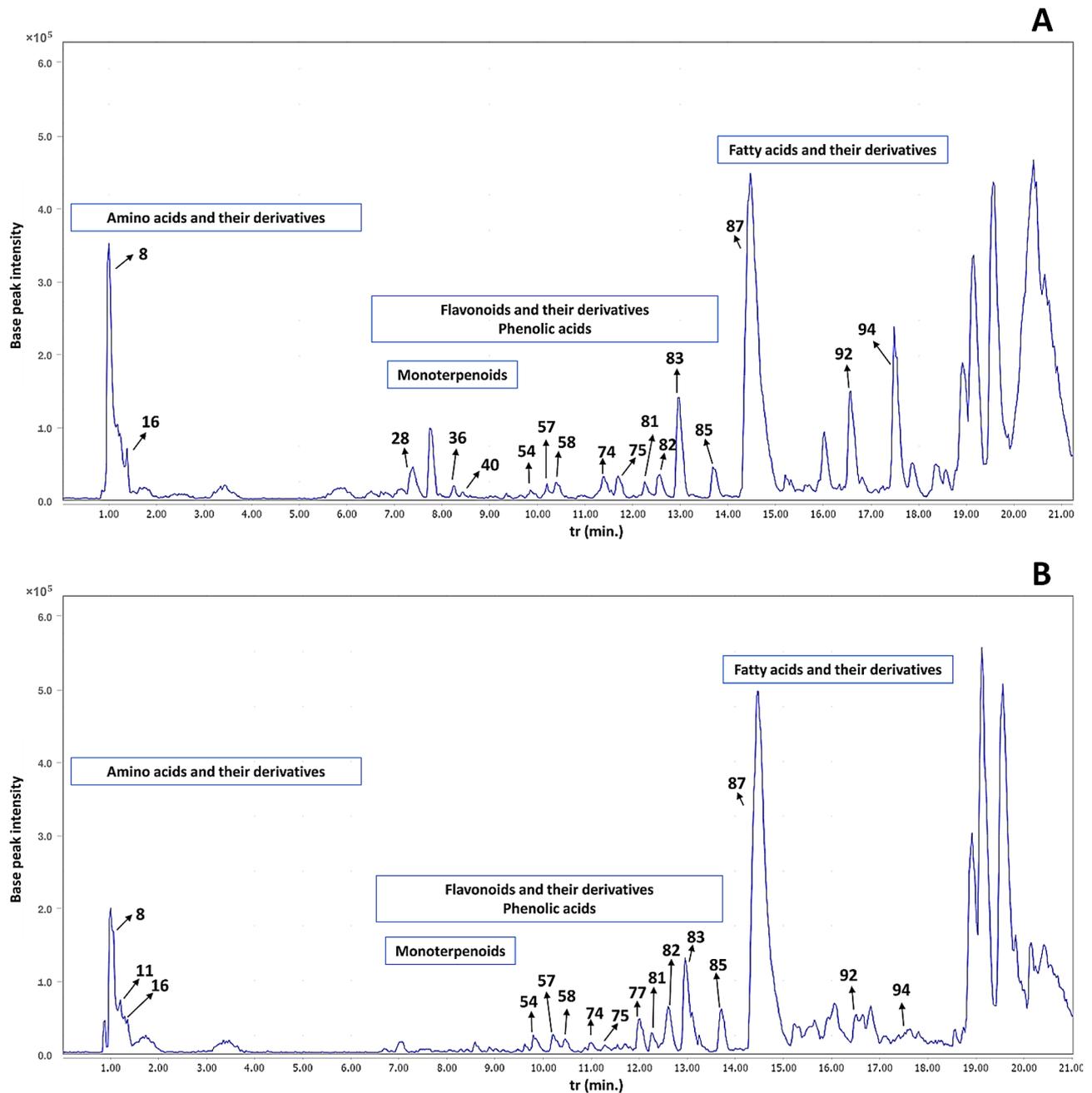
**Fig. 1** Base peak chromatograms of *Pisum sativum* L. (green pea, PS) waste extracts in the negative ESI mode. **(A)** PLS (leaves and stems extract) and **(B)** PSP (peels extract). Metabolite numbers are as listed in Table 1

processes at both the molecular and systems level [56]. PPI network results are represented by Fig. 6 and Table S2.

#### GO and KEGG pathway enrichment analysis

The identified protein targets (Tables S1 and S2) underwent enrichment analysis, with the top 20 related functions displayed in Fig. 7. Fold enrichment, defined as the ratio between the relative frequency of genes belonging

to a given pathway in the library and in the reference genome [57], indicates the extent to which genes of a particular pathway are overrepresented. The most enriched biological processes were associated with cellular response to nitrogenous and organo-nitrogenous compounds (Fig. 7A). Meanwhile, the most enriched KEGG pathways involved EGFR tyrosine kinase inhibitor resistance and the hypoxia-inducible factor 1 (HIF-1) signaling pathway (Fig. 7B). The target genes were



**Fig. 2** Base peak chromatograms of *Pisum sativum* L. (green pea, PS) waste extracts in the positive ESI mode. (A) PSLS (leaves and stems extract) and (B) PSP (peels extract). Metabolite numbers are as listed in Table 1

related to membrane components (Fig. 7C), with molecular functions enriched in the binding of tyrosine kinases (Fig. 7D).

#### MTT cytotoxicity assay

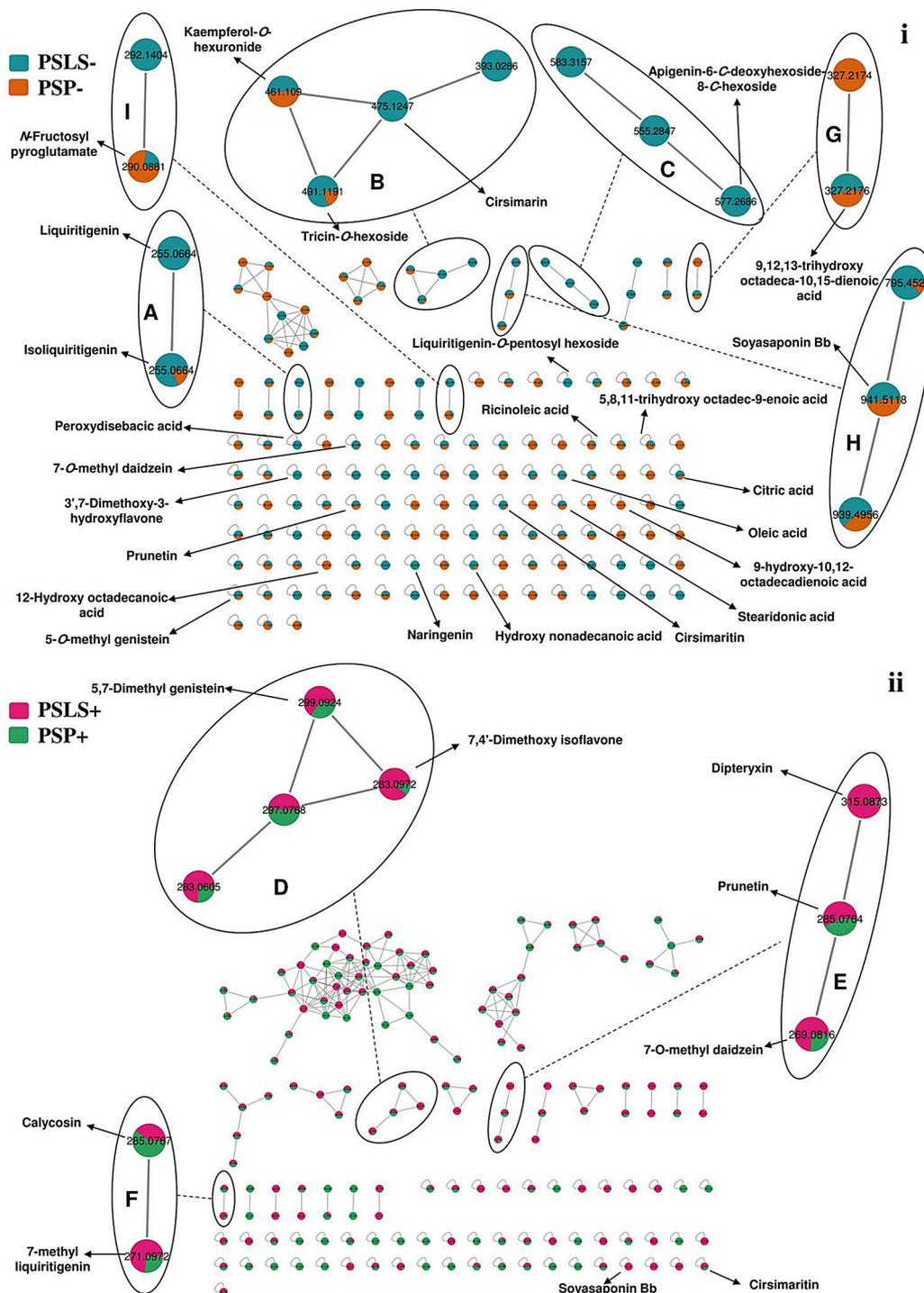
The MTT assay is a reliable, quantitative, and sensitive colorimetric method to measure cell viability and proliferation. Regarded as one of the simplest and most cost-effective cytotoxic assays, it requires minimal reagents and equipment. Additionally, it offers a combination of

accuracy and rapid results [31]. The results of the cytotoxic assay are represented by Table 3 and Figures S31 and S32.

#### Discussion

##### LC-MS/MS untargeted metabolomics

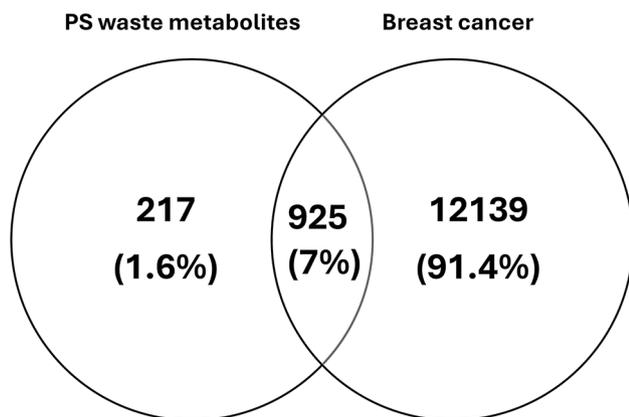
In total 84 and 38 metabolites were identified in both extracts in negative and positive ESI modes, respectively. The negative ESI mode allowed the identification of a larger number of metabolites, with only 14



**Fig. 3** Molecular network (showing clusters of metabolites of interest) based on LC-MS/MS data in *Pisum sativum* L. (green pea, PS) wastes. (i) Negative ESI mode and (ii) Positive ESI mode. Clusters **A** to **F**: flavonoids and their derivatives, **G**: fatty acids and their derivatives, **H**: saponins, **I**: amino acid and their derivatives. The network is displayed as a pie chart to reflect the relative abundance of each metabolite precursor ion in the analyzed samples

metabolites exclusively detected in positive ESI mode. These included some flavonoids, phenolic acids, amino acids, and fatty acids. This can be explained by the higher sensitivity of the negative ESI mode compared to positive one in the case of flavonoids and phenolic acids [58],

which represented the major class of identified metabolites. Regarding the metabolite profile of both extracts, a total of 93 and 84 metabolites were identified in both ESI modes in PSP and PSLs extracts, respectively. The investigated extracts showed high similarity, with 79



**Fig. 4** Venn's diagram representing the intersection of protein targets common to both the identified metabolites in *Pisum sativum* L. (green pea, PS) wastes and breast cancer

metabolites detected simultaneously in both. They presented differences in certain metabolites, such as flavonoids and monoterpenoids, which had higher abundance in the PSLs extract, as can be observed from Table 1; Fig. 3.

#### Flavonoids and their derivatives

Flavonoids are plant-produced secondary metabolites responsible for protecting plants from UV radiation, attracting insects, and aiding growth and fertility [59]. They have a basic skeleton containing 15 carbons, two benzene rings (A and B) connected through a bridge of three carbons forming a heterocyclic pyran ring (C). The difference in the arrangement of these rings, as well as the attachment of various substituents to the rings result in a wide range of flavonoid subclasses [60]. Flavonols, flavones, flavanones, isoflavones, chalcones, and some corresponding derivatives, such as glycosides, acyl glycosides, and prenylated derivatives, were identified in PSP and PSLs extracts.

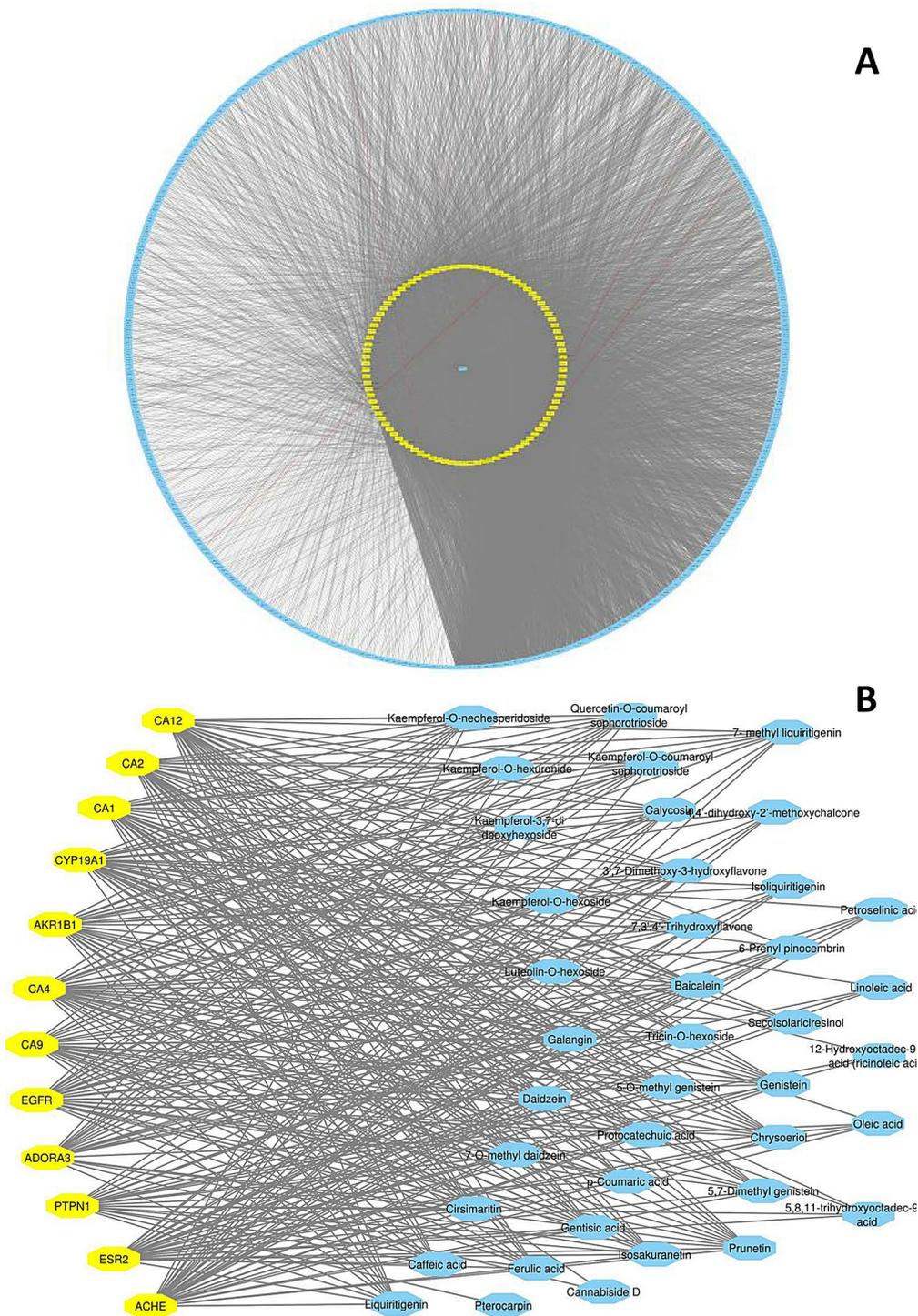
**Flavonols** Kaempferol-*O*-hexuronide (Peak 33), kaempferol-*O*-hexoside (48), and galangin (63) were identified in PSP and PSLs extracts. Whereas kaempferol-*O*-neohesperidoside (37), kaempferol-3,7-*O*-di deoxyhexoside (42, Figure S1), kaempferol-*O*-deoxyhexoside (46) were identified only in PSLs extract. The aforementioned kaempferol derivatives showed a common fragment ion at 285 *m/z* in the negative ESI mode and/or 287 in the positive ESI mode, which represents the kaempferol aglycone. Kaempferol-*O*-hexuronide (33) was identified based on the loss of 176 mass units, in the negative ionization mode, which corresponds to hexuronic acid [33]. Similarly, kaempferol-*O*-hexoside (48) and kaempferol-*O*-deoxyhexoside (46) showed losses of 162 and 146 mass units, upon fragmentation, which represent *O*-hexosyl and *O*-deoxyhexosyl moieties, respectively [33]. Kaemp-

ferol-*O*-neohesperidoside (37) (deoxyhexosyl hexoside) and kaempferol-3,7-*O*-di deoxyhexoside (42) were identified on the same basis. Besides, two kaempferol and two quercetin sophorotrioside acylated derivatives were detected in PSP and PSLs extracts and were identified as follows: quercetin-*O*-coumaroyl sophorotrioside (43), quercetin-*O*-feruloyl sophorotrioside (44), kaempferol-*O*-feruloyl sophorotrioside (45), and kaempferol-*O*-coumaroyl sophorotrioside (47). Upon fragmentation, the coumaroyl and feruloyl derivatives showed a loss of 146 or 176 mass units, corresponding to coumaroyl or feruloyl moieties, respectively, followed by the sequential loss of three hexosyl moieties, which confirms their identification [32]. Kaempferol glycosides and galangin (63) were reported to exhibit antiproliferative and apoptotic activity against breast cancer cell lines, respectively [61, 62].

**Flavones** Five flavones were identified in both PSP and PSLs extracts. They were identified as 7,3',4'-trihydroxyflavone (60), chrysoeriol (62, Figure S2), cirsimaritin (75), 3',7-dimethoxy-3-hydroxyflavone (79), and baicalein (88). Moreover, five flavone *O*-glycosides and one biflavone were identified in PSLs extract only. Tricin-*O*-hexoside (35) and luteolin-*O*-hexoside (56) showed a fragment ion peak at 329 *m/z* and 285 *m/z*, respectively, corresponding to tricetin and luteolin. In the same way, cirsimarin (54, Figure S3) exhibited a fragment ion peak, in the positive ESI mode, at 315 *m/z*, which represents its aglycone, cirsimaritin. In addition, apigenin-6-*C*-pentoside-8-*C*-hexoside (34, Figure S4) and apigenin-6-*C*-deoxyhexoside-8-*C*-hexoside (36) were identified based on the loss of 90 and 120 mass units, the common fragmentation pattern of *C*-hexosides, in addition to 60 and 90 mass units in case of *C*-pentoside, or 74 and 104 mass units for *C*-deoxyhexoside [63]. Finally, 3,8'-biapigenin (73) was a biflavone exclusively identified in PSLs extract. Different flavones, including cirsimarin (54), chrysoeriol (62), and baicalein (88), are reported to possess a potent cytotoxic activity against breast cancer cell lines [64, 65].

**Flavanones** Three flavanones, naringenin (69), isosakuranetin (74), 7-methyl liquiritigenin (82), along with a prenylated flavanone, 6-prenyl pinocembrin (28), were identified in PSP and PSLs extracts. In addition, liquiritigenin (58, Figure S5) along with two liquiritigenin glycosides were identified in PSLs extract only, namely liquiritigenin-*O*-pentosyl hexoside (39, Figure S6) and liquiritigenin-*O*-hexoside (41). Among the identified flavanones, liquiritigenin (58) and naringenin (69) were proven to be potent inhibitors of proliferation, migration, and progression of breast cancer [66, 67].

**Isoflavones** Daidzein (57, Figure S7), calycosin (64, Figure S8), dipteryxin (66), genistein (72), and prunetin (85),



**Fig. 5** (A) Compound-target-disease network and (B) sub-network showing the key metabolites and their interactions with the key protein targets

along with the methylated derivatives, 5-methyl genistein (59), 7-methyl daidzein (78), 5,7-dimethyl genistein (81), and 7,4'-dimethoxy isoflavone (84), were identified in both PSP and PSLs extracts. Isoflavones, especially daidzein (57) and genistein (72), were shown to exhibit a potent cytotoxic activity against breast cancer cell lines

[68]. In addition, the dietary consumption of isoflavones was reported to reduce the risk of breast cancer in postmenopausal women [68].

**Chalcones** Two chalcones, 4,4'-dihydroxy-2'-methoxychalcone (65) and isoliquiritigenin (77, Figure S9),

**Table 2** Metabolites identified in *Pisum sativum* L. (green pea, PS) waste extracts by LC-MS/MS untargeted metabolomics, ranked according to the degree of involvement in the compound-target-breast cancer network

Identified metabolite	Degree	Betweenness centrality	Closeness centrality
Kaempferol- <i>O</i> -neohesperidoside (37)	143	0.014612	0.379599
Kaempferol- <i>O</i> -hexuronide (33)	132	0.014563	0.376581
Kaempferol-3,7-di deoxyhexoside (42)	122	0.008718	0.373878
Kaempferol- <i>O</i> -coumaroyl sophorotrioside (47)	97	0.001701	0.367289
Quercetin- <i>O</i> -coumaroyl sophorotrioside (43)	97	0.001610	0.367289
4,4'-dihydroxy-2'-methoxychalcone (65)	96	0.004110	0.366772
<i>p</i> -Coumaric acid (40)	95	0.002978	0.366772
Isoliquiritigenin (77)	95	0.002378	0.366772
3,7-Dimethoxy-3-hydroxyflavone (79)	95	0.001493	0.366772
Protocatechuic acid (32)	94	0.002878	0.366514
Gentisic acid (25)	94	0.002591	0.366514
5,7-Dimethyl genistein (81)	94	0.003810	0.366514
Pterocarpin (70)	93	0.003866	0.366256
Secosolariciresinol (49)	93	0.004012	0.365742
12-Hydroxyoctadec-9-enoic acid (91)	93	0.002506	0.365999
Ferulic acid (80)	93	0.002618	0.366256
Prunetin (85)	93	0.002286	0.366256
Genistein (72)	93	0.001700	0.365742
Calycosin (64)	93	0.002372	0.366256
7-Methyl liquiritigenin (82)	93	0.003146	0.366256
6-Prenyl pinocembrin (28)	93	0.003302	0.366256
Chrysoeriol (62)	93	0.001338	0.366256
7,3,4'-Trihydroxyflavone (60)	93	0.001391	0.365485
Galangin (63)	93	0.001402	0.366256
Oleic acid (98)	92	0.002182	0.365999
Petroselinic acid (94)	92	0.002368	0.365999
5,8,11-trihydroxyoctadec-9-enoic acid (71)	92	0.003067	0.365742
Caffeic acid (83)	92	0.002446	0.365999
7-Methyl daidzein (78)	92	0.003118	0.365999
5-Methyl genistein (59)	92	0.002063	0.365999
Daidzein (57)	92	0.001942	0.365485
Isosakuranetin (74)	92	0.002490	0.365999
Baicalein (88)	92	0.001738	0.365999
Luteolin- <i>O</i> -hexoside (56)	92	0.001566	0.365999
Kaempferol- <i>O</i> -hexoside (48)	92	0.001887	0.365999
Cannabidiol (38)	91	0.003498	0.365742
Linoleic acid (96)	91	0.002348	0.365742
Liquiritigenin (58)	91	0.002350	0.365742
Cirsimaritin (75)	91	0.001508	0.365742
Tricin- <i>O</i> -hexoside (35)	91	0.002004	0.365742

Betweenness centrality is a measure of centrality in a graph based on shortest paths, while closeness centrality is an indication of how close a node is to all other nodes in the network

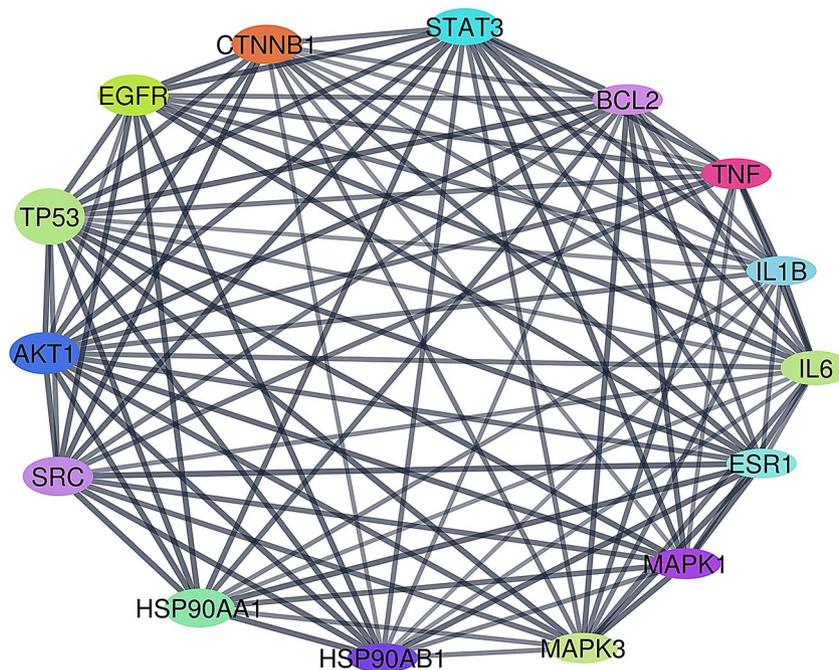
Metabolite numbers are as listed in Table 1

were identified in PSP and PSLs extracts. Reports concerning the effect of isoliquiritigenin (77) against breast cancers revealed its cytotoxic, migration inhibitory, and neoangiogenic suppressive effects [69, 70].

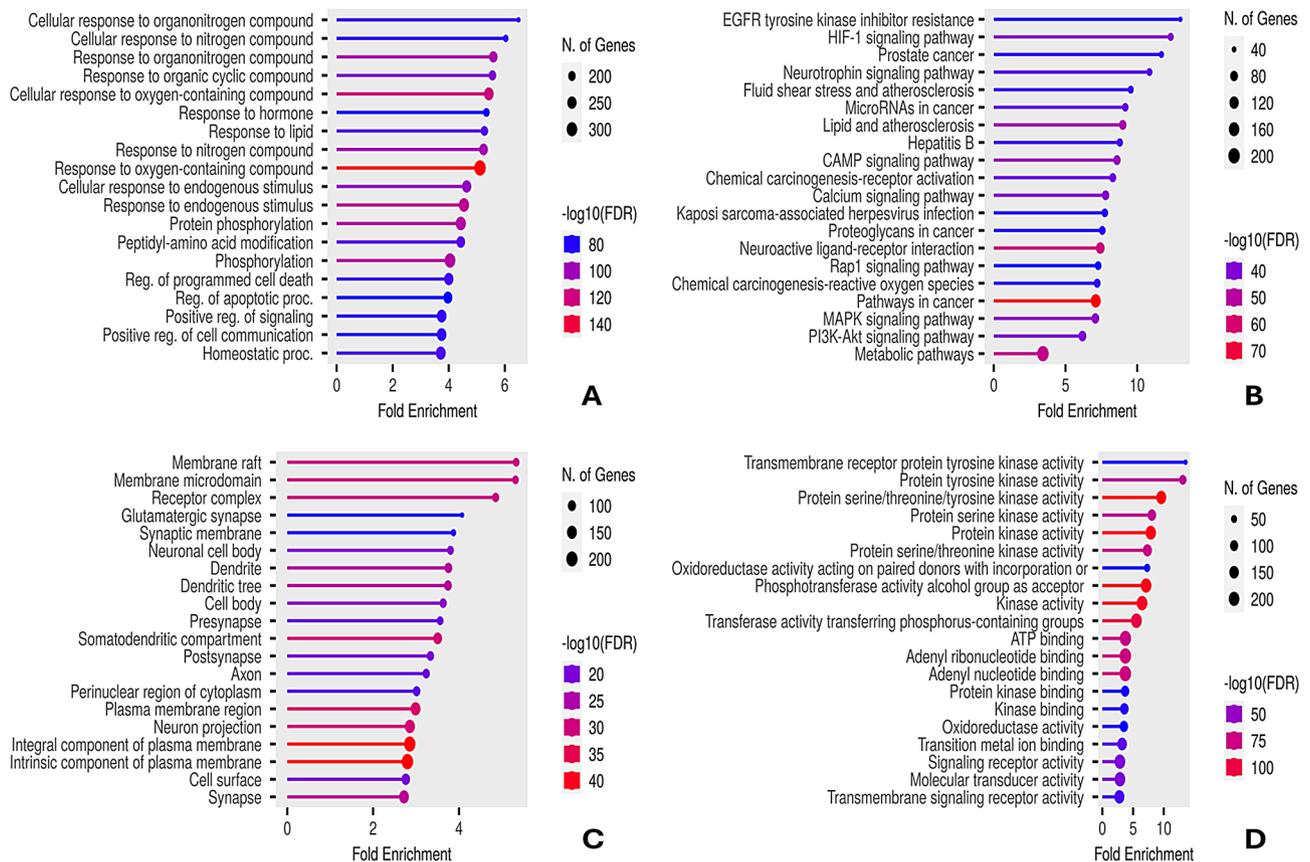
### Phenolic acids

Phenolic acids are abundant in different plants as they play an important role in the plant-microbe symbiosis and plant defense against microbial attacks [71]. They

consist of a phenolic ring bearing a carboxylic group along with one or more hydroxyl groups [71]. They are found in many plants either in free or bound forms, that include glycosides, esters, and amides [72]. Eight phenolic acids were identified in PSP and PSLs extracts, namely gentisic acid (25), protocatechuic acid (32), *p*-coumaric acid (40, Figure S10), veratric acid (53), salicylic acid (55), *p*-hydroxyphenyl acetic acid (61), ferulic acid (80), and caffeic acid (83, Figure S11). Whereas hydroxyphenyl



**Fig. 6** PPI network of the most interconnected targets involved in breast cancer



**Fig. 7** Functional analysis of the target genes identified by the compound-target-disease network. **(A)** Top GO biological processes, **(B)** top KEGG pathways, **(C)** top GO cellular components, and **(D)** top GO molecular functions. The circle size and the color density represent the number of involved genes and p-value, respectively. The fold enrichment is plotted on the abscissa while names are represented on the ordinate

**Table 3** IC<sub>50</sub> values of *Pisum sativum* L. (green pea, PS) waste extracts and the isolated components from the leaves and stems (PSLS) extract on MCF-7 and MCF-10a cell lines

Extract/Isolated compound	IC <sub>50</sub> µg/mL		IC <sub>50</sub> µM		Selectivity index IC <sub>50</sub> MCF-10a/ IC <sub>50</sub> MCF-7
	MCF-7	MCF-10a	MCF-7	MCF-10a	
PSLS	17.67 ± 1.12	62.07 ± 5.86			3.51
PSP	32.92 ± 1.64	53.30 ± 3.48			1.62
Liquiritigenin (58)/7-methyl liquiritigenin (82) mixture	11.10 ± 0.46	31.26 ± 1.49			2.82
Trans <i>p</i> -coumaric acid (40)	4.49 ± 0.19	22.67 ± 1.08	28.20 ± 9.6	146.46 ± 55.24	5.05
Methyl cis <i>p</i> -coumarate	1.18 ± 0.05	32.36 ± 2.09	6.91 ± 2.73	181.97 ± 21.23	27.42
Doxorubicin	2.69 ± 0.11	14.22 ± 0.68	5.13 ± 1.62	25.93 ± 2.31	5.28

PSLS *Pisum sativum* leaves and stems extract. PSP: *Pisum sativum* peels extract. Metabolite numbers are as listed in Table 1

propanoic acid (23) was exclusively identified in PSLS extract. Phenolic acids have a general fragmentation pattern in the negative ESI mode showing the loss of 44 mass units corresponding to a CO<sub>2</sub> group [73]. Various phenolic acids, including caffeic (83), ferulic (80), and protocatechuic (32) acids, were reported to exhibit a potent growth inhibitory effect on breast cancer cell lines, possibly through the interaction with different intracellular and membrane receptors [74].

#### Amino acids and their derivatives

Amino acids are organic compounds having a central carbon atom ( $\alpha$ -carbon) to which a carboxyl group (COOH), an amino group (NH<sub>2</sub>), and a specific side chain are attached [75]. They play several crucial roles for plants. Their major importance relies in fighting the hazardous effects of drought and salinity by changing the plant's osmotic pressure. They also play a key role in enhancing the plant yield and quality through the production of some natural growth hormones and stimulation of cell division [75]. Eleven amino acids were identified in both PSP and PSLS extracts, specifically threonine (1), aspartic acid (2, Figure S12), arginine (4), proline (7, Figure S13), homoserine (8), valine (11), leucine (15), tyrosine (16, Figure S14), isoleucine (17), phenylalanine (18, Figure S15), and tryptophan (21), along with three amino acid derivatives; O-acetyl homoserine (10), N-acetyl leucine (31), and N-fructosyl pyroglutamate (9, Figure S16). Amino acids have common fragments corresponding to the loss of NH<sub>3</sub>, H<sub>2</sub>O, CO, CO<sub>2</sub>, and CH<sub>2</sub>CO groups (17, 18, 28, 44, and 42 mass units, respectively) [76]. Homoserine (8) was the major amino acid identified. It was previously found at high concentration (214.9 mg/g dry extract) in PS flowering shoots and isolated therefrom [77]. Amino acids are essential for human health as they act as precursors for the synthesis of proteins and neurotransmitters, such as serotonin and dopamine, hormones, such as thyroxine, epinephrine, and norepinephrine, and antioxidants, such as glutathione and taurine. They also act as key regulators of body homeostasis,

including metabolic processes, transport and storage of nutrients, regulation of gene expression, immune response, growth, repair, and organ development [78].

#### Organic acids

Some plants accumulate certain organic acids in response to abiotic stress conditions such as drought and salinity. Organic acids, particularly succinic (13), galacturonic, and malic acids, increase the plant resilience in case of long-term drought [79]. Gluconic acid (3, Figure S17), threonic acid (5), citric acid (12), succinic acid (13, Figure S18), 3-hydroxy-3-methylglutaric acid (14), and 2-isopropylmalic acid (22) were identified in PSP and PSLS extracts.

#### Fatty acids and their derivatives

Fatty acids are long chain organic compounds (up to 24 carbons) having a carboxylic group at one end. They may be unsaturated or saturated according to the presence or absence of double bonds [80]. Fatty acids and their corresponding derivatives, such as fatty amides, esters, or amines, are produced by various plants to perform several regulatory roles in intracellular and extracellular signaling [81]. Twelve fatty acid were identified in PSP and PSLS extracts, namely tricladic acid A (29), peroxydisubacic acid (30), azelaic acid (51), 9,12,13-trihydroxyoctadeca-10,15-dienoic acid (67, Figure S19), undecanedioic acid (68), 5,8,11-trihydroxyoctadec-9-enoic acid (71, Figure S20), octadecanedioic acid (86), stearidonic acid (89), 9-hydroxy-10,12-octadecadienoic acid (90), hydroxyhexadecanoic acid (95), linoleic acid (96), hydroxynonadecanoic acid (97), and oleic acid (98), together with lauryldiethanolamine (87, Figure S21). Whereas 12-hydroxyoctadec-9-enoic acid (91), chaulmoogric acid (92), 12-hydroxyoctadecanoic acid (93), and petroselinic acid (94) were exclusively identified in PSP extract. The identified fatty acids were detected to large extent in both ESI modes and shared characteristic fragments resulting from the loss of water and CO<sub>2</sub> molecules, 18 and 44 mass units [45].

### Monoterpenoids

Monoterpenoids constitute a vast group of natural products widely spread in plants. They are part of the defensive oleoresins and essential oils produced by aromatic plants. They play a major role in plant defense, attraction of pollinators, and allelopathy [82]. Structurally, the monoterpene skeleton consists of ten carbon atoms derived from two  $C_5$  isoprene units [83]. Sacranoside A (27), a monoterpene diglycoside, was identified in PSL extract. Whereas cannabidiol D (38), a monoterpene glycoside, and its aglycone (50), together with lololide (52, Figure S22), a monoterpene lactone, were identified in PSP and PSL extracts. Some monoterpenoids were reported to exhibit anti-inflammatory and anticancer activities [82].

### Miscellaneous compounds

Other miscellaneous compounds identified in PSP and PSL extracts included glucose (6), pantothenic acid (vitamin B5) (19, Figure S23), triacetin (triglyceride) (20), catechol (24), indoleacrylic acid (26), along with secoisolariciresinol (lignan) (49), pterocarpin (pterocarpan) (70), and soyasaponin Bb (saponin glycoside) (76).

### MN study

The constructed network from the LC-MS/MS data (Fig. 3) comprised 155 nodes, including 38 clusters with at least 2 connected nodes, along with 160 self-looped nodes. Clusters of particular interest were clusters A to F, consisting of different types of identified flavonoids such as flavonol and flavone glycosides, flavanones, and isoflavones. Clusters G, H, and I contained some of the identified fatty acids, saponins, and amino acid derivatives, respectively. The self-looped nodes encompassed mainly metabolites belonging to different classes including flavonoids, fatty acids, and some organic acids.

### Network pharmacology study of potential therapeutic activity against breast cancer

#### Compound-target-disease network

Analysis of the compound-target-disease network (Fig. 5; Table 2) revealed that kaempferol-*O*-neohesperidoside (37) showed the highest degree of involvement, betweenness, and closeness scores, indicating its interconnection to multiple nodes (protein targets) and suggesting its potential as key active metabolite. Kaempferol-*O*-hexuronide (33), kaempferol-3,7-*O*-di deoxyhexoside (42), and kaempferol-*O*-neohesperidoside (37) completed the top of the ranking. On the other hand, carbonic anhydrases, namely CA12, CA2, and CA1, were the major breast cancer protein targets interconnected to the identified metabolites according to their degree of involvement (Table S1). CAs are membrane-associated enzymes overexpressed in several types of cancer cells, including

breast, ovarian, renal, and colorectal carcinomas, as well as astrocytic glioma and non-small cell lung cancer. Their tumorigenic effect relies on the regulation of tumor cell intracellular pH and acidification of the extracellular environment. Consequently, this confers tumor cells growth and survival advantages over normal cells, which are unable to adapt to acidity [84, 85].

#### PPI network

PPI network analysis revealed that tumor protein p53 (TP53), serine/threonine kinase 1 (AKT1), and epidermal growth factor receptor (EGFR) were the main hub targets i.e., the most interconnected nodes within the network, based on the degree of nodes sizes TP53 is highly related to breast cancer, as mutations in TP53 are found in around 30% of breast carcinomas, resulting in the loss of its function as a tumor-suppressing protein [86]. Conversely, AKT1 is overactivated in breast cancer, leading to increased mammary tumor growth [87]. Additionally, EGFR is overexpressed in around 14% of breast carcinomas. Overexpression of EGFR induces cell growth, differentiation, angiogenesis, and blocks apoptosis [88].

#### GO and KEGG pathway enrichment analysis

The enrichment analysis suggested that the identified metabolites of PS wastes would inhibit tyrosine kinases which are key enzymes in breast and other types of cancers [89].

In summary, the identified metabolites had the potential to counteract breast cancer by interacting with key enzymes and protein biomarkers involved in the survival and proliferation of breast carcinoma. Therefore, it was crucial to confirm this potential bioactivity by evaluating cytotoxic activity using the *in-vitro* MTT assay in MCF-7 and MCF-10a cell lines.

#### MTT cytotoxicity assay

Results of the MTT cytotoxicity assay revealed that PSL extract exhibited a potent cytotoxic activity on MCF-7 cell lines with good selectivity ( $IC_{50} = 17.67$  and  $62.07 \mu\text{g/mL}$  for MCF-7 and MCF-10a cell lines, respectively, and selectivity index of 3.51), whereas PSP extract showed less potency and less selectivity ( $IC_{50} = 32.92$  and  $53.30 \mu\text{g/mL}$  for MCF-7 and MCF-10a cell lines, respectively, and selectivity index of 1.62). This could be explained by the higher abundance of phenolic metabolites in PSL extract compared to PSP as noticed from the LC-MS/MS analysis and MN. As for the isolated components, methyl cis *p*-coumarate showed the strongest cytotoxic activity on MCF-7 cell line followed by trans *p*-coumaric acid then the liquiritigenin/7-methyl liquiritigenin mixture with  $IC_{50}$  values of 1.18, 4.49, and  $11.10 \mu\text{g/mL}$ , respectively. Additionally, Methyl cis *p*-coumarate exhibited a significantly higher selectivity index compared to the

reference drug doxorubicin, the extracts, and the rest of isolated components, as shown in Table 3.

Regarding the potential mechanisms of action of the isolated components, the subnetwork in Fig. 5B indicates that *p*-coumaric acid, liquiritigenin, and 7-methyl liquiritigenin interacted with several protein targets associated with breast cancer including CA 1, 2, 4, 9, and 12, as well as aldo-keto reductase family 1 member B1 (AKR1B1), adenosine A3 receptor (ADORA3), protein tyrosine phosphatase non-receptor type 1 (PTPN1), and estrogen receptor 2 (ESR2). Additionally, liquiritigenin and 7-methyl liquiritigenin interacted with cytochrome P19A1 (CYP19A1) and acetyl choline esterase (ACHE), and *p*-coumaric acid with EGFR. This could explain the mechanism of their cytotoxicity against the studied cell lines.

## Conclusion

The metabolite profiles of the 70% ethanolic extracts of PSP and PSLs wastes were established using LC-MS/MS untargeted metabolomics and MN. The results led to the identification of 98 metabolites belonging to different classes including flavonoids and their derivatives, phenolic acids, amino acids and their derivatives, organic acids, fatty acids and their derivatives, monoterpenoids, and other miscellaneous compounds. The identified metabolites were subjected to network pharmacology, which indicated their promising potential for prevention and treatment of breast cancer through the interaction with different enzymes and proteins involved in breast cancer progression. Flavonoids and their derivatives, as well as phenolic acids prevailed, especially in the PSLs extract. These metabolite categories comprised almost half of the total number of identified metabolites and also were the key metabolites interacting with breast cancer targets. As a result, PSLs extract was targeted for the isolation of phenolic components, where *trans p*-coumaric acid, methyl *cis p*-coumarate along with a mixture of liquiritigenin and 7-methyl liquiritigenin, were isolated. The in-vitro cytotoxic activity of the waste extracts and the isolated components were investigated on MCF-7 breast carcinoma and MCF-10a human mammary epithelial cell lines. The results revealed the more potent cytotoxic activity and considerable selectivity of the PSLs extract ( $IC_{50}=17.67\pm 1.12$   $\mu\text{g}/\text{mL}$  and  $62.07\pm 5.86$   $\mu\text{g}/\text{mL}$  on MCF-7 and MCF-10a, respectively) compared to the PSP extract ( $IC_{50}=32.92\pm 1.64$   $\mu\text{g}/\text{mL}$  and  $53.30\pm 3.48$   $\mu\text{g}/\text{mL}$  on MCF-7 and MCF-10a, respectively). This could be attributed to the higher abundance in PSLs extract of metabolites belonging to phenolic categories (flavonoids and phenolic acids). The isolated PSLs components also showed a potent cytotoxic activity with high selectivity index compared to the reference drug doxorubicin. Methyl *cis p*-coumarate showed the strongest activity as

well as the highest selectivity among the isolated components with  $IC_{50}$  values of  $1.18\pm 0.05$  and  $32.36\pm 2.09$   $\mu\text{g}/\text{mL}$ , which correspond to  $6.91\pm 2.73$  and  $181.97\pm 21.23$   $\mu\text{M}$ , on MCF-7 and MCF-10a, respectively. This could be ascribed to their interaction with various breast cancer-associated protein targets such as CAs, AKR1B1, ADORA3, PTPN1, ESR2, and EGFR. These results suggest the possibility for valorizing PS wastes, especially PSLs extract, as potential breast cancer therapeutic agents and as a source of several bioactive phenolic compounds. However, the efficiency and safety of the investigated extracts and isolated components require further confirmation through in-vivo studies and clinical trials.

## Abbreviations

LC-ESI-QTOF-MS/MS	Liquid chromatography-electrospray ionization-quadrupole time-of-flight tandem mass spectrometry
PS	<i>Pisum sativum</i>
LC-MS	Liquid chromatography-mass spectrometry
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
PSP	<i>Pisum sativum</i> peels
PSLS	<i>Pisum sativum</i> leaves and stems
GNPS-MN	Global Natural Products Social Molecular Networking
MN	Molecular networking
TLC	Thin layer chromatography
NMR	Nuclear magnetic resonance
SEA	SwissTargetPrediction database
PPI	Protein-protein interaction
GO	Gene ontology
KEGG	Kyoto encyclopedia of genes and genome
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide
$IC_{50}$	The half-maximal inhibitory concentration
CA	Carbonic anhydrase
TP53	Tumor protein p53
AKT1	Serine/threonine kinase 1
AKR1B1	Aldo-keto reductase family 1 member B1
ADORA3	Adenosine A3 receptor
PTPN1	Protein tyrosine phosphatase non-receptor type 1
ESR2	Estrogen receptor 2
CYP19A1	Cytochrome P19A1
ACHE	Acetyl choline esterase

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12906-024-04669-x>.

Supplementary Material 1

Supplementary Material 2

## Acknowledgements

Basma Eltanany would like to thank the Egyptian Ministry of Higher Education for funding her postdoctoral research stay at the Bioanalysis group of the University of Barcelona (UB), Barcelona, Spain. The Bioanalysis group of the UB is part of the INSA-UB Maria de Maeztu Unit of Excellence (Grant CEX2021-001234-M) funded by MCIN/AEI/FEDER, UE.

## Author contributions

(A) K.: Conceptualization, Methodology, LC-MS/MS data interpretation, Sample collection, extraction, components isolation and identification, and performance of the MTT cytotoxicity assay, Writing original draft, Writing – review & editing. O. S., H.E.: Supervision, Conceptualization, Methodology, Writing – review & editing. S.E.: Supervision, Conceptualization, Methodology, Writing original draft, Writing – review & editing. (B) E.: Methodology; LC-MS/MS analyses, Formal analysis, Validation, Data curation, Writing – review &

editing. L. P.: Methodology; LC-MS/MS analyses, Writing – review & editing. F. B.: Supervision; LC-MS/MS analyses, Writing – review & editing. A. M.: Methodology; Network pharmacology, Writing – review & editing. N. F.: Supervision, Conceptualization, Methodology; Molecular networking study, Writing original draft, Writing – review & editing.

#### Funding

Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

#### Data availability

The data that support the findings of this study will be made available on reasonable request.

#### Declarations

#### Ethical approval

Ethics approval was not required for this research.

#### Clinical trials and informed consents

No human subjects were involved in this article. Informed consents and clinical trial number are not applicable.

#### Supporting information

MS/MS spectra of representative members from each metabolite class, schematic representation of extraction, fractionation, and component isolation steps, structures, NMR charts of the isolated components, along with figures of MTT cytotoxicity assay results are attached in a supplementary file.

#### Competing interests

The authors declare no competing interests.

#### Author details

<sup>1</sup>Department of Pharmacognosy, Faculty of Pharmacy, Cairo University, Cairo 11562, Egypt

<sup>2</sup>Department of Pharmacognosy, Faculty of Pharmacy, Heliopolis University, Cairo 4645241, Egypt

<sup>3</sup>Department of Pharmaceutical Analytical Chemistry, Faculty of Pharmacy, Cairo University, Cairo 11562, Egypt

<sup>4</sup>Department of Chemical Engineering and Analytical Chemistry, Institute for Research on Nutrition and Food Safety (INSA-UB), University of Barcelona, Barcelona 08028, Spain

<sup>5</sup>Serra Hunter Program, Generalitat de Catalunya, Barcelona 08007, Spain

<sup>6</sup>Department of Pharmacology and Toxicology, Faculty of Pharmacy, Cairo University, Cairo 11562, Egypt

<sup>7</sup>Faculty of Pharmacy, King Salman International University (KSU), Ras Sedr 46612, Egypt

Received: 10 July 2024 / Accepted: 30 September 2024

Published online: 31 October 2024

#### References

- Zhang SJ, Hu TT, Chen YY, Wang S, Kang YF. Analysis of the polysaccharide fractions isolated from pea (*Pisum sativum* L.) at different levels of purification. *J Food Biochem*. 2020;44:1–11.
- Dahl WJ, Foster LM, Tyler RT. Review of the health benefits of peas (*Pisum sativum* L.). *Br J Nutr*. 2012;108 SUPPL. 1.
- Zilani MNH, Sultana T, Asabur Rahman SM, Anisuzzman M, Islam MA, Shilpi JA, et al. Chemical composition and pharmacological activities of *Pisum sativum*. *BMC Complement Altern Med*. 2017;17:1–9.
- Kumari T, Deka SC. Potential health benefits of garden pea seeds and pods: a review. *Legum Sci*. 2021;3:1–13.
- Dong X, Guo S, Wang H, Wang Z, Gao X. Physicochemical characteristics and FTIR-derived structural parameters of hydrochar produced by hydrothermal carbonisation of pea pod (*Pisum sativum* Linn.) Waste. *Biomass Convers Biorefinery*. 2019;9:531–40.
- Abd El-Galil AA, Negm El-Dein A, Awad HM, Helmy WA. Chemical composition and biological activities of aqueous extracts and their sulfated derivatives of pea peel (*Pisum sativum* L.). *Biocatal Agric Biotechnol*. 2021;35 June.
- Torres-Valenzuela LS, Ballesteros-Gómez A, Rubio S. Green Solvents for the extraction of high added-value compounds from Agri-food Waste. *Food Eng Rev*. 2020;12:83–100.
- Ben-Othman S, Jöudu I, Bhat R. Bioactives from agri-food wastes: present insights and future challenges. *Molecules*. 2020;25:1–32.
- Abdelghffar EA, Obaid WA, Elgamel AM, Daoud R, Sobeh M, El Raey MA. Pea (*Pisum sativum*) peel extract attenuates DOX-induced oxidative myocardial injury. *Biomed Pharmacother*. 2021;143:112120.
- Hadrich F, Arbi M, El, Boukhris M, Sayadi S, Cherif S. Valorization of the peel of pea: *Pisum sativum* by evaluation of its antioxidant and antimicrobial activities. *J Oleo Sci*. 2014;63:1177–83.
- Patra JK, Das G, Shin HS. Facile green biosynthesis of silver nanoparticles using *Pisum sativum* L. outer peel aqueous extract and its antidiabetic, cytotoxicity, antioxidant, and antibacterial activity. *Int J Nanomed*. 2019;14:6679–90.
- Dixit V, Pandey V, Shyam R. Differential antioxidative responses to cadmium in roots and leaves of pea (*Pisum sativum* L.). *J Exp Bot*. 2001;52:1101–9.
- Edwards EA, Rawsthorne S, Mullineaux PM. Subcellular distribution of multiple forms of glutathione reductase in leaves of pea (*Pisum sativum* L.). *Planta*. 1990;180:278–84.
- Chahbani A, Fakhfakh N, Balti MA, Mabrouk M, El-Hatmi H, Zouari N, et al. Microwave drying effects on drying kinetics, bioactive compounds and antioxidant activity of green peas (*Pisum sativum* L.). *Food Biosci*. 2018;25:32–8.
- Fahim JR, Attia EZ, Kamel MS. The phenolic profile of pea (*Pisum sativum*): a phytochemical and pharmacological overview. *Phytochem Rev*. 2019;18:173–98.
- Fahmi R, Ryland D, Sopiwnyk E, Malcolmson L, Shariati-Ilevari S, McElrea A, et al. Effect of Revtech thermal processing on volatile organic compounds and chemical characteristics of split yellow pea (*Pisum sativum* L.) flour. *J Food Sci*. 2021;86:4330–53.
- Awad K, Abdel-Nabey AA, Awney H. Phenolic composition and antioxidant activity of some agro-industrial Wastes Khadega. *Alexandrian J Food Sci Technol*. 2018;15:33–44.
- Bray F, Ferlay J, Soerjomataram I. Global Cancer statistics 2018: GLOBOCAN estimates of incidence and Mortality Worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68:1–31.
- Momenimovahed Z, Salehiniya H. Epidemiological characteristics of and risk factors for breast cancer in the world. *Breast Cancer Targets Ther*. 2019;11:151–64.
- Khalil MNA, Afifi SM, Eltanany BM, Pont L, Benavente F, El-Sonbaty SM et al. Assessment of the effect of drying on *Brassica* greens via a multiplex approach based on LC-QTOF-MS/MS, molecular networking, and chemometrics along with their antioxidant and anticancer activities. *Food Res Int*. 2024;180 February:114053.
- Zhang M, Otsuki K, Li W. Molecular networking as a natural products discovery strategy. *Acta Mater Med*. 2023;2:126–41.
- Otify AM, ElBanna SA, Eltanany BM, Pont L, Benavente F, Ibrahim RM. A comprehensive analytical framework integrating liquid chromatography-tandem mass spectrometry metabolomics with chemometrics for metabolite profiling of lettuce varieties and discovery of antibacterial agents. *Food Res Int*. 2023;172 June:113178.
- Ibrahim RM, Eltanany BM, Pont L, Benavente F, Elbanna A, Otify AM. Unveiling the functional components and antivirulence activity of mustard leaves using an LC-MS / MS, molecular networking, and multivariate data analysis integrated approach. *Food Res Int*. 2023;168:112742.
- Ibrahim RM, Fayed S, Eltanany BM, Abu-Elghait M, El-Demerdash A, Badawy MSEM, et al. Agro-byproduct valorization of radish and turnip leaves and roots as new sources of antibacterial and antivirulence agents through metabolomics and molecular networking. *Sci Hortic (Amsterdam)*. 2024;328:112924.
- Yuan H, Ma Q, Cui H, Liu G, Zhao X, Li W, et al. How can synergism of traditional medicines benefit from network pharmacology? *Molecules*. 2017;22:1–19.
- Khalil AM, Sabry OM, El-Askary HI, El Zalabani SM, Eltanany BM, Pont L, et al. Identification of cyclooxygenase-II inhibitory saponins from fenugreek wastes: insights from liquid chromatography-tandem mass spectrometry metabolomics, molecular networking, and molecular docking. *Phytochem Anal*. 2024;35:690–707.

27. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, et al. Identifying small molecules *via* high resolution mass spectrometry: communicating confidence. *Environ Sci Technol*. 2014;48:2097–8.
28. Hamed AA, El-Shiekh RA, Mohamed OG, Aboutabl EA, Fathy FI, Fawzy GA, et al. Cholinesterase inhibitors from an endophytic fungus *aspergillus niveus* Fv-er401: Metabolomics, isolation and molecular docking. *Molecules*. 2023;28:2559.
29. Kort R, Vonk H, Xu X, Hoff WD, Crielard W, Hellingwerf KJ. Evidence for *trans-cis* isomerization of the *p*-coumaric acid chromophore as the photochemical basis of the photocycle of photoactive yellow. *FEBS Lett*. 1996;382:73–8.
30. Daayf F, Bel-Rhliid R, Bélanger RR. Methyl ester of *p*-coumaric acid: a phytoalexin-like compound from long english cucumber leaves. *J Chem Ecol*. 1997;23:1517–26.
31. Senthilraja P, Kathiresan K. In vitro cytotoxicity MTT assay in vero, HepG2 and MCF-7 cell lines study of marine yeast. *J Appl Pharm Sci*. 2015;5:80–4.
32. Cuyckens F, Claeys M. Mass spectrometry in the structural analysis of flavonoids. *J Mass Spectrom*. 2004;39:1–15.
33. Pascale R, Acquavia MA, Cataldi TRI, Onzo A, Coviello D, Bufo SA, et al. Profiling of quercetin glycosides and acyl glycosides in sun-dried peperoni di Senise peppers (*Capsicum annum* L.) by a combination of LC-ESI(-)-MS/MS and polarity prediction in reversed-phase separations. *Anal Bioanal Chem*. 2020;412:3005–15.
34. Mišić D, Šiler B, Gašić U, Avramov S, Živković S, Živković JN, et al. Simultaneous UHPLC/DAD/(+/-)HESI-MS/MS analysis of phenolic acids and nepetalactones in methanol extracts of *nepeta* species: a possible application in chemotaxonomic studies. *Phytochem Anal*. 2015;26:72–85.
35. Benayad Z, Gómez-Cordovés C, Es-Safi NE. Identification and quantification of flavonoid glycosides from fenugreek (*Trigonella foenum-graecum*) germinated seeds by LC-DAD-ESI/MS analysis. *J Food Compos Anal*. 2014;35:21–9.
36. Zhang X, Liao M, Cheng X, Liang C, Diao X, Zhang L. Ultrahigh-performance liquid chromatography coupled with triple quadrupole and time-of-flight mass spectrometry for the screening and identification of the main flavonoids and their metabolites in rats after oral administration of *Cirsium japonicum* DC. *Ex. Rapid Commun Mass Spectrom*. 2018;32:1451–61.
37. Wu W, Yan C, Li L, Liu Z, Liu S. Studies on the flavones using liquid chromatography-electrospray ionization tandem mass spectrometry. *J Chromatogr A*. 2004;1047:213–20.
38. Wollenweber E, Fritz H, Henrich B, Jakupovic J, Schilling G, Roitman JN. Rare flavonoid aglycones from *Anaphalis margaritacea* and two *Gnaphalium* Species. *Z fur Naturforsch - Sect C J Biosci*. 1993;48:420–4.
39. Fabre N, Rustan I, de Hoffmann E, Quetin-Leclercq J. Determination of Flavone, Flavonol, and Flavanone aglycones by Negative Ion Liquid Chromatography Electrospray Ion Trap Mass Spectrometry. *Am Soc Mass Spectrom*. 2012;12:707–15.
40. Nakata R, Yoshinaga N, Teraiishi M, Okumoto Y, Huffaker A, Schmelz EA, et al. A fragmentation study of isoflavones by IT-TOF-MS using biosynthesized isotopes. *Biosci Biotechnol Biochem*. 2018;82:1309–15.
41. Tai Y, Pei S, Wan J, Cao X, Pan Y. Fragmentation study of protonated chalcones by atmospheric pressure chemical ionization and tandem mass spectrometry. *Rapid Commun Mass Spectrom*. 2006;20:994–1000.
42. Robbins RJ. Phenolic acids in foods: an overview of analytical methodology. *J Agric Food Chem*. 2003;51:2866–87.
43. Piraud M, Vianey-Saban C, Petritis K, Elfakir C, Steghens JP, Morla A, et al. ESI-MS/MS analysis of underivatized amino acids: a new tool for the diagnosis of inherited disorders of amino acid metabolism. Fragmentation study of 79 molecules of biological interest in positive and negative ionisation mode. *Rapid Commun Mass Spectrom*. 2003;17:1297–311.
44. Delgado-Povedano M, del Sánchez de Medina M, Bautista V, Priego-Capote J, Luque de Castro F. Tentative identification of the composition of *Agaricus Bisporus* aqueous enzymatic extracts with antiviral activity against HCV: a study by liquid chromatography-tandem mass spectrometry in high resolution mode. *J Funct Foods*. 2016;24:403–19.
45. Thomas MC, Dunn SR, Altvater J, Dove SG, Nette GW. Rapid identification of long-chain polyunsaturated fatty acids in a marine extract by HPLC-MS using data-dependent acquisition. *Anal Chem*. 2012;84:5976–83.
46. Zakharenko AM, Razgonova MP, Pikula KS, Golokhvast KS. Simultaneous determination of 78 compounds of *Rhodiola rosea* Extract by Supercritical CO<sub>2</sub>-Extraction and HPLC-ESI-MS/MS spectrometry. *Biochem Res Int*. 2021;2021:9957490.
47. Wang Y, Zhu H, Wang D, Cheng R, Yang C, Xu M, et al. A new phloroglucinol glycoside from the whole plants of *Glochidion Eriocarpum*. *Bull Korean Chem Soc*. 2014;35:631–4.
48. El-Mekkawy S, Hassan AZ, Abdelhafez MA, Mahmoud K, Mahrous KF, Meselhy MR, et al. Cytotoxicity, genotoxicity, and gene expression changes induced by methanolic extract of *Moringa Stenopetala* leaf with LC-qTOF-MS metabolic profile. *Toxicol*. 2021;203:40–50.
49. Rodríguez-Pérez C, Quirantes-Piné R, Fernández-Gutiérrez A, Segura-Carrettero A. Comparative characterization of phenolic and other polar compounds in Spanish melon cultivars by using high-performance liquid chromatography coupled to electrospray ionization quadrupole-time of flight mass spectrometry. *Food Res Int*. 2013;54:1519–27.
50. Nebel B, Mittelbach M, Uray G. Determination of the composition of acetylglycerol mixtures by <sup>1</sup>H NMR followed by GC investigation. *Anal Chem*. 2008;80:8712–6.
51. Kiprop AK, Pourtier E, Kimutai SK. GC-MS and ESI-MS detection of catechol. *Int J Educ Res*. 2013;1:1–12.
52. Hanhineva K, Rogachev I, Aura AM, Aharoni A, Poutanen K, Mykkänen H. Identification of novel lignans in the whole grain rye bran by non-targeted LC-MS metabolite profiling. *Metabolomics*. 2012;8:399–409.
53. Aryal B, Adhikari B, Aryal N, Bhattarai BR, Khadayat K, Parajuli N. LC-HRMS profiling and antidiabetic, antioxidant, and Antibacterial activities of *Acacia catechu* (L.f.) Willd. *Biomed Res Int*. 2021;2021:7588711.
54. Ohana P, Delmer DP, Carlson RW, Glushka J, Azadi P, Bacic T, et al. Identification of a novel triterpenoid saponin from *Pisum sativum* as a specific inhibitor of the diguanylate cyclase of *Acetobacter Xylinum*. *Plant Cell Physiol*. 1998;39:144–52.
55. Noor F, Qamar MTU, Ashfaq UA, Albutti A, Alwashmi ASS, Aljasir MA. Network Pharmacology Approach for Medicinal Plants: Review and Assessment. *Pharmaceuticals*. 2022;15:1–33.
56. Tomkins JE, Manzoni C. Advances in protein-protein interaction network analysis for Parkinson's disease. *Neurobiol Dis*. 2021;155:105395.
57. Avina-Padilla K, Ramirez-Rafael JA, Zambada-Moreno, Herrera-Oropeza O, Emilio G, Romero G, et al. Deciphering the Tissue-Specific Regulatory Role of Intronless Genes across Cancers. In: Jin L, Durand D, editors. *Comparative Genomics*. Springer Nature Switzerland AG; 2022. pp. 311–39.
58. Ana Plazonić, Bucar F, Maleš Željanić, Nigović B, Kujundžić N. Identification and quantification of flavonoids and phenolic acids in burr parsley (*Caucalis platycarpus* L.), using high-performance liquid chromatography with diode array detection and electrospray ionization mass spectrometry. *Molecules*. 2009;14:2466–90.
59. Falcone Ferreyra ML, Rius SP, Casati P. Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Front Plant Sci*. 2012;3:1–15.
60. Dias MC, Pinto DCGA, Silva AMS. Plant flavonoids: Chemical characteristics and biological activity. *Molecules*. 2021;26:1–16.
61. Wang J, Fang X, Ge L, Cao F, Zhao L, Wang Z, et al. Antitumor, antioxidant and anti-inflammatory activities of kaempferol and its corresponding glycosides and the enzymatic preparation of kaempferol. *PLoS ONE*. 2018;13:1–12.
62. Liu D, You P, Luo Y, Yang M, Liu Y. Galangin induces apoptosis in MCF-7 human breast cancer cells through mitochondrial pathway and phosphatidylinositol 3-Kinase/Akt inhibition. *Pharmacology*. 2018;102:58–66.
63. Vukics V, Guttman A. Structural characterization of flavonoid glycosides by multi-stage mass spectrometry. *Mass Spectrom Rev*. 2010;29:1–16.
64. Lin CH, Chang CY, Lee KR, Lin HJ, Chen TH, Wan L. Flavones inhibit breast cancer proliferation through the Akt/FOXO3a signaling pathway. *BMC Cancer*. 2015;15:1–12.
65. Serpeloni JM, de Oliveira LCB, Fujiike A, Tuttis K, Ribeiro DL, Camara MBP et al. Flavone cirsimarin impairs cell proliferation, migration, and invasion in MCF-7 cells grown in 2D and 3D models. *Toxicol Vitro*. 2022;83 June.
66. Liang Y, Besch-Williford C, Hyder SM. The estrogen receptor beta agonist liquiritigenin enhances the inhibitory effects of the cholesterol biosynthesis inhibitor RO 48-8071 on hormone-dependent breast-cancer growth. *Breast Cancer Res Treat*. 2022;192:53–63.
67. Harmon AW, Patel YM. Naringenin inhibits glucose uptake in MCF-7 breast cancer cells: a mechanism for impaired cellular proliferation. *Breast Cancer Res Treat*. 2004;85:103–10.
68. Basu P, Maier C. Phytoestrogens and breast cancer: in vitro anticancer activities of isoflavones, lignans, coumestans, stilbenes and their analogs and derivatives. *Biomed Pharmacother*. 2018;107:1648–66.
69. Wang KL, Hsia SM, Chan CJ, Chang FY, Huang CY, Bau DT, et al. Inhibitory effects of isoliquiritigenin on the migration and invasion of human breast cancer cells. *Expert Opin Ther Targets*. 2013;17:337–49.
70. Wang Z, Wang N, Han S, Wang D, Mo S, Yu L, et al. Dietary compound isoliquiritigenin inhibits breast Cancer neoangiogenesis via VEGF/VEGFR-2 signaling pathway. *PLoS ONE*. 2013;8:e68566.

71. Mandal SM, Chakraborty D, Dey S. Phenolic acids act as signaling molecules in plant-microbe symbioses. *Plant Signal Behav.* 2010;5:359–68.
72. Kumar N, Goel N. Phenolic acids: natural versatile molecules with promising therapeutic applications. *Biotechnol Rep.* 2019;24:e00370.
73. Gruz J, Novák O, Strnad M. Rapid analysis of phenolic acids in beverages by UPLC-MS/MS. *Food Chem.* 2008;111:789–94.
74. Kampa M, Alexaki VI, Notas G, Nifli AP, Nistikaki A, Hatzoglou A, et al. Antiproliferative and apoptotic effects of selective phenolic acids on T47D human breast cancer cells: potential mechanisms of action. *Breast Cancer Res.* 2004;6:63–74.
75. Baqir HA, Zeboon NH, Al-Behadili AAJ. The role and importance of amino acids within plants: a review. *Plant Arch.* 2019;19:1402–10.
76. El Aribi H, Orlova G, Hopkinson AC, Siu KWM. Gas-phase fragmentation reactions of protonated aromatic amino acids: concomitant and consecutive neutral eliminations and radical cation formations. *J Phys Chem A.* 2004;108:3844–53.
77. Grobbelaar N, Steward FC. The isolation of amino acids from *Pisum sativum*. Identification of l(-)-homoserine and l(+)-O-acetylhomoserine and certain effects of environment upon their formation. *Phytochemistry.* 1969;8:553–9.
78. Wu G. Functional amino acids in Growth, Reproduction, and Health. *Adv Nutr.* 2010;1:31–7.
79. Khan N, Ali S, Zandi P, Mehmood A, Ullah S, Ikram M, et al. Role of sugars, amino acids and organic acids in improving plant abiotic stress tolerance. *Pakistan J Bot.* 2020;52:355–63.
80. Wiktorowska-Owczarek A, Berezinińska M, Nowak JZ. PUFAs: structures, metabolism and functions. *Adv Clin Exp Med.* 2015;24:931–41.
81. Weber H. Fatty acid-derived signals in plants. *Trends Plant Sci.* 2002;7:217–24.
82. Parvin R, Shahrokh KO, Mozafar S, Hassan E, Mehrdad B. Biosynthesis, regulation and properties of plant monoterpenoids. *J Med Plants Res.* 2014;8:983–91.
83. Weidenhamer JD, Macias FA, Fischer NH, Williamson GB. Just how insoluble are monoterpenes? *J Chem Ecol.* 1993;19:1799–807.
84. Supuran CT, Winum JY. Designing carbonic anhydrase inhibitors for the treatment of breast cancer. *Expert Opin Drug Discov.* 2015;10:591–7.
85. Franke CM, Gu VW, Grimm BG, Cassady VC, White JR, Weigel RJ, et al. TFAP2C regulates carbonic anhydrase XII in human breast cancer. *Oncogene.* 2020;39:1290–301.
86. Dumay A, Feugeas JP, Wittmer E, Lehmann-Che J, Bertheau P, Espié M, et al. Distinct tumor protein p53 mutants in breast cancer subgroups. *Int J Cancer.* 2013;132:1227–31.
87. Riggio M, Perrone MC, Polo ML, Rodriguez MJ, May M, Abba M, et al. AKT1 and AKT2 isoforms play distinct roles during breast cancer progression through the regulation of specific downstream proteins. *Sci Rep.* 2017;7:1–12.
88. Maennling AE, Tur MK, Niebert M, Klockenbring T, Zeppernick F, Gattenlöhner S, et al. Molecular targeting therapy against EGFR family in breast cancer: Progress and future potentials. *Cancers (Basel).* 2019;11:1826.
89. Vlahovic G, Crawford J. Activation of Tyrosine Kinases in Cancer. *Oncologist.* 2003;8:531–8.

### Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.