Iran. J. Biotechnol. April 2024;22(2): e3762 DOI: 10.30498/ijb.2024.407148.3762

Research Article

A Meta-analysis of Transcriptome Data to Investigate the Effect of Soy Isoflavones on Breast Cancer Cell

Elham Ashrafi-Dehkordi [1], Ahmad Tahmasebi [1,2], Habil Zare [3,4], Seyed Mohammad Mazloomi[1*]

[1] Nutrition Research Center, Department of Food Hygiene and Quality Control, School of Nutrition and Food Sciences, Shiraz University of Medical Sciences, Shiraz, Iran
[2] Biotechnology Institute, College of Agriculture, Shiraz University, Shiraz, Iran
[3] Department of Computer Science, Texas State University, San Marcos, Texas, 78666, USA
[4] Department of Cell Systems & Anatomy, The University of Texas Health Science Center, San Antonio, Texas, 78229, USA

[*]*Corresponding author*: Seyed Mohammad Mazloomi, Nutrition Research Center, Department of Food Hygiene and Quality Control, School of Nutrition and Food Sciences, Shiraz University of Medical Sciences, Shiraz, Iran. Tel: 071-37251001, Fax: 71348-14336, E-mail: mazloom@sums.ac.ir

Received: 2023/09/11 ; **Accepted:** 2024/03/03

Background: Breast cancer ranks as the second highest cause of cancer-linked deaths in women, with varying rates between Western and Asian countries. The consumption of phytoestrogens can influence breast cancer occurrence.

Objective: To comprehend how soy isoflavones impact breast cancer cells, we conducted a meta-analysis, combining gene expression data from multiple studies. This approach aimed to identify crucial transcriptional characteristics driving breast cancer cell response to soy phytoestrogens.

Materials and Methods: The gene expression profiles obtained from the Gene Expression Omnibus and Array Express and were grouped into control and isoflavones exposure conditions. We performed a meta-analysis based on the effect size combination method to identify the differentially expressed genes (DEGs). In addition, we performed Gene Ontology (GO) enrichment analysis, pathway analysis, weighted gene co-expression network analysis (WGCNA) and recursive support vector machine (R-SVM) algorithm.

Results: Based on this meta-analysis, we identified 3,890 DEGs, of which 2,173 were up-regulated and 1,717 were down-regulated. For example, *SGCG*, *PLK2*, and *TBC1D9* were the most highly down-regulated genes and *EGR3*, *WISP2*, and *FKBP4* were the most highly expressed genes in the isoflavones exposure condition. The functional enrichment and pathway analysis were revealed “cell division” and “cell cycle” among the most enriched terms. Among the identified DEGs, 269 transcription factor (TF) genes belonged to 42 TF families, where the C_2H_2 ZF, bZIP, and bHLH were the most prominent families. We also employed the R-SVM for detecting the most important genes to classify samples into isoflavones exposure and control conditions. It identified a subset of 100 DEGs related to regulation of cell growth, response to estradiol, and intermediate ribonucleoside monophosphate in the purine (IMP) metabolic process. Moreover, the WGCNA separated the DEGs into five discrete modules strongly enriched for genes involved in cell division, DNA replication, embryonic digit morphogenesis, and cell-cell adhesion.

Conclusion: Our analysis provides evidence suggesting that isoflavone affects various mechanisms in cells, including pathways associated with NF-κB, Akt, MAPK, Wnt, Notch, p53, and AR pathways, which can lead to the induction of apoptosis, the alteration of the cell cycle, the inhibition of angiogenesis, and interference in the redox state of cells. These findings can shed light on the molecular mechanisms that underlie the response of breast cancer cells to isoflavones.

Keywords: Breast cancer, Coexpression analysis, Isoflavones, Meta-analysis, Microarray studies.

1. Background

Breast cancer is a significant global health concern for women. Lifestyle choices and diets play a role in breast cancer risk (1). Disparities in breast cancer rates between Asian and Western countries are attributed to soy consumption, which is linked to lower breast cancer risk in Asian populations (2, 3). However, the impact of isoflavones (found in soy) on breast cancer risk is debated. Some studies suggest isoflavones decrease risk (1), while others, including lab and animal studies, find no such association. A survey by Maskarinec *et al*. (4) found no change in mammographic density among premenopausal women taking isoflavone supplements, hinting that soy's influence might be linked to long-term or early-life exposure. Hence, investigating the cellular signaling related to soy-based breast cancer treatments could be valuable.

Soy foods contain isoflavones, a type of phytoestrogen found in plants like legumes, fruits, and vegetables. These compounds resemble the human hormone 17β-estradiol and act like estrogen in the body, potentially leading to various effects (5, 6). The main soy isoflavones, daidzein and genistein, can activate estrogen receptors (a and b), influencing estrogen-responsive genes in a dose-dependent manner (2, 5). These isoflavones compete with natural estrogens (2). Genistein has anti-breast cancer properties, inhibiting certain tumor-related activities, such as growth factor receptors and enzymes, while also exhibiting antioxidant, cell cycle-arresting, apoptotic, and anti-inflammatory effects (7, 8). Daidzein, another key phytoestrogen in soy, enhances the effectiveness of tamoxifen at normal levels in a rat model (9).

2. Objectives

The availability of extensive transcriptomic data has made it possible to study how exactly isoflavones affect cancer cells. Moreover, it provides researchers with large datasets, hence increasing statistical power (10, 11). As a result, statistical methods have become powerful tools for identifying genes suitable for testing under experimental conditions (12-15). In this comprehensive meta-analysis, we used microarray gene expression data to figure out how differentially expressed genes (DEGs) involved in breast cancer cells respond to isoflavones. In other words, we intended to identify core genes' regulatory responses to isoflavones regarding breast cancer. We also analyzed the DEGs by conducting functional enrichment analysis of Transcription factor (TF) families, identifying hub genes using topological analysis. Moreover, a pathway enrichment analysis was employed to understand gene differences in response to isoflavones regarding to breast cancer cells.

3. Materials and Methods

3.1. Data Collection

We obtained microarray datasets from two public databases, the Gene Expression Omnibus (GEO, www.ncbi.nlm.nih.gov/geo/) and ArrayExpress (www.ebi.ac.uk/arrayexpress). We used the search terms *isoflavone*, *soy*, *genistein*, *daidzein*, *breast cancer*, as well as combinations of these terms. We narrowed down the findings to those related to *Homo sapiens* to analyze MCF-7 cell lines (**Table 1**). Then, we used the robust multi-array average (RMA) and Expression Console software to normalize the data (16) and correct the background (17). Finally, we visualized a box plot for each sample to ensure identical selection among the samples.

Table 1. Characteristics of used datasets in the meta-analysis.

GEO Accession	Samples	
	Isoflavone exposure	Control
GSE9936	33	18
GSE69845	9	9
GSE59345	4	4
GSE5258	6	83
GSE5200	9	3
GSE50705	82	11

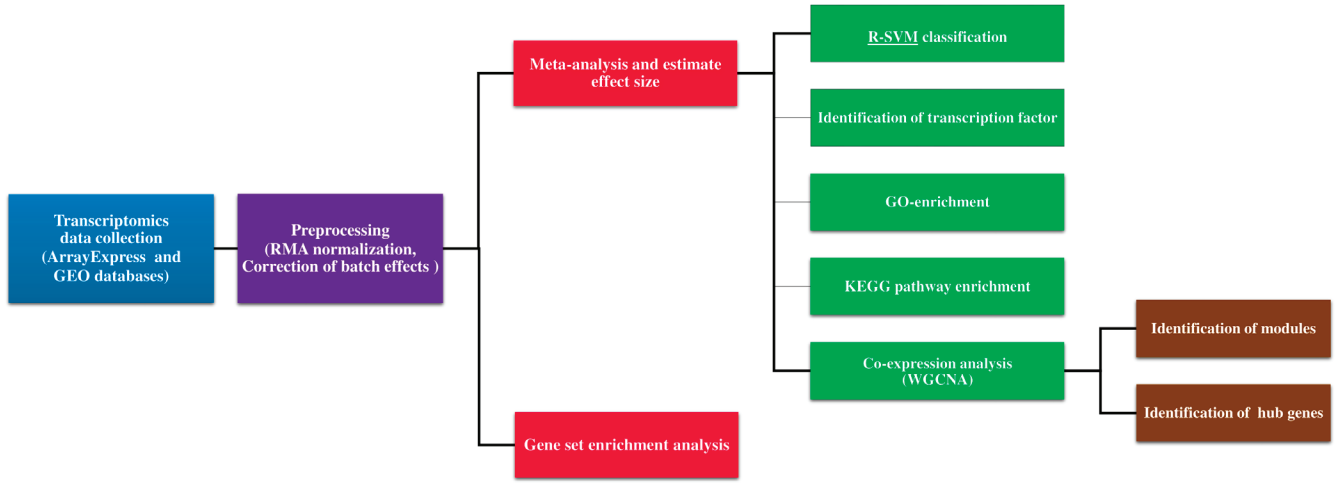

Figure 1. Schematic overview of multistep strategy for understanding aspects of response of breast cancer cells to isoflavones.

3.2. Meta-Analysis

We used the empirical Bayes method (CombBat) to adjust for batch effects (18). Further, we split datasets containing more than two conditions into sub-datasets in order to reduce heterogeneity within studies. We performed a meta-analysis based on the effect size combination method, and computed the effect sizes using the metaMA (Version 3.1.2) R package. We used the false discovery rate (FDR) to distinguish DEGs in both the control and isoflavones exposure conditions and other genes, so that genes with an FDR< 0.001 were considered to be DEGs.

3.3. Classification Analysis

We conducted a PCA and utilized a recursive SVM (R-SVM) to obtain a global perspective and to evaluate the performance of DEGs in the classification of control conditions versus isoflavones exposure conditions. The k-fold cross-validation (k=10) was also used to calculate the error rate of the R-SVM predictor. The PCA and SVM were implemented using the e1071 R package and MetaboAnalyst, respectively (19).

3.4. The Analysis of Pathway Enrichment of DEGs and Gene Ontology

To analyze gene ontology (GO) of DEGs in detail, we annotated them, using DAVID (https://david.ncifcrf.gov/home.jsp). We further used Benjamini-Hochberg method to determine the specific GO terms and applied REVIGO to make significant GO term lists less redundant. We used the Cluster profiler R package to estimate the statistical enrichment of DEGs in KEGG pathways. We retrieved TFs and their families from the CIS-BP database (20). We established a corrected *p-value* of 0.05 to test the significance of the effect.

3.5. Co-Expression Module Detection

To cluster highly correlated DEGs, we analyzed weighted gene co-expression networks (WGCN) (21) on normalized expression values of DEGs. Additionally, utilizing the determination of intra-modular connectivity values (kWithin), we also detected hub genes. We conducted GO functional analysis of the lists of genes that correspond to modules using the DAVID Web tool with a *p*-value < 0.05 as the cutoff.

4. Results

4.1. Identification of DEGs Associated with Isoflavones Exposure

We selected datasets based on the inclusion criteria. From these datasets, a total of 271 samples were retrieved for a meta-analysis (**Table S1**). **Figure 1** shows a flowchart of our analysis pipeline. There

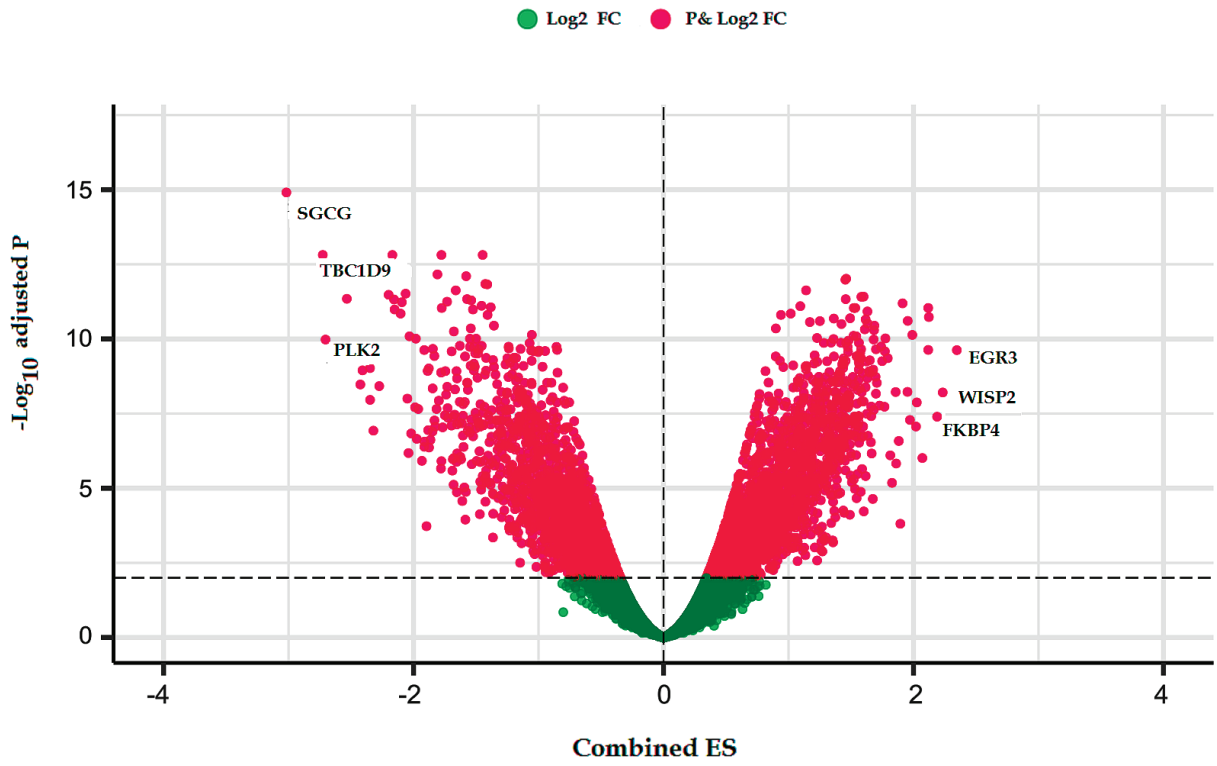

Figure 2. **Gene expression comparison between control and isoflavone exposure conditions.** Volcano plot showing combined effect size (x-axis) and significance level (−log10-adjusted p-value; y-axis). The significant up and down-regulated genes are plotted as red dots.

were 12,462 genes shared across all datasets. Each dataset was grouped into the control condition or the isoflavones exposure condition to identify DEGs according to a random-effects model. This meta-analysis identified a total of 3,890 DEGs, consisting of 2,173 up-regulated and 1,717 down-regulated genes (**Table S2**). While *SGCG*, *PLK2*, and *TBC1D9* were the most highly down-regulated genes; *EGR3*, *WISP2*, and *FKBP4* were the most highly expressed genes in the isoflavones exposure condition (**Fig. 2**).

4.2. Classification Performance of DEGs

The PCA indicated that the DEGs are effective in distinguishing between the control and isoflavones exposure conditions (**Fig. S1**). Based on the expression data from 3,890 DEGs, the R-SVM identified 1,556 genes (40% of 3,890) that could classify the control and isoflavones exposure conditions with 95.4% accuracy. Additionally, the R-SVM identified 100 genes (2.6%) with 88.1% accuracy and eight genes (0.2%) with 70.3% accuracy (**Fig. S2**).

The group of 100 genes identified with 88.1% accuracy are involved in a number of cell growth regulation processes, responses to estradiol, and IMP metabolic processes. Among the genes identified above, multiple genes encode a number of signal transductions and unsaturated fatty acid biosynthetic processes. Zhang *et al*. (22) reported that polyunsaturated fatty acids inhibit the growth of colon cancer cells. The *TMPRSS3*, *ELOVL2,* and *S100A6* genes scored the highest in the R-SVM. The *TMPRSS3* gene relates to the development and progression of tumors, and it is known as a prognostic marker of poor survival of breast cancer. The *ELOVL2* gene encodes an enzyme responsible for the production of 24 carbon-chain fatty acids, the depletion of which enhances the metastatic characteristics of breast cancer cells. The *S100A6* gene is a calcium-binding protein associated with estrogen receptor expression. This gene has been introduced as a biomarker of the early stages of breast cancer in humans (23).

4.3. Functional Enrichment Analysis

The analysis of up-regulated genes biological processes showed a significant overrepresentation of terms related to cell division, mitosis, and viral processes (**Fig. 3A**). These results also indicate that genistein might impact cell growth by affecting signaling pathways. The recombination-based telomere maintenance was also significant (**Table S3**). Genistein treatment (50–100 μM) suppressed the activity of telomerase in MCF-7 cells by altering the expression of crucial TFs and the methylation status (3). Furthermore, the down-regulated genes were found to have many terms associated with negative regulation of protein transport and cell proliferation. (**Fig. 3B**). Moreover, while the main classes of molecular functions could both bind protein and ATP for the upregulated genes, they could not bind ATP for the down-regulated genes (**Table S3, S4**). Accordingly, while the nucleus was highly represented in the cellular class component for up-regulated genes, the cytoplasm was the same for down-regulated genes (**Table S3, S4**).

4.4. KEGG Pathway

We conducted an enrichment analysis of KEGG pathway to identify pathways that were responsive to isoflavones. We found a link between up-regulated genes and the DNA replication, cell cycle, and mismatch repair pathways (**Fig. 4**). There are reports suggesting that soy isoflavones impede cancer progression by regulating genes that are associated with the cell cycle. Interestingly, we found many up-regulated genes involved in the p53 signaling pathway (**Table S5**). Isoflavones have been reported to inhibit activation of NF-κB. Moreover, a number of the identified genes involved in NF-κB signaling pathway, namely *NF-Kb2*, *NF-κB1A* and *GSK3B* were down-regulated (**Table S1**). The PI3K/Akt/mTOR intracellular pathway is a determining factor in cell proliferation, cellular survival, apoptosis, and protein synthesis (24). Moreover, this pathway is the most frequently activated signaling in breast cancer. In this study, a number of genes family, namely, *BAD*, *ERBB2*, *ERBB3*, *ERBB4*, *IGF2R*, *IGFBP2*, *IGFBP3*, *IGFBP5*, *PDGF*, *VEGF*, *FGFR1*, *FGFR4*, *EPHA4*, *DDR*, and *ROR* were down-regulated in this study (**Table S1**).

It has been found that isoflavones block the activation of the p38 MAPK pathway and lead to the induction of apoptosis. It also inhibits cancer cell invasion and metastasis in ER-negative and ER-positive breast cancer cells, though apparently by different mechanisms. Also, our result shows that TGF-beta and MAPK signaling pathways were down-regulated in isoflavones in the exposure condition (**Table S1**). In cancer cells, the Wnt signaling pathway interacts with the Akt signaling pathway to promote cell proliferation and prevent apoptosis. Isoflavones could inactivate the Wnt signaling pathway to induce apoptosis and inhibit cell growth. Also, our result shows that Wnt genes were down-regulated in isoflavones in the exposure condition (**Table S1**). Previous studies have shown that genistein can inhibit the expression of Notch genes (25, 26), which is consistent with our findings. Additionally, isoflavones can induce apoptosis by the up regulating the p53 pathway. The p53 gene has been shown to function as a key tumor suppressor. It directly activates gene expression in regions containing p53 binding sites by acting as a transcription factor. It regulates cancer cell progression via multiple mechanisms (induction of cell cycle arrest and cellular senescence or apoptosis) (27). In this study, several genes involved in the p53 signaling pathway, such as *CASP2*, *CASP3*, and *CASP8*, were found to be up-regulated, apoptosis increased, CDK genes (*CDK5* and *CDK17*) were down-regulated, and the cell cycle was arrested (**Table S1, S5**). Also, p53 activation is induced by a number of activated genes (i.e., *MYC*, *E2F1*, and *NRAS*), all of which were up-regulated (**Table S1**). The androgen receptor (AR) belongs to the steroid hormone receptor family of ligand-activated nuclear TFs. It is heavily involved in cancer development. Previous studies have reported that genistein treatment down regulates AR proteins and decreases the binding of nuclear proteins to androgen-responsive elements (28). A number of the identified genes that are AR activators (i.e., *PIAS1*, *PNRC1, 2*, *PRMT2*, and *Smad3*) were down-regulated, and a number of the identified AR repressors genes (i.e., *AKT*, *TGIF*, and *P53*) were up-regulated (**Table S1**).

Many of the identified genes involved in the TGF-beta signaling pathway were down-regulated. TGF-beta is an important regulator of cell proliferation. Genistein inhibits TGF-beta signaling and thus angiogenesis (5). Our results show that *CYP2J2*, *CYP4B1*, and *CYP1A1* were down-regulated after treatment with genistein and daidzein. These genes have been shown to modulate the CYP450 system (5).

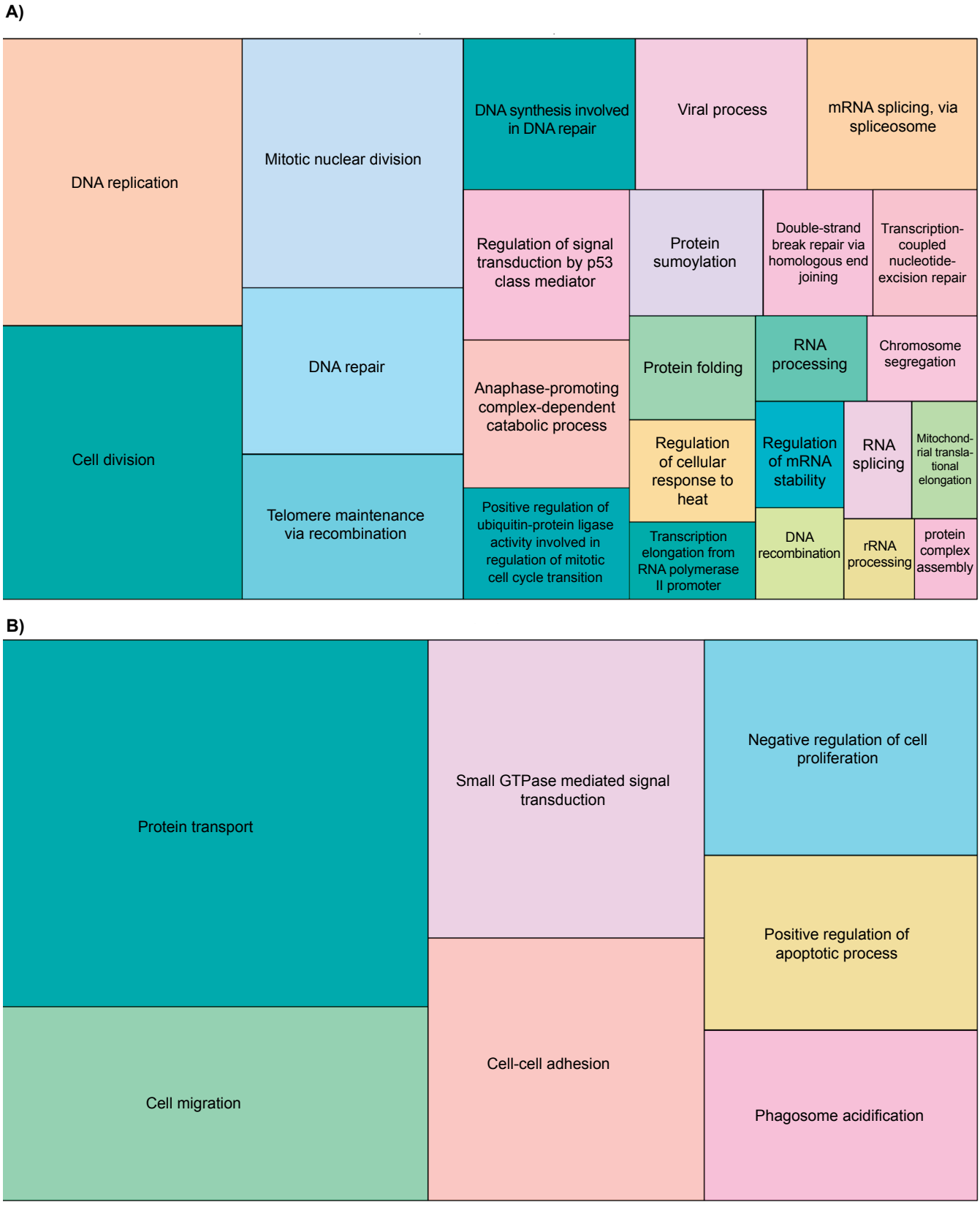

Figure 3. Gene Ontology of biological processes. A) The up-regulated and **B)** the down-regulated DEGs.

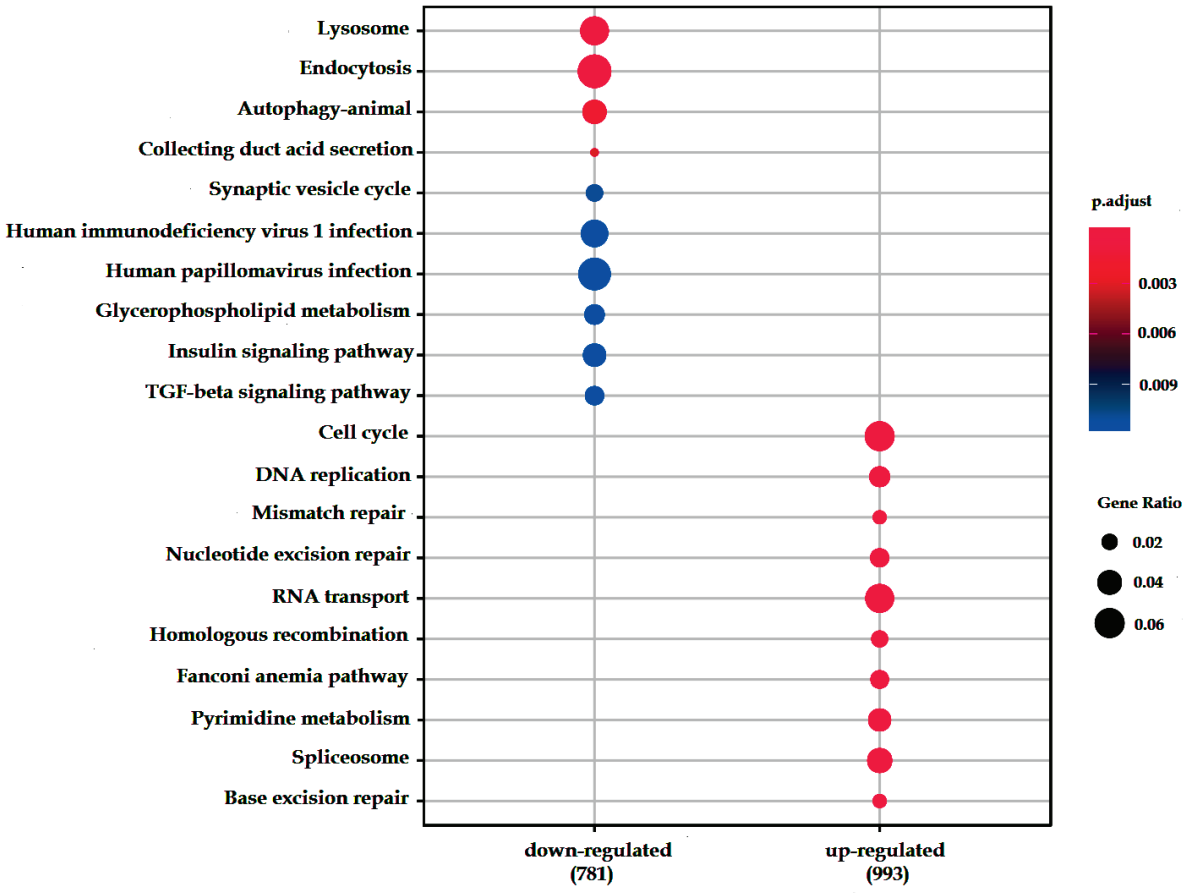

Figure 4. KEGG pathway analysis performed on genes differentially expressed. Top 10 pathways for the up-regulated and down-regulated DEGs have been represented in diagram. The size and color bubbles show the FDR and gene ratio of each pathway.

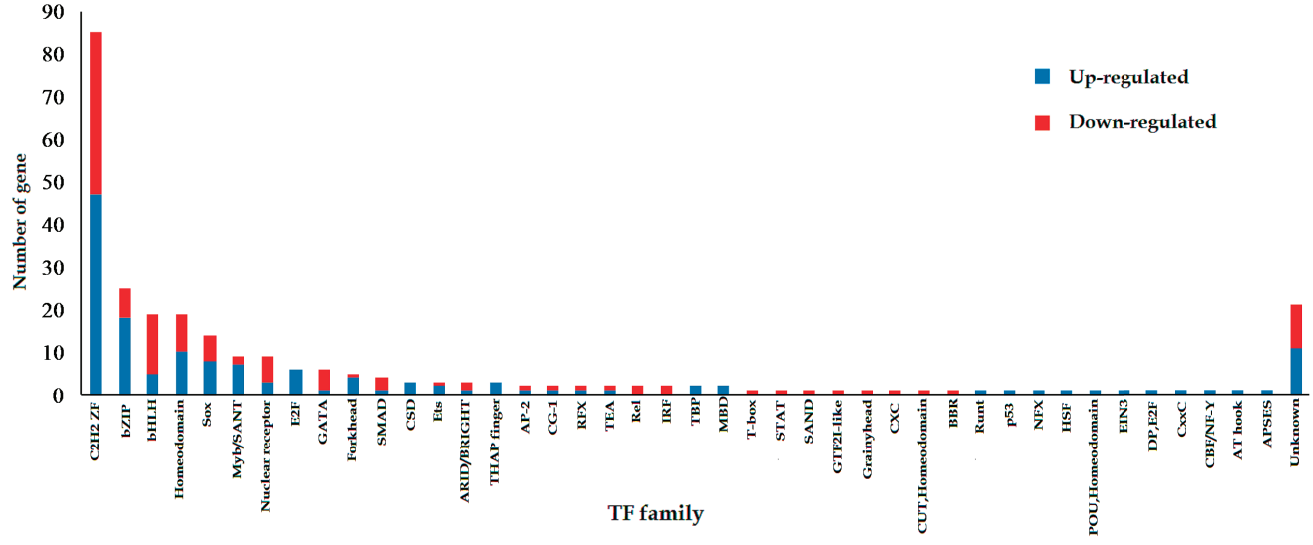

Figure 5. Distribution of TF families in the DEGs. The number of up- or down-regulated are shown for each transcription factor family.

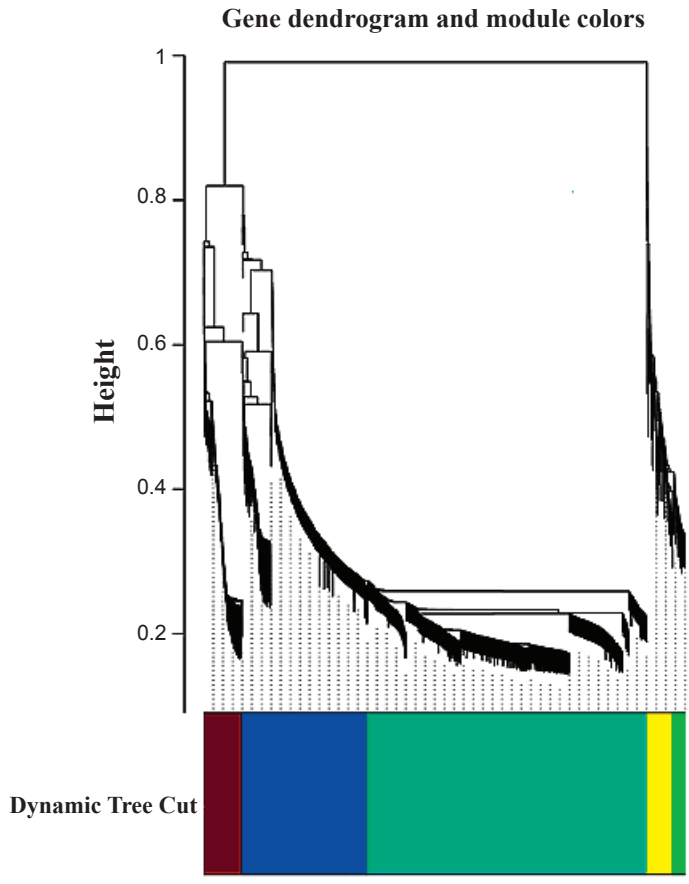

Figure 6. Weighted gene co-expression network analysis (WGCNA) of DEGs. Cluster dendrogram showing co-expression modules identified by WGCNA. The modules are denoted in the color bar.

Genistein has also been shown to be an effective inhibitor of DNA damage in MCF-7 by inhibiting the expression of *CYP1A1* and *CYP4B1* that result from inflammatory response and angiogenic activity.

4.5. Identification of TFs Related to Isoflavones Exposure

Transcription factors play have been linked to a broad spectrum of diseases and phenotypes. They have an effect on genome expression regulation. Out of the DEGs, 269 TFs were recognized, and these are part of 42 families. With 85, 25 and 19 members, the C_2H_2 ZF, bZIP and bHLH families were the most abundant. Overall, 55.4% of the TFs were found to be up-regulated, while 44.6% were found to be down-regulated. Current studies have identified zinc finger proteins as important TFs involved in cancer progression throught regulating the transcription of downstream genes involved in migration, apoptosis, proliferation, and invasion. Under the isoflavones exposure condition TFs of the E2F family were up-regulated (**Fig.5, Table S6**). Previous studies have shown that during much of the G1 phase of the cell cycle, pRb binds to E2F1, E2F2, and E2F3, silencing gene expression by recruiting HDACs or HMTs. The pRb affects cell growth by interacting with other factors, such as activating E2F1, E2F2, and E2F3 (29).

4.6. A Network Analysis of DEG's Weighted Gene Co-Expression

The WGCNA was conducted to discover modules and hub genes within the modules. This WGCNA analysis was utilized for all the DEGs. The genes were organized into five distinct modules, each represented by a distinct color code. There were a range 101 - 2,257 genes in each module. A module with 2,257 genes was

called *turquoise*, and another one with 1,013 genes was called *blue* (**Fig. 6**). The turquoise module significantly enhances cell division and mitosis, the blue module most of which were up-regulated for isoflavones, helped DNA replication (**Table S7**). Additionally, the blue module contained 39 genes associated with cancer pathways, such as *BAD*, *HSP90B1*, and *E2F3* (**Table S7**). There are two distinct mechanisms by which BAD regulates cell growth. Hsp90B1 has been shown to help ovarian cells survive and apoptosis (30). The *E2F3* transcription factor helps control cell cycle progression. Expression of certain members of the *E2F* family has been implicated in predicting breast cancer (29).

A highly enriched concentration of genes involved in embryonic digit morphogenesis was found in the brown module, being annotated in response to estrogen. It also contained a highly enriched concentration of hippo signaling pathway (**Table S7**). The Hippo signaling pathway is important in breast cancer proliferation and metastasis (31). In the brown module, the *GATA3* gene was also found. The *GATA3* required for normal mammary development and relatively high susceptibility to breast cancer. The expression level of *GATA3* is down-regulated after soy exposure in immature lobules (32). Properties such as cell-cell adhesion, translation and protein folding were characterized in the yellow module (**Table S7**). However, no significant enriched biological process term was found in the green module. A KEGG analysis of these modules revealed an overrepresentation of KEGG pathways related to the cell cycle. The DEGs were significantly enriched in the cell cycle and cancer pathways in the blue module (**Table S8**).

4.7. Identification of Hub Genes in Network Modules

The hub genes in each module were identified which may play an important role in responding to exposure to isoflavones. A total of 10 genes with the highest intramodular connectivity were detected as hub genes (**Table S9**). The results indicate a high degree of interconnection between *PKP4* and *BYSL* in the turquoise module, between *APBB2* and *ASNS* in the blue module, between *PRDX6* and *NDUFB1* in the yellow module, between *CNNM2* and *PLGRKT* in the brown module, and between *TPBG* and *PGAM1* in the green module. The results indicated in the isoflavones exposure condition, the expression of *PGAM1*, *TPBG*, *PLGRKT*, *NDUFB1*, *PKP4*, and *APBB2* were down-regulated and *BYSL*, *ASNS*, *PRDX6*, and *CNNM2* were up-regulated.

5. Discussion

Recent research has revealed that soy isoflavones have multiple pathways of action, contributing to their pleiotropic effects on cancer. These mechanisms include regulation of cell proliferation, apoptosis, and survival, along with the inhibition of angiogenesis and metastasis, as well as their antioxidant properties (25). In the current study, we have developed a comprehensive pipeline to examine transcriptional alterations induced by soy isoflavones, aiming to elucidate the molecular mechanisms underlying their impact on breast cancer. Through meta-analysis, we successfully pinpointed 3,890 genes that exhibit significant changes in expression upon exposure to isoflavones. Notably, certain genes, such as *EGR3*, *WISP2*, and *FKBP4*, demonstrated particularly substantial alterations in their expression levels. The *EGR3* exhibited the highest expression level. This was expected because *EGR3* plays an important role in induction of the immune evasion system in ER-positive breast cancer (33). The *WISP2* is a marker of estrogen exposure that is involved in the regulation of tumor cell proliferation (23). *FKBP4* may have an important role in immunoregulation processes and endocrine-responsive breast cancer (34).

The pathway analysis indicates that the DEGs are primarily linked to key pathways, including the p53 signaling, cell cycle, DNA replication, and lysosome pathways (**Fig. 4**). Soy isoflavones have been reported to inhibit cancer development, which may be caused by the regulation of genes related to the cell cycle (27). Furthermore, a recent study highlighted that minor defects within DNA repair systems raise the breast cancer risk. The loss of some genes in this pathway (e.g., *MLH1* and *MSH2*) may lead to breast cancer progression (35). The induction of cell death in MCF-7 cell cultures after treatment with genistein has been reported due to upregulation of the proapoptotic protein p53 (5). Generally, the lysosome pathway contributes significantly to oncogenic transformation in cancer cells (36). In addition, the genes related to the insulin signaling pathway are also abundant (**Table S5**). Normal homeostasis is maintained primarily through the process that involves insulin, IGFs, and IGF binding proteins. Any abnormal activity in these key components or any aberrant expression of these

components increases the risk for both type II diabetes and cancer. Future clinical opportunities include blocking insulin-mediated signalling and developing potential anti-cancer therapies (37).

Within the set of DEGs, we have identified a total of 269 TFs, with the majority falling into the C_2H_2 ZF, bZIP, and bHLH families (**Fig. 5, Table S6**). Recent studies suggest that zinc finger proteins play a role in cancer progression, as they modulate the transcription of downstream genes involved in invasion, mitigation, apoptosis and proliferation (38). ZNF217, for instance, contribute to breast cancer metastasis (39). We found that the expression level of *E2F1*, *E2F2*, *E2F3*, *E2F4*, *E2F6*, and *E2F8* were up-regulated, when exposed to isoflavones. Likewise, pRb has been found to silences gene expression by recruiting HDACs or HMTs as it binds to *E2F1*, *E2F2*, and *E2F3* during G1 phase of the cell cycle. It has also been reported that by activating E2F1, E2F2, and E2F3, pRb influences cell growth (40).

We conducted a WGCNA, to explore the interactions among the DEGs to gain insights into the mechanisms underlying the effects of isoflavones on cancer cells. We grouped the DEGs into five distinct modules (**Fig. 6**). Subsequent functional analysis revealed that the genes within these modules were primarily linked to processes such as mitotic nuclear division, cell division, and DNA replication (**Table S7**). Furthermore, we found 39 genes, linked to cancer pathways, including key genes such as *BAD*, *HSP90B1*, and *E2F3* in the blue module (**Table S7**). BAD regulates cell growth via two distinct mechanisms. There is evidence in the literature for the role of Hsp90B1in ovarian cell survival and cell apoptosis (30). The E2F3 transcription factor helps control cell cycle progression. E2F3a, which differs from E2F3b in its N-terminal sequence, has been linked to the transcriptional activation of E2F-responsive genes. These TF families are linked cycling E in the cell life. In breast carcinomas, the expression of some the E2F family members has been linked to prognosis (29). Interestingly, it was found that the brown module exhibited enrichment for genes related to the Hippo signaling pathway (**Table S7**), which plays a crucial part in breast cancer metastasis and cell proliferation (31). *GATA3*, which was found in the brown module, is to some extent sensitive to breast carcinomas because it needs the mammary gland to develop normally (41). The expression level of *GATA3* is down-regulated after soy exposure in immature lobules (32).

The highly connected genes within the modules, which include *PKP4*, *BYSL*, *APBB2*, *ASNS*, *PRDX6*, *NDUFB1*, *CNNM2*, *PLGRKT*, *TPBG* and *PGAM1* were found to be significantly enriched in the PDGF signaling pathway. A distant relative of plakophilins and a member of the armadillo subfamily member, *PKP4* has been described as a component of adherents junctions, as it clusters and stabilizes cadherins to control intercellular adhesion (31). *BYSL* encodes bystin, an essential protein component that contributes to the attachment of the embryo to the uterus (42). The BYSL protein can be found in human prostatic carcinoma cells in areas of perineural invasion in in substantial amount, playing an adhesive role (43). Overexpression of amyloid beta A4 precursor protein binding, family B, member 2 (APBB2) in PC12 cells inhibits cell cycle progression, as it translocates the nucleus and inhibits thymidylate synthase, thus delaying the cell cycle. The CNNM family increases intracellular magnesium levels that contribute to oncogenic transformation, and PRL-2 interacts with the magnesium transporter to promote oncogenesis (44). These families have essential functions, such as the regulation of the circadian rhythm and reproduction (44). *PRDX6* encodes a crucial antioxidant enzyme with various roles in cellular function. For instance, *PRDX6* reduces a wide range of peroxide substrates in the cell, thus it helps maintain redox homeostasis in mammalian cells (45). Moreover, *PRDX6* takes part in intracellular and intercellular signal transduction due to its phospholipase and peroxidase activity, thus facilitating the initiation of regenerative processes in the cell, the activation of cell proliferation, and the suppression of apoptosis (46). NDUFB1 relates mainly to mitochondrial ATP synthesis coupled electron transport, the electron transport chain, and mitochondrial electron transport. It also plays a role in lung adenocarcinoma carcinogenesis and could potentially serve as a diagnostic marker for lung cancer (47).

6. Conclusion

Our analysis provides evidence suggesting that isoflavones affect the NF-κB, Akt, MAPK, Wnt, Notch, p53, and AR pathways. These effects cause the induction of apoptosis, the alteration of the cell cycle, the inhibition of angiogenesis, and interference in the

redox state of the cell. In addition, our analysis expands the understanding of the role played by transcription factors such as C_2H_2, bZIP, and bHLH, which are essential for the development of breast cancer.

Financial support

This work was supported by grant from Shiraz University of Medical Sciences [No. 97-01-106-19209].

Competing interests

The authors declare no competing interests.

References

1. Trock BJ, Hilakivi-Clarke L, Clarke R. Meta-analysis of soy intake and breast cancer risk. *J Natl Cancer Inst.* 2006;**98**(7):459-471. doi:10.1093/jnci/djj102
2. Satih S, Chalabi N, Rabiau N, Bosviel R, Fontana L, Bignon YJ, Bernard-Gallon DJ. Gene expression profiling of breast cancer cell lines in response to soy isoflavones using a pangenomic microarray approach. *Omics: J Integr Biol*. 2010;**14**(3):231-238. doi:10.1089/omi.2009.0124
3. Ashrafi-Dehkordi E, Mazloomi SM, Hemmati F. A comparison of DNA extraction methods and PCR-based detection of GMO in textured soy protein. *JCPFS*. 2021;**16**:51-57. doi:10.1007/s00003-020-01300-2
4. Maskarinec G, Williams A, Carlin L. Mammographic densities in a one-year isoflavone intervention. *Eur J Cancer Prev*. 2003;**12**(2):165-169. doi:10.1097/00008469-200304000-00011
5. Satih S, Rabiau N, Bignon Y-J, Bernard-Gallon DJ. Soy phytoestrogens and breast cancer chemoprevention: molecular mechanisms. Curr. Nutr. *Food Sci*. 2008;**4**(4):259-264. doi:10.2174/157340108786263658
6. Sirtori CR, Arnoldi A, Johnson SK. Phytoestrogens: end of a tale? *Ann Med*. 2005;**37**(6):423-438. doi:10.1080/07853890510044586
7. Hsieh C-Y, Santell RC, Haslam SZ, Helferich WG. Estrogenic effects of genistein on the growth of estrogen receptor-positive human breast cancer (MCF-7) cells in vitro and *in vivo*. *Cancer Res*. 1998;**58**(17):3833-3838.
8. Bouker KB, Hilakivi-Clarke L. Genistein: does it prevent or promote breast cancer? *Environ Health Perspect*. 2000;**108**(8):701-708. doi:10.1289/ehp.00108701
9. Constantinou AI, White BE, Tonetti D, Yang Y, Liang W, Li W, Van Breemen RB. The soy isoflavone daidzein improves the capacity of tamoxifen to prevent mammary tumours. *Eur J Cancer.* 2005;**41**(4):647-654. doi:10.1016/j.ejca.2004.12.005
10. Ashrafi-Dehkordi E, Alemzadeh A, Tanaka N, Razi H. Meta-analysis of transcriptomic responses to biotic and abiotic stress in tomato. *PeerJ*. 2018;**6**:e4631. doi:10.7717/peerj.4631
11. Davati N, Ghorbani A, Ashrafi-Dehkordi E, Karbanowicz TP. Gene Networks Analysis of Salmonella Typhimurium Reveals New Insights on Key Genes Involved in Response to Low Water Activity. *Iran J Biotechnol*. 2023;**21**(4):71-82. doi:10.30498/IJB.2023.387696.3640
12. Piryaei Z, Salehi Z, Tahsili MR, Ebrahimie E, Ebrahimi M, Kavousi K. Agonist/antagonist compounds' mechanism of action on estrogen receptor-positive breast cancer: A system-level investigation assisted by meta-analysis. *IMU*. 2022;**31**:100985.
13. Katiyar A, Sharma S, Singh TP, Kaur P. Identification of shared molecular signatures indicate the susceptibility of endometriosis to multiple sclerosis. *Front Genet*. 2018;**16**(9):42.
14. Jha PK, Vijay A, Sahu A, Ashraf MZ. Comprehensive Gene expression meta-analysis and integrated bioinformatic approaches reveal shared signatures between thrombosis and myeloproliferative disorders. *Sci Rep*. 2016;**6**(1):37099.
15. Hojjati F, Roointan A, Gholaminejad A, Eshraghi Y, Gheisari Y. Identification of key genes and biological regulatory mechanisms in diabetic nephropathy: Meta-analysis of gene expression datasets. *Nefrologia*. 2023;**43**(5):575-586.
16. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostat*. 2003;**4**(2):249-264. doi:10.1093/biostatistics/4.2.249
17. Dziuda DM. Data mining for genomics and proteomics: analysis of gene and protein expression data. *John Wiley & Sons*. 2010;**16**.
18. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostat*. 2007;**8**(1):118-127. doi:10.1093/biostatistics/kxj037
19. Xia J, Sinelnikov IV, Han B, Wishart DS. MetaboAnalyst 3.0—making metabolomics more meaningful. *Nucleic Acids Res*. 2015;**43**(W1):W251-W257. doi:10.1093/nar/gkv380
20. Weirauch MT, Yang A, Albu M, Cote AG, Montenegro-Montero A, Drewe P, *et al*. Determination and inference of eukaryotic transcription factor sequence specificity. *Cell*. 2014;**158**(6):1431-1443. doi:10.1016/j.cell.2014.08.009
21. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform*. 2008;**9**(1):559. doi:10.1186/1471-2105-9-559
22. Zhang C, Yu H, Ni X, Shen, S, Das U. N. Growth inhibitory effect of polyunsaturated fatty acids (PUFAs) on colon cancer cells via their growth inhibitory metabolites and fatty acid composition changes. *PLoS One*. 2015;**10**(4):e0123256. doi:10.1371/journal.pone.0123256
23. Desai KV, Simmons JL, Fargiano A, Van Den Eynden G, Vermeulen PB, Dirix LY, Merino M, Green JE. S100A6 as a biomarker in human breast cancer. *In Proc Amer Assoc Cancer Res*. 2005;**46**:448.
24. Li Y, Liu L, Andrews LG, Tollefsbol TO. Genistein depletes telomerase activity through cross-talk between genetic and epigenetic mechanisms. I*nterl J Cancer*. 2009;**125**(2):286-296. doi:10.1002/ijc.24398
25. Varinska L, Gal P, Mojzisova G, Mirossay L, Mojzis J. Soy and breast cancer: focus on angiogenesis. *Int J Mol Sci*. 2015;**16**(5):11728-11749. doi:10.3390/ijms160511728
26. Rizzo P, Osipo C, Foreman K, Golde T, Osborne B, Miele L. Rational targeting of Notch signaling in cancer. *Onco Gene*. 2008;**27**(38):5124-5131.
27. Zhang Z, Wang CZ, Du GJ, Qi LW, Calway T, He TC, Du WE, Yuan CS. Genistein induces G2/M cell cycle arrest and apoptosis via ATM/p53-dependent pathway in human colon cancer cells. *Int J Oncol*. 2013;**43**(1):289-296. doi:10.3892/ijo.2013.1946
28. Bennett NC, Gardiner RA, Hooper JD, Johnson DW, Gobe GC. Molecular cell biology of androgen receptor signalling. *Int J Bioch Cell Biol*. 2010;**42**(6):813-827. doi:10.1016/j.biocel.2009.11.013
29. Zacksenhaus E, Liu JC, Jiang Z, Yao Y, Xia L, Shrestha M, Ben-

David Y. Transcription factors in breast cancer—lessons from recent genomic analyses and therapeutic implications. *Adv Protein Chem Struct Biol*. 2017;**107**:223-273. doi:10.1016/bs.apcsb.2016.10.003

30. Li L, Mo H, Zhang J, Zhou Y, Peng X, Luo X. The role of heat shock protein 90B1 in patients with polycystic ovary syndrome. *PloS One*. 2016;**11**(4):e0152837. doi:10.1371/journal.pone.0152837
31. Wei C, Wang Y, Li X. The role of Hippo signal pathway in breast cancer metastasis. *Onco Targets Ther*. 2018;**11**:2185. doi:10.2147/ott.s157058
32. Dewi FN, Wood CE, Willson CJ, Register TC, Lees CJ, Howard TD, Huang Z, Murphy SK, Tooze JA, Chou JW, Miller LD. Effects of pubertal exposure to dietary soy on estrogen receptor activity in the breast of cynomolgus macaques. *Cancer Pre Res*. 2016;**9**(5):385-395. doi:10.1158/1940-6207.capr-15-0165
33. Inoue A, Omoto Y, Yamaguchi Y, Kiyama R, Hayashi SI. Transcription factor EGR3 is involved in the estrogen-signaling pathway in breast cancer cells. *J Mol Endocrinol*. 2004; **32**(3):649-661. doi:10.1677/jme.0.0320649
34. Banerjee S, Saxena N, Sengupta K, Tawfik O, Mayo MS, Banerjee SK. WISP-2 gene in human breast cancer: estrogen and progesterone inducible expression and regulation of tumor cell proliferation. *Neoplasia*. 2003;**5**(1):63-73.
35. Malik SS, Masood N, Asif M, Ahmed P, Shah ZU, Khan JS. Expressional analysis of MLH1 and MSH2 in breast cancer. *Curr Probl Cancer*. 2019;**43**(2):97-105. doi:10.1016/j.currprobl cancer. 2018.08.001
36. Fennelly C, Amaravadi RK. Lysosomal biology in cancer. Lysosomes: *Springer*. 2017. **p**. 293-308. doi:10.1007%2F978-1-4939-6934-019
37. Ray A, Alalem M, Ray BK. Insulin signaling network in cancer. Indian. *J Biochem Biophys*. 2014;**51**:493-498
38. Jen J, Wang Y-C. Zinc finger proteins in cancer progression. *J Biomed Sci*. 2016;**23**(1):53. doi:10. 1186 /s12929-016-0269-9
39. Bellanger A, Donini CF, Vendrell JA, Lavaud J, Machuca-Gayet I, Ruel M, Vollaire J, Grisard E, Győrffy B, Bièche I, Peyruchaud O. The critical role of the ZNF217 oncogene in promoting breast cancer metastasis to the bone. *J Pathol*. 2017;**242**(1):73-89. doi:10.1002/path.4882
40. Dyson NJ. RB1: a prototype tumor suppressor and an enigma. *Genes Dev*. 2016;**30**(13):1492-1502. doi:10.1101/gad.282145.116
41. Shaoxian T, Baohua Y, Xiaoli X, Yufan C, Xiaoyu T, Hongfen L, Rui B, Xiangjie S, Ruohong S, Wentao Y. Characterisation of GATA3 expression in invasive breast cancer: differences in histological subtypes and immunohistochemically defined molecular subtypes. *J Clin Pathol*. 2017;**70**(11):926-934. doi:10.1136/jclinpath-2016-204137
42. Kimber SJ. Molecular interactions at the maternal-embryonic interface during the early phase of implantation. *Semin Reprod Med*. 2000;**18**(3):237-254. doi:10. 1055/s-2000-12562
43. Ayala GE, Dai H, Li R, Ittmann M, Thompson TC, Rowley D, Wheeler TM. Bystin in perineural invasion of prostate cancer. *The Prostate*. 2006;**66**(3):266-272. doi:10.1002/pros.20323
44. Giménez-Mascarell P, Oyenarte I, Hardy S, Breiderhoff T, Stuiver M, Kostantin E, Diercks T, Pey AL, Ereño-Orbea J, Martínez-Chantar ML, Khalaf-Nazzal R. Structural Basis of the Oncogenic Interaction of Phosphatase PRL-1 with the Magnesium Transporter CNNM2. *JBC*. 2017;**292**(3):786-801. doi:10.1074/jbc.m116.759944
45. Manevich Y, Fisher AB. Peroxiredoxin 6, a 1-Cys peroxiredoxin, functions in antioxidant defense and lung phospholipid metabolism. *Free Radic Biol Med*. 2005;**38**(11):1422-1432. doi:10.1016/j.freeradbiomed.2005.02.011
46. Sharapov MG, Novoselov VI, Gudkov SV. Radioprotective role of peroxiredoxin 6. *Antioxidants*. 2019;**8**(1):15. doi:10.3390/antiox8010015
47. Xu H, Ma J, Wu J, Chen L, Sun F, Qu C, Zheng D, Xu S. Gene expression profiling analysis of lung adenocarcinoma. Braz. *J Med Biol*. 2016;**49**. doi:10.1590%2F1414-431X20154861