Trop J Nat Prod Res, December 2024; 8(12):9476 - 9485
ISSN 2616-0684 (Print)
ISSN 2616-0692 (Electronic)

Tropical Journal of Natural Product Research

Available online at https://www.tjnpr.org

Original Research Article

Differentially Expressed Genes (DEGs) Analysis and In Silico Studies Identify Tumor Necrosis Factor (TNF) Inhibition and Peroxisome Proliferator-Activated Receptor Alpha (PPARA) Activation as Targets for Gallic Acid Derivatives in Insulin Resistance

Dwi Anita Suryandari[1], Aryo Tedjo[2,3,4*], Fadilah Fadilah[2,3,4]

[1]*Department of Medical Biology, Faculty of Medicine, Universitas Indonesia, Salemba Raya 6, Jakarta 10430, Indonesia*
[2]*Department of Medical Chemistry, Faculty of Medicine, Universitas Indonesia, Salemba Raya 6, Jakarta 10430, Indonesia*
[3]*Drug Development Research Cluster, Indonesian Medical Education and Research Institute, Faculty of Medicine, Universitas Indonesia, Salemba Raya 6, Jakarta 10430, Indonesia*
[4]*Bioinformatics Core Facility, Indonesian Medical Education and Research Institute, Faculty of Medicine, Universitas Indonesia, Salemba Raya 6, Jakarta 10430, Indonesia*

ARTICLE INFO

Article history:
Received 23 September 2024
Revised 09 November 2024
Accepted 15 November 2024
Published online 01 January 2025

ABSTRACT

Insulin resistance is a critical factor in developing metabolic disorders like type 2 diabetes, posing challenges for effective treatment. Identifying molecular targets to reverse or mitigate insulin resistance is a key focus in therapeutic research. Advances in genomics and bioinformatics have enabled researchers to explore differentially expressed genes (DEGs) as potential biomarkers and therapeutic targets. This study aims to identify potential therapeutic targets for overcoming insulin resistance based on the analysis of (DEGs). Gallic acid (GA) and its derivatives were then tested against these identified targets using in silico methods. DEGs were analyzed from two Gene Expression Omnibus (GEO) datasets: GSE13070 (human adipose tissue with insulin resistance and insulin sensitivity) and GSE24422 (TNF-induced and non-induced adipocyte cell culture). The identified DEGs were then compared to find common DEGs, which were subsequently analyzed to identify hub-genes. Cross-validation using neural network and principal component analysis (PCA) on gene expression values revealed that the identified hub-genes, including IRS1, PCK1, GYS1, PTRPF, ACACB, and PIK3R2, can serve as biomarkers for insulin resistance (area under the curve, AUC 0.956 and sensitivity 1.00). The search for upstream regulatory proteins (URPs) of the hub-genes in the Comparative Toxicogenomics Database (CTD) indicated that the activities of TNF, PPARA, and AHR could influence the expression of several hub-genes, namely IRS1, PCK1, and ACACB. The activity prediction analysis, which was based on SkelSpheres molecular descriptors and confirmed by molecular docking, suggests that caffeoyl gallic acid may be a candidate compound for overcoming insulin resistance by inhibiting TNFA and activating PPARA.

***Keywords*:** Differentially Expressed Genes, Hub-genes, Upstream regulator, Gallic acid.

Introduction

Insulin resistance is a condition in which the body's cells become less sensitive to the effects of insulin, a hormone responsible for regulating blood sugar. It is a hallmark of type 2 diabetes (T2D) and metabolic syndrome. This condition can be influenced by various factors, including genetics, environment, and lifestyle.[1,2] Recent research indicates that the expression of certain genes plays a crucial role in the development of insulin resistance. Under normal conditions, insulin binds to insulin receptors on the cell surface, activating a series of signals that allow glucose to enter cells for energy use. In insulin resistance, this signaling is impaired, causing glucose to remain in the bloodstream and ultimately leading to elevated blood sugar levels.[3] Gene expression plays a significant role in insulin resistance.

*Corresponding author. E mail: aryo.tedjo@ui.ac.id
Tel: +62-8998452818

Citation: Suryandari DA, Tedjo A, Fadilah F. Differentially Expressed Genes (DEGs) Analysis and In Silico Studies Identify Tumor Necrosis Factor (TNF) Inhibition and Peroxisome Proliferator-Activated Receptor Alpha (PPARA) Activation as Targets for Gallic Acid Derivatives in Insulin Resistance. Trop J Nat Prod Res. 2024; 8(12): 9476 – 9485 https://doi.org/10.26538/tjnpr/v8i12.19

Official Journal of Natural Product Research Group, Faculty of Pharmacy, University of Benin, Benin City, Nigeria

The Insulin Receptor (IR) gene encodes the insulin receptor located on the cell surface, and mutations or decreased expression of this gene can reduce cellular sensitivity to insulin, resulting in insulin resistance.[4] IR substrate (IRS-1 and IRS-2), the two major isoforms of the IR substrate, help mediate insulin signaling after insulin binds to its receptor. Reduced expression or mutations in IRS genes can disrupt the insulin signaling pathway, contributing to insulin resistance.[5] Peroxisome Proliferator-Activated Receptor Gamma (PPARγ), a transcription factor that regulates the expression of genes related to lipid and glucose metabolism, enhances insulin sensitivity by promoting fatty acid storage in adipose tissue and reducing lipotoxicity. Genetic variations that reduce PPARγ activity are associated with insulin resistance.[6] AKT serine/threonine kinase 2 (AKT2) is a protein kinase involved in the insulin signaling pathway, and mutations or decreased expression of AKT2. It can reduce the phosphorylation and activity of downstream proteins necessary for glucose uptake by cells, contributing to insulin resistance.[7] Glucose Transporter Type 4 (GLUT4), an insulin-induced glucose transporter primarily found in muscle and adipose tissues, also plays a crucial role in regulating glucose uptake. Reduced expression or function of GLUT4 leads to decreased glucose uptake by these cells, which is a key characteristic of insulin resistance.[8] These findings collectively highlight the complexity of gene expression involved in insulin resistance and emphasize the need for targeted gene expression-based therapeutic strategies. Based on these facts, research on gene expression markers or Differentially Expressed Genes (DEG) analysis is essential to determine effective therapeutic targets for insulin resistance. Gallic acid (GA) is a phenolic compound found in various

plants, including tea, grapes, and nuts. GA and its derivatives, such as methyl gallate, propyl gallate, and epigallocatechin gallate (EGCG), exhibit various biological activities, including antioxidant, antimicrobial, anti-inflammatory, and antidiabetic effects.[9–12] GA is known to ameliorate insulin resistance, partly through the modulation of gene expression involved in insulin signaling and glucose metabolism. GA activates AMP-activated protein kinase (AMPK), which in turn stimulates peroxisome proliferator-activated receptor gamma coactivator 1-alpha (PGC1α), enhancing mitochondrial function and energy expenditure.[13] GA is also known to enhance adipocyte differentiation and adiponectin expression, a hormone that increases insulin sensitivity, by upregulating Fatty Acid-Binding Protein 4 (FABP4).[14] Protocatechuic acid (PCA), a polyphenol similar to GA, significantly increases levels of GLUT-4, IRS-1, IRS-2, PPAR-γ, P-AMPK, and P-Akt in C2C12 myotubes, HepG2 cells, and 3T3-L1 adipocytes, thereby improving glucose uptake and insulin signaling.[15] Additionally, GA has been observed to ameliorate symptoms of polycystic ovary syndrome (PCOS) by reducing testosterone, LH, and inflammatory cytokines, while increasing estrogen levels and the expression of mRNA for Cyp11a1, Cyp19A1, KITL, PTGS2, and Adipo R1, which are associated with insulin sensitivity and metabolic balance.[16] Therefore, GA and its derivatives are predicted to provide therapeutic effects for T2D by modulating the expression of genes associated with insulin resistance.

The objective of this research is to analyze differentially expressed genes (DEGs) associated with insulin resistance and identify key gene targets that could be modulated to restore insulin sensitivity. This study aims to address the gap in knowledge regarding specific gene expression markers for insulin resistance, offering a novel insight into potential therapeutic targets. In this study, DEG analysis was performed on human genome mRNA microarray data from adipose tissues of human subjects with insulin resistance and insulin sensitivity (GSE13070) available in the Gene Expression Omnibus (GEO) database. The identified DEGs were then compared with DEGs obtained from gene expression data of adipocytes derived from Human Mesenchymal Stem Cells (hMSC) treated with tumor necrosis factor (TNF) to induce insulin resistance (GSE24422). This comparison aimed to find common DEGs between the two datasets so that gene expression targets affected by new therapeutic agents at the *in vitro* level could be extrapolated to tissue levels in human subjects. Based on the common DEGs found, hub-genes were determined as targets to be influenced by the activity of GA and its derivatives through in silico methods.

Materials and Methods

In general, this study was divided into three stages. The first stage involved identifying common DEGs from two GEO datasets: GSE13070, which was human genome mRNA microarray data from adipose tissue of subjects with insulin resistance and insulin sensitive, and GSE24422, which was gene expression profiling data from mRNA microarrays of adipocyte cell cultures derived from Human Mesenchymal Stem Cells (hMSC) treated with TNF. The second stage involved analyzing hub-genes from the common DEGs to identify target proteins whose activities would be affected. The third stage involved predicting the protein activity values (IC_{50} or EC_{50}) using molecular descriptors and machine learning methods, and confirming the results with molecular docking.

Identification of Common DEGs

The DEG analysis between the two GEO datasets was performed using GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r/).[17] Each dataset was analyzed separately, with each sample group arranged according to its experimental conditions. In GEO2R, p-value correction was performed using the Benjamini and Hochberg method [18] with adjusted P value < 0.05.The results of the DEG analysis from both datasets were then used to find common DEGs with the same regulation using Orange v3.37.0 (University of Ljubljana, 2024) in the 'Merge Data' widget.[19] Subsequently, the involvement of the common DEGs in the Kyoto Encyclopedia of Genes and Genomes (KEGG) signaling pathways related to insulin resistance (hsa04931) was analyzed using Enrichr.[20,21]

Hub-Genes Analysis and Validation

From the DEGs mapped to each signaling pathway, hub-genes were determined using Cytoscape v3.10.2 (National Human Genome Research Institute, 2024).[22] Hub-genes are genes that have many interactions with other genes.[23] The raw expression data of the resulting hub-genes were then retrieved from the 'Series Matrix File(s)' in each GSE. This raw data was then analyzed using Orange v3.37.0 for cross-validation with Principal Component Analysis (PCA) and machine learning methods. Cross-validated hub-genes were then analyzed for their upstream regulatory proteins (URPs) using the Comparative Toxicogenomics Database (CTD).[24] Upstream regulators are any molecules (including proteins) that can influence the expression of other molecules (e.g., genes/mRNA).

Prediction of Activity of GA and its Derivatives on Hub-gene-related Proteins and In Silico Interaction Confirmation

Various active compounds related to URPs can be obtained from the ChEMBL database using DataWarrior v6.1.0 (openmolecules, 2024).[25] ChEMBL is a manually curated database containing bioactive chemical entities with drug-like properties.[26] In DataWarrior, active compound data can include IC_{50} or EC_{50} values tested on URPs related to common DEGs or hub-genes. To predict the activity values of GA against these URPs, a predictive model must first be created using the activity data of compounds against URPs. Each active and inactive compound against URPs from ChEMBL is clustered into one several groups based on structural similarity using SkelSpheres descriptors. SkelSpheres-based similarity essentially measures how many identical circular fragments or atom groups (hybridization) exist between two molecules compared to the total number of circular fragments in both molecules.[26] In this study, a SkelSpheres-based similarity of 90% was set. A representative compound from each group was then selected for use as training data. The remaining compounds were used as testing data, whose activity was predicted using support vector regression (SVR) in DataWarrior. The correlation coefficient (*R*-square, R2) value of the actual vs. predicted $\log IC_{50}$ or $\log EC_{50}$ plot was then determined, and the model was considered valid if $R2 > 0.5$.[27] Once the model is valid, the activity prediction of GA derivatives against URPs is performed in the same manner as the activity prediction on testing data. The structures of GA and its derivatives were obtained from the PUBCHEM database.[28] GA derivatives predicted to have the best activity can then have their interactions with URPs confirmed using molecular docking. The crystal structures of URPs, retrieved from the RCSB PDB database, are imported into Molegro Virtual Docker (MVD v.7.0.0, Molexus Aps, 2019, free trial). For molecular docking purposes, all water molecules were removed, and corrections are made to any mismatched amino acid residues. Ligand preparation was carried out in Molegro Virtual Docker (MVD) by minimizing the ligand using the "Ligand Energy Inspector" tool to find the most stable conformation or the one with the lowest energy. The validity of the model and molecular docking parameters is determined with a Root Mean Square Deviation (RMSD) value < 2Å.[29] Subsequently, molecular docking was performed on the 3D structure of the ligand (GA derivatives) against the crystal structure of the protein (upstream regulators protein, URP) repeated 20 times using the MVD. In this docking process, parameters such as energy values, including MolDock Score, Rerank Score, and Hbond, were measured. To evaluate the binding strength between the ligand and the receptor protein, the Rerank Score is often used as a common parameter.[30]

Statistical Analysis

The p-values for the DEG analysis were derived using the GEO2R tool accessible via the GEO platform. Metrics such as the Area Under the Curve (AUC), Classification Accuracy (CA), Precision (Prec), Sensitivity (Recall), and Specificity were calculated using Orange v3.37.0.

Results and Discussion

Identification of Common DEGs

The results of the Differentially Expressed Genes (DEG) analysis using GEO2R from the two datasets (GSE13070 and GSE24422) can be visualized with a volcano plot, as shown in Figure 1. A volcano plot is

ISSN 2616-0684 (Print)
ISSN 2616-0692 (Electronic)

a visualization tool used to display the statistical significance and the magnitude of effect (fold change, FC) of each gene. In Figure 1(a), it can be observed that the DEG analysis effectively distinguishes adipose tissue between subjects with insulin resistance and insulin sensitive (GSE13070), although the number of DEGs is smaller compared to non-DEGs (adj.P.Val < 0.05). In contrast, the results of the DEG analysis for the hMSC with TNF vs. hMSC (GSE24422) shown in Figure 1(b) reveal a larger number of DEGs compared to non-DEGs (adj.P.Val < 0.05).

A total of 2,831 DEGs from the GSE13070 dataset and 281 DEGs from the GSE24422 dataset were then analyzed to identify common DEGs. Common DEGs are those with the same regulation across two or more datasets and are often used to identify shared cellular signaling pathways occurring between the datasets.[31] The results of the common DEG search are presented in Table 1. Table 1 shows that there are 88 common DEGs with consistent regulation between the two datasets. It can be observed that most of the common DEGs are upregulated (log2FC > 0). Assuming no post-transcriptional modifications occur, the protein expression corresponding to these common DEGs is also likely to be overexpressed in the insulin-resistant condition. The mapping results of the 88 common DEGs to the KEGG pathway related to insulin resistance (hsa04931) are shown in Table 2.

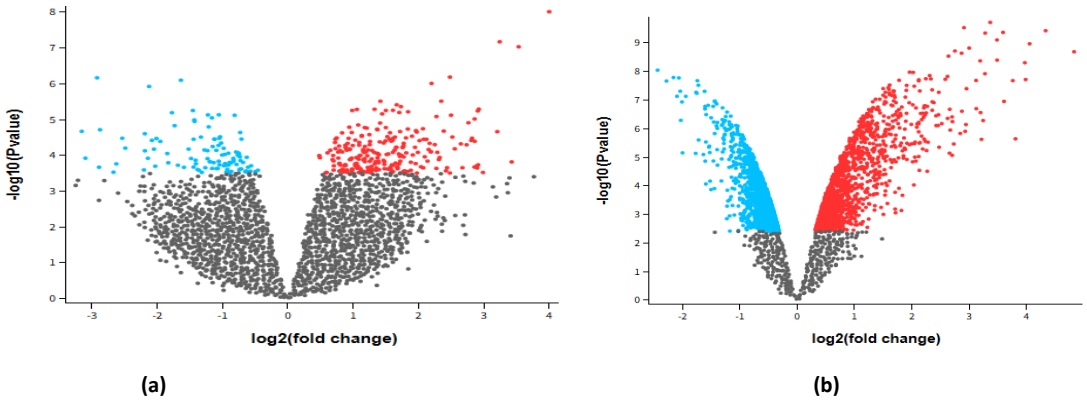


Figure 1. Volcano plot of GSE13070 (a); and volcano plot of GSE24422 (b)

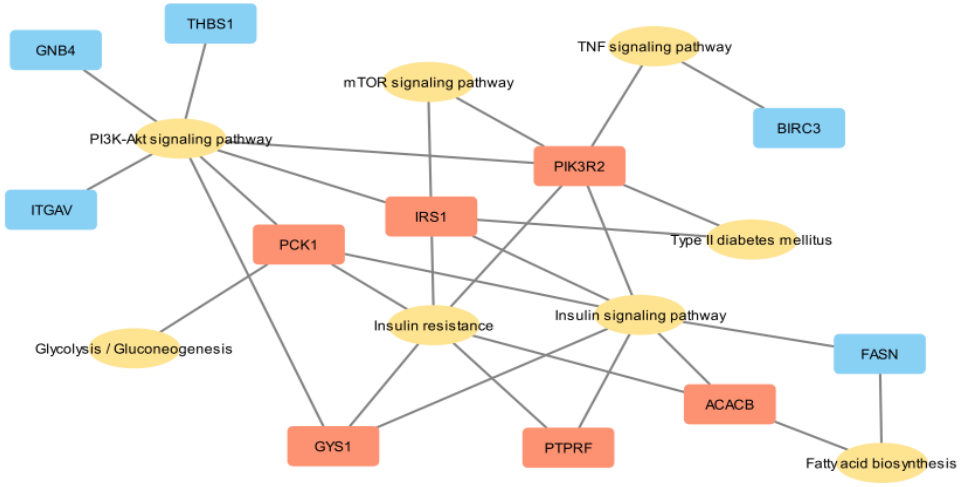


Figure 2. The interaction network of common DEGs in the KEGG pathway related to insulin resistance (hsa04931). Nodes in red represent common DEGs that connect two or more signaling pathways (hub-genes).

Hub-Genes Analysis and Validation

If the genes listed in Table 2 are depicted within the common DEG interaction network, hub-genes are identified as shown in Figure 2. Figure 2 reveals that insulin receptor substrate 1 (IRS1), phosphatidylinositol 3-kinase regulatory subunit beta (PIK3R2), phosphoenolpyruvate carboxykinase (PCK1), glycogen [starch] synthase (GYS1), protein tyrosine phosphatase-receptor type F (PTRPF), and acetyl-CoA carboxylase 2 (ACACB) are hub-genes. These can be further explored as potential markers and therapeutic targets for insulin resistance. Based on Table 1, it is observed that, under insulin-resistant conditions, IRS1, PCK1, GYS1, PTRPF, and ACACB are underexpressed, while PIK3R2 is overexpressed. To assess whether these hub-genes could serve as markers for insulin resistance, the PCA analysis, confusion matrix, and cross-validation results are presented in Figure 3, Table 3, and Table 4. Table 4 shows the cross-validation results on the GSE13070 dataset using neural network model. The area under the curve (AUC) is 0.956, the classification accuracy (CA) is 0.905, the sensitivity (Recall) is 1.000, and the specificity is 0.667. These values indicate that the six common DEGs can accurately predict the insulin-resistant condition 100% of the time (sensitivity = 1.000), but they can only predict 66.7% of the subjects with insulin sensitive (specificity = 0.667).

Figure 4 displays the results for URP searching of the five hub-genes, excluding PIK3R2, in the Comparative Toxicogenomics Database (CTD). As shown in Figure 4, several URPs can influence the expression of multiple genes simultaneously (marked by red and green

ISSN 2616-0684 (Print)
ISSN 2616-0692 (Electronic)

Table 1. Regulation of Common DEGs between GSE24422 and GSE13070

Symbol	log2(FC) GSE24422	log2(FC) GSE13070	Symbol	log2(FC) GSE24422	log2(FC) GSE13070	Symbol	log2(FC) GSE24422	log2(FC) GSE13070
AZGP1	-0.825	-2.921	PTPRF*	-0.615	-0.685	NEXN	0.836	1.034
CASQ2	-1.265	-2.535	TJP2	-0.487	-0.562	SLC15A3	0.846	1.281
GPAT3	-2.433	-2.188	LRPPRC	-0.636	-0.496	SLCO2B1	0.896	2.859
AZGP1	-0.715	-2.124	ADAM9	0.341	0.886	NRP2	0.922	1.151
FASN	-0.433	-1.683	MAP1B	0.341	0.91	CCL13	0.937	1.572
PHGDH	-0.553	-1.427	SAMHD1	0.363	1.181	LTBP2	0.952	1.539
YME1L1	-1.277	-1.407	RNASET2	0.399	1.109	ITGAV	0.969	0.794
LRIG1	-1.013	-1.33	MSC-AS1	0.408	0.931	OSBPL3	1.004	0.987
ACACB*	-0.37	-1.263	BMP2K	0.417	1.291	TYMS	1.081	1.256
PXMP2	-0.554	-1.223	CD28	0.421	1.397	BIRC3	1.092	0.944
CLMN	-0.823	-1.211	GNB4	0.436	0.609	P2RX7	1.115	1.725
PCK1*	-0.644	-1.194	SLC18B1	0.436	0.749	PIK3R2*	1.131	2.422
CDKN2C	-0.948	-1.19	SEC24D	0.437	0.651	LY96	1.14	1.181
SLC16A7	-1.538	-1.094	RTN2	0.447	1.049	HLA-DRA	1.188	1.208
IRS1*	-0.575	-1.073	PALLD	0.492	1.249	CYBA	1.223	1.112
STBD1	-0.685	-1.049	SHTN1	0.548	1.093	TFPI	1.247	0.796
PPP1R16A	-0.573	-1.037	UCHL1	0.552	1.611	PRSS23	1.263	0.88
ADAMTS9	-0.503	-1.021	SLAMF8	0.564	2.861	KYNU	1.263	1.884
ACADM	-0.66	-0.949	CALU	0.566	0.855	ALDH1A3	1.331	1.719
CYB5A	-0.37	-0.941	HLA-DQB1	0.567	0.95	SRPX2	1.341	1.81
TM7SF2	-0.989	-0.926	IL1RN	0.622	3.428	PTGFR	1.434	1.611
GYS1*	-0.529	-0.892	CCL19	0.632	1.68	TIMP1	1.525	1.177
ALDH6A1	-0.522	-0.871	DAB2	0.644	0.903	C1S	1.698	0.966
CS	-0.435	-0.83	VMP1	0.657	2.082	CYP1B1	1.838	1.132
NPR1	-0.561	-0.767	GLIPR1	0.673	1.565	EFEMP1	1.888	1.043
CENPV	-0.427	-0.754	FAM20A	0.773	2.483	CTSS	2.122	1.583
TXLNG	-0.421	-0.741	ITIH5	0.803	1.209	MYOF	2.239	0.988
SMURF1	-0.565	-0.735	THBS1	0.812	2.891	FHL2	2.246	0.835

* The hub-gene symbols obtained from the hub-gene analysis are shown in Figure 2.

Table 2. Mapping of Common DEGs to the KEGG Pathway Related to Insulin Resistance

Term	Genes
Insulin signaling pathway	GYS1, IRS1, FASN, PIK3R2, PCK1, ACACB,PTPRF
Insulin resistance	GYS1, IRS1, PIK3R2, PCK1, ACACB, PTPRF
PI3K-Akt signaling pathway	GYS1, IRS1, GNB4, ITGAV, PIK3R2, PCK1, THBS1
Fatty acid biosynthesis	FASN, ACACB
Type II diabetes mellitus (T2D)	IRS1, PIK3R2
TNF signaling pathway	PIK3R2, BIRC3
mTOR signaling pathway	IRS1, PIK3R2
Glycolysis / Gluconeogenesis	PCK1

nodes), including peroxisome proliferator-activated receptor alpha (PPARA), aryl hydrocarbon receptor (AHR), tumor necrosis factor (TNF), insulin (INS), and noggin (NOG). Since all five common differentially expressed genes (IRS1, PCK1, GYS1, PTRPF, and ACACB) are similarly regulated (downregulated) under insulin resistance conditions, only URPs that exert an upregulatory effect on the common DEGs will be further analyzed, specifically TNF, AHR, and PPARA. The URP targets on the common DEGs (hub-genes) are

ISSN 2616-0684 (Print)
ISSN 2616-0692 (Electronic)

shown in Table 5. Among the three URPs, based on the availability of data in ChEMBL for predictive model construction and validation, PPARA and TNF were selected as the URP targets whose activity changes will be predicted using GA and its derivatives.

Table 3. Confusion Matrix of Insulin Resistance vs Sensitive Clustering Using Neural Networks Based on Hub-Genes Expression

		Predicted		
		insulin resistant	insulin sensitive	∑
Actual	insulin resistant	15	0	**15**
	insulin sensitive	2	4	**6**
	∑	**17**	**4**	**21**

Table 4. Cross Validation of Data Clustering (Insulin Resistance vs Sensitive) Using Neural Networks Based on Hub-Genes Expression

Model	AUC	CA	Sensitivity	Specificity
Neural Network	0.956	0.905	1.000	0.667

Prediction of Activity of GA and its Derivatives on Hub-gene-related Proteins and In Silico Interaction Confirmation

Based on the information in Table 5, GA and its derivatives will be predicted in silico for their activity as TNF inhibitors and PPARA activators (agonists). Prior to this, the IC_{50} prediction model for TNFA and the EC_{50} prediction model for PPARA were validated using compound data obtained from ChEMBL. The results of the model validation are presented in Figure 5. It is shown that the R-square (R2) values obtained for both predictive models are greater than 0.5. With R2 > 0.5, the $LogIC_{50}$ prediction model based on SkelSpheres descriptors using the SVR model is considered valid.[24] Therefore, this predictive model can be used to predict the activity of gallic acid (GA) and its derivatives as TNFA inhibitors and PPARA activators (agonists). A total of 98 GA compounds and their derivatives from PUBCHEM were predicted for their IC_{50} against TNFA and EC_{50} against PPARA. Table 6 shows that the 10 GA derivative compounds with the lowest IC_{50} and EC_{50} values predominantly fall within the 'Good activity' criteria (EC_{50} or IC_{50} < 1-20 µM).[32] It can be concluded that these compounds have the potential to influence the expression of hub-genes (ACACB, IRS1, and PCK1) involved in insulin resistance. The interactions of five active compounds from Table 6 with TNFA and PPARA were then confirmed using molecular docking. The docking method validation and the GA-based docking process can be found in Table 7. As shown in Table 7, the redocking of the native ligand (UTJ_201 [A]) as a TNFA inhibitor (PDB ID: 6X81 [A]) and AZ2_469 [A] as a PPARA activator (PDB ID: 1I7G_[A]) yielded RMSD values of < 2Å. Therefore, the docking method is considered valid. Subsequently, the molecular docking method can be used to predict the interaction between compounds/ligands listed in Table 6 with TNFA protein (PDB ID: 6X81 [A]) and PPARA (PDB ID: 1I7G_[A]). The results of this molecular docking process can be seen in Table 8.

As shown in Table 8, based on docking scores and rerank scores, Gallic acid 5,6-dihydroxy-3-carboxyphenyl ester (CID_44592636), hamamelofuranose-2'5'-digallate (CID_44584241), and Caffeoyl gallic acid (CID_140567676) are confirmed to have a higher affinity for the TNFA protein than the native ligand (UTJ_201).

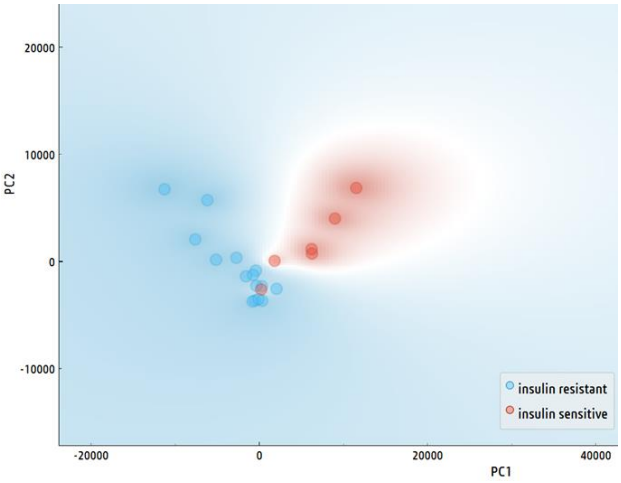


Figure 3. Principal Component 1 (PC1) and PC2 constructed from the expression values of hub- genes (IRS1, PIK3R2, PCK1, GYS1, PTRPF, and ACACB).

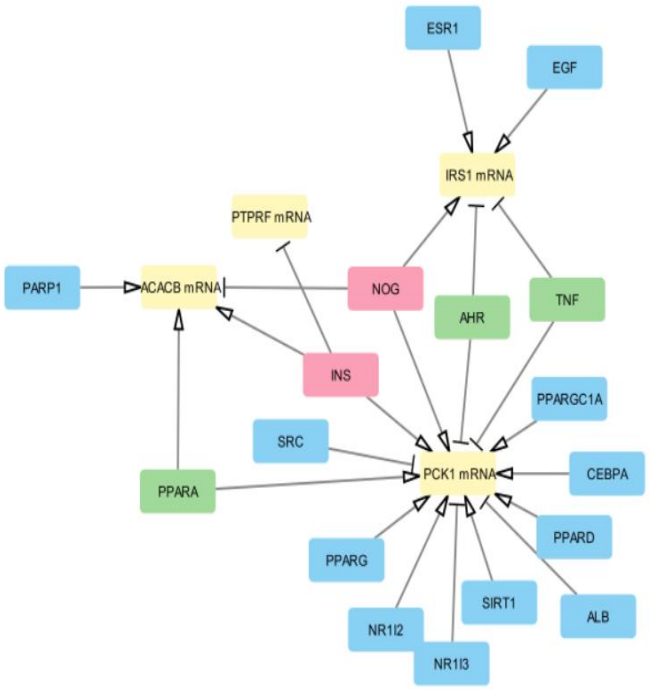
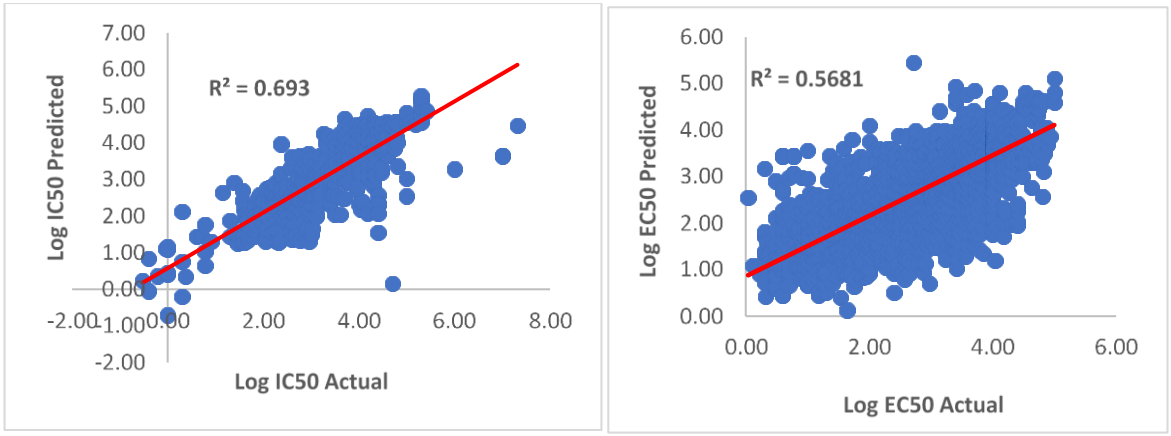


Figure 4. The URPs are represented by green, red, and blue nodes. Yellow nodes indicate shared DEG that are the targets of regulatory changes (→ increase expression; —| inhibit/decrease expression)

ISSN 2616-0684 (Print)
ISSN 2616-0692 (Electronic)

IC_{50}Table 5. Upstream Regulatory Protein (URPs) of DEG and Their Target Activities

URP	Genes expression target (upregulation)	URP activity target	References (PMID)
PPARA	ACACB, PCK1	Activator (agonist)	19124612; 16197558
AHR	PCK1, IRS1	inhibitor	20181658; 34848246
TNF	PCK1, IRS1	inhibitor	17327424; 12732648



(a) (b)

Figure 5. Validation of the predictive model for compounds in the ChEMBL database as TNFA inhibitors (a) and PPARA activators (b)

Table 6. Predicted activity of GA and its derivatives against TNF-alpha and PPARA proteins

Activity Rangking (TNFA)	Compound name	IC_{50} Predicted (μM)	Criteria [32]
1	Caffeoyl gallic acid (CID_140567676)*	2.68	Good activity
2	Methylenedigallic acid (CID_68372)	2.80	Good activity
3	Gallic acid 5,6-dihydroxy-3-carboxyphenylester (CID_44592636)	3.03	Good activity
4	2-C-((Galloyloxy)methyl)-D-ribose 5-gallate (CID_21145076)	3.20	Good activity
5	Hamamelofuranose 2'5'-digallate (CID_44584241)	3.25	Good activity
90	Gallic Acid (CID_370)	45.85	Moderate activity
Activity Rangking (PPARA)	**Compound name***	**EC_{50} Predicted (μM)**	**Criteria [32]**
1	Methylenedigallic acid (CID_68372)	0.92	Excellent activity
2	3,4,5-Tris(acetyloxy)benzoic acid (CID_95088)	1.22	Good activity
3	Rhodanine-gallic acid (CID_129848640)	1.22	Good activity
4	Maloyl gallic acid (CID_155490657)	1.30	Good activity
5	Caffeoyl gallic acid (CID_140567676)*	1.41	Good activity
80	Gallic Acid (CID_370)	7.18	Good activity

* Compounds with confirmed good activity criteria have a docking score lower than or close to the native ligand docking score of the TNFA and PPARA proteins

Meanwhile, as PPARA activators, none of the GA derivatives in Table 8 have docking scores or rerank scores better than the native ligand (AZ2_469), which is a PPARA agonist. However, all five compounds have negative docking and rerank scores, indicating that their interaction with PPARA may occur spontaneously, although not as favorably as AZ2_469. Caffeoyl gallic acid (CID_140567676) is the GA derivative that most closely approaches the docking score and rerank score of AZ2_469. This compound also has better docking and rerank scores compared to the native TNFA ligand (PDB ID: 6X81 [A]). These results were further supported by examining the interactions

ISSN 2616-0684 (Print)
ISSN 2616-0692 (Electronic)

between Caffeoyl gallic acid and TNFA and PPARA. As shown in Figure 6, the interaction of Caffeoyl gallic acid with TNFA and PPARA occurs through hydrogen bonding and electrostatic interactions, indicating a strong affinity of Caffeoyl gallic acid for both proteins.

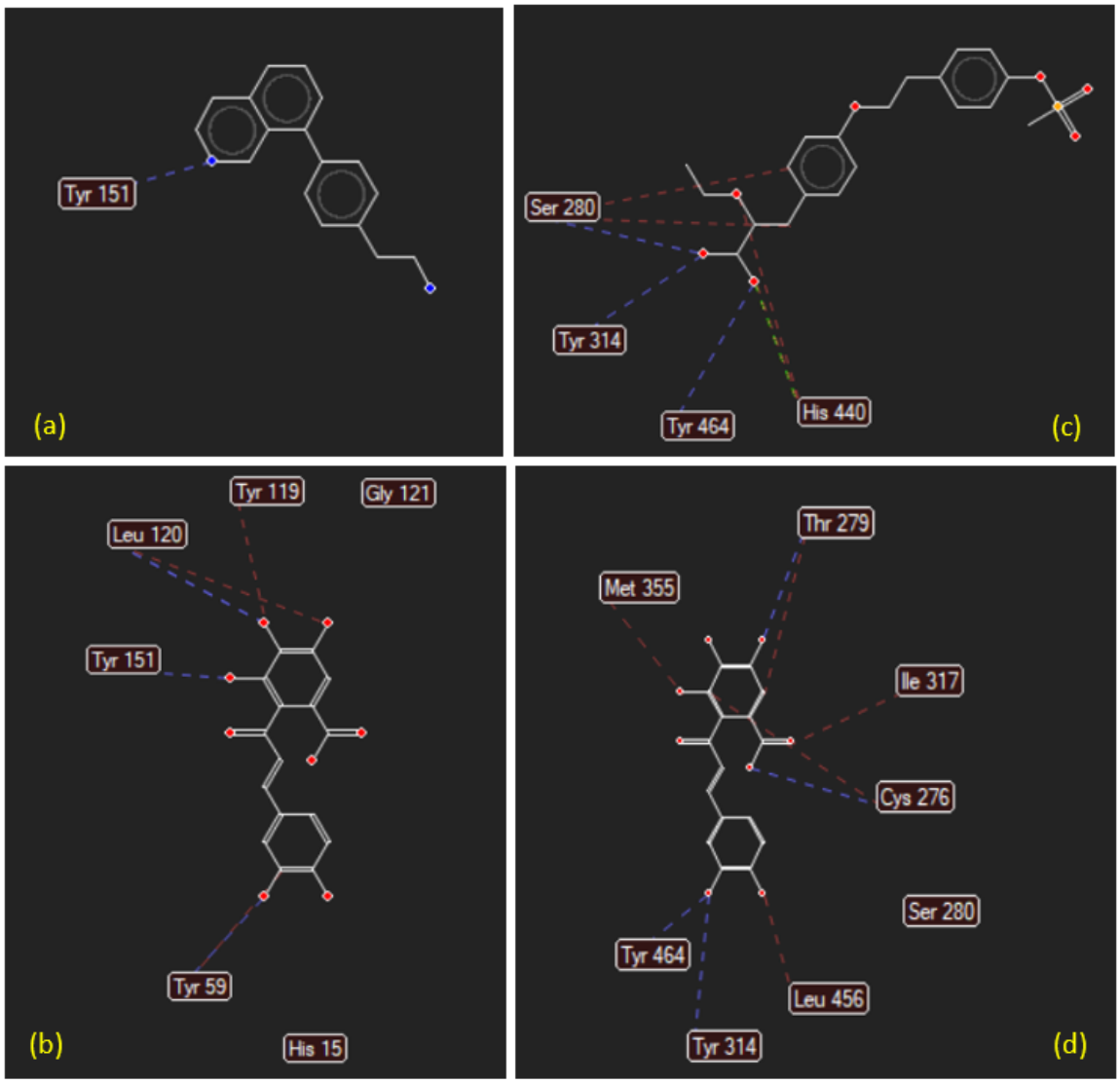


Figure 6. Amino acid residue interactions of TNFA with the native ligand UTJ_201 (a) and Caffeoyl gallic acid (c); Amino acid residue interactions of PPARA with the native ligand AZ2_469 (c) and Caffeoyl gallic acid (d). Blue dashed lines represent hydrogen bonds, and red dashed lines represent steric-electrostatic interactions

Table 7. Validation of the Molecular Docking Method

Pose Ligand	Protein	MolDock Score	Rerank Score	RMSD	HBond
[05]UTJ_201[A]	TNFA (PDB ID: 6X81 [A])	-56.737	-43.254	1.548	0.000
[02]AZ2_469 [A]	PPARA (PDB ID: 1I7G [A])	-133.106	-110.911	1.069	-6.831

With its dual potential as a TNFA inhibitor and PPARA activator, Caffeoyl gallic acid (CID_140567676) is a promising GA derivative that could be explored as a drug candidate to combat insulin resistance. Hub-genes refer to genes that play a central role in regulation and interaction within gene networks. They are often highly connected with other genes and significantly influence various biological processes and diseases. For instance, in Adamantinomatous craniopharyngioma (ACP), hub-genes have been specifically identified as important markers for early diagnosis and as new therapeutic targets.[33] In non-small cell lung cancer (NSCLC), eight hub-genes associated with the cell cycle pathway have been identified as significant markers for poor prognosis in lung adenocarcinoma.[34] In this study, six DEG (IRS1, PIK3R2, PCK1, GYS1, PTRPF, and ACACB) were identified as hub-genes and have the potential to be used as therapeutic targets.

Assuming no post-transcriptional modifications occur, only PIK3R2 from the six hub-genes would have higher protein levels in the adipose tissue of subjects with insulin resistance compared to those with insulin sensitive. Phosphoinositide-3-Kinase Regulatory Subunit 2 (PIK3R2) is a regulator of the enzyme Phosphoinositide 3-kinase (PI3K), which is known to play a crucial role in regulating glucose uptake, suggesting its potential involvement in the development of diabetes mellitus. Metabolic signaling is largely mediated by the recruitment of dimeric

(p85-p110) PI3K to IRS, stimulating PKC activation through PDK1, and enhancing glucose uptake.[35] The p85β protein, the regulatory subunit of PI3K encoded by PIK3R2, plays a complex role in modulating PI3K activity and signaling. Under normal conditions, p85 stabilizes the catalytic subunit p110 (encoded by the isoform genes PIK3CA, PIK3CB, PIK3CD, and PIK3CG) in class IA PI3K and induces its activity.[36] Overexpression of p85β can disrupt this balance. Excess free p85β protein competes with the p85-p110 heterodimer to bind phosphotyrosine residues on insulin receptor substrates (e.g., IRS1), thereby reducing the stability of the interaction.[37] However, the regulatory mechanisms of PIK3R2 expression remain largely unknown.[38] This makes PIK3R2 unsuitable as a therapeutic target for insulin resistance in this study.

Table 8. Docking Score and Rerank Score of the Native Ligand and GA Derivatives against TNFA and PPARA

TNFA (PDB ID: 6X81 [A])			
Pose Ligand	**MolDock Score**	**Rerank Score**	**HBond**
[00]Gallic acid 5,6-dihydroxy-3-carboxyphenylester (CID_44592636)	-77.031	-67.582	-7.300
[01]Hamamelofuranose 2'5'-digallate (CID_44584241)	-94.557	-67.110	-6.668
[01]Caffeoyl gallic acid (CID_140567676)	-67.281	-59.612	-7.864
[01UTJ_201 [A]	-68.348	-58.145	-2.500
[0]2-C-((Galloyloxy)methyl)-D-ribose 5-gallate (CID_21145076)	-80.008	-58.118	-6.635
[03] Methylenedigallic acid (CID_68372)	-60.960	-55.215	-12.089
PPARA (PDB ID: 1I7G [A])			
Pose Ligand	**MolDock Score**	**Rerank Score**	**HBond**
[00]AZ2_469 [A]	-142.551	-118.758	-5.971
[00]Caffeoyl gallic acid (CID_140567676)	-123.253	-104.969	-7.860
[09]Methylenedigallic acid (CID_68372)	-88.937	-81.796	-13.270
[00]Rhodanine-gallic acid (CID_129848640)	-96.035	-78.597	-5.8101
[03]3,4,5-Tris(acetyloxy)benzoic acid (CID_95088)	-89.586	-75.182	-3.373
[09]Maloyl gallic acid (CID_155490657)	-76.132	-43.229	-4.8591

IRS1 (Insulin Receptor Substrate 1) is the primary mediator of insulin action, transmitting signals from the insulin receptor to downstream effectors such as AKT, where its dysregulation can reduce glucose uptake.[5] ACACB (Acetyl-CoA Carboxylase Beta) is involved in fatty acid metabolism, and its dysregulation may contribute to lipid accumulation and insulin resistance.[39] Meanwhile, the dysregulation of Phosphoenolpyruvate Carboxykinase 1 (PCK1) in adipose tissue is known to increase plasma free fatty acids, further enhancing insulin resistance.[40] Caffeoyl gallic acid has the potential to downregulate these three genes through inhibition of TNFA activity and activation (agonism) of PPARA. The interaction of Caffeoyl gallic acid with TNFA is even stronger compared to UTJ_201 or [4-(isoquinolin-8-yl)phenyl]acetonitrile, its native ligand. UTJ_201 is an isoquinoline alkaloid, and several isoquinoline alkaloids, such as fangchinoline and isotetrandrine, have been found to inhibit the production of proinflammatory cytokines, including TNFα.[41]

Conclusion

This study identified IRS1, PCK1, GYS1, PTRPF, ACACB, and PIK3R2 as hub-genes and potential biomarkers for insulin resistance with high accuracy and sensitivity. The regulatory proteins TNF, PPARA, and AHR were found to influence these hub-genes, and activity prediction suggests that caffeoyl gallic acid may address insulin resistance by inhibiting TNFA and activating PPARA. Further research should focus on experimental validation through *in vitro* and *in vivo* testing to confirm the effectiveness of Caffeoyl gallic acid in combating insulin resistance.

Conflicts of interest

The authors have declared that there is no conflict of interest

Authors' Declaration

The authors hereby declare that the work presented in this article is original and that any liability for claims relating to the content of this article will be borne by them.

Acknowledgement
Authors appreciate the technical and theoretical assistance from the Bioinformatics Core Facilities, Indonesia Medical Education and Research Institute, Faculty of Medicine, Universitas Indonesia.

References

1. Szablewski L. Changes in Cells Associated with Insulin Resistance. Int J Mol Sci. 2024; 25:2397-2424. https://doi.org/10.3390/ijms25042397.
2. Chaoji SA. Insulin Resistance. Vidarbha J Intern Med. 2023; 33:27-31. https://doi.org/10.25259/VJIM_41_2022.
3. Singh A, Ghai N, Mohinder Singh Bedi P. Molecular Mechanisms Involved in Insulin Resistance: Recent Updates and Future Challenges. Evolving Concepts in Insulin Resistance, IntechOpen; 2022. https://doi.org/10.5772/intechopen.104806.
4. Calcaterra V, Zuccotti G, Mari A, Iafusco F, Maione G, Iafusco D, Tinto N. Can obesity exacerbate hyperinsulinaemia in the presence of the mutation of an insulin receptor gene? Clin Obes. 2023; 13:e12619-e12627. https://doi.org/10.1111/cob.12619.
5. Bhattacharjee A, Pranto SMAM, Ahammad I, Chowdhury ZM, Juliana FM, Das KC, Keya CA, Salimullah M. High risk genetic variants of human insulin receptor substrate 1(IRS1) infer structural instability and functional interference. J Biomol Struct Dyn. 2023; 41:15150–64. https://doi.org/10.1080/07391102.2023.2187232.
6. Song Y, Li S, He C. PPARγ Gene Polymorphisms, Metabolic Disorders, and Coronary Artery Disease. Front Cardiovasc Med. 2022; 9:808929-808941. https://doi.org/10.3389/fcvm.2022.808929.
7. Elangeeb ME, Elfaki I, Elkhalifa MA, Adam KM, Alameen AO, Elfadl AK, Albalawi IA, Almasoudi KS, Almotairi R, Alsaedi BSO, Alhelali MH, Mir MM, Amle D, Mir R. In Silico Investigation of AKT2 Gene and Protein Abnormalities Reveals Potential Association with Insulin Resistance and Type 2 Diabetes. Curr Issues Mol Biol. 2023; 45:7449–7475. https://doi.org/10.3390/cimb45090471.
8. van Gerwen J, Shun-Shion AS, Fazakerley DJ. Insulin signalling and GLUT4 trafficking in insulin resistance. Biochem Soc Trans. 2023; 51:1057–1069. https://doi.org/10.1042/BST20221066.
9. Baheej MAA, Haniffa HM, Siddiqui H, Jabeen A. Synthesis and biological evaluation of gallic acid esters as phagocyte oxidative burst inhibitors. J Natl Sci Found. 2023; 51:415–421. https://doi.org/10.4038/jnsfsr.v51i3.11199.
10. Saini A, Saini VD, Saharan R, Kumar S. Gallic Acid: A Wonderful Remedy in Medicinal Field. Curr Tradit Med. 2023; 10:1-17 https://doi.org/10.2174/0122150838266828231005055236.
11. Errabiti B, El Aabedy A, Er-Rahmani S, El Abed S, Koraichi SI. Antimicrobial Effects of Tannic and Gallic Acids: A Study on 3D-Printed Polylactic Acid Surfaces Against *P. aeruginosa* and *S. aureus*. Trop J Nat Prod Res. 2024; 8(10):8753 -8763. https://doi.org/10.26538/tjnpr/v8i10.18.
12. Vu D. Assessing the Differences in Phenolics, Antioxidant and Anti-Tyrosinase Activities of Spent Coffee Ground Fractions . Trop J Nat Prod Res. 2022; 6(4):558-562
13. Doan KV, Ko CM, Kinyua AW, Yang DJ, Choi YH, Oh IY, Nguyen NM, Ko A, Choi JW, Jeong Y, Jung MH, Cho WG, Xu S, Park KS, Park WJ, Choi SY, Kim HS, Moh SH, Kim KW. Gallic acid regulates body weight and glucose homeostasis through AMPK activation. Endocrinology. 2015; 156:157–168. https://doi.org/10.1210/en.2014-1354.
14. Makihara H, Koike Y, Ohta M, Horiguchi-Babamoto E, Tsubata M, Kinoshita K, Akase T, Goshima Y, Aburada M, Shimada T. Gallic Acid, the Active Ingredient of *Terminalia bellirica*, Enhances Adipocyte Differentiation and Adiponectin Secretion. Biol Pharm Bull. 2016; 39:1137–1143. https://doi.org/10.1248/bpb.b16-00064.
15. Shakoor H, Hussein H, Al-Hassan N, Alketbi M, Kizhakkayil J, Platat C. The Muscle-Conditioned Medium Containing Protocatechuic Acid Improves Insulin Resistance by Modulating Muscle Communication with Liver and Adipose Tissue. Int J Mol Sci. 2023; 24:9490-9511. https://doi.org/10.3390/ijms24119490.
16. Ul Haq Shah MZ, Soni M, Shrivastava VK, Mir MA, Muzamil S. Gallic acid reverses ovarian disturbances in mice with letrozole-induced PCOS via modulating Adipo R1 expression. Toxicol Rep. 2022; 9:1938–1949. https://doi.org/10.1016/j.toxrep.2022.10.009.
17. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S, Soboleva A. NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Res. 2013; 41:D991-D995. https://doi.org/10.1093/nar/gks1193.
18. Sweeney TE, Haynes WA, Vallania F, Ioannidis JP, Khatri P. Methods to increase reproducibility in differential gene expression via meta-analysis. Nucleic Acids Res. 2017; 45:e1–e14. https://doi.org/10.1093/nar/gkw797.
19. Demšar J, Curk T, Erjavec A, Gorup C, Hocevar T, Milutinovic M, Možina M, Polajnar M, Toplak M, Starič A, Štajdohar A, Umek L, Žagar L, Žbontar J, Žitnik M, Zupan B. Orange: Data Mining Toolbox in Python. J Mach Learn Res. 2013; 14:2349–2353
20. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M. KEGG for representation and analysis of molecular networks involving diseases and drugs. Nucleic Acids Res. 2010; 38:D355-D360. https://doi.org/10.1093/nar/gkp896.
21. Xie Z, Bailey A, Kuleshov MV, Clarke DJB, Evangelista JE, Jenkins SL, Lachmann A, Wojciechowicz ML, Kropiwnicki E, Jagodnik KM, Jeon M, Ma'ayan A. Gene Set Knowledge Discovery with Enrichr Curr Protoc. 2021; 1:e90-e174. https://doi.org/10.1002/cpz1.90.
22. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003; 13:2498–2504. https://doi.org/10.1101/gr.1239303.
23. Yu D, Lim J, Wang X, Liang F, Xiao G. Enhanced construction of gene regulatory networks using hub gene information. BMC Bioinformatics 2017; 18:186-206. https://doi.org/10.1186/s12859-017-1576-1.
24. Davis AP, Wiegers TC, Johnson RJ, Sciaky D, Wiegers J, Mattingly CJ. Comparative Toxicogenomics Database (CTD): update 2021. Nucleic Acids Res. 2021; 49:D1138–D1143. https://doi.org/10.1093/nar/gkaa891.
25. Sander T, Freyss J, von Korff M, Rufener C. DataWarrior: an open-source program for chemistry aware data visualization and analysis. J Chem Inf Model. 2015; 55:460–473. https://doi.org/10.1021/ci500588j.
26. Mendez D, Gaulton A, Bento AP, Chambers J, De Veij M, Félix E, Magariños MP, Mosquera JF, Mutowo P, Nowotka M, Gordillo-Marañón M, Hunter F, Junco L, Mugumbate G, Rodriguez-Lopez M, Atkinson F, Bosc N, Radoux CJ, Segura-Cabrera A, Hersey A, Leach AR. ChEMBL: towards direct deposition of bioassay data. Nucleic Acids Res. 2019; 47:D930–D940. https://doi.org/10.1093/nar/gky1075.
27. Frimayanti N, Yam ML, Lee HB, Othman R, Zain SM, Rahman NA. Validation of quantitative structure-activity relationship (QSAR) model for photosensitizer activity prediction. Int J Mol Sci. 2011; 12:8626–8644. https://doi.org/10.3390/ijms12128626.
28. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, Connor R, Funk K, Kelly C, Kim S, Madej T, Marchler-Bauer A, Lanczycki C, Lathrop S, Lu Z, Thibaud-Nissen F, Murphy T, Phan L, Skripchenko Y, Tse T, Wang J, Williams R, Trawick BW, Pruitt KD, Sherry ST. Database resources of the national center for biotechnology information. Nucleic Acids Res. 2022; 50:D20–D26. https://doi.org/10.1093/nar/gkab1112.

29. Tuccinardi T, Poli G, Romboli V, Giordano A, Martinelli A. Extensive consensus docking evaluation for ligand pose prediction and virtual screening studies. J Chem Inf Model. 2014; 54:2980–2986. https://doi.org/10.1021/ci500424n.
30. Molexus ApS. molegro virtual docker user manual. P. Baatrupsvej 41, DK-8300 Odder, Denmark: 2019.
31. Gabrielli AP, Manzardo AM, Butler MG. GeneAnalytics Pathways and Profiling of Shared Autism and Cancer Genes. Int J Mol Sci. 2019; 20(5):1166-1180 https://doi.org/10.3390/ijms20051166.
32. Indrayanto G, Putra GS, Suhud F. Validation of in-vitro bioassay methods: Application in herbal drug research. Profiles Drug Subst Excip Relat Methodol. 2021; 46:273–307. https://doi.org/10.1016/bs.podrm.2020.07.005.
33. Zou YF, Zhang SY, Li LW, Jing K, Xia L, Sun CX, Wu B. Hub genes for early diagnosis and therapy of adamantinomatous craniopharyngioma. Medicine 2022; 101:37-47. https://doi.org/10.1097/MD.0000000000030278.
34. Chen M, Pan J, Wang G. Identification and Functional Analysis of Hub Genes in Cell Cycle Pathway Linked to the Development of Lung Cancer with Non-Small Cells. Highlights Sci. Eng. Technol. 2023; 74:1501-1512.
35. Copps KD, White MF. Regulation of insulin sensitivity by serine/threonine phosphorylation of insulin receptor substrate proteins IRS1 and IRS2. Diabetologia. 2012; 55:2565–2582. https://doi.org/10.1007/s00125-012-2644-8.
36. Geering B, Cutillas PR, Vanhaesebroeck B. Regulation of class IA PI3Ks: is there a role for monomeric PI3K subunits? Biochem Soc Trans. 2007; 35:199–203. https://doi.org/10.1042/BST0350199.
37. Luo J, Cantley LC. The negative regulation of phosphoinositide 3-kinase signaling by p85 and it's implication in cancer. Cell Cycle. 2005; 4:1309–1312. https://doi.org/10.4161/cc.4.10.2062.
38. Vallejo-Díaz J, Chagoyen M, Olazabal-Morán M, González-García A, Carrera AC. The Opposing Roles of PIK3R1/p85α and PIK3R2/p85β in Cancer. Trends Cancer. 2019; 5:233–244. https://doi.org/10.1016/j.trecan.2019.02.009.
39. Ma L, Murea M, Snipes JA, Marinelarena A, Krüger J, Hicks PJ, Langberg KA, Bostrom MA, Cooke JN, Suzuki D, Babazono T, Uzu T, Tang SC, Mondal AK, Sharma NK, Kobes S, Antinozzi PA, Davis M, Das SK, Rasouli N, Kern PA, Shores NJ, Rudel LL, Blüher M, Stumvoll M, Bowden DW, Maeda S, Parks JS, Kovacs P, Hanson RL, Baier LJ, Elbein SC, Freedman BI. An ACACB variant implicated in diabetic nephropathy associates with body mass index and gene expression in obese subjects. PLoS One. 2013; 8:e56193-e56203. https://doi.org/10.1371/journal.pone.0056193.
40. Millward CA, Desantis D, Hsieh CW, Heaney JD, Pisano S, Olswang Y, Reshef L, Beidelschies M, Puchowicz M, Croniger CM. Phosphoenolpyruvate carboxykinase (Pck1) helps regulate the triglyceride/fatty acid cycle and development of insulin resistance in mice. J Lipid Res. 2010; 51:1452–1463. https://doi.org/10.1194/jlr.M005363.
41. Onai N, Tsunokawa Y, Suda M, Watanabe N, Nakamura K, Sugimoto Y, Kobayashi Y. Inhibitory effects of bisbenzylisoquinoline alkaloids on induction of proinflammatory cytokines, interleukin-1 and tumor necrosis factor-alpha. Planta Med. 1995; 61:497–501. https://doi.org/10.1055/s-2006-959356.