In Silico ANALYSIS OF PLANT *Carthamus tinctoris* AGAINST INSULIN RESISTANCE

Thesis Submitted
In Partial Fulfilment of the Requirements for the
Degree of

MASTER OF SCIENCE
in
BIOTECHNOLOGY

by

SHIKHA MISHRA
(2k23/MSCBIO/73)

Under the Supervision of
Dr. YASHA HASIJA

DEPARTMENT OF BIOTECHNOLOGY
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahabad Daulatpur, Main Bawana Road, Delhi-110042,
India

JUNE, 2025

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahabad Daulatpur, Main Bawana Road,
Delhi-110042, India

CANDIDATE'S DECLARATION

I, SHIKHA MISHRA (2k23/MSCBIO/73), hereby certify that the work which is being presented in the dissertation entitled **— "*In Silico* ANALYSIS OF PLANT *Carthamus tinctoris* AGAINST INSULIN RESISTANCE"** in partial fulfilment of the requirements for the award of the Degree of Masters in Science, submitted in the Department of Biotechnology, Delhi Technological University is an authentic record of my work carried out during the period from Aug 2024 to Mar 2025 under the supervision of Prof Yasha Hasija.

We have not submitted the matter presented in this dissertation for the award of any other degree from this or any other institute.

PLACE :

DATE : **CANDIDATE'S SIGNATURE**

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahabad Daulatpur, Main Bawana Road,
Delhi-110042, India

CERTIFICATE BY THE SUPERVISOR

This is to certified that SHIKHA MISHRA (2k23/MSCBIO/73) carried out their research work presented in this thesis entitled **— "*In Silico* ANALYSIS OF PLANT *Carthamus tinctoris* AGAINST INSULIN RESISTANCE"** for the award of Master of Science from the Department of Biotechnology, Delhi Technological University, Delhi, under my supervision. The dissertation embodies results of original work, and studies are carried out by the student herself, and the contents of the dissertation do not form the basis for the award of any other degree to the candidate or anybody else from this or any other University/Institution.

PLACE:

DATE:

PROF. YASHA HASIJA
SUPERVISOR
Department of Biotechnology
Delhi Technological University

PROF. YASHA HASIJA
HEAD OF DEPARTMENT
Department of Biotechnology
Delhi Technological University

ACKNOWLEDGEMENT

The success and outcome of this dissertation required a lot of guidance and assistance from many people, and I am extremely fortunate to have received this throughout the completion of this dissertation work. To begin with, I want to thank my supervisor, Prof. Yasha Hasija, Head of the department, Department of Biotechnology, Delhi Technological University, for her supportive advice and support throughout the study program.

I would like to thank Prof. Prakash Chandra and all the professors of the Department of Biotechnology for their support and guidance throughout my journey.

In particular, I'd want to thank PhD scholar Ms. Khushi and Mr. Nakul for their significant assistance at various points throughout this project. Her insights and knowledge have expanded this work scope and improved its overall quality.

My parents and my family have my deepest thanks for their unwavering support and love during this journey. My greatest strength and most significant source of motivation has been their encouragement and belief in my abilities.

A special and sincere thank you should go out to my friends, Mr. Aayush Ratna Bajpai and Mr. Ram Avatar Singh, for their exceptional assistance and support with my dissertation throughout every step of the way.

Finally, I want to express my gratitude to Delhi Technological University for providing the resources, academic environment, and facilities that allowed me to complete my research. The university's commitment to quality in research and education has had a significant impact on my academic experience.

Thank you,

Shikha Mishra

In Silico ANALYSIS OF PLANT *Carthamus tinctoris* AGAINST INSULIN RESISTANCE

SHIKHA MISHRA

ABSTRACT

Insulin resistance disease is rising exponentially worldwide, emerging as the major cause of type II diabetes. Insulin resistance makes the insulin-sensitive tissues unresponsive towards insulin, these muscles include the liver, skeletal muscles and adipose tissue. This condition causes a hyperglycaemic condition in the cells. The available drugs that are used in the treatment of Insulin Resistance have some limitations, and they are less effective as well, therefore, phytochemicals from natural sources can be used to derive potential treatment. In this project, we made an effort to treat insulin resistance by using natural compounds that have fewer side effects. We did an in-silico study on the plant *Carthamus tinctoris,* commonly known as safflower. Carthamus tinctoris, or safflower is a plant of the Compositae or Asteraceae family. This plant has numerous active components such as flavonoids, phenylethanoid glycosides, fatty acids, and steroids. This study is made by using computational tools like Biovia Discovery Studio, PyRx virtual screening tool, Cytoscape and Autodock tools. In this research, we examined potential bioactive compounds of the plant safflower that can be used in the treatment of insulin resistance by stimulating the activity of Akt1. Docking results gave three potential ligands which are N-Coumaroyl serotonin (5458879), N-Feruloyl Serotonin (5969616) and Serotobenine (11725426) showing lowest binding energies -7.1 kcal/mol, -7.3 Kcal/mol and -7.5 Kcal/mol respectively thus showing higher binding affinities which were comparable to the binding energies of the reference (2S)-2-(4-chlorobenzyl)- 3-oxo-3-[4-(7H-pyrrolo[2,3-d]pyrimidin-4-yl)piperazin1-yl]propan-1-amine (46870040) showing binding energy of -7.0 Kcal/mol. In addition, these docking results also showed good results in bioavailability and toxicity analysis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

IR – Insulin Resistance

T2DM – Type II Diabetes

CUPS-7 – Chennai Urban Population Study-7

IMPPAT – Indian Medicinal Plants, Phytochemistry and Therapeutics

ADMET – Absorption, Distribution, Metabolism, Excretion, and Toxicity

PGC-1α – Peroxisome proliferator-activated receptor gamma coactivator 1-alpha

SGLT2 – Sodium-glucose cotransporter-2

ADRs – Adverse drug reactions

OMIM– Online Mendelian Inheritance in Man.

STRING– Search Tool for Retrieval of Interacting Genes/Proteins

IRS – Insulin receptor substrate

AMPK – AMP-activated protein kinase

PI3K – Phosphatidylinositol 3-kinase

ROS – Reactive oxygen species

NF-κB – Nuclear factor-kappa B

GLP-1 - Glucagon-like peptide-1

CHAPTER 1

INTRODUCTION

BACKGROUND

Insulin resistance (IR) is a metabolic condition where the insulin-sensitive muscles of the body, like skeletal muscle, liver muscle and adipose tissue, become less responsive towards insulin. Insulin is the hormone produced and released by the beta cells of the pancreas, which is further responsible for regulating blood glucose levels. Therefore, the pancreas compensates for this insensitivity by producing more insulin to maintain normal glucose levels. With time, this compensatory mechanism usually fails, which results in elevated blood glucose levels and the initial development of type 2 diabetes mellitus (T2DM) and cardiovascular diseases.

The complex interplay between environmental and genetic factors plays a major role in the advent of IR. Free fatty acids, along with pro-inflammatory cytokines released by visceral fat, interfere with insulin signalling pathways, leading to a decrease in peripheral tissues' absorption of glucose and an increase in the release of hepatic glucose.

Globally, the prevalence of IR differs widely due to differences in population characteristics and the criteria of diagnosis. In Asia, a meta-analysis reported an overall IR prevalence of 44.3%. In India, the Chennai Urban Population Study-7 (CUPS-7) found an overall IR prevalence of 11.2%, with higher rates observed in the middle-income group (18.7%) as compared to the low-income group (6.5%).

Phytochemicals are bioactive compounds that are not essential for the basic metabolic functioning of the plant but have been shown to provide health benefits as they contain medicinal properties. These compounds are secondary metabolites produced by plants to defend which defend them against pests, pathogens, and environmental stresses. Phytochemicals are of several types, which include alkaloids, flavonoids, phenolic acids, terpenoids and glycosides. Safflower is also such a reservoir of diverse phytochemicals, which makes this plant highly potential for therapeutic use. Till now, over 100 compounds have been isolated from different parts of the plant, including seeds, flowers, and stems. Quinochalcones and flavonoids are said to be the characteristic and bioactive constituents of safflower.

The phenolic compounds which is present in Safflower make this plant a potential antioxidant source, which has been demonstrated through various assays.

The *Carthamus tinctorius* plant is rich in phytochemicals, which contribute to its properties like antioxidant nature, anti-inflammatory components, and perhaps potential anticancer properties. These bioactive compounds highlight the therapeutic potential of safflower in both conventional and non-conventional treatment.

Several computational tools are used in computational drug design, with the major help of a molecular docking tool to find out new compounds for selectivity and effectiveness that require further in vitro and in vivo investigation. This technique is significant in the identification of new drugs against IR. In this thesis, I have worked on the possibilities of *Carthamus tinctorius,* a medicinal plant, in the treatment of insulin resistance using a network pharmacological approach. By merging bioinformatics, computational biology, and pharmacological research, I am trying to identify potential target pathways that can be targeted by *Carthamus tinctorius* and elucidate its mechanisms of action in insulin resistance. This work may contribute to advancing our understanding of insulin resistance and can help in uncovering novel therapeutic strategies.

OBJECTIVES

1. To analyse the potential of *Carthamus tinctorius* therapeutic use for the management of Insulin resistance
2. To examine the possible synergistic effects of multiple bioactive phytochemicals of *Carthamus tinctorius* against the potential targets in the IR metabolic pathway.
3. To find out core target genes and important phytochemicals in treating insulin resistance with the help of Network Pharmacology.
4. To evaluate the affinity of *Carthamus tinctorius* active phytochemicals for the core targets against the disease using molecular docking tools.

CHAPTER 2

REVIEW OF LITERATURE

Insulin resistance (IR) is an impaired metabolic condition where insulin-sensitive tissues of the body stop responding to insulin inside the cells, leading to hyperinsulinemia in the body and glucose surge in the major tissues. If this condition prevails for a prolonged time, then it causes T2DM and fatty liver disease.

2.1 Pathogenesis of Insulin Resistance

The intricate relation of Environmental and genetic factors involves the pathophysiology of IR. Ectopic lipids build up in tissues due to physical inactivity and excessive calorie intake, and reduced glucose intake by the tissues. This lipid deposit causes increased hepatic glucose production due to impairment in the IRS-1/2 and AMPK pathways. These metabolic changes constitute the pathogenesis of Insulin Resistance.

Visceral fat, which is a part of adipose tissue, plays an important role in IR development in the body. It releases several bioactive substances which including adipokines and inflammatory cytokines, which interfere with insulin signalling and increase systemic inflammation. Insulin resistance is exacerbated due to these chronic inflammatory conditions. This further accelerates the development of metabolic diseases.

Risk Factors

Several factors play a role in the development of IR:

- **Obesity**: Abdominal obesity is a major risk factor for IR.
- **Physical Inactivity**: Lack of regular exercise in a sedentary lifestyle makes tissue more insensitive towards insulin.
- **Poor Diet**: IR is further accelerated by taking food with high sugar content and a High glycaemic index.
- **Age**: Earlier, it was seen that it affects more to older individuals but nowadays, children are also developing IR from an early age.
- **Genetics**: Families having previous cases of diabetes are more likely to develop IR.

2.2 Molecular Biology of Insulin Resistance

Insulin resistance (IR) is a pathological condition where major body tissues become less responsive to insulin, leading to impaired insulin uptake and metabolism. At the molecular level, IR involves an intricate complex of alterations in insulin signalling

pathways, Lipid deposition mechanisms, mitochondrial function, inflammatory responses, and gene expression.

2.2.1. Impaired Insulin Signalling Pathway

The insulin receptor (IR) initiates signalling by autophosphorylation upon insulin binding, activating downstream effectors such as IRS, PI3K, and protein kinase B (Akt). Additionally, the p85 regulatory subunit of PI3K can negatively regulate insulin signalling, contributing to IR.

2.2.2. Mitochondrial Dysfunction and Oxidative Stress

Mitochondria play a crucial role in energy metabolism. In IR, mitochondrial dysfunction leads to increased production of reactive oxygen species (ROS), which activate stress kinases like JNK and IKKβ. These kinases phosphorylate IRS proteins on serine residues, further impairing insulin signalling. Oxidative stress also activates nuclear factor-kappa B (NF-κB), promoting inflammatory cytokine production that exacerbates IR.

2.2.3. Inflammatory Cytokines and Adipokines

Adipose tissue secretes many cytokines and adipokines that influence insulin sensitivity. Elevated levels of resistin, an adipokine, have been associated with increased inflammatory conditions and IR. Resistin can activate inflammatory pathways, leading to insulin resistance in peripheral tissues.

2.2.4. Epigenetic Regulation

Epigenetic modifications, like DNA methylation and histone acetylation, can alter gene expression without making any changes in the DNA sequence. These modifications affect genes that are involved in insulin signalling and metabolism, contributing to the development of IR. For instance, alterations in chromatin-modifying enzymes have been linked to obesity and IR.

2.2.5. Physical Inactivity and Muscle Insulin Resistance

Physical inactivity causes a decrease in mitochondrial content and function in skeletal muscle, interfering with insulin sensitivity. Reduced expression of PGC-1α, a key regulator of mitochondrial biogenesis, has been observed following muscle inactivity, contributing to IR.

Conclusion

Insulin resistance is a multidimensional condition involving disruptions in the insulin signalling pathway, mitochondrial dysfunction, inflammation, epigenetic changes, and lifestyle factors. Understanding and analysing these molecular mechanisms provides insights into potential therapeutic targets for managing and treating insulin resistance and its associated metabolic disorders with the help of natural resources.

2.3 Treatment Approaches

2.3.1 Changes in Lifestyle

The pillar of IR management is major lifestyle modifications. Insulin sensitivity can be improved by having a balanced diet that includes portions of whole grains, lean proteins, and healthy fats, complemented with regular exercise. Intensive lifestyle changes, balanced diet intake and increased physical activity have shown a significant decrease in the risk of acquiring type 2 diabetes.

2.3.2 Pharmaceutical Treatments

Pharmacological medications are the only option when lifestyle changes are not enough and in cases where the individual is not able to perform physical activity. Medications aim to reduce hepatic glucose synthesis and increase peripheral glucose absorption. Metformin, a primary treatment for type 2 diabetes, tries to restore insulin sensitivity in affected parts. Pioglitazone, part of a group, thiazolidinediones, also works by increasing insulin sensitivity in adipose tissue by activation of PPAR-γ. Additionally, glucagon-like peptide-1 (GLP-1) receptor agonists, along with sodium-glucose cotransporter-2 (SGLT2) inhibitors, have shown significant efficacy in decreasing insulin resistance and controlling glycaemic load.

2.3.3. New Therapeutic Targets

New Potential treatment targets have been identified through investigations into the molecular pathways involved in causing IR. Strategies that include methods to target inflammatory pathways, increase AMPK activation, and alter mitochondrial function have been identified. Mitochondrial biogenesis through exercise has shown enhanced insulin sensitivity, signifying the role of physical activity in IR management.

Insulin resistance is a multifactorial condition that requires a comprehensive approach that involves lifestyle modifications, pharmacological interventions, and research of novel therapeutic targets. Early detection is crucial in preventing the progression to T2DM, as it can be managed in early stages and mitigate associated complications.

2.3.4 Pharmacological Drugs involved in the treatment of Insulin resistance are shown in Table 1.

S.No.	Drugs	Mechanism of Action	Reference
1	Metformin	Decreases hepatic glucose production, enhances glucose uptake by peripheral organs.	(L Currati et al.,2000)
2	Thiazolidinediones (e.g., Pioglitazone)	PPAR-γ agonist, which increases insulin sensitivity in adipose tissue	(Salt et al.,2000)
3	GLP-1 Receptor Agonists (e.g., Liraglutide)	Stimulates insulin release, delays gastric emptying, and reduces body weight	(Jabbour S et al.,2024)
4	SGLT2 Inhibitors (e.g., Empagliflozin)	Inhibits renal glucose reabsorption, which leads to glycosuria and reduces blood glucose levels	(Yun Wu Li et al.,2014)

Table 1: Showing drugs which are involved in IR treatment along with their mechanism of action

2.4 Limitations of drugs used in the treatment of IR

- **Metformin-** Gastrointestinal side effects like nausea, diarrhoea, risk of lactic acidosis, less effective in lean IR.
- **Thiazolidinediones-** They result in weight gain, fluid retention, and increased heart failure risk and possibility of bladder cancer.
- **GLP-1 Receptor Agonists-** Costly; commonly cause gastrointestinal issues, require injection for administration.
- **SGLT2 Inhibitors-** Increased risk of genital infections, dehydration, and diabetic ketoacidosis; expensive.

2.5 *Carthamus tinctoris*

Carthamus tinctoris, also known as Safflower, is used as a traditional medicine in Asia, the Middle East and the African region. Its extract contains bioactive compounds like linoleic acid, oleic acid, flavonoids and alkaloids. The following are the medicinal benefits of this plant:

Metabolic Syndrome Management

A clinical trial showed that safflower oil supplementation causes significant reductions in blood pressure, fasting blood sugar, and insulin resistance without making any lifestyle modifications.

Antidiabetic Effects

Safflower seed oil inhibits enzymes, which shows its activity in carbohydrate digestion and fat metabolism, like α-amylase and porcine pancreatic lipase. These activities suggest their potential in managing diabetes and obesity.

Anti-Adipogenic Effects

Hot water extracts from safflower seeds have been demonstrated to suppress lipid accumulation in adipocytes, indicating potential for use in anti-obesity therapies.

Antifungal Activity

Extracts from safflower seeds, which are rich in alkaloids, flavonoids, and terpenoids, showed significant antifungal activity against *Aspergillus* species, suggesting their potential as a natural antifungal agent.

2.6 Exploration of the network pharmacology approach and its importance in studying complex diseases like Insulin Resistance

Network pharmacology is a computational science that applies network science, system biology, and pharmacology approaches to identify the drug-target interactions and form networks of biological systems. This strategy has been shown to be extremely beneficial in the research of complex diseases, offering details about drug targets, treatment methods, and disease mechanisms.

Anticipating negative drug response: Adverse drug reactions (ADRs) are a major issue in patient safety and drug development. Therefore, the concept of network pharmacology provides an approach to get insights into the network context of drug-target interactions, which is helpful to understand the mechanisms of Adverse Drug Reactions (ADRs) at the systems level. That could also be used to improve and fine-tune dosages or structures for safer drugs.

Network pharmacology and personalised medicine: Systems Biology is considered to be in alignment with Network Pharmacology. Network pharmacology integrates genomics, proteomics, metabolomics, and clinical data to provide a holistic understanding of diseases and drug actions at a systems level. This strategy can thus be used to develop personalised medicine, providing individual treatment based on the genetic background and disease network features of each patient.

Understand disease mechanisms: Multiple molecular connections and abnormal pathways are shared components in complex diseases. Network pharmacology explores various aspects of disease creation and evaluation, using networks to represent interactions between molecules in diseased states. It provides a holistic knowledge of the disease manifestation through the integration of omics data, drug-target interactions, protein protein-protein-protein interaction networks.

CHAPTER 3

MATERIALS AND METHODOLOGY

3.1 Tools & Software

3.1.1 IMPPAT

IMPPAT is Indian Medicinal Plants, Phytochemistry and Therapeutics, is largest database managed and released by the Central Council for Research in Ayurvedic Sciences, Government of India. This carefully curated collection includes 1124 therapeutic uses, 9596 phytochemicals, and 1742 Indian medicinal plants. 9596 phytochemicals with 2-D and 3-D chemical structures of phytochemicals (Mohanraj et al., 2018). One can easily access a plant‘s phytochemicals and also get the location where is particular phytochemical is located, and a number of plants as well as associations in therapeutic use for different diseases.

3.1.2 Swiss ADMET analysis

It is software which is widely used in computational biology for identifying druglikeness score and molecular property prediction, such as molecular formula, molecular weight and Blood-Brain-Barrier score of phytochemicals.

ADMET analysis assesses five key aspects of a drug's behaviour:

- **Absorption**: The less drug takes time to get absorbed into the bloodstream.
- **Distribution**: The pathway through which the drug spreads throughout the body and reaches target tissues.
- **Metabolism**: The way how drug is metabolised in the body, typically in the liver, and determines whether its metabolites are active or toxic.
- **Excretion**: The process through which the components of drug metabolites are excreted from the body, primarily via urine or faeces.
- **Toxicity**: Potential harmful effects on organs or systems, including mutagenicity, carcinogenicity, and cardiotoxicity.

ADMET analysis is a critical component of drug development that helps in identifying and optimising drug candidates, reducing development risks, enhancing patient safety, and increasing the overall efficiency of the drug discovery process.

3.1.3 Swiss Target Prediction

For mapping predictions based on homology within and across species to match near paralogs and orthologs, this server utilises 2D and 3D similarity metrics that are specific to the targets of active compounds. This server results in five distinct organisms that can be used to validate predictions (Safran et al., 2010). Using a ligand-based approach, this software compares the chemical structure of a drug to the database of known ligands and the protein targets that are linked to them. Additionally, by comparing the chemical characteristics of the material with those of known protein structures, a target-based method is used to identify possible binding sites and interactions.

3.1.4 Gene Card

It has been widely used for ~15 years and provides a detailed, comprehensive set of gene annotations for human genes. Over 73,000 human gene entries have their gene-centric materials automatically mined & integrated from > 80 digital resources to generate this web-based. It compiles genetic data from extensive data sources such as NCBI, Abcam, Alobono Labs and many more. The data for each gene is presented in a detailed ‘card ' type format with separate functional components and various features, including text and links to other dedicated and genome-wide databases. Gene Cards Version 5: an updated platform with new search and support tools. It provides various information about each human gene, which includes gene information (gene name, location), genomic data, protein information, biological pathways, disease pathways and gene ontology.

3.1.5 OMIM

OMIM stands for Online Mendelian Inheritance in Man. It gives an excellent, reliable repository for human genes and genetic phenotypes as described in the scientific literature, and enhances the understanding of genetic variations, but it cannot provide data for analysis or diagnosis of disease conditions. This database was initiated as a catalogue of Mendelian traits and disorders, entitled Mendelian Inheritance in Man MIM (Shannon et al., 2003). One convenient and fast way to explore the growing body of knowledge surrounding human genetics is through the site OMIM. 3.1.7 Bioinformatics and Evolutionary Genomics. It is used to draw custom Venn diagrams. The full-text, referenced overviews in the database contain information on all known Mendelian disorders and information about more than 16,000 genes. OMIM is primarily designed on the basis of the relationship between phenotype and genotype.

3.1.6 KNApSAcK

The KNApSAcK database is a comprehensive, easily accessible resource that integrates information on plant metabolites, their biological activities, and their applications in traditional medicines. It was designed by the Laboratory of Bioinformatics at Nara

Institute of Science and Technology (NAIST), Japan. This database acts as a tool of primary importance for researchers in fields such as pharmacology, metabolomics, and ethnobotany. It provides detailed information on plant metabolites, and this database aids in the identification of potential phytochemicals. The database's data on human disease biomarkers supports studies on genetic variations influencing drug action. Researchers can get information about the traditional uses of plants, which contributes to the preservation of indigenous knowledge and the discovery of novel therapeutic agents using natural sources. The informative metabolite profiles facilitate studies on plant metabolism and its ecological implications. The KNApSAcK database is freely accessible online, with user-friendly interfaces for searching and visualising data. It helps in various search criteria, including plant species, metabolite names, and biological activities, making it a valuable resource for researchers worldwide. (http://kanaya.naist.jp/KNApSAcK_Family/)

3.1.7 Cytoscape

Cytoscape software helps in providing flexibility for network research, data import, and visualisation. It is a network tool that has to import its network from external sources. However, it can be used in any system for analysing molecular components and interactions. The function is most effective when utilised in harmony with large databases of protein-protein, protein-DNA DNA and genetic interactions (Chin et al., 2014). The following are some functions of Cytoscape:

1. Network Visualisation & Analysis: Cytoscape is a tool for creating graphics and organising detailed network visualisations. Network data may also be utilised by users and rendered in Edge and Node format. This tool includes layout and arrangement settings for the network's components, as well as can change the colour, size, and style to highlight or isolate specific node and edge features, respectively. Along with clustering, network motif identification, and topology analysis, users can also compute network centrality metrics

2. Network Simulation: It contains a model of a network's simulation as well as a module analysis that analyses network dynamics. It must be mentioned that users can alter dynamic models according to their requirements for better illustrations. When studying a biological system or comprehending how a certain network affects biological functionality, this trait is helpful.

3. Data Visualization: Network data can be combined with other biological data formats by users using cytoscape, including gene expression, protein annotation, and functional annotation, using the Cytoscape program. Through the link, other networks can also be visualised, showing details like expression levels and functional annotations of the network's paths.

3.1.8 Cytoscape- Cytohubba

The Cytohubba is a plugin of Cytoscape. It can be used for the visualisation of biological networks in the form of PPI networks. Cytohubba helps in the identification of hub nodes in the network analysis. It is particularly used for forming molecular mechanisms on the basis of genes or proteins. It works on multiple centrality indices like Degree, Betweenness, Bottleneck, Maximum Neighbourhood Components and many more. Additionally, researchers can use Cytohubba by combining it with other data like STRING and BioGRID. It enables the extraction of required regions of the network, which enables further study to be more feasible. It can also be used in the analysis innate immune system, complex 26 biofilm communities (von Mering et al., 2003). Hub genes identified in the Cytohubba can be of high metabolic importance which as they regulate crucial metabolic pathways in the biological system. They are highly coupled nodes that are identified by Cytohubba. Its users may customise the examination by adjusting parameters and settings, such as the number of top-ranked nodes to display, filtering criteria, and choosing the scoring method or collection of techniques. The most important advantage of Cytohubba is that it can be used with the other plugins of Cytoscape as well like ClueGO, which is used for enrichment analysis. It aids in the understanding of biological pathways on the basis of PPI networks and allows researchers to manipulate networks and examine them in various dimensions (Chin et al. 2014).

3.1.9 Bioinformatics and Evolutionary Genomics

This particular tool is used for the identification of overlapping content from different sets of data. It represents different sets and their overlapping content in the graphical visualisation that is a Venn diagram. It generates a Venn diagram when the sets of different data are fewer than seven. Its major advantage is that the overlapping terms or data can be extracted in the text form, which is used in further studies.

3.1.10 STRING Database

STRING could predict PPI using comprehensive biological data in various orthologs as well. It has worked using 700 unique algorithms, which made its prediction more accurate. The orthologs for which prediction is done by the STRING database are more than 1000 organisms, these organisms include both model and non-model organisms. Its network can be used for multiple other analyses, like functional enrichment analysis, which includes enriched gene ontology and KEGG pathways. The STRING database allows flexible extraction of the data, ranging from high-resolution images to a TSV Cytoscape-compatible format.

3.1.11 Protein Databank (PDB)

It is a database that stores information about the structure of biological molecules, which are experimentally determined, such as proteins and nucleic acids (Kim et al., 2016). It contains all the information about a particular biological molecule, such as its 3D and 2D structures, the inhibitors or ions with which it is in bound state, its crystallographic details and details regarding how the structure is obtained as well. Biologists and scientists from

all over the world utilise the PDB data for their research and also deposit the newly identified structure to strengthen the repository of the database. It provides major structural details, atomic coordinates, citations, and links with several scientific databases. These structures help in the design of newer therapeutic approaches on the initial level using bioinformatics tools.

3.1.12 PubChem

It provides all the information about most of the identified chemicals and molecules. It provides data like molecular weight, canonical smiles, 2D and 3D structure, and the corrosive power of the chemical. It has information on almost 157 million chemical substances, 60 million distinct chemical structures, and one million biological test descriptions from depositors, which together account for over 10,000 distinct protein target sequences and make PubChem the largest online repository available. Structure-Data File (SDF), Comma Separated Values (CSV) and other file formats are available while downloading 2D or 3D structure.

3.1.13 PyRx

It's a computational biology tool that finds data by screening a lot of database compounds. It allows the analysis of binding ability between a target and ligand, saving lots of time and money that goes into the in vivo studies. This tool integrates a number of open-source programmes, including Open Babel, Autodock, and Autodock Vina. In this, molecular docking is performed, and coordinates can be customised according to the active site and consequently learn about binding energy. The most advantageous features of PyRx are that multiple ligands against a particular target can be examined for binding in one go.

3.1.14 Biovia Discovery Studio

With Biovia Discovery Studio's structure-based drug design tools, researchers may predict binding affinities and perform virtual screening. This tool provides a comprehensive method for organising large data of different formats appropriately, which enables researchers to have easy access. The major functions that it can perform are molecular docking, homology modelling, pharmacophore modelling, and investigation of protein-ligand interactions.

3.2 Methodology

3.2.1 Exploration of phytochemicals of *Carthamus tinctoris*

The phytochemicals are downloaded from IMPPAT, which is openly accessible at: https://cb.imsc.res.in/imppat. The IMPPAT identifiers were also noted. More phytochemicals were extracted using the KNApSAcK database.

3.2.2 Screening ligands based on Pharmacokinetic Properties

The Canonical Smiles of phytochemicals and 3D structures are retrieved using the PUBCHEM database. These canonical smiles and structures were used for analysis of pharmacokinetic properties like drug likeness, mutagenicity, bioavailability, and toxicity. 3D structures were further used for ADMET analysis; the ligands that showed positive results in these analyses were utilised in further steps.

3.2.3 Identification of target genes of Insulin Resistance

The target genes were acquired from the Gene Card and OMIM database using the keyword 'Insulin Resistance'. These databases give the list of genes reported to date that are associated with a particular disease. A list of genes along with their approved symbols is organised in an Excel format.

3.2.4 Swiss Target Prediction

The genes that are targeted by the shortlisted ligand group are acquired from the tool Swiss Target Prediction. The list of ligands is uploaded on the software, and a comprehensive table containing all the details of ligands is generated, which has information of ligand gene targets, uniport ID, common name of the target, ChEMBL ID and probability of targeting a particular gene by a particular ligand.

3.2.5 Identifying intersection gene targets

To get more accurate target genes closely related to disease pathways, the target genes of Insulin resistance and genes acquired from Swiss Target Prediction are compared. The target genes for *Carthamus tinctoris* in Insulin resistance were mapped using the tool Bioinformatics and Evolutionary Genomics, which shows putative target genes in the form of a Venn diagram. The list of intersecting entities can also be downloaded in excel format.

3.2.6 Protein-protein intersection analysis

The intersecting gene targets were uploaded to the STRING database. Tab-separated values (tsv) files containing the PPI analysis results from STRING were then imported into Cytoscape to investigate possible anti-IR core targets. Targets with a confidence score > 0.4 were considered only.

3.2.7 Core Targets identification using Cytohubba

The Cytoscape plugin "Cytohubba" was put in action for the identification of the top 10 targets. The intersection of the achieved goals may be found by applying the four approaches of Degree, Maximum Neighbourhood Component (MNC), Maximal Clique Centrality (MCC), and Closeness. The core targets were identified using criteria of the shortest pathway with maximum interaction (Ilievska-Poposka et al., 2018).

3.2.8 Molecular Docking

1. Selection and Preparation of Ligand: Three bioactive compounds, N-Coumaroyl serotonin, Serotobenine and N-Feruloyl Serotonin, which are shortlisted through network pharmacology, are downloaded from PubChem in 3D SDF format.

2. Selection and Preparation of Target Gene: The 3D structure of the Akt1 kinase domain with pyridopyrimidine inhibitor with PDB ID: 3ocb was retrieved from the RCSB PDB. Target protein is then prepared using Biovia Discovery Studio, removal of water entities and the addition of polar hydrogen were done and saved in a PDBQT format.

3. Docking Analysis: After the target gene and ligand were prepared, molecular docking was performed using PyRx. The target gene was loaded and converted into a macromolecule, and also converted into a PDBQT format. Ligands were loaded one by one using Open Babel within PyRx and converted from 3D (three dimensional) structures to PDBQT format. Proteins and phytochemicals that have been uploaded are chosen as ligands and targets, respectively, and then stored in the PDBQT format by the Vina wizard within PyRx. Then, a grid box was set up for the active site in order to dock by making only the active sites accessible. Eventually, the result was presented in a table displaying different ligands along with their binding affinities with the target gene.

CHAPTER 4

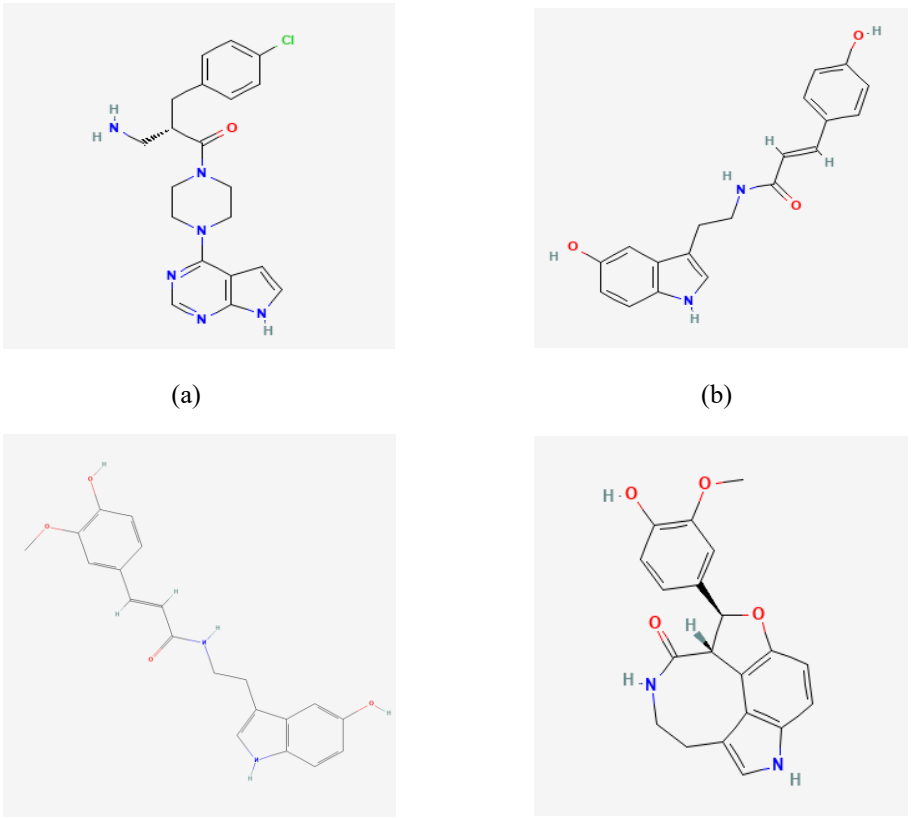
RESULTS

4.1 Screening of *Carthamus tinctoris* phytochemicals

The *Carthamus tinctoris* plant produced 190 phytochemicals in total, according to the IMPPAT database. Out of which 3 were shortlisted based on Lipinski's rule. In essence, 3 drugs were continued for the rest of the network pharmacology study as shown in Fig.4.1 (a, b, c, d).

Table 2: The list of phytochemicals of *Carthamus tinctoris* which were used in this study

COMPOUND NAME	PUBCHEM ID	BINDING ENERGY (KJ/MOL)	ESOL (logS)	H bond acceptors	H bond acceptors	NUMBER OF H-ATOMS	Intersecting residues
Pyrrolopyrimidine	46870040	**-7**	-3.57	4	4	2	LYS389, LYS385, LYS386 ASP324, GLU319, GLU322, LEU362
N-Coumaroyl serotonin	5458879	-7.1	-3.80	3	3	1	LYS386, ASP325, ASP387, PRO388
N-Feruloyl Serotonin	5969616	-7.3	-3.86	4	4	4	LYS389, PRO388, ASP325 GLU322, LEU322, LYS386, ASP387, GLN390
Serotobenine	11725426	-7.5	-3.71	4	4	2	LYS389, ASP324, ASP387, PRO388, GLY327



(a) (b)

(c) (d)

Fig 4.1 Final compounds selected from network pharmacology for molecular docking (a) Pyrrolopyrimidine (b) N-Coumaroyl serotonin, (c) N-Feruloyl Serotonin, (d) Serotobenine

4.2 Genes related to insulin resistance

Genes associated with insulin resistance are isolated from the Genecard database and the OMIM database. A total of 11149 gene targets were identified using these databases.

4.3 Intersection gene target analysis

According to Fig. 4.2, Bioinformatics & Evolutionary Genome identified 102 gene targets that were shared by the 112 putative gene targets for the active phytochemicals in *Carthamus tinctoris,* as well as 11149 gene targets connected to insulin resistance.

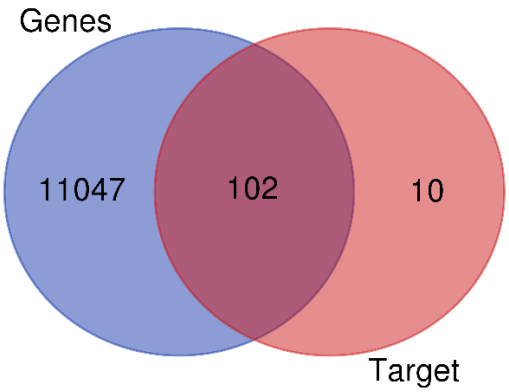


Fig. 4.2 Venn Diagram showing intersecting gene targets

4.4 PPI network analysis

Protein-protein interaction is obtained by using STRING. The PPI interaction that was formed has 102 nodes and 548 edges, as shown in fig. 4.3, this interaction has an average node degree of 10.7 and an average local clustering coefficient is 0.566.

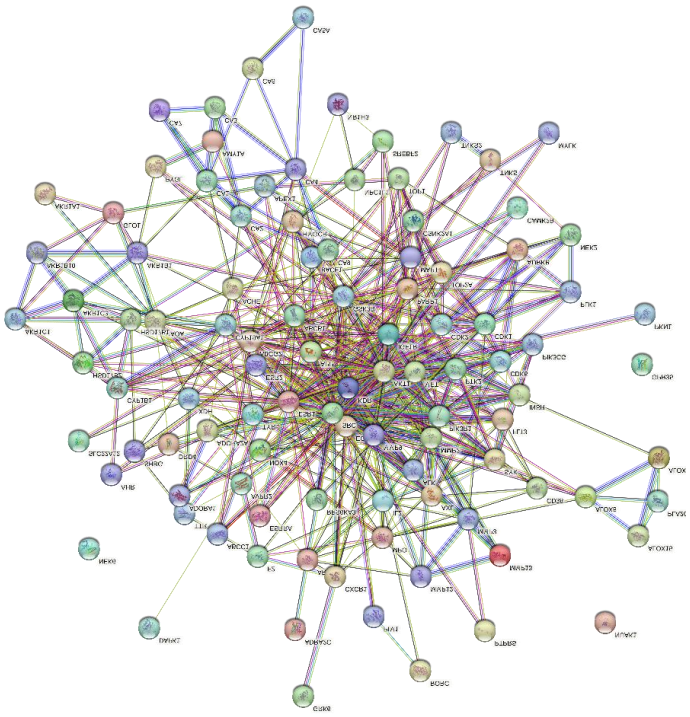


Fig. 4.3 PPI interaction formed by STRING

4.5 Core Target Analysis

Cytoscape Cytohubba is used to identify 10 core target proteins. The intersection of the required data can be founded by applying the four methods of Degree, Maximum Neighbourhood Component (MNC), Maximal Clique Centrality (MCC), and Closeness as shown in Fig 4.4.

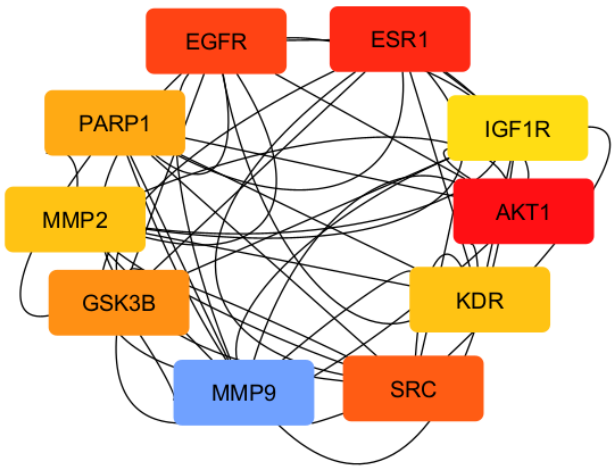
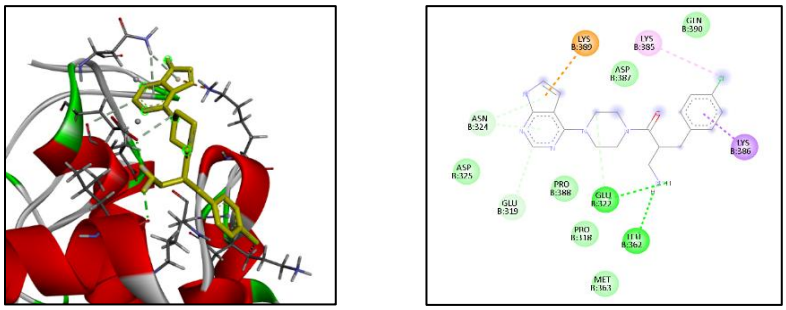


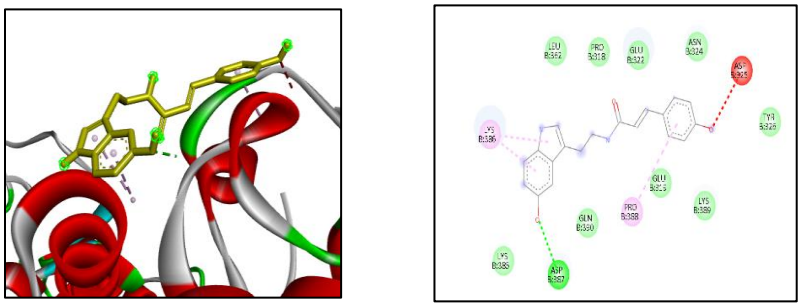
Fig. 4.4 Hub genes identified using Cytohubba

4.6 Molecular Docking Analysis

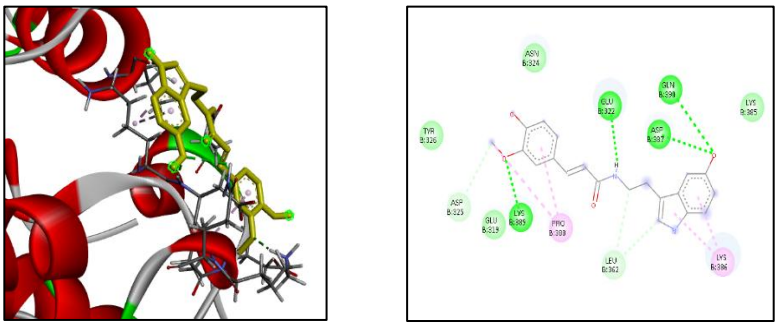
The top gene target, Akt1, is selected as the target protein. This particular protein is docked with the shortlisted ligand molecules, which are (2S)-2-(4-chlorobenzyl)-3-oxo-3-[4-(7H-pyrrolo[2,3-d] pyrimidin-4-yl) piperazin-1-yl] propan-1-amine, N-Coumaroyl serotonin, N-Feruloyl Serotonin and Serotobenine. The results of docking are shown in the table, analysis is done on the basis of binding affinity(kcal/mol) scores; the lower the binding affinity, more is the binding ability of the ligand towards the protein target. Fig. 4.5 (a), (b), (c) and (d) show the 2D and 3D structure of the docked complexes. The ligands show significant results in binding, which shows that *Carthamus tinctoris* can act as a potential alternative in the treatment of insulin resistance.



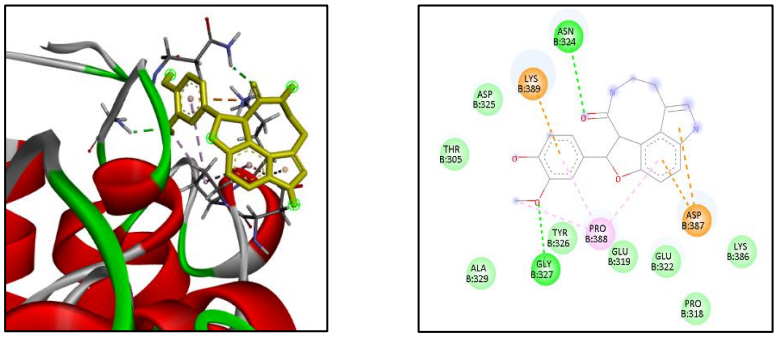
(a)



(b)



(c)



(d)

Fig. 4.5 Showing top docking results in 2D and 3D structure (a) (2S)-2-(4- chlorobenzyl)-3-oxo-3-[4-(7H-pyrrolo[2,3-d]pyrimidin-4-yl)piperazin-1- yl]propan-1-amine (46870040), (b) N-Coumaroyl serotonin (5458879), (c) N-Feruloyl Serotonin (5969616) and (d) Serotobenine (11725426)

DISCUSSION

This project is currently studying the potential of a natural source, Carthamus tinctoris, against the disease insulin resistance. The treatment that is available has multiple limitations and side effects on the body, as can be seen in Table 4. IR cause an increase in T2DM cases all over the world because of a changed lifestyle, intake of ultra-processed food, which causes obesity in individuals.

The sedentary lifestyle of individuals has caused a drastic increase in the number of Insulin resistance cases all over the world, and thus cases of Diabetes Mellitus as well.

In the condition of IR body becomes partially insensitive towards insulin, so we targeted the molecular pathway that can reverse the condition. Akt1 is a protein that is present inside the cell and plays a major role in insulin uptake. Akt1 protein activity is inhibited in insulin-resistant conditions. We analyzed phytochemicals that show effective binding towards Akt1. In this project, potential ligands were identified using molecular docking and ADME analysis. Ligands were identified against the reference ligand PubChem ID 46870040, showing binding energy (-7.0 kcal/mol). The molecular docking analysis has shown three ligands having PubChem IDs as 5458879, 5969616, and 11725426 showing lower binding energies than reference ligands, thus exhibiting higher binding affinity towards the Akt 1 protein. The ADME analysis of these ligands further confirmed their efficacy as a potential ligand of the target protein. All these ligands showed good results in toxicity analysis.

CONCLUSION

The result of this project shows that *Carthamus tinctoris* can act in the formulation of a potential treatment against the insulin resistance disease. The current study determines the evidence using pharmacological evidence, which shows good results, in order to determine that this natural source can be used for further analysis for formulating potential therapy. This can also be said about the importance of natural sources, and traditional knowledge can be utilized effectively in order to deal with modern age disorders like cancer, cardiovascular diseases, and lifestyle diseases. Also, natural sources are safer alternatives to chemically derived drugs, as they show fewer side effects in the biological system.

REFERENCES

1. Zyoud SH. Mapping the landscape of research on insulin resistance: a visualization analysis of randomized clinical trials. *J Health Popul Nutr*. 2024 Jan 9;43(1):6. doi: 10.1186/s41043-024-00497-4. PMID: 36631365; PMCID: PMC9259665.
2. Das SK, Mangaraj M, Nayak S, Satapathy S, Mahapatro A, Goyal RK. Prevalence of insulin resistance in urban Indian school children who are overweight/obese: a cross-sectional study. *J Health Popul Nutr*. 2021 Jan 8;39(1):1. doi: 10.1186/s41043-020-00248-2. PMID: 33644095; PMCID: PMC7907002.
3. Das SK, Mangaraj M, Nayak S, Satapathy S, Mahapatro A, Goyal RK. Prevalence of metabolic syndrome among urban Indian adolescents and its relation with insulin resistance (HOMA-IR). *J Health Popul Nutr*. 2013 Jun;31(2):169-74. doi: 10.3329/jhpn.v31i2.15783. PMID: 23751385; PMCID: PMC3687807.
4. Srinivasan SR, Myers L, Berenson GS. Predictability of obesity and insulin resistance in young adulthood from childhood and adolescent measurements: the Bogalusa Heart Study. *Diabetes*. 2002 Jul;51(7):2179-85. doi: 10.2337/diabetes.51.7.2179. PMID: 12086988.
5. Misra A, Khurana L. Obesity and the metabolic syndrome in developing countries. *J Clin Endocrinol Metab*. 2008 Apr;93(4):S9-30. doi: 10.1210/jc.2008-1595. PMID: 18492788.
6. International Diabetes Federation. IDF Diabetes Atlas, 9th edition. Brussels, Belgium: International Diabetes Federation; 2019. Available from: https://www.diabetesatlas.org
7. American Diabetes Association. Global prevalence of prediabetes. *Diabetes Care*. 2023 Jul;46(7):1388-1395. doi: 10.2337/dci23-0017. PMID: 37499223.
8. Srinivasan SR, Myers L, Berenson GS. Predictability of obesity and insulin resistance in young adulthood from childhood and adolescent measurements: the Bogalusa Heart Study. *Diabetes*. 2002 Jul;51(7):2179-85. doi: 10.2337/diabetes.51.7.2179. PMID: 12086988.
9. Chennai Urban Population Study-7 Group. Prevalence of insulin resistance syndrome in a selected south Indian population--the Chennai urban population study-7 (CUPS-7). *
10. Zhao, Z., et al. (2013). Phytochemistry, pharmacology and medicinal properties of Carthamus tinctorius L. *Phytochemistry Reviews*, 12(3), 537–548. https://doi.org/10.1007/s11101-013-9302-5
11. Zhao, Z., et al. (2016). Towards a better understanding of medicinal uses of Carthamus tinctorius L. in traditional Chinese medicine: a phytochemical and pharmacological review. *Journal of Ethnopharmacology*, 183, 1–10. https://doi.org/10.1016/j.jep.2016.01.022
12. Zhao, Z., et al. (2014). Phytochemical composition and antidiabetic, anti-obesity, antioxidant, and cytotoxic activities of Carthamus tinctorius seed oil.

Journal of Agricultural and Food Chemistry, 62(51), 12294–12301. https://doi.org/10.1021/jf5038356

13. Zhao, Z., et al. (2011). Variation in phenolic composition and antioxidant activity during flower development of safflower (Carthamus tinctorius L.). *Journal of Agricultural and Food Chemistry*, 59(24), 13057–13063. https://doi.org/10.1021/jf1049936
14. Zhao, Z., et al. (2014). Antioxidant and pro-oxidant properties of Carthamus tinctorius, Hydroxy Safflor Yellow A, and Safflor Yellow A. *Phytochemistry Reviews*, 13(1), 1–15. https://doi.org/10.1007/s11101-013-9330-3
15. Morino K, Petersen KF, Shulman GI. Molecular Mechanisms of Insulin Resistance in Humans and Their Potential Links With Mitochondrial Dysfunction. *Diabetes*. 2006 Dec;55(Suppl 2):S9–S15. doi: 10.2337/db06-S002. PMID: 17130651; PMCID: PMC2995546. Link
16. Choi K, Kim YB. Molecular Mechanism of Insulin Resistance in Obesity and Type 2 Diabetes. *Korean J Intern Med*. 2010 Jun;25(2):119–129. doi: 10.3904/kjim.2010.25.2.119. PMID: 20526383; PMCID: PMC2880683. Link
17. Pei J, Wang B, Wang D. Current Studies on Molecular Mechanisms of Insulin Resistance. *J Diabetes Res*. 2022 Dec 23;2022:1863429. doi: 10.1155/2022/1863429. PMID: 36589630; PMCID: PMC9803571. Link
18. Li, M., Chi, X., Wang, Y., Setrerrahmane, S., Xie, W., & Xu, H. (2022). Trends in insulin resistance: insights into mechanisms and therapeutic strategy. *Signal Transduction and Targeted Therapy*, 7(1), 216. https://doi.org/10.1038/s41392-022-01073-0
19. Li, M., Chi, X., Wang, Y., Setrerrahmane, S., Xie, W., & Xu, H. (2021). Insulin resistance: From mechanisms to therapeutic strategies. *Frontiers in Endocrinology*, 12, 757. https://doi.org/10.3389/fendo.2021.757
20. Li, M., Chi, X., Wang, Y., Setrerrahmane, S., Xie, W., & Xu, H. (2023). Recent advances in the treatment of insulin resistance targeting molecular and metabolic pathways: Fighting a losing battle? *Journal of Clinical Medicine*, 58(4), 472. https://doi.org/10.3390/jcm58040472
21. Li, M., Chi, X., Wang, Y., Setrerrahmane, S., Xie, W., & Xu, H. (2023). Insights into the development of insulin resistance: Unraveling the interaction of physical inactivity, lipid metabolism and mitochondrial biology. *Frontiers in Physiology*, 14, 1151389. https://doi.org/10.3389/fphys.2023.1151389
22. Li, M., Chi, X., Wang, Y., Setrerrahmane, S., Xie, W., & Xu, H. (2024). Lifestyle intervention in reducing insulin resistance and preventing type 2 diabetes in the Asia Pacific region: A systematic review and meta-analysis. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 18(1), 11303493. https://doi.org/10.1016/j.dsx.2024.11303493
23. Li, M., Chi, X., Wang, Y., Setrerrahmane, S., Xie, W., & Xu, H. (2023). Molecular mechanisms for the vicious cycle between insulin resistance and the inflammatory response in obesity. *International Journal of Molecular Sciences*, 24(12), 9818. https://doi.org/10.3390/ijms24129818

PLAGIARISM REPORT

Shikha Dessertation.pdf

Delhi Technological University

Document Details

Submission ID
trn:oid:::27535:99130492

Submission Date
Jun 3, 2025, 1:51 PM GMT+5:30

Download Date
Jun 3, 2025, 1:52 PM GMT+5:30

File Name
Shikha Dessertation.pdf

File Size
1.2 MB

35 Pages

7,805 Words

50,346 Characters

10% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Cited Text
- Small Matches (less than 10 words)

Match Groups

41 Not Cited or Quoted 10%
Matches with neither in-text citation nor quotation marks

0 Missing Quotations 0%
Matches that are still very similar to source material

0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

8% Internet sources
4% Publications
7% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- **41** Not Cited or Quoted 10%
  Matches with neither in-text citation nor quotation marks
- **0** Missing Quotations 0%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

8% Internet sources
4% Publications
7% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	
dspace.dtu.ac.in:8080		2%
2	Internet	
dtu.ac.in		1%
3	Submitted works	
Delhi Technological University on 2025-05-06		<1%
4	Submitted works	
Koc University on 2016-01-20		<1%
5	Internet	
ligand-depot.rutgers.edu		<1%
6	Submitted works	
Roehampton University on 2025-05-31		<1%
7	Internet	
pubmed.ncbi.nlm.nih.gov		<1%
8	Internet	
www.coursehero.com		<1%
9	Publication	
Md Arju Hossain, Md Sohel, Tayeba Sultana, Md Imran Hasan et al. "Study of kae...		<1%

11	Internet	
www.uniprot.org		<1%
12	Internet	
repository.ias.ac.in		<1%
13	Internet	
www.tandfonline.com		<1%
14	Internet	
eprints.uniss.it		<1%
15	Internet	
www.evitachem.com		<1%
16	Internet	
www.rxeconsult.com		<1%
17	Publication	
S. Sinha, A. Hazarika, S. Johari, B. Neog. "IMPDB: Indian Medicinal Phytochemical ...		<1%
18	Internet	
doaj.org		<1%
19	Submitted works	
KIET Group of Institutions, Ghaziabad on 2025-06-02		<1%
20	Publication	
Trueman, Mollie. "A Computational and Atomic Force Microscopy Study on the Cr...		<1%
21	Submitted works	
Victoria University of Wellington on 2024-06-03		<1%
22	Internet	
fastercapital.com		<1%
23	Internet	

AI DETECTION

Shikha Dessertation.pdf

Delhi Technological University

Document Details

Submission ID
trn:oid:::27535:99130492

Submission Date
Jun 3, 2025, 1:51 PM GMT+5:30

Download Date
Jun 3, 2025, 2:46 PM GMT+5:30

File Name
Shikha Dessertation.pdf

File Size
1.2 MB

35 Pages

7,805 Words

50,346 Characters

*% detected as AI

AI detection includes the possibility of false positives. Although some text in this submission is likely AI generated, scores below the 20% threshold are not surfaced because they have a higher likelihood of false positives.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Disclaimer

Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (it may misidentify writing that is likely AI generated as AI generated and AI paraphrased or likely AI generated and AI paraphrased writing as only AI generated) so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.

Frequently Asked Questions

How should I interpret Turnitin's AI writing percentage and false positives?

The percentage shown in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was either likely AI-generated text from a large-language model or likely AI-generated text that was likely revised using an AI-paraphrase tool or word spinner.

False positives (incorrectly flagging human-written text as AI-generated) are a possibility in AI models.

AI detection scores under 20%, which we do not surface in new reports, have a higher likelihood of false positives. To reduce the likelihood of misinterpretation, no score or highlights are attributed and are indicated with an asterisk in the report (*%).

The AI writing percentage should not be the sole basis to determine whether misconduct has occurred. The reviewer/instructor should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in accordance with their school's policies.

What does 'qualifying text' mean?

Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be likely AI-generated will be highlighted in cyan in the submission, and likely AI-generated and then likely AI-paraphrased will be highlighted purple.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.