NETWORK PHARMACOLOGY APPROACH: BIOACTIVE PHYTOCHEMICALS FROM *SALVIA OFFICINALIS L.* TARGETING SRC FOR THE TREATMENT OF COLORECTAL CANCER

A Dissertation Submitted
In Partial Fulfillment of the Requirement for the
Degree of

MASTER OF SCIENCE

in

BIOTECHNOLOGY

by

POOJA
2K22/MSCBIO/37

Under the supervision of
Dr. ASMITA DAS

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

JUNE, 2024

ACKNOWLEDGEMENT

I would like to offer my sincere gratitude to my faculty advisor, Dr. Asmita Das, who has extended her support and guidance through all discourses of the ideation and execution stages. Her expertise, unique insights, and positive critiquing have been invaluable to me and have played a crucial role in the success of this thesis.

I am grateful to all the faculty members of the Department of Biotechnology at Delhi Technological University, who shared their expertise through lectures during the coursework of my master's program, which helped me construct and refine my ideas.

Additionally, I am grateful to Delhi Technological University for providing me with the opportunity and technical knowledge base to conduct this research.

I would also like to thank my family and friends for their emotional and intellectual support throughout the completion of my master's program.

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly, Delhi College of Engineering)

Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, Pooja, Roll no. 2k22/MSCBIO/37 student of M.Sc. Biotechnology, hereby declare that the project Dissertation titled "Network Pharmacology approach: Bioactive phytochemicals from Salvia officinalis L. targeting SRC for the treatment of Colorectal Cancer" which is submitted by me to the Department of Biotechnology, Delhi Technological University, Delhi, in partial fulfillment of the requirement for the award of the degree of Master of Science, is original and not copied from any source without proper citation. This work has not been previously formed the basis for the award of any Degree, Diploma associateship, Fellowship, or other similar title or recognition.

Pooja

Candidate's Signature

This is to certify that the student has incorporated all the corrections suggested by the examiners in the thesis and the statement made by the candidate is correct to the best of our knowledge.

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

CERTIFICATE BY THE SUPERVISOR

Certified that **POOJA** (2K22/MSCBIO/37) has carried out their search work presented in this thesis entitled **"Network Pharmacology approach: Bioactive phytochemicals from Salvia officinalis L. targeting SRC for the treatment of Colorectal Cancer"** for the award of **Master of Science** from Department of Biotechnology, Delhi Technological University, Delhi, under my supervision. The thesis embodies results of original work, and studies are carried out by the student herself and the content of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

DR. ASMITA DAS

SUPERVISOR
Department of Biotechnology
Delhi Technological University

PROF YASHA HASIJA

HEAD OF DEPARTMENT
Department of Biotechnology
Delhi Technological University

Network Pharmacology approach: Bioactive phytochemicals from *Salvia officinalis L.* targeting SRC for the treatment of Colorectal Cancer

POOJA

ABSTRACT

Colorectal cancer (CRC) is one of the most common disease across the globe, around millions of cases are being diagnosed each year. Targeted therapy using bioactive compounds obtained from medicinal plants for various disease has gained a lot interest in the past decade. There is a rising need for identification for more safe and effective treatment strategies to reduce to mortality rate of cancer. Treatment using traditional medicinal herbs leads to less side effects as compared to conventional cancer therapies. Network pharmacology approach provides insights into pharmacological mechanisms of phytochemicals obtained from therapeutic herbs. This method utilizes various databases software for predicting the relationship between bioactive compound and target protein involved in pathogenesis of disease, then its affinity is determined by bioinformatics analysis such as Molecular docking simulations.

Network pharmacology is a useful approach for discovery of novel drug candidates. *Salvia officinalis L.,* also known as sage plant have shown to contain around 749 active phytochemicals in different parts according to IMPPAT and KNApSAcK databases. The OMIM (Online Mendelian Inheritance in Man) and Genecards databases provides information regarding gene targets involved in CRC. The venn plot between common target genes of phytochemicals and the disease target genes showed 12 common targets, that act as potential key players in promoting CRC carcinogenesis. These 12 common genes are modulated by phytochemicals of *salvia officinalis L.*, these were subjected to KEGG (Kyoto Encyclopedia of Genes and Genomics) and GO (Gene Ontology) enrichment analysis, which predicts the involvement of target gene in specific pathway of disease. String database predicts the protein-protein interaction, explaining the relationship ranking between bioactive compound and the target genes. This network of drug-target pathway and PPI is visualized by using Cytoscape. Protein-Protein interactions provides information complex interaction network between these proteins in CRC. SRC gene which encodes for protein tyrosine kinase came out to be highest scoring gene in Cytoscape analysis. Lastly, Molecular docking simulations was performed to analyze the affinity of interaction between selected bioactive phytochemicals and its target gene. We obtained structure of Src kinase (2H8H) from RCSB protein databank. The results showed that bioactive phytochemicals (Hispidulin, 6-Epi-beta-bisabolol) present in *Salvia officinalis* L. can be potential inhibitor of Src kinase which is encoded by SRC gene. This comprehensive method demonstrates the effectiveness of medicinal plants in treatment of cancer and lays the groundwork for comprehending the efficaciousness of herbal remedies.

The use of medicinal plant is increasing worldwide, many FDA approved drugs are obtained from medicinal plants. Hence, the need for new potential bioactive compound that functions as potential drug is in demand in healthcare. The significance and relevance of these studies are underscored by the growing interest in complementary medicine. By combining science with traditional knowledge, new treatment options for cancer can be investigated, providing hope for safer and more effective treatments.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

RTKs: receptor tyrosine kinase

NRTKs: non-receptor tyrosine kinases

CRC: Colorectal Cancer

OMIM: Online mendelian inheritance in Man

IMPPAT: Indian Medicinal Plants, Phytochemistry And Therapeutics

DL: Drug-Likeness

BBB: Blood-Brain Barrier

OB: Oral bioavailability

STRING: Search Tool for the Retrieval of Interacting Genes/Proteins

PPI: Protein-Protein interaction

GO: Gene Ontology

KEGG: Kyoto Encyclopedia of Genes and Genomes

CC: Cellular component

MF: Molecular function

BC: Biological component

SRC: Tyrosine-protein kinase

PTK2: Protein tyrosine kinases

EGFR: Epithelial growth factor receptor

P-gp: P-glycoprotein

BOILED: Brain Or Intestinal EstimateD Permeation Method

PDB: Protein Data Bank

CHAPTER 1

INTRODUCTION

This study aims to provide a network pharmacology approach along with bioinformatics analysis for the prediction of potential bioactive compounds present in *Salvia officinalis L.* effective against Src kinase encoded by SRC gene. It plays an important role in colorectal cancer carcinogenesis. Colorectal cancer is third common cancer affecting individuals of different age groups and causing cancer death worldwide. Emerging research has shown that network pharmacological approaches can be used in many different fields of study, particularly in figuring out how herbal remedies and ayurvedic formulations might work to treat different disorders [1]. This computational pharmacology research has been conducted with the assistance of various online prediction tools and servers. The introductory section alludes to the motivation and rationale behind the research, followed by abridgment of the objectives, the research hypothesis, and the thesis outline.

1.1 Rationale:

Colorectal cancer (CRC) ranks as the third most common type of cancer and is leading cause of death globally [2]. Development of CRC is a complex multistep process which is initiated by gene and epigenetic changes that causes transformation of normal epithelial cells in colorectal region to abnormally proliferating cancerous cells. Its carcinogenesis initiates with formation of polyps, a small outgrowth on the colon and rectum internal lining. Polyps which are considered as precursors of CRC are generally arise from epithelial tissues. This polyp formation is followed by well characterized genetic events that transform it into malignant cancerous cells. Current treatment strategy such as chemotherapy and radiotherapy used for treating colorectal cancer have some limitations such as off target effects, adverse side effects [3]. To overcome this limitations, innovative treatment methods are desperately required to improve the therapeutic strategy for individuals with this illness. Identifying phytochemicals that specifically targets particular molecular pathways that is linked to the colorectal cancer carcinogenic events is one promising area of research with a great potential of developing innovative therapeutic strategies [4]. The involvement of SRC gene which encodes for Src protein kinase aids in the carcinogenesis of colorectal cancer is becoming more frequently recognized. It has been seen that Src kinase has been linked with enhancing proliferation of tumor cell, survival, and metastasis and it is overexpressed in a considerable proportion of colorectal cancer cases. Overexpression and hyper activation of SRC gene in colorectal cells can

cause its transformation into malignant cells. Hence, inhibiting the expression of Src kinase is one possible strategy to treat CRC [5].

Salvia officinalis L. (sage) is herbal medicinal herb which has been traditionally used due to its unique qualities such as antibacterial, antioxidant and anti-inflammatory qualities. Its shows potential anticancer effects due to the presence of potential bioactive phytochemicals. *Salvia officinalis L.* contains phytochemicals that may provide an effective natural source of Src kinase inhibitors, which may be developed into potent CRC therapeutic agents [6].

1.2 Objectives:

The main purpose of this study is to discover potential phytochemicals from *Salvia officinalis L.* that shows potential inhibitory activity against Src kinase, thus can be used as promising therapeutic agent against Colorectal cancer. This objective can be achieved step by step analysis involved in network pharmacology approach along with bioinformatics analysis.

- Determine the potential bioactive phytochemicals present within *Salvia officinalis L.* (Medicinal plant)
- Identify effective targets of these bioactive substances in disease which promotes carcinogenesis in colorectal cancer
- Examine the linkages that exist between the phytochemicals and the targets that they are expected to affect.
- Conducting the toxicity analysis of bioactive compounds to demonstrate its drug likeness property, i.e., whether the bioactive compound in query qualify to be used as oral drug medications.
- Performing lipinski rule of five. These five rules of Lipinski help in determining which substances have higher chances of getting absorbed by human body via gastrointestinal tract when consumed orally. It is a crucial step for drug discovery and development.
- Assess the binding affinity and effectiveness of this phytochemicals against Src kinase via bioinformatics analysis
- Predict the most promising compound for further in vitro and in vivo research.

1.3 Research hypothesis and pipeline:

The research pipeline involves the prediction of bioactive compounds obtained from medicinal plant which is effective against specific disease by using Network pharmacology and Bioinformatics analysis (see fig . IMPPAT and KNApSAcK are online plant databases that were used to determine all of the available phytochemicals in various parts of *Salvia officinalis L.* Among the 729 phytochemicals, we filtered out effective phytochemicals based on their drug likeness value and bioavailability score. Toxicity

analysis was performed to analyze immunotoxicity, mutagenicity, hepatotoxicity and cytotoxicity of a compound. The bioactive compound must be inactive in all of these four parameters. The phytochemicals predicted by this approach should not violate Lipinski rule of five that includes having molecular weight less than 500 Daltons, Hydrogen bind donor must be less than five, hydrogen bond acceptor must be less than ten and logP value (partition coefficient) should be below five. These criteria determine whether a particular bioactive drug can be used as a viable oral drug. Then, OMIM and Genecard databases were used to find the disease target of CRC. Venn diagram was constructed that showed 12 targets as the common targets present between compounds and colorectal cancer disease. String database helps in visualization and understanding of protein-protein interaction network. Then, we created and visualized network of drug-target pathway and PPI using Cytoscape, these were subjected to KEGG and GO enrichment analysis. Key targets involved biological mechanisms are revealed by GO enrichment analysis, and their functional pathway annotations are revealed by KEGG enrichment analysis. This step is followed by, molecular docking simulations to analyze the affinity of interaction between potential phytochemicals and its target Src kinase which was done using Pyrx software. Pyrx used for performing molecular docking is a powerful computational tool for new drug discovery as it predicts the most promising bioactive agent against the target protein and how it binds with its target thereby saving both resources and time.

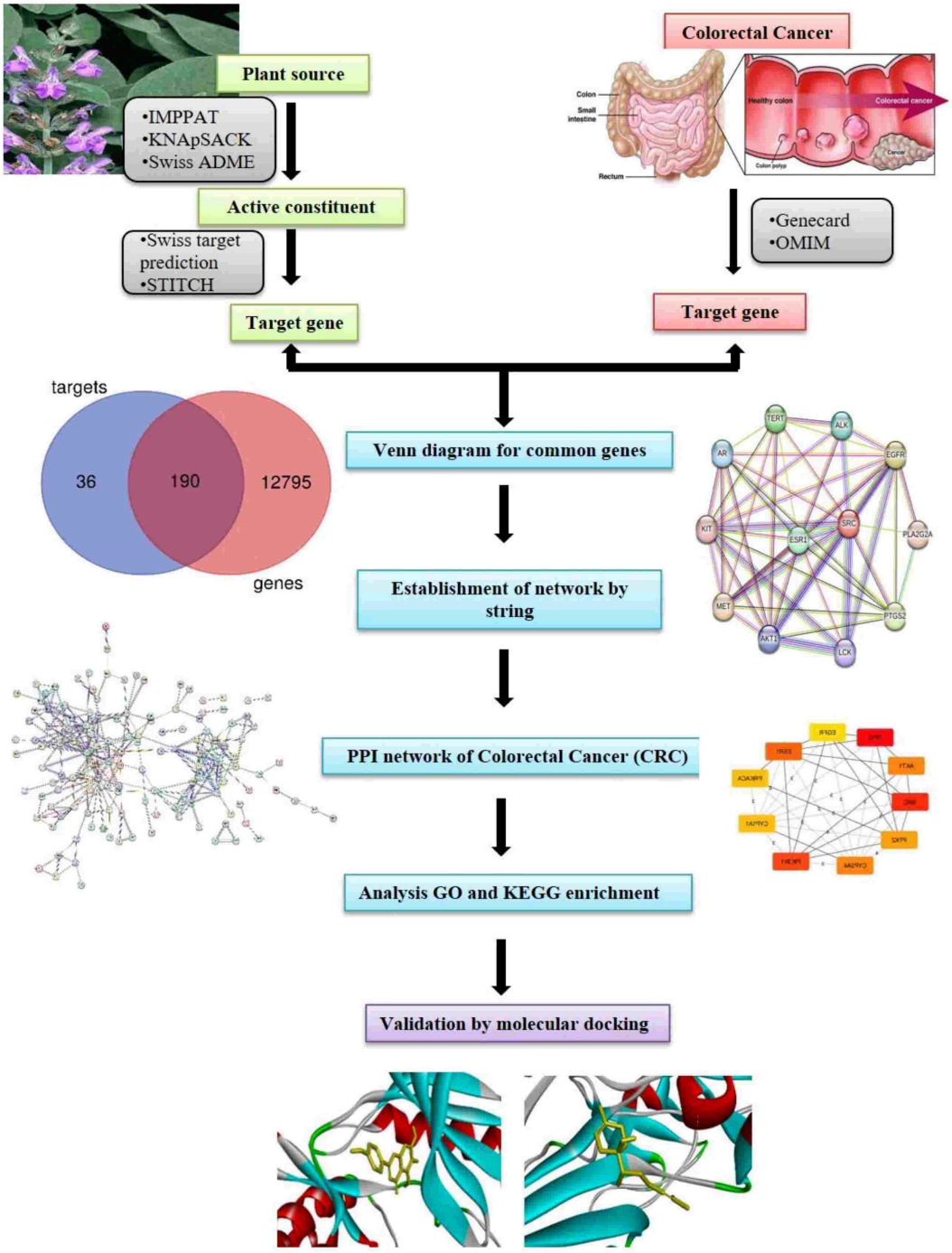


Figure 1: Flowchart depicting the steps involved in network pharmacology method followed by molecular docking

Table 1: List of databases utilized in Network pharmacology approach

S. no.	Database and software	website
1.	IMPPAT	https://cb.imsc.res.in/imppat/
2.	KNApSAcK	http://www.knapsackfamily.com/KNApSAcK/
3.	Molsoft L.L.C.	https://www.molsoft.com
4.	SwissADME	http://www.swissadme.ch
5.	Swiss Target Prediction	https://www.swisstargetprediction.ch
6.	Genecards	https://www.genecards.org
7.	OMIM	https://www.omim.org
8.	Vennplot	https://bioinformatics.psb.ugent.be/webtools/Venn/
9.	STRING	https://string-db.org
10.	Cytoscape 3.10.1	https://cytoscape.org
11.	Protein Data Bank	https://www.rcsb.org

1.4 Thesis Outline

Chapter 2: A brief description of theoretical aspects explaining the malignancy rate of CRC, importance of advancement in drug development and strategy for the therapy of this disease.

Chapter 3: This chapter provides explanation of methodology and various online platforms and web-tools used for identifying and validating effectiveness of proposed bioactive compounds against CRC. This chapter describes the online tools used for prediction and validation of ligand-protein interaction. Description of the mode of

operation used by each online server and its threshold value is also provided in each subsection.

Chapter 4: This chapter shows the results of each prediction and evaluation steps.

Chapter 5: The discussion, importance, and interference of the achieved technical results are given in this chapter.

Chapter 6: This section provides a summary of the results obtained from this study and mentions the future perspective and limitations of the current study.

CHAPTER-2

LITERATURE REVIEW

2.1 Colorectal Cancer

The third common type of cancer in the US is colorectal cancer and over the coming decades, an increase in its occurrence is expected worldwide. The risk of colorectal cancer in individual is 4%-5% globally (see Table 1). Predictions regarding patient survival or response to therapy are not possible due to the mutational landscape of colorectal cancer (CRC). Significant progress has been achieved in comprehending the molecular processes that causes the development of cancer and adenomatous polyps [7]. Colorectal cancer is a multistep carcinogenesis, basically it is a combined impact of several successive genetic changes. These changes can be inherited, as in genetic cancer predisposition syndromes, or acquired, like the sporadic forms. The formation of CRC involves occurrence of histopathological changes in epithelial cells present in mucosal lining of colorectal cells. Initially, hyperplasia develops which causes thickening of mucosal surface, this is followed by atypical hyperplasia which shows abnormal features of cells along with excessive proliferation [8]. Persistence of this atypical hyperplasia leads to the formation of adenomas in colorectal mucosa epithelial cells that breaks down normal regulatory mechanism completely can ultimately progresses into carcinoma. Genetic mutations either somatic (also known as acquired) or germline (also known as inherited) are primary factors responsible for progression of cells from precancerous to carcinoma stage, cells at carcinoma stage are highly invasive and can move to the distant parts of the body. This process begins with the factors that induces changes in the structural makeup of DNA by binding to it. Traditional cancer treatment options like radiotherapy and chemotherapy involves utilization of high energy radiation and cytotoxic drugs which pose some limitations such as off-target effects, meaning they cannot distinguish between the normal cells and affected or diseased cells. Various symptoms such as nausea, vomiting, immunosuppression and hair loss are caused by the off targets effects of therapy. While chemotherapeutic drugs primarily target cancer cells, they can also inadvertently harm surrounding healthy cells. [9].

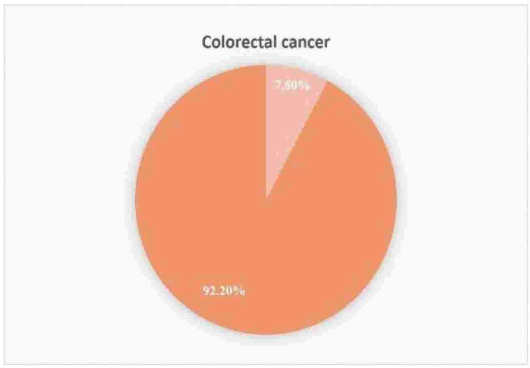


Figure 2.1: CRC accounts for 7.8% of newly diagnosed cancer cases in year 2023 in the US

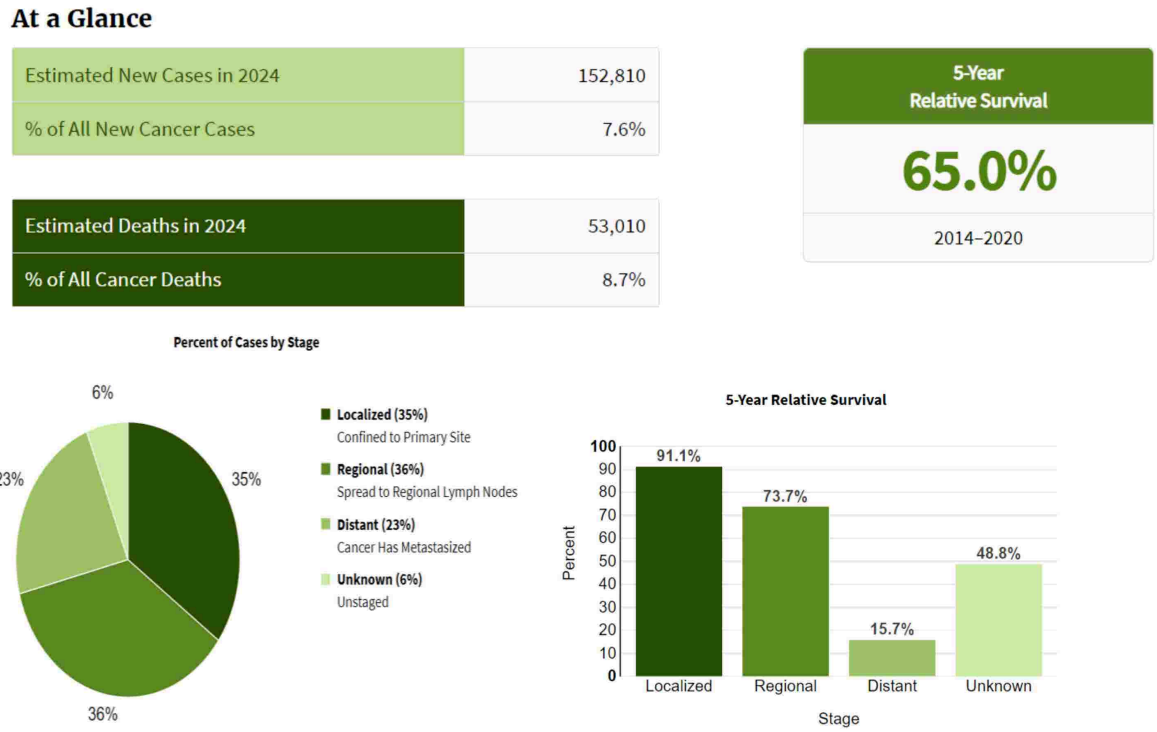


Figure 2.2: a) According to data obtained from National Cancer Institute, it is estimated that around 53,010 deaths will occur due to colorectal cancer and 1,52,810 new CRC cases will emerge in 2024. Overall survival rate of individuals suffering from CRC is 65% from the year 2014 to the year 2020 b) Pie chart and Bar graph is showing the survival by stage which interprets the extent of the colorectal cancer in the body during diagnosis, or at a stage of cancer, which influences both the length of survival and the available treatment options. This data is obtained from National Cancer Institute, NIH.

Different stages of Cancer:

- Localized: Localized cancer refers to the cancer which is present in tissue or organ where it began and has not yet invaded or reached to nearby body parts. It is also referred to as stage 1 cancer. In figure b) 91.1% are of localized cases, which remain confined to primary sites
- Regional/Distant Cancer: It represents the advanced stages of cancer progression. Cancer which has spread to the nearby lymph node and distant locations within the body such as liver, lungs and bone is known as regional or distant cancer.

2.2 *Salvia officinalis L.*

Salvia officinalis L. is a medicinal plant or Sage plant is a member of Labiatae/Lamiaceae family. Bioactive compounds have been abundantly available in this medicinal plants. This herb has been used for many different applications, such as medicine and diet. Although this genus of plants is found all over the world, its species are regional to Mediterranean and Middle East regions. [10]. Salvia species are rich in biologically active substances majorly classified into phenolic components, monoterpenes, diterpenes, and triterpenes. These compounds possess unique pharmacological activities, including anticancer, antioxidative, antimicrobial, and antimutagenic properties, making them potential treatments for various disorders. [2]. When examining the phytochemical content of *S. officinalis L*, it can be observed that linalool is the most abundant in the stem, α-pinene and cineole and some others are highest in the flowers, and bornyl acetate, limonene, camphene, humulene, and thujone etc. are the most abundant in the leaves [11].

There are many common names for S. officinalis. The most well-known names for this herb are culinary sage, dalmatian sage, true sage, garden sage, golden sage, and broadleaf sage. Purple sage and red sage are two types that are grown.

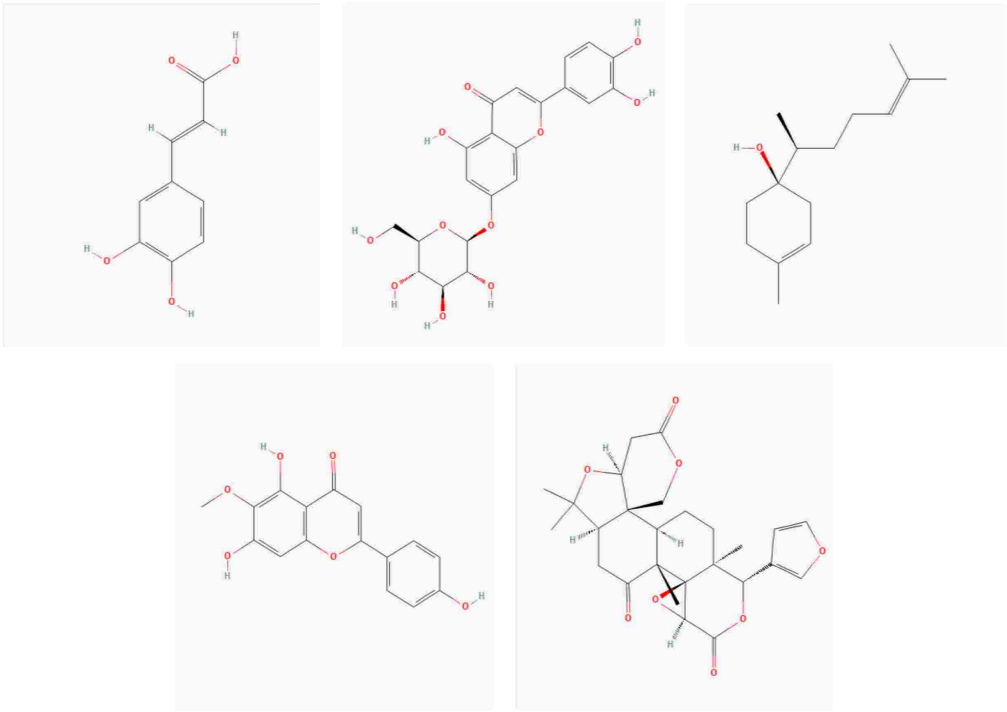


Figure 2.3: 2D structure of a) Caffeic acid b) Cynaroside c) 6 Epi-beta bisabolol d) Hispidulin e) Limonin

Studies have shown that extract of sage plant inhibits the process of angiogenesis (formation of new blood vessels) in vivo at optimum concentrations, suggesting that this could be a novel avenue for the development of a new and novel anti-angiogenic medication [12]. New blood vessels development from previous vessels is referred as angiogenesis, this promotes development and survival of cancer cells by providing the factors and oxygen required for its growth.

Biologically active compound present in Sage that can be effective against colorectal cancer include caffeic acid, luteolin, quercetin, limonin etc (see fig 2.3). The polyphenol content was present in high concentration in sage leaf extracts, then in aerial parts followed by its other parts. *S. officinalis L.* has been studied for potential anticancer effects on a variety of malignant cell lines and animal cancer models. Results against these cell lines showed pro-apoptotic and growth inhibitory effects promoting cell destruction and avoiding its proliferation in tissue. It has been seen that drinking sage tea can stop CRC in its early stages. [13]. In rats, the impact of consuming sage herbal tea on the prevention of CRC carcinogenesis was investigated. The DNA damage induced by oxidative H2O2 within in-vitro condition was found to be significantly minimized due to exposure of extracts obtained from *S. officinalis L.* Phytochemicals exhibit anti-carcinogenic effects by inhibiting the process of mitosis and promoting apoptosis (programmed cell death) during the early stages of cancer progression. [14]. Network pharmacology has been

recently adopted as practical tool to understand the mechanisms of herbal medicine to identify potential targets based on multi-compound and multi-target theory.

2.3 Effect of SRC gene in colorectal cancer carcinogenesis

Sample analysis revealed in 80% of CRC patients, SRC gene have shown overexpression. SRC kinase (a proto oncogene) is encoded by SRC gene. This kinase protein has tyrosine kinase domain at its C-terminus. Tyrosine kinase is categorized into two types: receptor tyrosine kinases (RTKs) and non-receptor tyrosine kinases (NRTKs). SRC gene forms non-receptor tyrosine kinase.

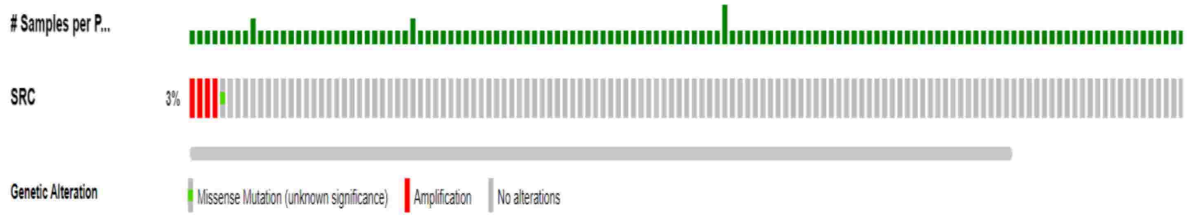


Figure 2.4: This data presented in this study interprets the result of 165 samples obtained from 161 patients to analyze query of SRC gene in CRC using cBioPortal platform. Among all the patients 3% of patients showed amplification and missense mutation in SRC gene resulting in overexpression of Src kinase protein.

cBioPortal serves as an extensive online tool or portal for analyzing multidimensional cancer genomics data, which enables researchers to look into intricate genetic changes in a range of different cancer types. By analyzing the query gene in this portal, it provides a wide range of interpretation including pathways involved, which genes is over expressed, reason behind the overexpression of a particular gene by analyzing the sample obtained from patients suffering from particular disease. Our findings in this portal showed that the SRC gene had genetic changes in about 3% of the patients (Total sample of patients-195) (see fig 2.4). In particular, these changes comprised:

- Gene amplifications: Gene amplification results due to multiple amplification of particular gene segment, which causes increase in its number. An increase in copy no. of SRC gene occurs in CRC. Overexpression of SRC gene increases the amount of Src kinase protein, which induces and causes dysregulation of multiple downstream pathways as a result leading to carcinogenic processes. Its overexpression causes significant changes within normal cells, converting it into cancerous cells
- Missense mutations: Missense mutation occurs when one amino acid is replaced by another amino acid in the polypeptide chain affecting its structure and function. These are single nucleotide alterations in the SRC gene's DNA sequence that cause

the SRC protein to contain a different amino acid. These mutations have the potential to change the function of protein, which may cause development of cancer.

Thus, overexpression of SRC gene causes uncontrolled division of cells, resistance to cell death, increased angiogenesis. These findings suggest that SRC holds promise as a both therapeutic target as well as a biomarker in colorectal cancer, offering new strategies for personalized treatment therapy. Targeting and inhibiting SRC to treat CRC has full therapeutic potential, but more research investigations and clinical validation are needed [15]. SRC gene has over more than 30 substrates, which includes transcription factors, RNA-binding proteins, and signal transducers. It is involved in various signaling pathways including PI3K/Akt pathway, MAPK, Stat3, IL-8, and VEGF, as well as pathways for cytoskeleton formation, which control cellular activities and migration [16]. Despite the tight regulation of their activity in normal cells, alternations such as mutations in target gene, its overexpression, and autocrine paracrine stimulation can induce normal cells to acquire transforming functions, which can result in cancer. SRC also upregulates p-gp causing increase in resistance to drugs. P-gp stands for P-glycoprotein which removes the absorbed drug from the intestine back into the lumen.

As compared to normal intestinal tissues, colorectal tumors express higher levels of SRC gene expressing Src protein kinase, which induces carcinogenesis. Protein tyrosine kinase phosphorylates tyrosine residues in target protein which transfers its signal. The majority of signal transduction cascades depends on NRTKs to control vital biological functions such as adhesion of cells, its proliferation, survival, cell division, gene expression regulation, and prevention of cell development [17]. Tyrosine kinases make up a significant fraction of all oncoproteins, which are involved in the transformation of normal cells into numerous cancers [18].

2.3.1 Structure of SRC gene:

Src is made up of the following segment: 14-carbon myristoyl group at N-terminus, an SH3 domain, an SH2 domain around N-terminus, a protein-tyrosine kinase domain at C-terminus and a C-terminal regulatory tail [19]. The small amino-terminal domain of Src kinase is composed of 267–337 aa residues, while its large carboxyl-terminal domain contains 341–520 aa residues. The primary function of smaller amino-terminal lobe of protein kinases is to anchor and orient ATP; the G-rich loop is a part of nucleotide-phosphate binding region. The β-sheet structure is primarily antiparallel in this smaller lobe. Overexpression of Src kinase occurs due to mutations in the SH3 and SH2 domains at N-terminus; Tyr527Phe mutation also increases Src enzyme activity. Tyrosine 416, which causes activation due to autophosphorylation, and tyrosine 527 causes inhibition due to phosphorylation by Src kinase at C-terminal, these two are main phosphorylation region of Src kinase protein. SH3 and SH2 are Src homology (SH) domains. SH2 function as protein-protein interaction domain and phosphor-tyrosine binding. SH3 function as

both recognition domain for substrate and linker interacting domain. The phosphotyrosine 527 and SH2 domain form salt bridge, while the polyproline type II left-handed helix present at SH3 domain attaches to the kinase domain. A dormant enzyme conformation is stabilized by both SH2 and SH3 domains, which are located away from the active site of enzyme usually at the backside of the kinase domain [20] (see fig 2.5).

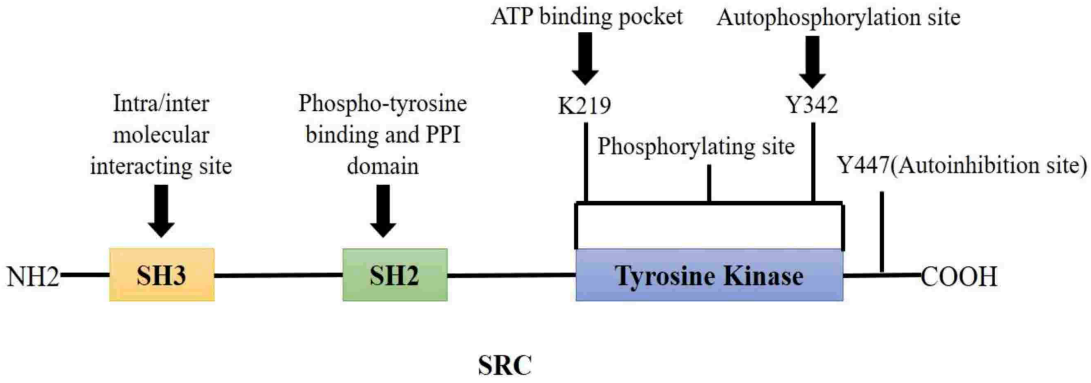


Figure 2.5: SRC kinase structure

Various cytokines and growth hormones binds to target receptor and positively stimulates Src kinase protein which in turn activates or inhibits multiple signaling pathways such as mitogen activated protein kinase (MAPK), phosphatidylinositide 3 kinase, signal transducer and activator of transcription 3, endothelial growth pathways, vascular endothelial growth factor etc. It also regulates pathways involved in formation of cytoskeleton.

2.3.2 Pathways regulated by Src kinase

Src regulates the target protein by kinase activity, i.e., by phosphorylating the target protein. Src phosphorylates p85 subunit present within PI3K which causes its activation. Increase in expression level of PI3K in turn activates many downstream effector proteins such as mTor C1, NF-κB, GSK that play an important role in various cellular events. The second major pathway controlled by Src is Ras-Raf-MEK-ERK1/2 pathways. This pathway involves series of activation of downstream protein. Src activates Ras protein, this protein is a major regulator of Raf. Ras activates Raf, which in turn activate MEK. MEK activates ERK1/2. This ERK1/2 is a transcription factor which binds to DNA, thereby increasing the cell activity. Third major signaling pathway controlled or regulated by this gene include STAT3 pathway. Src phosphorylates STAT3 and induces the expression of STAT3 protein. This activated STAT3 acts as a transcription factor and enters inside the nucleus. Once it is inside the nucleus, it activates various effector proteins by binding with the DNA such as c-Myc, VEGF, Bcl2, IL-6 (see fig 2.6)

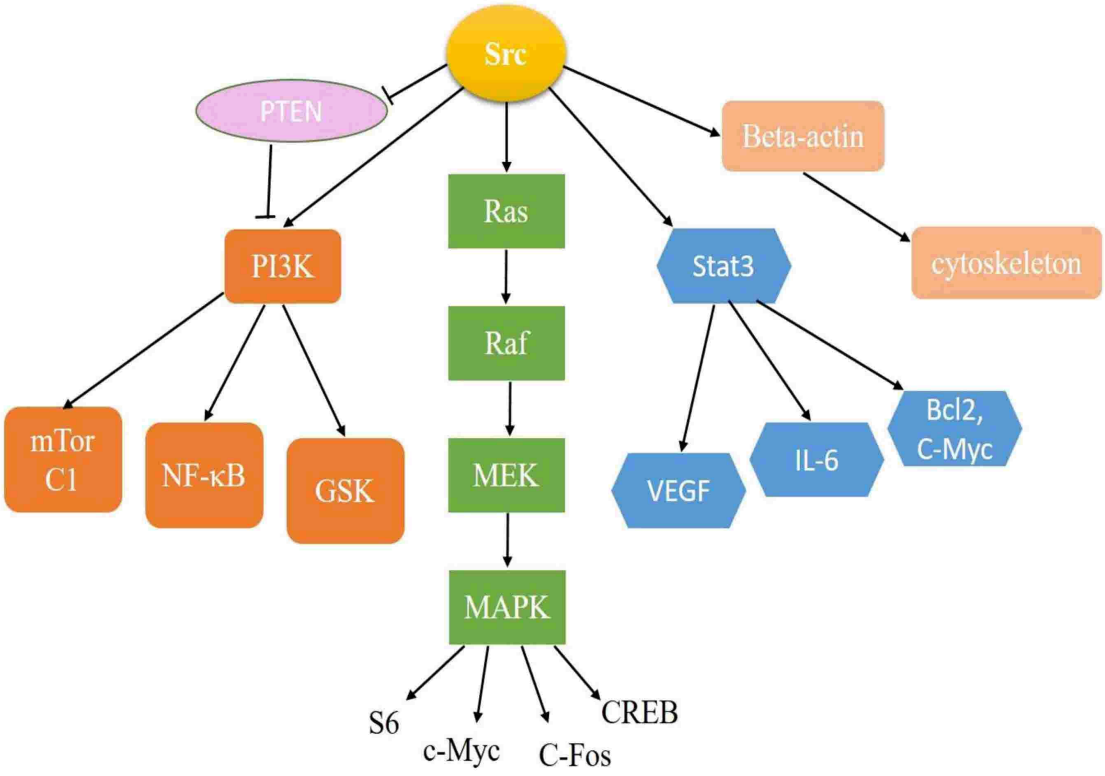


Figure 2.6: Major pathways of Src

Research indicates that overexpression of SRC gene in colon cancer can speed up the spread of the disease [21]. Its overexpression increases cell adhesion, migration and invasion indicating its role in cancer progression. Selective tyrosine kinase inhibitors can block constitutive oncogenic activation of Src in cancer cells, which makes it a practical approach that can be used for novel genome-based therapeutics [22]. Therefore, inhibiting SRC (non-receptor tyrosine kinase) could be helpful in treating colorectal cancer.

2.4 Lipinski Rule of five

Lipinski rule of five, also known as Ro5 evaluates the drug-likeness property of a bioactive compound. It was originally published in the year 1997 by Christopher A. Lipinski who developed this rule after noticing that the majority of medications are taken orally and it is small, somewhat lipophilic molecules [23]. It provides general guideline for assessing drug-likeness of a chemical compound, it also asses its pharmacological or biological activity that would likely make it qualify as oral drug to be used by humans.

Bioactive compounds that fulfills all the five criteria of Lipinski rule is considered as orally active drug for humans. This analysis should be carried out to evaluate whether a particular bioactive compound is a drug or non-drug. For the biologically active molecule

or compound to have the potential to be used as an oral medication, it needs to have molecular weight less than 500 Daltons (<500 Da), a compound must have less than five hydrogen bond donor (<5) and less than ten hydrogen bond acceptor (<10), An extensive analysis of ligands with several hydrogen bond donor groups indicates that it is very tough to accommodate more than two or three of these H-groups in a single ligand [24]. It is very difficult to obtain the exact geometry needed for the best possible hydrogen bonding. The accurate 3D alignment of the donor atoms and acceptor atoms is necessary for directional interactions like hydrogen bonds. One important factor in H-bonding is conformational flexibility of ligand molecule because flexible ligands have the ability to take on various conformations, not all of which are favorable for hydrogen bonding. On the other hand, stiff ligands show restricted degree of conformational freedom, which would make it harder to achieve the required conformational geometry. It is also intrinsically difficult to maintain the precise alignment required for multiple strong hydrogen bonds by balancing these factors [25]. Because of this, it increases the possibility that potential hydrogen bond donors will not positively contribute to ligand binding and shows the opposite effect as ligand binding increases in number. The calculated logP (partition coefficient) value should be below 5. The hydrophobicity of a compound is measured by its logP which shows how well the compound gets dissolved in fats vs water. If a compound has logP value above 5, it means compound is hydrophobic, get easily soluble in fat but not in water. This can affect its absorption by the GI tract. They may also be more likely to become stuck in lipid membranes, which would make it more difficult for them to get to the intended locations inside the body. According to the rule, drug is more likely to have inappropriate permeation and absorption if compound violates any of these five criteria [23]. We can estimate and analyze all these parameters using SwissADME online webserver.

2.5 Network pharmacology

By investigating into anti-CRC compounds and the related molecular mechanism of Salvia officinalis L. in teating CRC, the researchers hope to provide a theoretical framework for finding and creating novel drugs against diseases [22]. In order to understand the mechanisms behind the complementary therapeutic actions of herbal medicines, network pharmacology is an integrative in-silico method that uses a protein-ligand /disease-gene network. It basically elucidates the pharmacological effect of a phytochemical by targeting specific disease gene. As a result, the paradigm has changed from "one target, one drug" to "network target, multiple component therapeutics" [26] Because network pharmacology approaches address a drug effectiveness against multiple target key players in disease or its associated pathway involved in a disease. This approach allows us to identify multiple drug candidates against a particular target gene. In-silico toxicity analysis and efficacy testing of phytochemicals against the target protein is crucial

pre-step that saves time. Several databases have been established that provides on drugs, targets, pathways, interactions, and disease associations. Drug databases such as PubChem is used to obtain information about phytochemicals such as Canonical SMILES, its molecular weight, structural details etc. List of phytochemicals present in particular plant is acquired from IMPPAT ad KnapSacK databases. OMIM and Genecard databases are comprehensive database to understand human genes and genomes. This database provides a list of genes linked to specific diseases. The Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology are databases that covers diseases, drugs, chemicals, biological pathways, and genomes. A common tool for comprehending molecular networks and medication action mechanisms is the KEGG pathway and GO analysis. STRING database predicts interaction and its affinity between proteins. Output of protein interaction analysis is depicted with nodes and edges. It helps us to understand the broader aspects of drug targets.

The analysis of Salvia officinalis L. bioactive compounds with their anti-cancer effect through the use of molecular-docking, simulation, and network pharmacology sparked interest in the plant as a potential drug source [22]. Additionally, more wet lab testing is required to investigate the compound's potential pharmacological uses. Although computational techniques provide significant insights into prediction, but validation by wet lab experimentation is still essential for verification of these predictions and converting them into clinical applications. To validate the bioactivity of Salvia officinalis L. compounds against cancer cells, clarify their mechanisms of action, evaluate their safety and efficacy profiles, and optimize their pharmacokinetic properties, wet lab experiments—such as in vitro and in vivo studies—are crucial.

CHAPTER 3

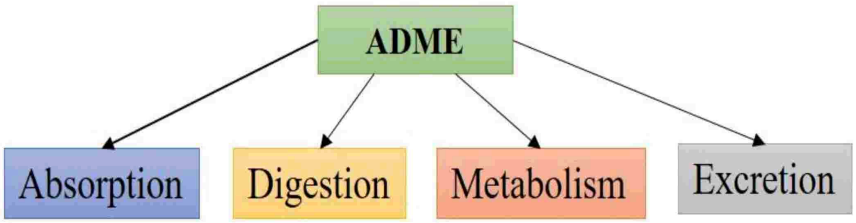
METHODOLOGY

3.1 Identification of active phytochemicals present in *Salvia officinalis L.*

In recent years, the research on plant metabolites has gained a lot of interest within the scientific community. Plant metabolites includes a wide range of bioactive compounds including Alkaloids, flavonoids, terpenoids and phenolics. These phytochemicals show immense potential for human health. In this study, we have taken Salvia officinalis L. to determine the effectiveness of its phytochemicals against CRC. The list of phytochemicals found in different portions of Salvia officinalis L. are all obtained from plant databases like KNApSACk and IMPPAT (Indian Medicinal Plants, Phytochemistry and Therapeutics) [27]. These plant databases provide wide range of information about bioactive phytochemicals present within the plants and biological sources associated with them. It provides information such as chemical structure of the phytochemicals, its molecular weight and physiochemical properties allowing scientists to carry out in depth analysis. These databases facilitate identification of potential therapeutic candidates by providing broad data on plant metabolites. Scientists can examine naturally occurring bioactive compound that exhibits medicinal properties and evaluate their use in drug discovery and development. Total 749 different phytochemicals are found in *Salvia officinalis L.* Software like Molsoft and SwissADME is used to analyze pharmacokinetics properties, bioavailability score and drug-likeness of these 749 phytochemicals by using their Canonical SMILES which are taken from PubChem database. In order to facilitate effective data processing and analysis, canonical SMILES offer a standardized representation of the molecular structure.

3.2 ADME analysis

ADME refers Absorption, Distribution, Metabolism and Excretion. Analysis of these four parameters play essential role in discovery and development of drugs, thereby determining the overall efficacy and safety profile of a drug candidate [22]. Drug must arrive at target site within the body in appropriate amount, must show improved bioavailability at target site and it should be in its bioactive form for sufficient amount of time to produce intended biological effect. SwissADME is an online tool available for free, offering rapid and reliable predictive models for pharmacokinetics, physicochemical properties, drug-likeness, and medicinal chemistry friendliness.[28] [29]. These models include in-house expertise from methods like the iLOGP, Bioavailability Radar, and BOILED-Egg.



Software SwissADME analyze pharmacokinetics properties, bioavailability score and drug-likeness of these 749 compounds by using their Canonical SMILES which is obtained from PubChem database. Drug likeness of a compound should be >0.18 and bioavailability score must be >0.55 for a compound to be qualify as viable oral drug. A compound must also qualify Lipinski's rule of five (Ro5). This analysis helps to identify promising drug candidates [30]. It was discovered that out of 749, only 22 phytochemicals met the requirement for drug-likeness and had appropriate ADME values.

3.3 Toxicity analysis

It is estimated that the chances of drug candidates are only 8% to become marketed drug medications, when they enter clinical trials, and the drug toxicities account for 20% of failures in late-stage drug development [31]. Computational prediction method provides multiple benefits during the drug development process. It saves time and resources, reduces the chances of drug failures due to prior predictions of its properties [32].

Using Protox 3.0 software, the toxicological properties of 30 non-toxic bioactive compounds were analyzed (refer to Appendix 2). This step is crucial to evaluate the effectiveness and safety of drug medications [33]. The main objective of this toxicity models is to save time and money invested in toxicity assessments by supplementing the current in vitro toxicity methods that predict the toxicity effects of chemicals [29]. The Protox 3.0 is an in-silico online tool that makes use of advanced machine learning algorithms to predict toxicity endpoints of a compound. We look for specific toxicity parameters in our drug candidates such as hepatotoxicity, mutagenicity, carcinogenicity and immunogenicity [22]. ProTox 3.0 models have demonstrated balanced ratio of performance taking into account specificity and sensitivity of the model respectively. Compared to commercial software, It provides better performance in all 61 models as indicated in the ProTox 3.0 web server. The 30 selected phytochemicals were subjected to toxicity analysis one by one using their Canonical smiles as input data on web page. Only 3 phytochemicals came out to be non-toxic in four essential parameters (Hepatotoxicity, Mutagenicity, Carcinogenicity, immunogenicity) [34].

3.4 Prediction of target genes associated with colorectal cancer

The next step is to identify the possible targets for these three phytochemicals so that we can treat colorectal cancer using them. The phytochemical targets of Salvia officinalis L. were predicted by Swiss Target Prediction website. This webtool makes it easier to estimate most probable targets of a phytochemical drugs and it contains over 3000 target proteins from different species and have an internal library of 3,70,000 known active compounds. Genecards and the OMIM (Online Mendelian Inheritance in Man) database were searched to find genes associated with colorectal cancer. [22]. Genecard provide extensive details about each gene that has been annotated and predicted, involving its functions, diseases that they are linked to, pathways, and its associated literature. It is an important tool for determining the no. of genes linked to particular illnesses, for eg: colorectal cancer. OMIM is an extensive, reliable, and up-to-date research resource. It provides carefully chosen descriptions of human genes, phenotypes, and their interactions. In January 2011, OMIM.org (http://omim.org), the organization's new official website, was launched. OMIM database is very useful database to interspersed and diverse communities of scientists, clinicians, molecular biologists, and genome scientists, as well as by students and educators in these fields [22]. It offers data that is derived from published, peer-reviewed biomedical literature. Both OMIM and genecards database display a list of genes associated with colorectal cancer. Comprehensive details about the linked genes, their mutations, along with their functions in colorectal cancer are included with each entry.

3.5 Establishment of Sage-CRC gene network

Venn diagram tools estimate common genes between disease and compound. It was used to identify which targets overlapped between the two. [22] [29]. Two datasets are required to generate vennplot: one dataset containing the genes associated with colorectal cancer and other datasets that represents target genes of the bioactive substances extracted from *Salvia officinalis L.* Vennplot showed 190 common genes indicating that these common genes are associated with colorectal cancer and are mainly targeted by phytochemicals obtained from *Salvia officinalis L.*

Using the STRING online tool, the protein-protein interaction network of 190 overlapped genes with the highest confidence (0.900) was created. After adding 190 common target genes to the DAVID database, the top five most likely pathways for colorectal cancer were found.

3.6 Biological functional analysis

Enrichr bioinformatics online tool, which analyze gene sets and perform functional enrichment analysis were used to find intersection targets of genes related to colorectal cancer and active phytochemicals were analyzed for enrichment analysis using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) [22]. Gene ontology data and KEGG enrichment analysis analyze genes and genomic data (refer to Appendix 2). Their examination helps in the identification and discovery of new biological roles, genotype-phenotype correlations, and illness mechanisms.

- Gene ontology: Multiple genes associated with biological processes, molecular functions, and cellular components are provided by GO. The hierarchical structure and careful regulation of these terms enable accurate and methodical gene annotation [35].
- KEGG pathway: KEGG (Kyoto Encyclopedia of Genes and Genomes) is an attempt to standardize gene annotations and computerize current understanding of cellular processes to establish a link between genomic data and higher order functional data [36].

There are various webtools available for GO and KEGG enrichment analysis. In our study, this enrichment analysis is done using Enrichr. This enrichment analysis tool, utilizing Fisher's exact test, is tailored for gene lists without ranking. It encompasses over 100 gene set databases, having a total of more than 180,000 gene sets across diverse categories that are known for its user-friendly interface, this tool offers extensive interactive reporting capabilities. [37].

3.7 Protein-Protein interaction network and critical subnetwork

Functional association between two proteins is the basic interaction unit displayed in STRING database, that collectively plays an important function in specific biological pathways. Protein interaction scores are represented as approximate confidence by STRING [38], [39]. Common genes obtained were used as an input in string database for homo sapiens. It predicts the interactions at different confidence level: 0.15 represents the low confidence, 0.4 shows medium confidence and 0.9 shows highest confidence score. We obtained our string interaction analysis at 0.9 (high confidence). Nodes represent protein, connected to each other via string that represents physical interaction and it provides result which is based on known and predicted protein-protein interaction. More significant a node is in the PPI network, the higher its degree value. As a result, the degree value of the network's nodes can be used to filter out hub genes. This database aims to set itself apart primarily by (i) having a high coverage range, (ii) being user-friendly, and (iii) having a uniform and accurate scoring system. Presently, it has largest collection of

proteins (24.6 million) and organisms (5090), along with a diverse array of benchmarked data sources, as well as intuitive online viewers for quick and easy access.

3.8 Construction of component target pathway network

We used this cytohubba plugin provided within Cytoscape to identify top hub genes in the PPI network, which was made up of different phytochemicals and the anti-CRC targets that went along with them. This cytohubba plugin identified the top 10 hub genes and phytochemicals involved in CRC. Hub genes are the genes that show high degree of interconnection between genes and other proteins present within the network. We can orient and arrange the interaction network in different patterns. Its result is shown in Cytoscape is a useful bioinformatics software platform available for free, which enables the visualization and interpretation of intricate networks of molecular interaction. [40], [41]. The main core of this software provide features to organize the data in different layouts and analyzing the network, integrating it visually with phenotypes, expression profiles, and other molecular states, and also connecting it to functional annotation databases. CytoHubba is plugin within Cytoscape software. It is an effective tool for locating hub genes, or important nodes within biological networks such as PPI networks [42]. These networks are essential for understanding the underlying molecular mechanisms of disease and locating its possible targets for treatment.

3.9 Molecular docking using Pyrx

Molecular docking was conducted to find out binding strength between the two top-ranking phytochemicals (Hispidulin and 6-Epi-beta-bisabolol) according to hub genes analysis and the most prominent target protein. PyRx software was utilized to determine affinity of interaction between the target protein Src kinase and phytochemical ligands. Protein-ligand binding mode and affinity between them is predicted by docking and drug discovery relies heavily on it. Structure of Src kinase encoded by SRC gene is taken in pdb format from RCSB Protein data bank. Removal of water molecules, ligands if present and removal of side chains is done by selecting and deleting to prepare target protein using Biovia Drug Discovery Studio Visualizer 2024…then target protein is saved in .pdb format. Resolution of this Src kinase is 2.2 Å resolution. The ligands (Hispidulin and 6-Epi-beta-bisabolol) were sourced from the PubChem database in 3D conformer provided in SDF format. These were then transformed into .pdbqt format using Open Babel, an open-source chemical toolkit available within PyRx software. Open Babel is an open source chemical toolbox that provides an easy interface to convert one chemical file format to another chemical file format. Proteins were further analyzed and docked within the PyRx software, here the energy of all ligands was minimized. The autogrid dimensions in Autodock Vina between the ligands and receptor tyrosine kinase atoms are X- 64.315 Å, Y- 68.403 Å and Z- 58.514 Å. To save computation time, a good scoring grid is

computed from the structure of ligands. Through docking simulations, various ligand binding affinities for the SRC gene are obtained, offering insights into their possible interactions and affinities. Different binding affinities of ligands for its receptor is then obtained. Subsequently, visual representations of the 2D and 3D interactions between these entities were generated utilizing Discovery Studio Visualizer. Visualization of this binding interaction helps to identify residues involved in bond formation, general structural features of ligand protein complexes, and binding mechanism. It offers prediction of significant affinity of phytochemicals against SRC gene.

CHAPTER 4

RESULTS

4.1 Screening of active compounds and potential targets

As per IMMPAT and KNApSACk databases, total 749 phytochemicals were discovered across different regions of *Salvia officinalis L.* After assessing the drug-likeness and bioavailability scores of each phytochemical using the Molsoft webtool and their canonical SMILES, we applied multiple filters to identify the most promising phytochemicals. Among them, 22 phytochemicals passed the filters of drug likeness >0.18 and oral bioavailability score of >0.55 (see table 4.1). Hence, approximately 4% of available phytochemicals qualified for toxicity analysis step. Oral Bioavailability depicts the degree of absorption of drug compound by the body and Drug likeness compares the bioactive compounds with known drugs and predicts the drug likeness of a compound. It is very crucial to conduct toxicity analysis of compounds to analyze its safety profile.

Table 4.1: Name of compounds that offers suitable drug likeness above 0.18 and have bioavailability score greater than 0.55. Name of molecule is provided with its Molecule ID.

MOL ID	Components	Drug-likeness	Bioavailability score
12300148	6-Epi-beta-bisabolol	0.48	0.55
161271	Salvigenin	0.45	0.55
5281792	Rosmarinic acid	0.37	0.56
5280445	Luteolin	0.38	0.55
5317284	Nepetin	0.44	0.55
188323	Cirsimaritin	0.47	0.55
5281617	Genkwanin	0.37	0.55
5281628	Hispidulin	0.46	0.55

Table 4.1 (continued)			
6918774	Corosolic acid	0.6	0.56
10494	Oleanolic acid	0.37	0.85
64945	Ursolic acid	0.66	0.85
5280704	Cosmosiin	0.59	0.55
222284	beta-Sitosterol	0.29	0.55
494501	7alpha-Acetoxyroyleanone	0.39	0.56
160237	Cirsiliol	0.56	0.55
1507228	5, 7-Dimethoxy-1-naphthol	0.48	0.56
11172	Octadecamethylcyclononasiloxane	0.61	0.55
442664	Vicenin-2	0.2	0.55
179482	2alpha,3alpha-Dihydroxyolean-12-en-28-oic acid	0.55	0.55
11869658	3-Epioleanolic acid	0.37	0.56
5280489	beta-Carotene	0.29	0.55
5280637	Cynaroside	0.6	0.55

Web tool such as Protox 3.0 software predict compounds toxicity that may be hazardous to humans, animals and plants. Comparing the ProTox 3.0 web server to current computational models reveals a number of advantages. Adverse outcomes pathways (AOPs), metabolism, and chemical and molecular target knowledge are all included in the ProTox webserver. This software can conduct toxicity analysis at different levels such as hepatotoxicity, carcinogenicity, mutagenicity, immunotoxicity. Out of 22 phytochemicals, Hispidulin, Cynaroside, 6-Epi-beta-bisabolol showed no toxicity in all these four parameters (see table 4.2). For a specific phytochemical to be deemed suitable for therapeutic purposes, it should show inactivity across all these parameters. This signifies that these three compounds didn't show any signs if inducing cancer, genetic

mutations, liver cells damage and unfavorable immune response. Table is showing the probability of inactivation of compounds in this four parameters.

Table 4.2: Only three phytochemicals out of 22 showed no toxicity in four essential parameters such as Hepatotoxicity, Carcinogenicity, Immunotoxicity and Mutagenicity

Phytochemicals	**Hepatotoxicity (Inactive probability)**	**Carcinogenicity (Inactive probability)**	**Immunotoxicity (Inactive probability)**	**Mutagenicity (Inactive probability)**
6-Epi-beta-bisabolol	0.81	0.70	0.86	0.75
Hispidulin	0.72	0.68	0.72	0.94
Cynaroside	0.82	0.85	0.74	0.76

Validation results of phytochemicals in the ProTox 3.0 (Toxicity prediction platform) showing probability of inactivation in terms of Hepatotoxicity, Carcinogenecity, Immunotoxicity and mutagenicity.

4.2 ADME analysis

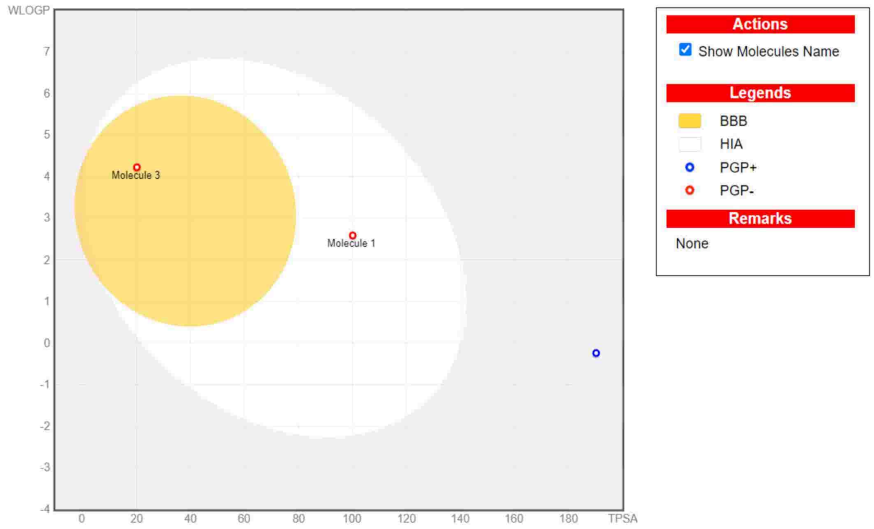


Figure 4.1: BOILED Egg analysis representing whether the bioactive compound will get absorbed effectively in GI tract and BBB. Molecule 1 is Hispidulin: it gets absorbed by the GI tract and Molecule 3 is 6-Epi-beta-bisabolol: it gets absorbed effectively and can even cross the BBB

ADME analysis represents evaluation of absorption, distribution, metabolism, excretion of drug compound that is crucial to evaluate during drug development process. It provides details regarding pharmacokinetics property of a compound using SWISS ADME web tool. The top scoring ligands Hispidulin, Cynaroside, 6-Epi-beta-bisabolol were subjected to ADME analysis. Its graphical output is represented as BOILED egg analysis that explains Hispidulin among the three absorbs effectively in gastrointestinal tract as shown in white region and 6-Epi-beta-bisabolol have the potential to cross blood-brain barrier as shown in yellow region of the egg (see fig 4.1). Cynaroside showed poor absorption that can't be absorbed in GI tract and it cannot cross BBB as shown as blue dot outside egg region (see fig 4.2).

Figure 4.2: Pharmacokinetic profile of ADME calculated using SWISSADME tool a) Cynaroside b) Hispidulin c) 6-Epi-beta-bisabolol

Hispidulin: The average lipophilicity (log po/w) of Hispidulin is 2.27 which indicates that it is moderately soluble in water. Skin permeation (log Kp) value is -6.01cm/s. 0.55 is bioavailability score and its synthetic accessibility is 3.12 respectively. It reveals no violation of lipinski's rule, therefore categorized as drug likeness compound. Xenobiotics and drugs are pumped back into the gastrointestinal lumen by P-gp, an efflux transporter. Hence, it reduces the concentrations of drugs in plasma and tissue (see table 4.3). Hispidulin is PGP- (shown as red color), means it has good bioavailability, gets absorbed in the tissue and plasma effectively and is not pumped back by P-gp transporter and it shows high absorption in the gastrointestinal tract.

6-Epi-beta-bisabolol: The average lipophilicity (log po/w) of 6-Epi-beta-bisabolol is 3.54 and it is moderately soluble in water. Its skin permeation (log Kp) is -4.08cm/s. The bioavailability score is 0.55 and synthetic accessibility is 4.06 respectively. It reveals no violation of lipinski's rule, therefore categorized as drug likeness compound. 6-Epi-beta-bisabolol is PGP- (shown as red color), means it has good bioavailability, gets absorbed in the tissue and plasma effectively and is not pumped back by P-gp transporter and it can cross blood brain barrier (BBB).

Cynaroside: Cynaroside does not qualify as viable drug delivery agent because it is violating Lipinski rule of five. It exhibits more than five hydrogen donor, i.e., 7 and more than 10 hydrogen acceptor, i.e., 11. Hence, it is violating more than one rule of Lipinski and cannot qualify as oral drug agent.

Table 4.3: ADME analysis

ADME analysis			
Parameters	**Hispidulin**	**6-Epi-beta-bisabolol**	**Cynaroside**
Molecular weight (<500 Da)	300.26 g/mol	222.37 g/mol	448.38 g/mol
LogP value (<5)	2.27	3.54	1.83
No. of H-bond donors (5)	3	1	7
No. of H-bond acceptor (<10)	6	1	11
No. of rotatable bonds (<10)	2	4	4

4.3 Target profiling of Salvia officinalis L. and CRC

Using the SwissTargetPrediction database, possible protein targets for *Salvia officinalis L.* is identified. A sum of 226 distinct target genes was displayed and collected for three bioactive phytochemicals derived from *Salvia officinalis L.,* each with a probability score greater than zero. An extensive search carried out using GeneCards, a database that offers thorough information about genes, their functions, and related diseases, to find possible therapeutic targets for Colorectal cancer. Vennplot showed the common targets for the CRC (see fig 4.3). It predicts possible target genes by narrowing down the list and identify genes that might be an important factor for a disease. 190 target genes in all were found to be viable targets and were subsequently used in the research that followed (refer to Appendix 1).

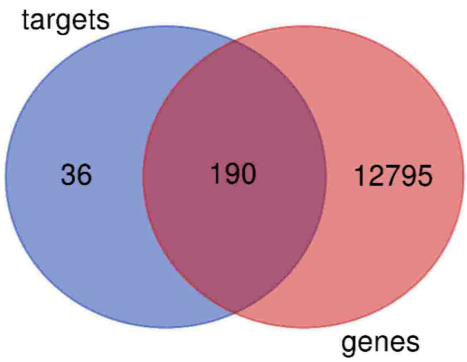


Figure 4.3: The overlapping targets were identified between the potential protein targets of two active phytochemicals in *Salvia officinalis L.* and the genes associated with Colorectal Cancer (CRC)

4.4 Network of protein-protein interaction and core subnetwork

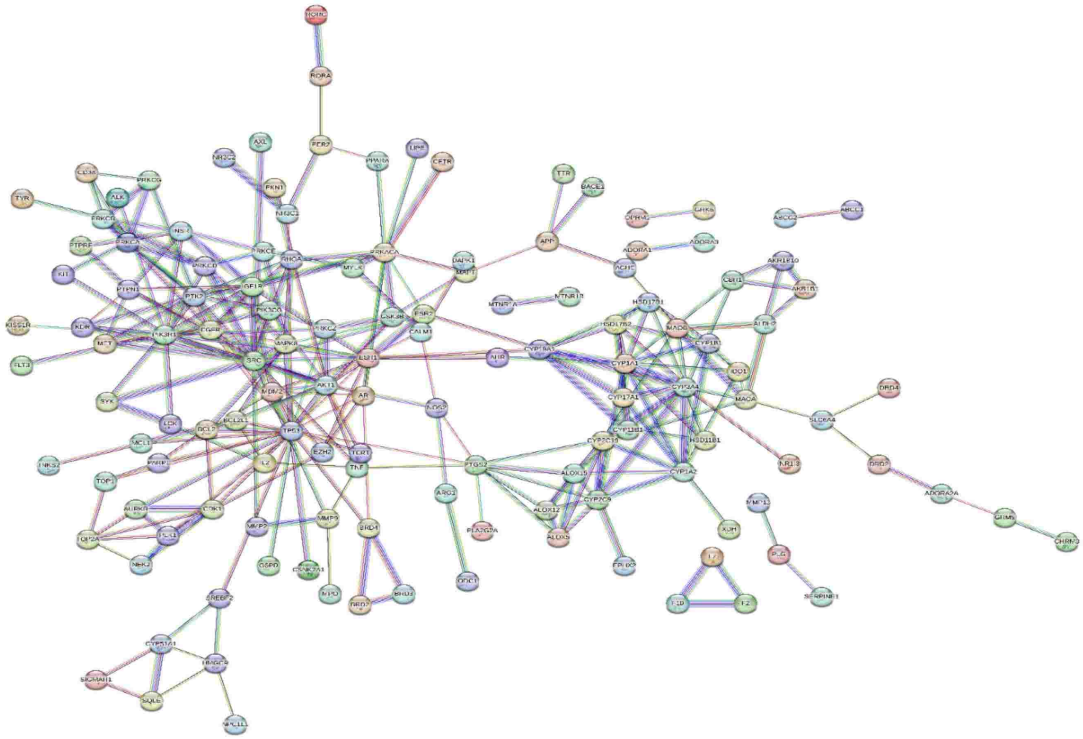


Figure 4.4: Network of Protein-Protein interaction of 190 intersecting gene target made using STRING database

190 common genes obtained in venn plot were used as an input in string database search. Homo sapiens were selected as an input so that it will show interactions which are specific to human proteins. The "highest" (0.900) confidence cutoff has been set for visibility of interaction link. With a list of the proteins as input, STRING can find the proteins that directly interact with the inputted proteins via physical interactions, or their neighbor interactors. From there, STRING can create a PPI network that includes every protein and every interaction that occurs between 285 edges, as seen between protein nodes. In fig 4.4 nodes display proteins and edges display interactions, where the degree of strength interactions or confidence can be indicated by thickness of the edges. The interactions between the target protein comprises 190 nodes, 285 edges and 3.02 average node degree see table 4.4. Degree is crucial in illustrating the interactions between proteins and network nodes.

Table 4.4: Network statistics obtained from String database

Total number of nodes	190
Total number of edges	285

Table 4.4 (continued)	
Average degree of node	3.02
Mean local clustering coefficient	0.406
p value of PPI enrichment	< 1.0e-16

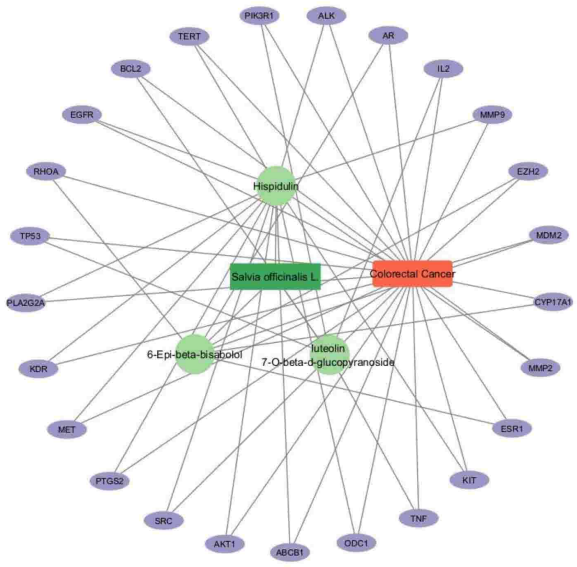


Figure 4.5: Network of plants, its active phytochemicals and common gene targets associated with in disease. Different categories are represented by different color of nodes. Light green is used for active phytochemicals, Dark green for herb *Salvia officinalis L.* Purple indicates common target and Colorectal cancer (disease) is represented by orange.

4.5 GO and KEGG enrichment analysis

To gain more knowledge about the biological activity and function of the targets which are targeted by the phytochemical drugs, an analysis of gene function annotation was done. KEGG and GO enrichment analysis was done using Enrichr online webtool. The 190 common genes obtained between the phytochemical and target disease were used as an input to pinpoint all the critical pathways involved. The highest and topmost 10 enriched annotations were selected based on the targeted genes paired with their most effective ligands, as they appeared within the top 10 enriched ontologies and the 10 pathways having the lowest p-value, which suggest a higher level of confidence within the data. This p value represents the interaction strength between the phytochemicals and the biological pathways.

In KEGG pathway analysis showed pathways in cancer, Chemical carcinogenesis, insulin resistance, proteoglycan, sphingolipid signaling pathway and so on as shown in fig 4.6 (a).

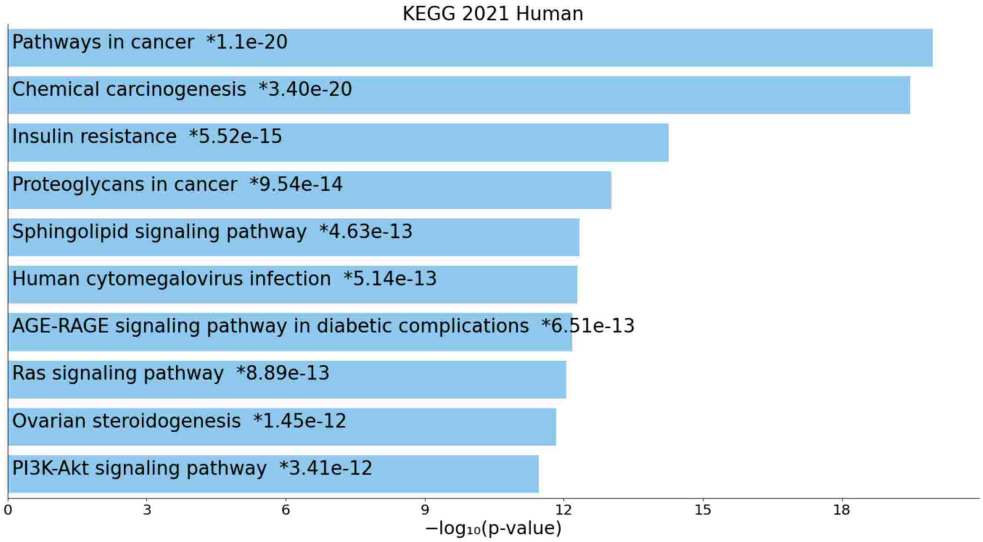


Figure 4.6 (a): A bar chart illustrating the top 10 enriched terms from the KEGG_2021 Human gene set library: The top 10 terms shows enriched pathways for the input gene set, are presented based on the -log10(p-value), with the respective actual p-values displayed alongside each term. The pathway ranking highest at the top shows the most substantial overlap with the input query gene set.

Biological process (BP), cellular component (CC) and molecular function (MF) are three categories described by GO enrichment analysis. According to output generated: In biological process, potential targets were primarily focused on responding to protein phosphorylation process, Negative regulation of apoptosis, protein modification process and so on. In cellular components, potential targets were involved in membrane raft, Neuron projection, Endoplasmic reticulum membrane, Intracellular membrane bound organelle and many more. In case of molecular function, potential targets played significant role in Protein tyrosine kinase, Protein serine/threonine kinase activity, Steroid hydroxylase activity and so on.

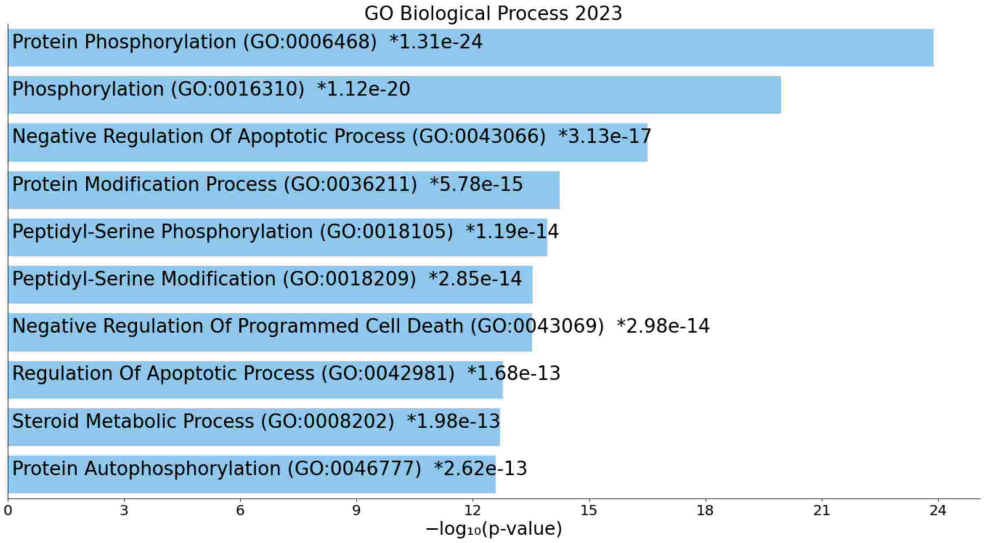


Fig 4.6 (b): A bar chart representing top enriched terms from the GO biological process 2023 gene set library: The topmost 10 terms were enriched for the input gene set which are presented based on the -log10(p-value), and its corresponding actual p-values indicated next to each process. The process positioned at the top indicates the most substantial overlap with the input query gene set.

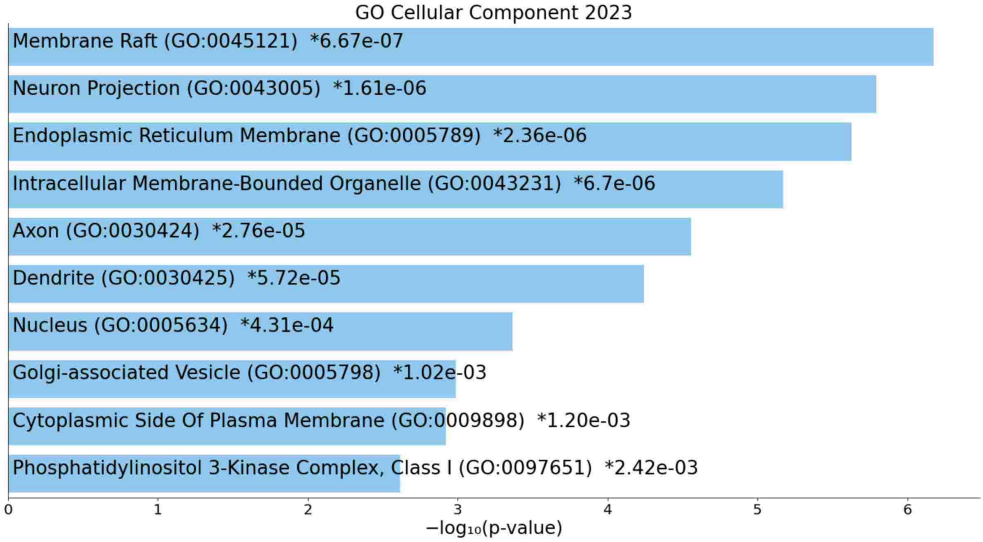


Fig 4.6 (c): A bar chart illustrating top 10 enriched cellular component from the GO Cellular Component 2023 gene set library. The top 10 terms enriched for the input gene set are presented based on their -log10(p-value) along with respective actual p-values provided after each component. The term positioned at the top indicates the most substantial overlap with the input set of query gene.

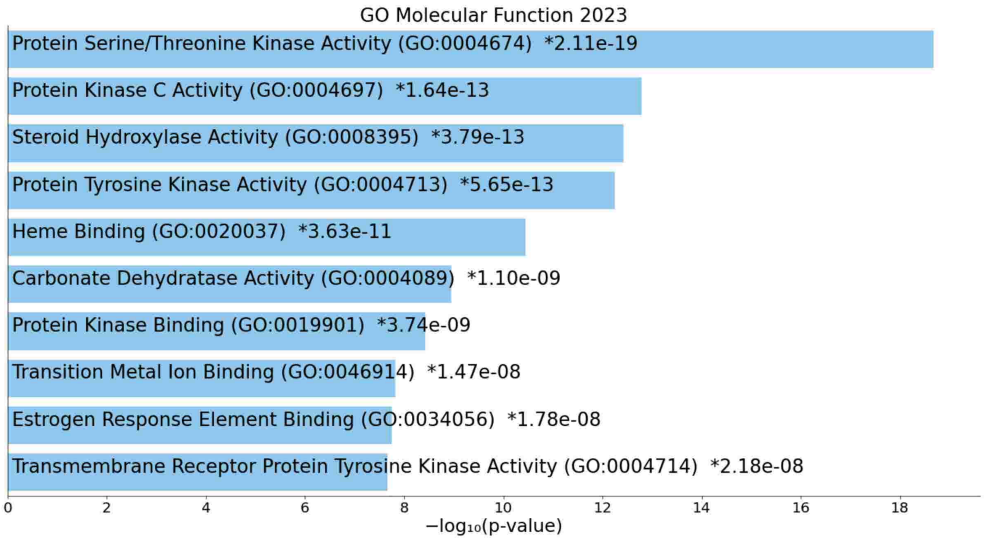


Figure 4.6 (d): A bar chart showing the top 10 enriched functions from the GO Molecular Function 2023 gene set library. The top 10 functions enriched for the input gene set are presented according to their -log10(p-value), with the corresponding actual p-values displayed next to each term. The term positioned at the top reflects the most substantial overlap with the input set of query gene.

The GO analysis results showed that the potential targets were predominantly enriched in the biological process related to responding to protein phosphorylation., Membrane raft in cellular function and protein tyrosine kinase activity in molecular function. Moreover, the KEGG results were directly related to the pathways in cancer (see fig. 4.6 (b,c,d))

4.6 Component target pathway network construction

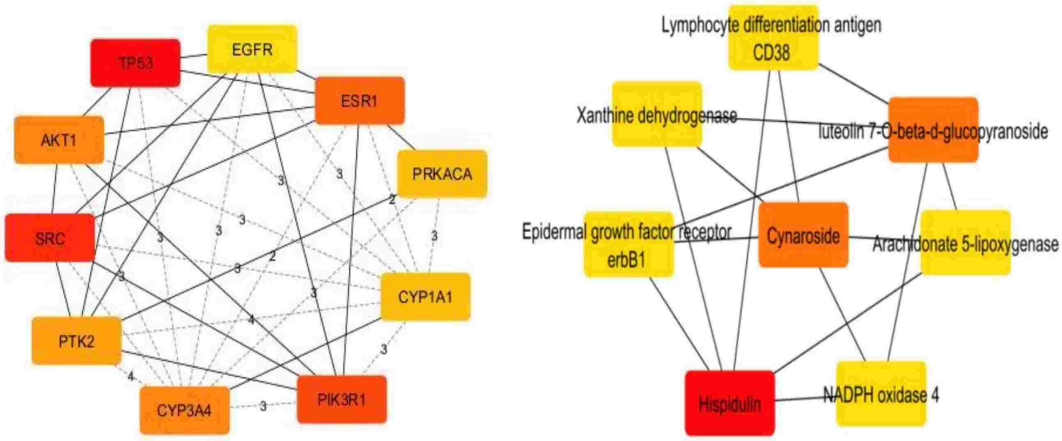


Figure 4.7: The interaction network of disease target and phytochemicals

Interaction network between active phytochemicals in *Salvia officinalis L.* and anti-CRC target genes was constructed and visualized by Cytoscape software, displaying the highest ranking gene as shown in red color (see fig 4.7). CytoHubba is a plugin present within the cytoscape software that provides the details on hub gene ranking and network. This was accomplished by using sophisticated network analysis made possible by the CytoHubba plugin, which is compatible with Cytoscape version 3.10.1.

Top ranking phytochemicals and target genes effective against a disease was determined using this software. Degree method showed ranking of topmost 10 genes among all the genes and top 9 phytochemicals. The Src gene which encodes for Src protein kinase ranked highest among the top 10 genes, with a score of 121., followed by TNF, TP53, EGFR, AKT1 and other genes as shown in bar graph (see fig 4.8).

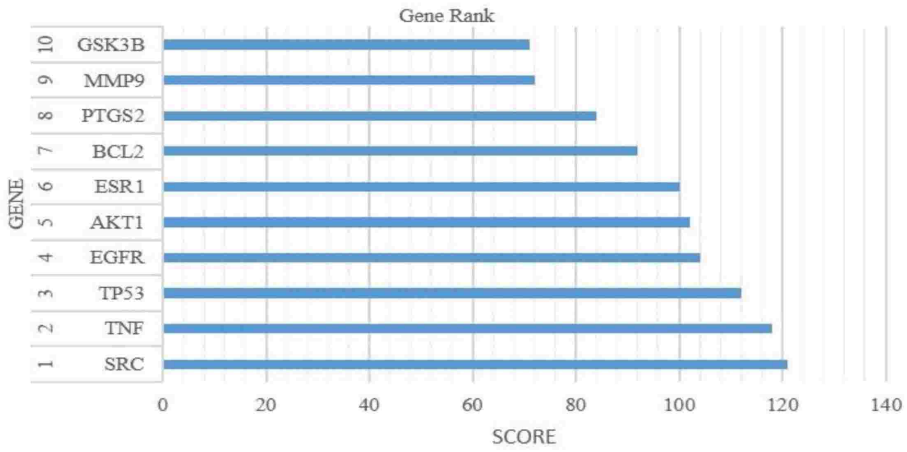


Figure 4.8: a) The top 10 target genes were selected based on their degree in the hub network. SRC achieved the highest score of 121, making it the top-ranking gene among all target genes.

4.7 Molecular Docking

Among all 749 phytochemicals in Sage plant, the top 2 selected phytochemicals Hispidulin and 6-Epi-beta-bisabolol showed significant affinity of interaction between the phytochemicals and Src kinase encode by SRC gene ranging between -7.9 to -9 kcal/mol (see table 4.5).

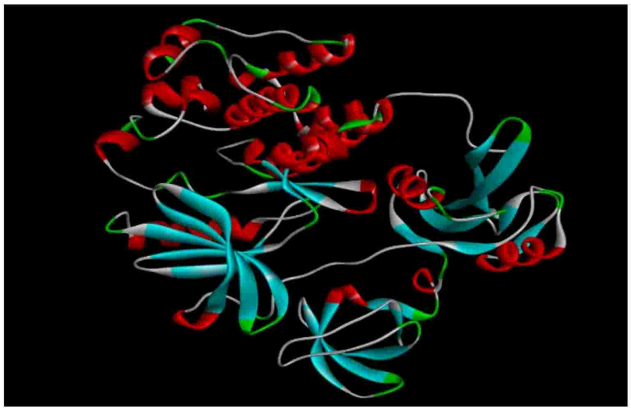


Figure 4.9: 3D structure of Src kinase (non-receptor tyrosine kinase) visualized in Biovia Discovery Studio Visualizer.

Hispidulin showed relatively highest binding affinity of -9 kcal/mol forming both 4 H-bonds and other non-polar interactions as compared to 6-Epi-beta-bisabolol (see table 4.5). Residues involved in H-bond formation includes Ile294, Val337, Asp404, Lys295. Nine residues are involved in making other non-covalent interactions including van der Waals, Pi-sigma bonds, Pi-Pi stacked bonds, alkyl and Pi-alkyl bonds by residues Ala403, Lys343, Ser342, Gly344, Tyr340, Ala293, Leu273, Val281, Leu393. 6-Epi-beta-bisabolol has binding affinity of -7.9 kcal/mol with target Src protein kinase forming weak interactions that includes van-der-waal interactions and other non-polar interactions. Gly344, Thr388, Asp404, Leu325, Ile336, Lys295, Met341 forms van der Waal interactions. Leu273, Tyr340, Ala293, Leu393, Val281, Val323 and Ala403 residues form alkyl and Pi-Alkyl bonds. 2D and 3D structure is visualized in BIOVIA discovery studio demonstrating interactions and amino acids taking part in interactions with phytochemicals (see fig 11, 12).

Table 4.5: It shows hydrogen bonds and hydrophobic interactions between ligands (Phytochemicals) and receptor (target protein).

Molecule ID	**Ligand (Phytochemical)**	**Molecular weight**	**Binding affinity (Kcal/mol)**	**Drug likeness**	**H-bond**	**Other non-polar interactions**	**Van-der-waal interactions**	**Chain**
5281628	Hispidulin	448.38	-10	0.6	I-336, K-295, D-404, M-341,	A-293, L-393, V-281, L-273, Y-340	K-343, S-342, G-344, A-403	A

12300148	6-Epi-beta-bisabolol	222.37	-7.9	0.48	-	L-273, L-393, Y-340, A-293, V-281, V-323, A-403	G-344, M-341, K-295, I-336, L-325, D-404,T-338	A

Figure 4.10: Docked conformations of bioactive compounds obtained from *Salvia officinalis L* a) Hispidulin b) 6-Epi-beta-bisabolol

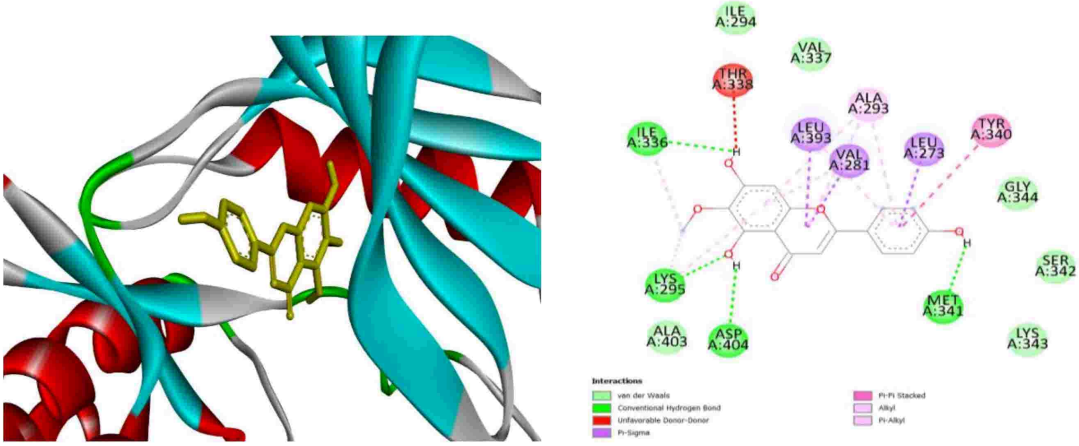


Figure 11: Visualization of molecular docking results in BIOVIA Discovery studio visualizer a) 2D diagram displaying residues involved in making H-bonds and non-polar interactions between Hispidulin and target protein Src kinase b) 3D structure displaying the regions of H-bonds donor and acceptor regions

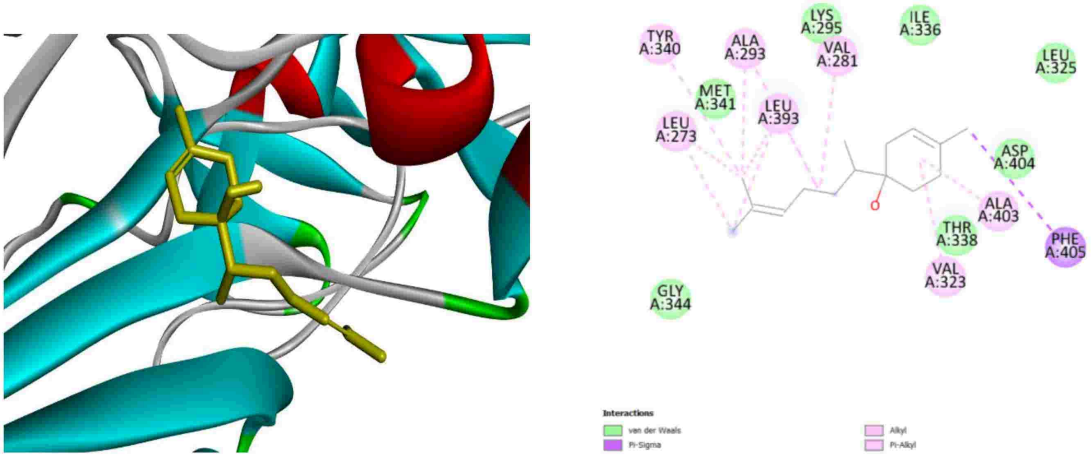


Figure 12: Visualization of molecular docking results in BIOVIA Discovery studio visualizer a) 2D diagram displaying residues involved in making H-bonds and non-polar interactions between 6-Epi-beta-bisabolol and target protein Src kinase b) 3D structure displaying the regions of H-bonds donor and acceptor regions

Docking results of Hispidulin and 6-Epi-beta-bisabolol against Src kinase revealed clearly that among the two, Hispidulin exhibits highest binding affinity and docks well with target protein. The results of this investigation suggested that phytochemicals obtained from *Salvia officinalis* L. can be effective against Colorectal cancer. Hence, Network pharmacology analysis combined with bioinformatics can be used as scientific evidence for use of traditional medicines. It helps us to understand new therapeutic leads derived from natural products and also reduces cost and time through in silico evaluation. We can also identify novel drug targets, cellular pathways and signal transduction responsible for causing a disease.

CHAPTER 5

DISCUSSION

Investigations of phytochemicals obtained from plants have gained lot of interest in recent years. doi.org/10.1016/j.biotechadv.2018.03.003 Network ph armacology is computational research which is well suited to the "multicomponent, multitarget, multipathway" characteristics of herbal medicine. This method integrates pharmacology, systemic biology and bioinformatics to analyze and understand the complex interaction network present between potential drugs and biological systems, providing a basic understanding of how herbal medicine can have therapeutic effects against particular disease.

This study investigated the principal active phytochemicals and molecular mechanisms of *Salvia officinalis L.* implicated in the treatment of colorectal cancer (CRC). After extensive screening out of the phytochemicals from this sage plant, hispidulin and 6-Epi-beta-bisabolol came out to be most effective phytochemicals effective against CRC. These two phytochemicals didn't violate Lipinski rule of five and can bind to and inhibit biological activity of Src kinase, thereby preventing its overexpression. Salvia officinalis L. exhibits a significant medicinal value in boosting human health because extracts of this plant have antioxidant, anti-inflammatory, anti-cancerous and antivenom properties, that provide hope for the advancement of medicine. The potential active phytochemicals present in Salvia officinalis L. have been identified as per their oral bioavailability value (OB) and drug likeness metrics. 22 phytochemicals meeting the requirement of OB > 0.55 and DL >0.18 were considered biologically active. Multiple disease-related signaling pathways have been mediated by a range of putative target genes. The progression of the disease can be reduced by focusing on these genes that regulate pathways linked to the disease. In our investigation, src gene which is a non-receptor tyrosine kinase came out to be the key player in mediating CRC carcinogenesis. According to KEGG and GO enrichment analysis, these target gene is a part of several pathways in the system. GO enrichment analysis comprehend the intricate biological processes underlying the disease like CRC. By classifying genes into molecular functions (MFs), cellular components (CCs), and biological processes (BPs), researchers can gain crucial insights into the pathogenesis of disease. Furthermore, the docking study results confirmed a significant binding affinity between both bioactive compounds and the target protein. Molecular docking is computational simulations technique to model interaction affinity between ligand and the target at molecular level, giving details about its affinity of interaction, orientation, binding site, and stability of ligands when complexed with target receptor protein. This robust interaction suggests that the biological activity of the receptors may be modulated with potential efficacy, which may lead to beneficial therapeutic effects. The binding energies of Hispidulin and 6-Epi-beta-bisabolool were -7.9 and -10.0 kcal/mol respectively. These results highlight the efficacy of bioactive component as

promising candidates for drugs, and they call for additional experimental validation and investigation in preclinical and clinical settings.

Hispidulin is a naturally occurring flavone compound found in several species of salvia and Artemisia. It shows potential anti-fungal, anti-inflammatory, anti-mutagenic and anti-carcinogenic properties. This phytochemical inhibits the epithelial to mesenchymal cell transition (EMT), thereby reducing the likelihood of cancer progression. [43]. 6-Epi-beta-bisabolol is a secondary derivative, which belongs to the class of sesquiterpenes, a terpenoid compound. It shows anti-cancerous, anti-inflammatory properties [44]. Due to their various biological activities and possible health advantages, 6-Epi-beta-bisabolol and hispidulin are both of great interest in pharmacological and therapeutic research.

CHAPTER-6

CONCLUSION

This study presents a novel therapeutic target 'SRC' gene which encodes for Src protein kinase that plays in important role in progression and development of CRC carcinogenesis. Src protein kinase can be targeted by potential phytochemicals obtained from *Salvia officinalis L.* named as Hispidulin and 6-Epi-beta-bisaabolol. Particularly hispidulin stands out as a promising substance that needs more research. Targeting of proposed gene can helps to reduce its over-expression in cell, thereby reducing overall carcinogenic events. We can identify the potential phytochemicals from plants and its target using network pharmacology approach. Then, the effectiveness and safety of proposed drug can be predicted by their toxicity analysis using protox 3.0 software which confirms that drug candidate is absent in four important toxicity parameters such as hepatotoxicity, carcinogenicity, mutagenicity and immunotoxicity. ADME analysis of drug candidate is crucial, it make sure that the drug does not exhibits any toxic effects against humans. Affinity of interaction between phytochemicals and target Src protein kinase is deremined by Molecular docking simulations. The integration of network pharmacology, toxicity analysis, ADME, molecular docking simulations, and profiling in this study has demonstrated a robust framework for the identification and validation of novel drug candidates. This study created a dynamic environment to predict the most likely drug-target protein interaction. With sufficient in vitro research, the proposed drug candidates from sage plant may provide revolutionary in treatment of colorectal cancer.

Network pharmacology approach also faces some limitations such as reliability on the usage of public databases. Investigating the effectiveness of phytochemicals of medicinal herbs from the standpoint of network pharmacology approach seems overly biased. In order to completely utilize the therapeutic effects of medicinal herbs, it is important to consider the availability of the phytochemicals in appropriate concentration in plants. Binding energy shown by the phytochemicals against the target will not necessarily be same in in-vivo conditions [45]. Furthermore, this study does not take into account chemical alterations that take place during different drug manufacturing processes. So, further in vitro and in vivo analysis is needed to validate safety and effectiveness of proposed drug against target gene. The results of this investigation supports the possible anti-cancerous effect of Hispidulin and 6-Epi-beta-bisabolol against the target protein Src kinase, thereby reducing the carcinogenesis of CRC.

REFERENCES

[1] Saima, S. Latha, R. Sharma, and A. Kumar, "Role of Network Pharmacology in Prediction of Mechanism of Neuroprotective Compounds," 2024, pp. 159–179. doi: 10.1007/978-1-0716-3662-6_13.

[2] P. Currais, I. Rosa, and I. Claro, "Colorectal cancer carcinogenesis: From bench to bedside," *World J. Gastrointest. Oncol.*, vol. 14, no. 3, pp. 654–663, Mar. 2022, doi: 10.4251/wjgo.v14.i3.654.

[3] P. Ashokkumar, T. Divya, K. Kumar, V. Dineshbabu, B. Velavan, and G. Sudhandiran, "Colorectal carcinogenesis: Insights into the cell death and signal transduction pathways: A review," *World J. Gastrointest. Oncol.*, vol. 10, no. 9, pp. 244–259, Sep. 2018, doi: 10.4251/wjgo.v10.i9.244.

[4] A. S. Choudhari, P. C. Mandave, M. Deshpande, P. Ranjekar, and O. Prakash, "Phytochemicals in Cancer Treatment: From Preclinical Studies to Clinical Practice," *Front. Pharmacol.*, vol. 10, Jan. 2020, doi: 10.3389/fphar.2019.01614.

[5] H. Allgayer, D. D. Boyd, M. M. Heiss, E. K. Abdalla, S. A. Curley, and G. E. Gallick, "Activation of Src kinase in primary colorectal carcinoma," *Cancer*, vol. 94, no. 2, pp. 344–351, Jan. 2002, doi: 10.1002/cncr.10221.

[6] M. Hamidpour, R. Hamidpour, S. Hamidpour, and M. Shahlari, "Chemistry, Pharmacology, and Medicinal Property of Sage (Salvia) to Prevent and Cure Illnesses such as Obesity, Diabetes, Depression, Dementia, Lupus, Autism, Heart Disease, and Cancer," *J. Tradit. Complement. Med.*, vol. 4, no. 2, pp. 82–88, Apr. 2014, doi: 10.4103/2225-4110.130373.

[7] C. N. Arnold, A. Goel, H. E. Blum, and C. Richard Boland, "Molecular pathogenesis of colorectal cancer," *Cancer*, vol. 104, no. 10, pp. 2035–2047, Nov. 2005, doi: 10.1002/cncr.21462.

[8] T. Armaghany, J. D. Wilson, Q. Chu, and G. Mills, "Genetic alterations in colorectal cancer.," *Gastrointest. Cancer Res.*, vol. 5, no. 1, pp. 19–27, Jan. 2012, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/22574233

[9] D. T. Debela *et al.*, "New approaches and procedures for cancer treatment: Current perspectives," *SAGE Open Med.*, vol. 9, p. 205031212110343, Jan. 2021, doi: 10.1177/20503121211034366.

[10] A. Ghorbani and M. Esmaeilizadeh, "Pharmacological properties of Salvia officinalis and its components," *J. Tradit. Complement. Med.*, vol. 7, no. 4, pp. 433–440, Oct. 2017, doi: 10.1016/j.jtcme.2016.12.014.

[11] D. Velickovic, N. Randjelovic, M. Ristic, A. Velickovic, and A. Smelcerovic, "Chemical constituents and antimicrobial activity of the ethanol extracts obtained from the flower, leaf and stem of Salvia officinalis L," *J. Serbian Chem. Soc.*, vol. 68, no. 1, pp. 17–24, 2003, doi: 10.2298/JSC0301017V.

[12] M. Keshavarz, A. Mostafaie, K. Mansouri, A. Bidmeshkipour, H. R. M. Motlagh, and S.

Parvaneh, “In vitro and ex vivo antiangiogenic activity of salvia officinalis,” *Phyther. Res.*, vol. 24, no. 10, pp. 1526–1531, Oct. 2010, doi: 10.1002/ptr.3168.

[13] M. Jakovljević *et al.*, “Bioactive Profile of Various Salvia officinalis L. Preparations.,” *Plants (Basel, Switzerland)*, vol. 8, no. 3, Mar. 2019, doi: 10.3390/plants8030055.

[14] D. F. N. Pedro, A. A. Ramos, C. F. Lima, F. Baltazar, and C. Pereira-Wilson, “Colon Cancer Chemoprevention by Sage Tea Drinking: Decreased DNA Damage and Cell Proliferation,” *Phyther. Res.*, vol. 30, no. 2, pp. 298–305, Feb. 2016, doi: 10.1002/ptr.5531.

[15] C. Lieu and S. Kopetz, “The Src Family of Protein Tyrosine Kinases: A New and Promising Target for Colorectal Cancer Therapy,” *Clin. Colorectal Cancer*, vol. 9, no. 2, pp. 89–94, Apr. 2010, doi: 10.3816/CCC.2010.n.012.

[16] J. Chen, A. Elfiky, M. Han, C. Chen, and M. W. Saif, “The Role of Src in Colon Cancer and Its Therapeutic Implications,” *Clin. Colorectal Cancer*, vol. 13, no. 1, pp. 5–13, Mar. 2014, doi: 10.1016/j.clcc.2013.10.003.

[17] K. K. Bhanumathy, A. Balagopal, F. S. Vizeacoumar, F. J. Vizeacoumar, A. Freywald, and V. Giambra, “Protein Tyrosine Kinases: Their Roles and Their Targeting in Leukemia,” *Cancers (Basel).*, vol. 13, no. 2, p. 184, Jan. 2021, doi: 10.3390/cancers13020184.

[18] S. Jerin, A. J. Harvey, and A. Lewis, “Therapeutic Potential of Protein Tyrosine Kinase 6 in Colorectal Cancer,” *Cancers (Basel).*, vol. 15, no. 14, p. 3703, Jul. 2023, doi: 10.3390/cancers15143703.

[19] R. Roskoski, “Src protein-tyrosine kinase structure, mechanism, and small molecule inhibitors,” *Pharmacol. Res.*, vol. 94, pp. 9–25, Apr. 2015, doi: 10.1016/j.phrs.2015.01.003.

[20] R. Roskoski, “Src protein–tyrosine kinase structure and regulation,” *Biochem. Biophys. Res. Commun.*, vol. 324, no. 4, pp. 1155–1164, Nov. 2004, doi: 10.1016/j.bbrc.2004.09.171.

[21] F. Dehkhoda, C. M. M. Lee, J. Medina, and A. J. Brooks, “The Growth Hormone Receptor: Mechanism of Receptor Activation, Cell Signaling, and Physiological Aspects,” *Front. Endocrinol. (Lausanne).*, vol. 9, Feb. 2018, doi: 10.3389/fendo.2018.00035.

[22] M. K. Paul and A. K. Mukhopadhyay, “Tyrosine kinase – Role and significance in Cancer,” *Int. J. Med. Sci.*, pp. 101–115, 2004, doi: 10.7150/ijms.1.101.

[23] L. Z. Benet, C. M. Hosey, O. Ursu, and T. I. Oprea, “BDDCS, the Rule of 5 and drugability.,” *Adv. Drug Deliv. Rev.*, vol. 101, pp. 89–98, Jun. 2016, doi: 10.1016/j.addr.2016.05.007.

[24] G. R. Bickerton, G. V Paolini, J. Besnard, S. Muresan, and A. L. Hopkins, “Quantifying the chemical beauty of drugs.,” *Nat. Chem.*, vol. 4, no. 2, pp. 90–8, Jan. 2012, doi: 10.1038/nchem.1243.

[25] M. Merski, J. Skrzeczkowski, J. K. Roth, and M. W. Górna, “A Geometric Definition of Short to Medium Range Hydrogen-Mediated Interactions in Proteins,” *Molecules*, vol. 25, no. 22, p. 5326, Nov. 2020, doi: 10.3390/molecules25225326.

[26] F. Noor *et al.*, "Integrating Network Pharmacology and Molecular Docking Approaches to Decipher the Multi-Target Pharmacological Mechanism of Abrus precatorius L. Acting on Diabetes," *Pharmaceuticals*, vol. 15, no. 4, p. 414, Mar. 2022, doi: 10.3390/ph15040414.

[27] K. Mohanraj *et al.*, "IMPPAT: A curated database of Indian Medicinal Plants, Phytochemistry And Therapeutics," *Sci. Rep.*, vol. 8, no. 1, p. 4329, Mar. 2018, doi: 10.1038/s41598-018-22631-z.

[28] J. G. M. Mvondo, A. Matondo, D. T. Mawete, S.-M. N. Bambi, B. M. Mbala, and P. O. Lohohola, "In Silico ADME/T Properties of Quinine Derivatives using SwissADME and pkCSM Webservers," *Int. J. Trop. Dis. Heal.*, pp. 1–12, Aug. 2021, doi: 10.9734/ijtdh/2021/v42i1130492.

[29] C. Ononamadu and A. Ibrahim, "Molecular docking and prediction of ADME/drug-likeness properties of potentially active antidiabetic compounds isolated from aqueous-methanol extracts of Gymnema sylvestre and Combretum micranthum," *BioTechnologia*, vol. 102, no. 1, pp. 85–99, 2021, doi: 10.5114/bta.2021.103765.

[30] A. Daina, O. Michielin, and V. Zoete, "SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules.," *Sci. Rep.*, vol. 7, p. 42717, Mar. 2017, doi: 10.1038/srep42717.

[31] W. Muster, A. Breidenbach, H. Fischer, S. Kirchner, L. Müller, and A. Pähler, "Computational toxicology in drug development," *Drug Discov. Today*, vol. 13, no. 7–8, pp. 303–310, Apr. 2008, doi: 10.1016/j.drudis.2007.12.007.

[32] M. D. Segall and C. Barber, "Addressing toxicity risk when designing and selecting compounds in early drug discovery," *Drug Discov. Today*, vol. 19, no. 5, pp. 688–693, May 2014, doi: 10.1016/j.drudis.2014.01.006.

[33] M. N. Drwal, P. Banerjee, M. Dunkel, M. R. Wettig, and R. Preissner, "ProTox: a web server for the in silico prediction of rodent oral toxicity," *Nucleic Acids Res.*, vol. 42, no. W1, pp. W53–W58, Jul. 2014, doi: 10.1093/nar/gku401.

[34] S. Venkateswaran, H. P. Manivannan, A. P. Francis, V. P. Veeraraghavan, G. R, and K. Sankaran, "Identification of Potential Phytochemical Inhibitors From Conium maculatum Targeting the Epidermal Growth Factor Receptor in Metastatic Colorectal Cancer via Molecular Docking Analysis," *Cureus*, Oct. 2023, doi: 10.7759/cureus.48000.

[35] M. Ashburner *et al.*, "Gene Ontology: tool for the unification of biology," *Nat. Genet.*, vol. 25, no. 1, pp. 25–29, May 2000, doi: 10.1038/75556.

[36] M. Kanehisa, "KEGG: Kyoto Encyclopedia of Genes and Genomes," *Nucleic Acids Res.*, vol. 28, no. 1, pp. 27–30, Jan. 2000, doi: 10.1093/nar/28.1.27.

[37] M. V. Kuleshov *et al.*, "Enrichr: a comprehensive gene set enrichment analysis web server 2016 update," *Nucleic Acids Res.*, vol. 44, no. W1, pp. W90–W97, Jul. 2016, doi: 10.1093/nar/gkw377.

[38] D. Szklarczyk *et al.*, "The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets," *Nucleic Acids Res.*, vol. 49, no. D1, pp. D605–D612, Jan. 2021, doi:

10.1093/nar/gkaa1074.

[39] C. v. Mering, "STRING: a database of predicted functional associations between proteins," *Nucleic Acids Res.*, vol. 31, no. 1, pp. 258–261, Jan. 2003, doi: 10.1093/nar/gkg034.

[40] A. Majeed and S. Mukhtar, "Protein–Protein Interaction Network Exploration Using Cytoscape," 2023, pp. 419–427. doi: 10.1007/978-1-0716-3327-4_32.

[41] Z. Mousavian, M. Khodabandeh, A. Sharifi-Zarchi, A. Nadafian, and A. Mahmoudi, "StrongestPath: a Cytoscape application for protein–protein interaction analysis," *BMC Bioinformatics*, vol. 22, no. 1, p. 352, Dec. 2021, doi: 10.1186/s12859-021-04230-4.

[42] C.-H. Chin, S.-H. Chen, H.-H. Wu, C.-W. Ho, M.-T. Ko, and C.-Y. Lin, "cytoHubba: identifying hub objects and sub-networks from complex interactome," *BMC Syst. Biol.*, vol. 8, no. S4, p. S11, Dec. 2014, doi: 10.1186/1752-0509-8-S4-S11.

[43] K. Liu, F. Zhao, J. Yan, Z. Xia, D. Jiang, and P. Ma, "Hispidulin: A promising flavonoid with diverse anti-cancer properties," *Life Sci.*, vol. 259, p. 118395, Oct. 2020, doi: 10.1016/j.lfs.2020.118395.

[44] M. A. Egbuta, S. McIntosh, D. L. E. Waters, T. Vancov, and L. Liu, "In Vitro Anti-Inflammatory Activity of Essential Oil and β-Bisabolol Derived from Cotton Gin Trash," *Molecules*, vol. 27, no. 2, p. 526, Jan. 2022, doi: 10.3390/molecules27020526.

[45] M. Ji, C. Wang, T. Yang, X. Meng, X. Wang, and M. Li, "Integrated Phytochemical Analysis Based on UPLC–MS/MS and Network Pharmacology Approaches to Explore the Effect of Odontites vulgaris Moench on Rheumatoid Arthritis," *Front. Pharmacol.*, vol. 12, Aug. 2021, doi: 10.3389/fphar.2021.707687.

APPENDIX 1

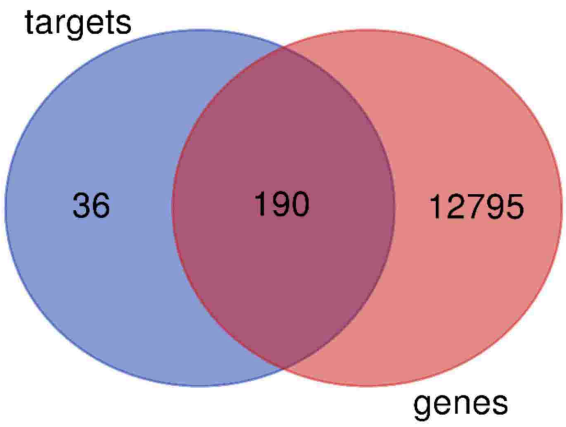


Figure III.I: List of 190 common genes

MMP2, XDH, CD38. MAOA. CDK1. CYP1A1, KISS1R, TNKS2, SQLE, ALOX12, AXL, MDM2, EGFR, KCNA3, G6PD, F10, ODC1, PLG, ACP1, PIK3R1, CHRM3, BRD3, CSNK2A1, NEK2, SLC6A2, MMP13, AKR1B10, EZH2, ABCB1, TNKS, ACHE, CA7, PRKCE, PTPN1, CCR1, PYGL, CYP11B1, KCNA5, PABPC1, HSD11B1, MPG, CBR1, HSD17B1, CYP2C19, ALDH2, ESRRA, KDM5A, OPRK1, NR1I3, AURKB, TNF, PTGS2, CYP2C9, ESR1, PGGT1B, MMP12, GLO1, MAPK8, GPR35, AKR1B1, SRC, CYP17A1, BRD2, PRKCH, PRKACA, ABCC1, PRCP, CA13, SLC6A4, GRK6, PLA2G2A, ADORA2A, MAPT, PIM1, NR3C1, CDK6, LCK, NOX4, PFKFB3, PER2, AVPR2, SLC10A2, EPHX2, PARP1, NPC1L1, FLT3, BCL2L1, CA2, CYP51A1, OPRM1, SERPINE1, PRKCA, TOP1, NOS2, PRKCG, MET, PRKCB, C5AR1, SHBG, BCHE, BCL2, F7, TACR1, NQO2, PTPN2, PTGER1, AHR, BRD4, CYP19A1, CA9, GSK3B, CA12, PPARD, IDO1, PDE5A, DRD4, PDE10A, TYR, ST6GAL1, TRPM8, CCR9, LIPE, ALOX15, PRKCD, MPO, TNK2, MTNR1A, SIGMAR1, MCL1, ABCG2, TTR, MAOB, CYP1A2, PIK3CG, SYK, GRM5, TOP2A, IGF1R, CFTR, PKN1, INSR, AKT1, AR, DAPK1, CYP3A4, APP, HSD17B2, MTNR1B, PTK2, NEK6, KIT, F2, SREBF2, HMGCR, NR3C2, PRKCZ, AMY1A, CA4, PLA2G1B, PTPRF, ICMT, MYLK, ADORA3, TERT, SLC29A1, TP53, RORC, CALM1, ESR2, CYP1B1, ALOX5, FASN, NR1H3, PLK1, ALK, RHOA, DRD2, ARG1, LIPA, KDR, PTPRS, PPARA, RORA, BACE1, ADORA1, IL2, PTAFR, CA1, RPS6KA3, MMP

APPENDIX 2

Table IV.I Toxicity analysis

Phytochemical	Hepatotoxicity	Cytotoxicity	Immunotoxicity	Mutagenicity
6-Epi-beta-bisabolol	No	No	No	No
Salvigenin	No	No	Yes	No
Rosmarinic acid	No	No	Yes	No
Luteolin	No	Yes	No	Yes
Nepetin	No	Yes	Yes	No
Cirsimaritin	No	No	Yes	No
Genkwanin	No	No	No	Yes
Hispidulin	No	No	No	No
Corosolic acid	No	Yes	Yes	No
Oleanolic acid	Yes	Yes	Yes	No
Ursolic acid	Yes	Yes	Yes	No
Cosmosiin	No	No	No	Yes
beta-Sitosterol	No	No	Yes	No
7alpha-Acetoxyroyleanone	Yes	No	No	Yes
beta-Sitosterol	No	No	Yes	No
Oleanolic acid	Yes	Yes	Yes	No
Cosmosiin	No	No	No	Yes
Cirsiliol	No	Yes	Yes	No
luteolin 7-O-beta-d-glucopyranoside	No	No	No	No
5, 7-Dimethoxy-1-naphthol	No	Yes	Yes	No
Octadecamethylcyclononasiloxane	No	No	No	No
beta-Sitosterol	No	No	Yes	No
Vicenin-2	No	No	No	Yes
2alpha,3alpha-Dihydroxyolean-12-en-28-oic acid	No	Yes	Yes	No
Ursolic acid	Yes	Yes	Yes	No
3-Epioleanolic acid	Yes	Yes	Yes	Yes
Oleanolic acid	Yes	Yes	Yes	No
beta-Carotene	No	No	No	Yes
Cynaroside	No	No	No	No
luteolin 7-O-beta-d-glucopyranoside	No	No	No	No

GO biological process

Table IV.II significant p values: Top 3 notable p values and q values given for GO biological process 2023

Biological process	P value	Q value	Overlapped genes
Protein Phosphorylation (GO:0006468	1.310345e-24	2.914208e-21	ALK, APP, GSK3B, TNKS, SRC, FLT3, PRKCZ, EGFR, AURKB, IGF1R, LIPE, RPS6KA3, MAPK8, GRK6, KDR, PIM1, AKT1, NEK2, PRKACA, PRKCG, BRD2, PRKCH, CSNK2A1, SYK, PRKCB, DAPK1, INSR, NEK6, PRKCE, PRKCD, PLK1, PRKCA, PTK2, CDK6, LCK, KIT, CDK1, TOP1, PKN1]
Phosphorylation (GO:0016310)	1.119443e-20	1.244820e-17	[ALK, APP, GSK3B, PRKCZ, AURKB, PIK3CG, LIPE, MAPK8, GRK6, PIM1, AKT1, NEK2, PRKACA, PRKCG, BRD2, PRKCH, CSNK2A1, SYK, PRKCB, DAPK1, INSR, NEK6, PRKCE, TNK2, PRKCD, PLK1, PRKCA, PTK2, CDK6, LCK, CDK1, TOP1, PKN1]
Negative Regulation of Apoptotic Process (GO:0043066)	3.126863e-17	2.318048e-14	[GSK3B, SRC, GLO1, ALOX12, PIK3R1, TNF, MPO, EGFR, IGF1R, RPS6KA3, MAPK8, KDR, PIM1, AKT1, CD38, MCL1, PRKCG, PRKCH, CSNK2A1, PRKCD, PLK1, PRKCA, MMP9, PTK2, IL2, AXL, MDM2, CDK1, BCL2, TP53, BCL2L1]

GO Cellular Process

Table IV.III: Topmost three p-values and q-values obtained for GO Cellular Component 2023

Cellular component	**p-value**	**q-value**	**Overlapped genes**
Membrane Raft (GO:0045121)	6.673594e-07	0.000105	APP, LCK, SRC, KDR, KCNA5, MAPT, PRKACA, TNF, EGFR, ABCG2, SLC6A4
Neuron Projection (GO:0043005)	1.614203e-06	0.000123	PRKCG, CHRM3, APP, GSK3B, PTPRS, INSR, KCNA3, OPRK1, OPRM1, PTGS2, SLC6A2, IGF1R, SLC6A4, MAPK8, ADORA3, ADORA1, MAPT, DRD2, DRD4
Endoplasmic Reticulum Membrane (GO:0005789)	2.355861e-06	0.000123	PTPN1, ICMT, PLA2G2A, CYP51A1, HMGCR, CYP2C19, CYP3A4, PTGS2, CYP19A1, SREBF2, RHOA, CYP17A1, HSD11B1, CYP2C9, SQLE, CYP1A2, HSD17B2, CYP1A1, CDK1, BCL2, CYP1B1, NOX4, CFTR

GO Molecular Function

Table IV.IV: Topmost notable p-values and q-values for GO Molecular Function 2023

Molecular function	P value	Q value	Overlapped genes
Protein Tyrosine Kinase Activity (GO:0004713)	5.647348e-13	5.463809e-11	ALK, SYK, FLT3, SRC, INSR, TNK2, PTK2, EGFR, IGF1R, LCK, AXL, KIT, KDR, MET
Protein Serine/Threonine Kinase Activity (GO:0004674)	2.107749e-19	8.156988e-17	GSK3B, PRKCZ, EGFR, AURKB, MYLK, RPS6KA3, MAPK8, GRK6, PIM1, AKT1, NEK2, PRKACA, BRD4, PRKCG, BRD2, PRKCH, CSNK2A1, SYK, PRKCB, DAPK1, NEK6, PRKCE, PRKCD, PLK1, PRKCA, CDK6, CDK1, TOP1, PKN1
Protein Tyrosine Kinase Activity (GO:0004713)	5.647348e-13	5.463809e-11	ALK, SYK, FLT3, SRC, INSR, TNK2, PTK2, EGFR, IGF1R, LCK, AXL, KIT, KDR, MET

LIST OF PUBLICATIONS

CONFERENCE

Presented conference paper entitled “Unveiling the potential of phytochemicals: Network pharmacology approach combined with bioinformatics analysis of active phytochemicals obtained from *Salvia officinalis L.* in treating Colorectal Cancer” Scopus indexed conference at ICAEM 2024 organized by International school of technology and sciences for women in collaboration with Samarkand state university, Uzbekistan held on 20 April 2024

INTERNATIONAL CONFERENCE ON ADVANCES IN ENGINEERING AND MEDICAL SCIENCES 2024

Accredited by

Approved by

ORGANISED BY

INTERNATIONAL SCHOOL OF TECHNOLOGY AND SCIENCES FOR WOMEN (AUTONOMOUS)

NH-16, EAST GONAGUDEM, RAJANAGARAM, EAST GODAVARI-533297 | WWW.ISTS.AC.IN

IN COLLABRATION WITH

SOCIETY OF PUBLIC HEALTH EXPERTS, INDIA

SAMARKAND STATE UNIVERSITY, UZBEKISTAN

UNIVERSITY OF TECHNOLOGY AND APPLIED SCIENCES, IBRI SULTANATE OF OMAN

IN ASSOCIATION WITH

ORGANIZATION OF SCIENCE AND INNOVATIVE ENGINEERING & TECHNOLOGY (OSIET)

Certificate of Presentation

This is to certify that Mr/Mrs/Dr Pooja *from* Delhi Technological University, Delhi *has presented a paper title* "Unveiling the potential of phytochemicals: Network pharmacology approach combined with bioinformatics analysis of active phytochemicals obtained from Salvia officinalis L. in treating Colorectal cancer" *in the "4th International Conference on Advances in Engineering and Medical Sciences" held on 19th & 20th April 2024.*

Dr. Y.Rajasree Rao
Principal & Conference Chair

K. Janani
CEO

Dr.Srinivasa Muthukaruppan
Founder and CEO

Dr.Christo Ananth
Professor

Dr. Akhatov Akmal Rustamovich
Vice Rector (International Cooperation)

PAPER NAME

Thesis_writeup-1-1_compressed.pdf

WORD COUNT	CHARACTER COUNT
11160 Words	**64521 Characters**
PAGE COUNT	FILE SIZE
46 Pages	**1.2MB**
SUBMISSION DATE	REPORT DATE
Jun 4, 2024 5:07 PM GMT+5:30	**Jun 4, 2024 5:08 PM GMT+5:30**

7% Overall Similarity

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

- 4% Internet database
- 4% Publications database
- Crossref database
- Crossref Posted Content database
- 6% Submitted Works database

Excluded from Similarity Report

- Bibliographic material
- Manually excluded sources