Combinatorial Therapy for Tumor Treatment

THESIS

Submitted to the Delhi Technological University For the award of the degree of

DOCTOR OF PHILOSOPHY

Submitted by

Sunil Kumar

Guide

Dr. Asmita Das

Delhi Technological University, DELHI

Department of Biotechnology
Delhi Technological University
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042, INDIA

July 2023

Combinatorial Therapy for Tumor Treatment

THESIS

Submitted to the Delhi Technological University For the award of the degree of

DOCTOR OF PHILOSOPHY

Submitted by

Sunil Kumar

Guide

Dr. Asmita Das

Delhi Technological University, DELHI

Department of Biotechnology
Delhi Technological University
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042, INDIA

July 2023

Dedicated to Cancer Patients

DECLARATION

I hereby declare that the thesis entitled **"Combinatorial Therapy for Tumor Treatment"** submitted by me, for the award of the degree of ***Doctor of Philosophy*** to **Delhi Technological University (Formerly DCE)** is a record of *bona fide* work carried out by me under the guidance of Dr. Asmita Das.

I further declare that the work reported in this thesis has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma in this Institute or any other Institute or University.

Sunil 31/7/2023

Name: Sunil Kumar
Reg No: 2K18/PHDBT/19
Department of Biotechnology
Delhi Technological University (DTU)
Shahbad Daulatpur, Bawana Road, Delhi-110042
Place: New Delhi
Date: 31.07.2023

CERTIFICATE

This is to certify that the thesis entitled **"Combinatorial Therapy for Tumor Treatment"** submitted by **Mr. Sunil Kumar** to **Delhi Technological University (Formerly DCE)**, for the award of the degree of "Doctor of Philosophy" in Biotechnology is a record of *bona fide* work carried out by him. Sunil Kumar has worked under my guidance and supervision and has fulfilled the requirements for the submission of this thesis, which to our knowledge has reached requisite standards.

The results contained in this thesis are original and have not been submitted to any other university or institute for the award of any degree or diploma.

31/7/23

Dr. Asmita Das
Department of Biotechnology
Delhi Technological University
Shahbad Daulatpur, Bawana Road, Delhi-110042
Place: New Delhi
Date: 31.07.2023

ABSTRACT

Cancer is a complex and multifaceted disease that continues to pose a significant challenge to global health. As the second leading cause of death worldwide. Early detection and noninvasive techniques of detecting cancer are necessary to improve treatment outcomes, save lives and improve the quality of life. Biopsies of tumors are often expensive and invasive and raise the risk of serious complications like infection, excessive bleeding, and puncture damage to nearby tissues and organs. Early detection biomarkers are often variably expressed in different patients and may even be below the detection level at an early stage. Hence PBMC that shows alteration in gene profile as a result of interaction with tumor antigens may serve as a better early detection biomarker. Also, such alterations in immune gene profile in PBMCs are more detectable in a wide variety of cancer patients despite their variability in different cancer mutants. Tumor cell biomarkers lack specificity, and tumor heterogeneity complicates accurate diagnosis and treatment. Changing biomarker expression affects treatment responses, and technical challenges impact utility. Synthetic drugs targeting tumor cells often trigger tumour cells to acquire resistance against them. Tumor progression is an outcome of tumor growth regulation in conjunction with tumor evasion by immune modulation. Therefore, understanding of immunological biomarkers is equally important. Hence, designing a prudent chemotherapeutic combination requires a detailed understanding of gene regulation altering cancer prognosis and its impact on immune regulation . Immunotherapy also has its side effects and does not provide an adequate response in all patients, and its inherent variability in patient response often makes them prohibitive. Hence, a concomitant targeting of tumour cells and modulation of immune cell function may be a particularly beneficial mechanism for cancer treatment.

Machine learning tools are crucial for early cancer detection and immune modulation due to their ability to analyze complex data and identify patterns that may not be apparent through

traditional methods. Potential diagnostic biomarkers were predicted for breast cancer using eXplainable Artificial Intelligence (XAI) on XGBoost machine learning (ML) models trained on a binary classification dataset containing the expression data of PBMCs from 252 breast cancer patients and 194 healthy women. After effectively adding SHAP values further into the XGBoost model, ten important genes related to breast cancer development were discovered to be effective potential biomarkers. It was discovered that SVIP, BEND3, MDGA2, LEF1-AS1, PRM1, TEX14, MZB1, TMIGD2, KIT, and FKBP7 are key genes that impact model prediction. These genes may serve as early, non-invasive diagnostic and prognostic biomarkers for breast cancer patients. The impact of concomitant intervention cancer progression and immune regulation therefore necessitated identification of such biomarkers that have dual impact. Gene expression data of HNSC tumor samples and PBMCs of tumor patient datasets were analysed for the identification of differentially expressed genes. 110 DEGs were found to be common in both datasets. Further, it was identified that these 110 DEGs were involved in biological processes related to tumor regulation. Potential Immunological biomarkers were identified for HNSC cancer. The Genes that play a role in both tumour growth and immune suppression were identified by enrichment analysis followed by gene expression analysis. 10 such genes were shortlisted, Foxp3, CD274, IDO1, IL-10, SOCS1, PRKDC, AXL, CDK6, TGFB1, FADD. CD274 and IDO1 were found to have the highest degree of interaction based on their network of interactions.

Synthetic drugs including many of FDA approved drugs might cause significant side effects, leading to adverse impacts on patients' quality of life. Additionally, some cancer cells may develop resistance to synthetic drugs over time, reducing treatment efficacy. Moreover, targeted therapies may only be effective in cancers with specific molecular characteristics, limiting their broad applicability. To address these limitations, ongoing research focuses on developing more targeted and personalized therapies, combining synthetic drugs with other

treatment modalities, and exploring alternative natural compounds with multi-target effects. Multi-target natural compounds offer the advantage of targeting multiple pathways involved in cancer progression without significant side effects. These compounds, derived from plants and other natural sources, hold promise in cancer treatment due to their diverse mechanisms of action and potential for reduced toxicity. Natural compounds that help in tumour suppression as well as functional immune modulation were identified for their dual roles. Np care and GEO databases were used for retravel of natural compounds. 102 potential anti-cancer natural compounds treatment gene expression data was analysed and key differentially regulated genes by them were identified. These 102 natural compounds were analysed for their ability to alter the expression of 110 commonly differentially expressed (identified in first objective). Salidroside was altering maximum number of 66 gene from them. Gallic acid and Shikonin were found to be the natural compounds that target CD274 and IDO1 respectively. Galic acid is extracted from leaves of bearberry, in pomegranate root bark, gallnuts, witch hazel, both in free-state and as part of the tannin molecule, whereas Shikonin is found in the extracts of dried root of the plant Lithospermum erythrorhizon. Studies have demonstrated that both Shikonin and Gallica acid exhibits anti-cancer properties.

Single drug treatment can lead to the development of drug resistance, where cancer cells become less responsive to the treatment over time. Some cancers may be inherently resistant to certain drugs, restricting their effectiveness. Moreover, high doses of a single drug can cause severe side effects, impacting patients' quality of life. Additionally, single-drug therapy may not be effective due to the heterogeneity of cancer cells, allowing potential tumor recurrence. Combination therapy targets cancer cells through multiple pathways, reduces drug resistance, and enhances efficacy of treatment outcomes. Synergistic interactions can improve efficacy while minimizing side effects, advancing personalized cancer care for better patient outcomes. A combination of Salidroside, Ginsenoside Rd, Oridonin, Britanin, and Scutellarein was

chosen such that they could alter the expression of 108 genes out of the selected 110 genes. The combination was further analyzed for regulating pathways and biological processes that were affected. Expression data analysis of HNSC cancer exhibited 1745 differentially expressed genes. Gallic acid treatment results in the downregulation of 120 genes and upregulation of 35 genes while Shikonin results in the downregulation of 660 genes and upregulation of 38 genes. Pathway analysis of these genes that were modulated by Gallic acid and Shikonin showed them to be crucially involved in pathways that were essential for cancer prognosis. Further Gallic acid and shikonin treatment impact on cancer cell line was analysed individually as well as in combination with the help of in vitro experiments. Gallic acid showed IC_{50} value of 46.87, 59.37, and 93.75 at 12h, 24h, and 48h treatment, respectively. Shikonin showed IC_{50} value of 13.86, 11.95, and 10.89 at 12h, 24h, and 48h treatment, respectively. Lowest percentage of cell viability was observed for combination of 80 µl Gallic acid and 16 µl of Shikonin. So, this combination of gallic acid and shikonin could be effective for the HNSC cancer treatment. Our studies showed a multifaceted, multi-dimensional tumor regression by altering autophagy, apoptosis, inhibiting cell proliferation, angiogenesis, metastasis and inflammatory cytokines production. Thus, the study has helped develop a unique combination of natural compounds that will markedly reduce the propensity of development of drug resistance in tumors and immune evasion by the tumors. This study is crucial to developing a combinatorial natural therapeutic cocktail with accentuated immunotherapeutic potential.

ACKNOWLEDGEMENT

First and foremost, I am thankful to all the cancer patients for share their data to the scientific community.

I would like to express my gratitude to ***Prof. Jai Prakash Saini****, Vice chancellor, Delhi Technological University, Delhi for providing this opportunity to carry out this work in this prestigious institute. Further, I express my gratitude to* ***Prof. Yogesh Singh****, former Vice chancellor Delhi Technological University, Delhi for providing me opportunity and infrastructure to carry out this work.*

With pleasure, I acknowledge my deep sense of gratitude and indebtedness to my guide and mentor ***Prof. Pravir Kumar****, Professor, DRC Chairman and Head, Department of Biotechnology, Delhi Technological University, Delhi, for their enlightening guidance, intelligent approach, constructive critique, whole-hearted and ever available help, which has been the primary impetus behind the research. Without the wise advice and able guidance, it would have been impossible to complete the thesis in this manner. Here, I would also like to thanks* ***Prof. Jai Gopal Sharma****, Former Head of the Department, department of Biotechnology, for providing me the infrastructure and smooth functioning of official work.*

The constant guidance and encouragement received from ***Dr. Asmita Das****, Guide, Assistant Professor, Delhi Technological University, Delhi, has been of great help in carrying out the present work and is acknowledged with reverential thanks.*

I would like to thank my fellow lab mates especially ***Shweta Gulia, Ritu,*** *and* ***Niharika Gupta*** *for helping and encouraging me throughout my research. This would be incomplete without saying thanks to my senior* ***Dr. Richa Sharma****, who corrected me several times during my initial days of career.*

I would also like to thank the Senior Management and technical staff of DTU, who helped me to carry out my research work.

Thanks are due to the wonderful friends in my life who were always on standby to bring me to positivity, hope and smiles when things didn't seem favoring, and it seemed a far-fetched journey. Special mention is deserved for all my friends, especially ***Dr. Rohan Gupta, Dr. Rajkumar****,* ***Kamesh Saroha****, and* ***Sidarth Sharma*** *who have always been my guiding light when I didn't believe in myself. I would also like to thank my dear and lovely niece,* ***Gugnu*** *for releasing my stress during my research work. Lastly, I want to express my gratitude to my brother* ***Sahil Boora****, who always stood beside mi in my ups and downs.*

Finally, I want to dedicate this thesis to my parents, ***Mr. Narender Singh*** *and* ***Mrs. Sudesh****,*

whose loving support has been my strongest inspiration. Their sacrifice allowed me to pursuit my dream and has made Ph.D. study seems completely painless.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

Introduction

CHAPTER I: INTRODUCTION

1.1.OVERVIEW

Cancer is a complex and multifactorial disease. Which is initiated by the uncontrolled growth of abnormal cells and their spread in the different body parts. Cancer has a history as ancient as human civilization itself without any sustainable cure. According to WHO, the death rate increases continuously from cancer and that reaches 9.9 million in 2020. Cancer, a term encompassing a vast array of diseases, arises from uncontrolled growth and division of abnormal cells in the body. It can affect any organ or tissue and has the potential to spread to other parts of the body, leading to life-threatening consequences. Despite significant advances in cancer research and treatment, it remains a major global health concern.

Cancer arises from a complex interplay of genetic, environmental, and lifestyle factors. Mutations in key genes, such as oncogenes and tumor suppressor genes, play a crucial role in initiating and promoting the development of cancer. Additionally, factors such as tobacco use, unhealthy diet, physical inactivity, exposure to carcinogens, infections, and genetic predisposition contribute to the risk of developing cancer.

Cancer cells possess distinct hallmarks that differentiate them from normal cells. These hallmarks include sustained proliferation, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, activating invasion and metastasis, and evading immune destruction. Furthermore, tumor heterogeneity, the presence of cancer stem cells, and the tumor microenvironment significantly influence cancer progression and response to treatment. The genetic variability across different cancer types has impeded the identification of therapeutic targets and the drug design and development against tumors[1]. The most commonly used cancer therapies include surgery, radiation, and chemotherapy, which can be used in isolation or different combinations [2]. However, these techniques have

been associated with a high morbidity rate and a significant decline in quality of life [3].

Additionally, despite monotherapy's specificity and efficiency, cancer cells' molecular flexibility renders ideal lethal effects challenging [4]. HNSC has a poor prognosis due to many patients' high likelihood of recurrence or metastasis following radiation or chemotherapy [5]. This high metastatic rate of HNSC is because of the tumor cell's interactions with the surrounding tissues and immune cells that will form the tumor microenvironment (TME) [6]. Host immune cells can recognize and eliminate the tumor cells, but an evasion of immunosurveillance generates an environment that accommodates the progression and survival of tumor cells [7]. Cancer-associated stromal fibroblasts, T cells, B cells, neutrophils, macrophages, myeloid-derived suppressor cells (MDSC), natural killer (NK) cells, and mast cells are all part of the TME [8]. These numerous cell subsets penetrate tumors and interact with cells and one another through multiple networks [9]. Tumors progress if they can evade and/or suppress antitumor immune responses [10]. Tumors frequently elude the immune system of their hosts by inhibiting cytotoxic T-cell activity or activating and increasing immunosuppressive cell populations [11].

Early detection plays a crucial role in improving cancer outcomes. Diagnostic techniques include imaging methods (e.g., X-ray, ultrasound, magnetic resonance imaging), laboratory tests (e.g., blood tests, tumor markers), and histopathological examination of tissue samples. Cancer screening programs for specific types of cancer have been implemented to identify pre-cancerous lesions or early-stage cancer in asymptomatic individuals. Due to the limitations of chemotherapy and radiation therapy, there is a critical need for early detection and prevention of high-risk premalignant lesions. Still, due to the complex nature of cancer, a unilateral therapeutic strategy often is insufficient and redundant.

Even though tumor sampling is frequently used to identify biomarkers, collecting tissue is difficult because of restricted accessibility, many lesions and heterogeneity of the biopsy site,

and patient conditions [12]. Biopsies of tumors are often expensive, invasive, and time-consuming, and they raise the risk of serious complications [13]. Most screening systems cannot detect and identify cancers until they have reached a particular stage of development [14]. Breast cancer, for example, may have been present for many years before it is discovered through palpation or mammography, and it has the potential to spread to other organs [15]. There is a pressing need to identify cancer at its earliest stages, particularly before the onset of clinical signs and symptoms. Early breast cancer detection is essential since it provides a more significant number of treatment choices, higher survival, and enhanced quality of life. While there is no fool proof way to avoid breast cancer, early diagnosis gives the most significant opportunity for successful treatment. Early detection and modern treatment are key to avoiding breast cancer fatalities. Early-stage breast cancer is simpler to treat. Regular screenings are the best method to detect breast cancer early [16].

Studies of biomarkers from blood, nipple aspirate fluid, perspiration, urine, tears, or breath may diagnose breast cancer early and in a non-invasive manner [17]. A simple blood or breath test may soon be able to identify breast cancer early[18]. Recent studies imply peripheral blood analyses might provide prognosis and treatment responsiveness [19]. Cancer detection using peripheral blood is more straightforward and less invasive [20]. As a result, generating clinically useful biomarkers requires the study of readily available peripheral blood [21]. The immune system relies on these PMBCs to combat infection and adapts to new threats. Oncogenic cells interact with normal stromal cells and the host immunological defense system to form tumors and prevent apoptosis [22]. The tumor's ability to evade the immune system also plays a significant role. Immune suppression in the tumor microenvironment by CD4+, CD25+, and FoxP3+ cells, regulatory T cells (Tregs), and other inhibitory peripheral blood mononuclear cells is the primary mechanism of tumor immune evasion [23]. Because of this, gene expression profiling of peripheral blood cells has the potential to identify early cancers

[24]. Michael E. Burczynski et al demonstrated that circulating monocytes of peripheral blood may be utilized as a surrogate monitor for difficult-to-biopsy tissues and/or as an extremely sensitive monitor to check for changes in the physiological condition of the organism [23]. Sharma et al. showed that PBMCs might be utilized to build gene expression assays for early diagnosis of breast cancer based on the properties of these cells [25]. The process by which malignant development induces distinctive alterations in the blood biochemical environment justifies the use of the PBMC transcriptome gene as a monitor for malignant solid tumors [26]. Tumor cells interact with immune cells and change their expression profiling of genes and can escape the immune system of the host easily [27]. The transcriptome gene expression of PBMCs may be used as a tumor screening marker since it is conveniently retrieved. Clinical pharmacogenomics might benefit from the use of PBMCs as predictive biomarkers because of the ease with which they can be obtained [28].

Cancer treatment approaches are diverse and depend on various factors such as cancer type, stage, and patient characteristics. Common treatment modalities include surgery, chemotherapy, radiation therapy, immunotherapy, targeted therapy, hormone therapy, and stem cell transplantation. Advances in precision medicine and personalized therapies hold great promise for improving treatment efficacy and reducing side effects.

Thousands of drugs have been used to treat cancer, but it is still the most abundant cause of fatality in the world. There are different types of therapy used for cancer treatment, such as radiation, surgery, chemotherapy, immunotherapy. However, many chemotherapeutic measures often result in the development of drug resistance in patients. Immune response in every individual is a complex array of immune functionality that are interrelated and regulated in a complex cascade of mechanisms that vary significantly in different individuals, hence patients often have variable tumor immunity and tumor prognosis. Therefore the same immunotherapy may have functional variability for every patient and even exhibit variable side

effects [29].

Drug resistance to cancer is a very complicated process and may depend on different factors, such as mutation at the drug's target site [30], any alteration in drug metabolism [31], resistance may be due to downregulation of pro-apoptotic signals and upregulation of anti-apoptotic signals [32], may lead to an increase in impaired DNA repair [33], and may result in a decrease in drug uptake or increase in drug efflux [34][35]. Moreover, in the vicinity of the tumor, there are not only uncontrollably proliferating cells, but over time there is an immense accumulation of divergent cells that modulate the surrounding environment that is known as the tumor microenvironment. It contains immune cells, extracellular matrix, blood vessels, fibroblasts and signaling molecules [36]. The immune system plays an important role in the development, establishment, and progression of HNSC. Better treatment for HNSC can be achieved by understanding the dysregulation and evasion of immune system. HNSC cells evade the host immune system through manipulation of their immunogenicity, production of immunosuppressive mediators, promotion of immunomodulatory cell types [37].

Tumor cells and their microenvironment are closely related and continuously interacting. Initially, immune cells try to eliminate tumor cells, but as the tumor grows, tumor cells over-express certain ligands that bind to immune cells and suppress the immune response. For example, PD-L1 present on tumor cells binds to T cell PD-1 so that the T cell response is suppressed [38][39]. Galetin-9 present on tumor cells binds to TIM-3 present on T cells to alter T cell function [40]. TIGIT and CD96 present on T cell bind to CD155 and CD112 present on tumor cells [41]. When these immune cells interact with other immune cells and also try to suppress the immune response. CTLA4, another receptor present on T cells that binds to CD80 and CD86 present on APCs and suppress T cell function [42]. MHC-II present on APCs binds to LAG3 present on T cell and block T cell recognition of tumor antigens [43]. There are many antibodies against these immune checkpoints, but they don't result in the same response in

every individual, and in some cases may result in severe side effects. Other immune checkpoints help in tumor suppression like DNAM-1 present on T cell [44] and NK cell binds to CD155 and CD112 present on tumor cell [45], ICOS present on T cell binds to ICOS-L present on APC [46], GITR present on T cell binds to GITR-L present on APC [47].

Further, to assess the efficacy of these chemo-preventive medicines, new biomarkers with predictive value for clinical disease and risk stratification can be employed for more disease specific strategy. From our literature survey, we shortlisted some drugs that show anti-cancer properties also work in immune modulation, such as mTOR inhibitor (rapamycin) reduces the expression of PD-L1 in HNSC cancer [48]. Drugs such as Statin, Metformin, and Antheracyclines can also enhance the immune system and cause tumor cells to kill more effectively [49][50][51]. Thalidomide and its derivative drugs, such as lenalidomide, were first used as a direct anti-cancer drug due to its cell-cycle arrest properties but later recognized for their role as immunomodulatory drugs due to their ability to stimulate T cells to secrete IL-2 and interferon-gamma [52].

Cancer research is a dynamic field, with ongoing efforts to deepen our understanding of cancer biology, identify novel therapeutic targets, and develop innovative treatment approaches. Advancements in genomics, proteomics, and artificial intelligence are revolutionizing cancer research and opening new avenues for precision medicine and personalized cancer care.

Here we try to explore those genes, whose expressions are functionally associated with HNSC disease and are involved in immune suppression. We used an unbiased approach for targeting these genes because of their upregulation associated with both immune suppression and tumor progression. Further, we have explored natural compounds that can inhibit these gene functions because natural compounds are cost effective and have fewer side effects as compared to synthetic compounds. These natural compounds have properties that interfere with initiation, development and progression of tumors through various mechanisms including apoptosis,

angiogenesis, metastasis, cell proliferation and cell differentiation [53]. So combinatorial use of natural compounds that interfere with multiple pathways, thus resulting in better therapeutic strategies, can also address the problem of drug resistance and hence may serve as a better therapeutic strategy [54].

Efforts to increase the efficacy of cancer treatment have primarily failed in recent decades, underlining the need for novel techniques such as complementary and alternative medicine [55]. Numerous natural herbal substances have caught the interest of academics and physicians due to their potential to prevent or improve the treatment of chronic diseases, including cancer[56]. Natural chemicals and combinations thereof may be a potential source of synergistic cancer treatments since they can interact with multiple biological targets involved in tumor growth, drug resistance, and metastasis [57]. Through their multitargeting action, natural chemicals may enhance the efficacy of already available cancer treatments or diminish treatment resistance [58]. Cancer treatment tries to eliminate or destroy tumor cells while sparing normal ones. The majority of natural substances are less poisonous, less expensive, have fewer side effects, and have been carefully researched for their carcinogenic potential [59]. Due to the adverse effects and drug resistance associated with conventional therapy, it was evident that natural substances can act as anticancer agents or adjuvants in chemotherapy [60].

Cancer chemoprevention reverses, suppresses, or prevents cancer initiation, propagation, or advancement using natural or synthetic medications [61]. To be effective in people, a chemopreventive medicine must have an acceptable safety profile and be efficacious at a low enough dose to avoid severe side effects [62]. Natural dietary interventions such as fruits and vegetables show tremendous promise for chemopreventive research due to their potential to prevent and reduce cancer [63]. The chemical diversity of natural chemicals suggests a range of cancer chemoprevention techniques. Chemoprevention appears to be a rational and

appealing strategy, as indicated by the success of several recent clinical trials aimed at cancer prevention in high-risk populations [61].

Combination therapy combines two or more therapeutic drugs and is a crucial component of cancer treatment [64]. In comparison to monotherapy, the combination of anticancer drugs is more effective because it targets important pathways in a synergistic or additive [65]. This method might reduce drug resistance while providing therapeutic anticancer benefits, such as inhibiting mitotically active cells, reducing cancer stem cell populations, and triggering death [66]. Most metastatic tumors still have poor 5-year survival rates, and creating a new anticancer medicine is expensive and time-consuming [67]. As a result, new techniques are being investigated that target survival pathways and give efficient and effective results at a low cost [68]. In TME, the expression of many genes is regulated, affecting cancer prognosis. Thus, designing combinatorial therapy required evidence to reverse those gene regulations and be free of side effects due to concomitant undesirable gene regulation. In this study, the different combinations of natural compounds have been studied for the treatment of HNSC through various computational approaches.

1.2.MOTIVATION OF RESEARCH

- Targeting tumor cell biomarkers for tumor diagnosis and treatment has limitations because of difficulty in heterogeneity of tumor, drug resistance and their immune modulation.
- Traditional FDA approved drugs showed drug resistance, side effects and are not much effective in cancer patients. Use of high dose of single drug showed toxicity.
- However, traditional cancer drugs have limitations, such as specificity and selectivity. Similarly, current therapeutic biomarkers failed to show effective treatment or diagnosis.
- Thus, there is a growing need to identify novel immune system mediated biomarkers and their therapeutic target needs to be identified.

- In addition, multi-target natural compounds cocktail might be designed for targeting tumors. Combinations of ligand targeting tumor microenvironment will provide less toxic therapeutic approach in immune therapeutics.

1.3.AIM AND OBJECTIVES

1.3.1. AIM:

- Designing a combinatorial therapy targeting tumor progression biomarkers in conjunction with immune modulatory markers for more effective tumor immunotherapy

1.3.2. OBJECTIVES:

- Identification of Immunologically regulated biomarkers as indirect therapeutic targets for combinatorial therapy.
- Mitigation of side effects of chemotherapeutic drugs using Natural compounds.
- Exploring combinatorial potential of natural compounds and their validation via *in vitro* experiments.

CHAPTER II

Objective 1

- Identification of immunologically regulated biomarkers as indirect therapeutic targets for combinatorial therapy

CHAPTER II: OBJECTIVE 1

2.1 RATIONALE OF THE STUDY

Numerous FDA approved drugs for HNSC and Breast cancer have been extensively studied for improving their efficacy and decreasing their side effects. However, their effectiveness is highly variable in different patients. FDA approved drugs for breast cancer (82 Drugs) and Head and Neck cancer (13 Drugs) were retrieved from NCI website and their molecular were identified from the drug bank as shown in **Table 2.1**.

Drug	Molecular Targets	Drug	Molecular Target	Drug	Molecular Target
Abemaciclib	CDK4 and CDK6	Exemestane	Aromatase	Pembrolizumab	PD-1 receptor
Abraxane	Microtubules	Fluorouracil Injection	Thymidylate synthase	Pertuzumab	HER2 receptor
Ado-Trastuzumab Emtansine	HER2	Fam-Trastuzumab Deruxtecan-nxki	HER2-targeted antibody-drug conjugate	Pertuzumab, Trastuzumab, and Hyaluronidase-zzxf	HER2 receptor
Afinitor	mTOR	Fareston	estrogen receptor	Piqray (Alpelisib)	PI3K
Afinitor Disperz	mTOR	Faslodex	estrogen receptor	Ribociclib	CDK4 and CDK6
Alpelisib	PI3K	Femara	Aromatase	Sacituzumab Govitecan-hziy	Trop-2 protein
Anastrozole	Aromatase	Fulvestrant	estrogen receptor	Soltamox	Estrogen receptor modulator
Aredia	Bisphosphonate	Gemcitabine Hydrochloride	DNA synthesis	Talazoparib Tosylate	PARP
Arimidex	Aromatase	Goserelin Acetate	GnRH receptor	Talzenna	PARP
Aromasin	Aromatase	Lapatinib Ditosylate	Dual EGFR and HER2 receptor inhibitor	Tamoxifen Citrate	estrogen receptors
Capecitabine	Thymidylate synthase, DNA polymerase	Lapatinib Ditosylate	EGFR, HER2	Taxotere	microtubule function
Cyclophosphamide	DNA crosslinking, DNA synthesis	Letrozole	Aromatase	Tecentriq	PD-L1
Docetaxel	Microtubules	Margetuximab-cmkb	HER2	Tepadina	Alkylating agent that crosslinks DNA strands
Doxorubicin Hydrochloride	Topoisomerase II, DNA intercalation	Megestrol Acetate	Progesterone receptor	Thiotepa	Alkylating agent that crosslinks DNA strands
Elacestrant Dihydrochloride	Estrogen receptor	Methotrexate Sodium	Dihydrofolate reductase inhibitor	Toremifene	estrogen receptors
Ellence	Topoisomerase II	Neratinib Maleate	HER2, EGFR	Trastuzumab	HER2 receptors
Enhertu	HER2-targeted antibody-drug conjugate	Olaparib	PARP	Trastuzumab and Hyaluronidase-oysk	HER2 receptors

Table 2.1: List of FDA approved drugs with their molecular targets.

Their polymorphism analysis was conducted by COSMIC database. It was found that most of these genes are polymorphic in cancer patients as shown in **Figure 2.1(A, B)**, and it was observed that the polymorphism increases upon treatment with drugs. This heterogeneity can

be attributed to variability of frequency of target gene products in different patients. Also, the approach of many of these drugs leads to a progressive development of resistance. For example, As shown in **Figure 2.1(C, D)** mutation rate in EGFR is 9.6% which increases to 65.51 % after inducing EGFR targeting drug.

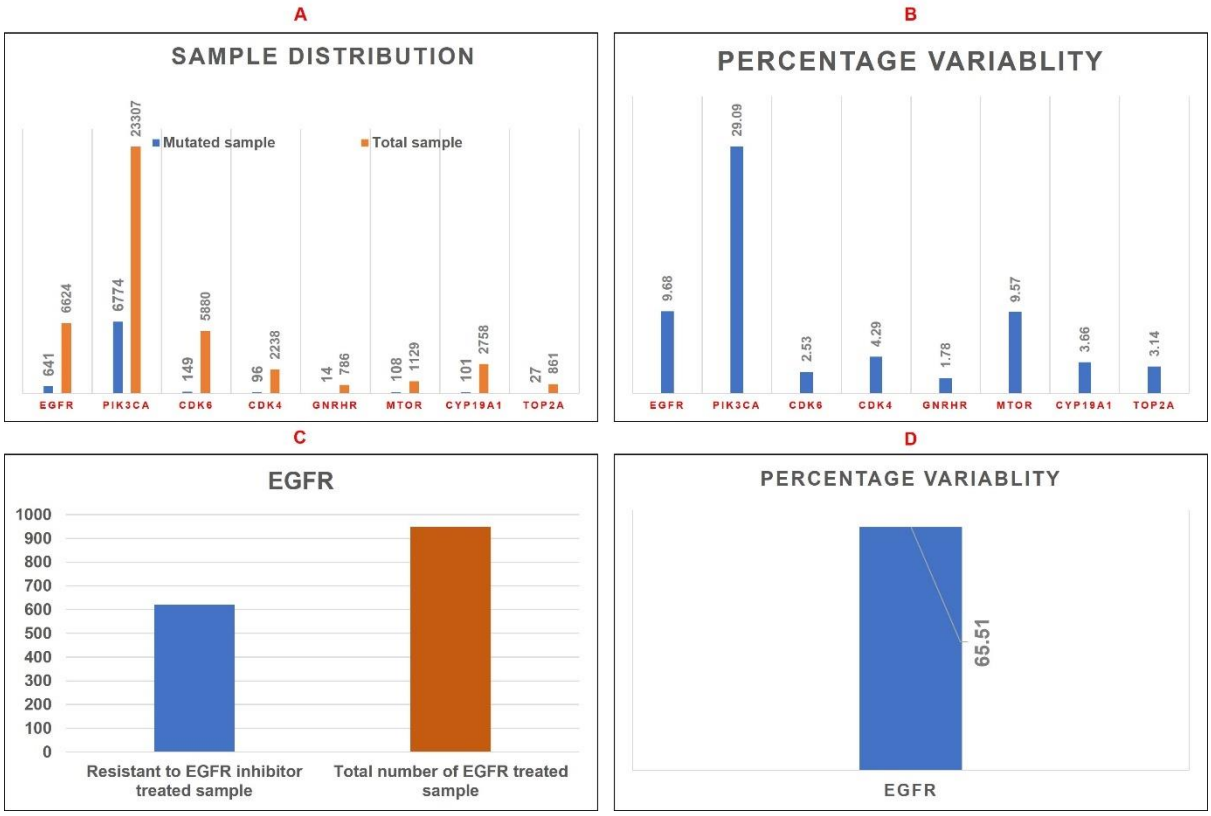


Figure 2.1: Following figure demonstrate the mutated sample distribution and their percentage variability of different targets in cancer patients in figure A and B. Polymorphism even increases after treatment with their inhibitor shown in figure C and D.

So, indirect targeting of cancer cell though immune system might be beneficial for tumor regression. So, targeting those genes which help cancer cells to escape immune response might be a better approach. Our novel approach of targeting immunological regulated biomarkers can prove the way for more widespread therapeutics that would be effective in wider cross section of cancer patients and instrumented in decreasing progression of cancer in all patients.

2.1.1 BENEFITS OF IMMUNOLOGICALLY REGULATED BIOMARKERS

Immunologically regulated biomarkers play a crucial role in cancer research and treatment due to the complex interactions between the immune system and cancer cells. Understanding and

monitoring these biomarkers can provide valuable insights into the tumor microenvironment, the effectiveness of immune responses against cancer, and the potential for immunotherapy interventions. Here are some key factors for using immunologically regulated biomarkers in cancer:

Immune Response Assessment: The immune system plays a critical role in recognizing and eliminating cancer cells. Immunologically regulated biomarkers, such as immune cell populations, cytokines, and chemokines, can provide information about the immune response within the tumor microenvironment. Assessing these biomarkers helps determine the presence and activity of immune cells, their infiltration into tumors, and the overall immune status. This knowledge is essential for understanding how the immune system is interacting with the tumor and can guide treatment decisions.

Predicting Treatment Response: Immunologically regulated biomarkers can help predict which patients are more likely to respond to specific immunotherapies. For example, the expression of programmed death-ligand 1 (PD-L1) on cancer cells has been used as a biomarker to select patients for immune checkpoint inhibitor therapies. Additionally, the presence of tumor-infiltrating lymphocytes (TILs) has been associated with better response rates to immunotherapy in some cancers. By identifying biomarkers associated with treatment response, healthcare providers can personalize treatment plans and optimize patient outcomes.

Monitoring Treatment Efficacy: Immunologically regulated biomarkers can be used to monitor the effectiveness of immunotherapy during the course of treatment. Changes in biomarker levels, such as tumor-specific antigens, immune cell populations, or cytokine profiles, can indicate the activation or suppression of immune responses against cancer cells. Monitoring these biomarkers over time allows clinicians to assess treatment efficacy, make adjustments if necessary, and determine the duration of therapy.

Prognostic Indicators: Certain immunologically regulated biomarkers have prognostic value

and can provide insights into the overall outcome and prognosis for cancer patients. For example, the presence of tumor-infiltrating lymphocytes (TILs) has been associated with better overall survival in various cancers. These biomarkers help stratify patients into different risk groups and guide treatment decisions accordingly. They also aid in estimating the likelihood of disease recurrence and identifying patients who may benefit from additional therapies.

In summary, immunologically regulated biomarkers in cancer provide valuable information about the interaction between the immune system and tumors. They enable the assessment of immune responses, prediction of treatment response, monitoring of treatment efficacy, prognostic evaluation, and development of novel therapeutic strategies. Incorporating these biomarkers into clinical practice improves patient selection for immunotherapies, enhances treatment outcomes, and advances the field of precision oncology.

2.2 METHODOLOGY AND MATERIALS REQUIRED

2.2.1 DEVELOPMENT OF EFFICIENT MACHINE LEARNING ALGORITHM FOR IDENTIFICATION OF PBMC BASED BIOMARKERS

2.2.1.1 DATA RETRIEVAL

The datasets for peripheral blood cells from breast tumor patients and normal samples were obtained from NCBI-GEO Database. Two datasets were identified with suitable numbers of samples and matching queries. GSE27562 contains 162 samples. Of them, 31 are from normal women, 57 are from malignant BC patients, 37 are from benign BC patients, and 37 are from patients of other cancers termed ectopic samples. GSE47862 contains 321 samples. Out of them, 52 are from BC patients who had no family history of BC, 43 are from normal women who had no family history either, 106 are from breast cancer patients with a family history, and 120 are from normal women who had a family history of BC.

2.2.1.2 DATA PRE-PROCESSING

GSE27562 and GSE47862 GEO datasets were integrated to construct the final dataset. The

quality of the dataset must be verified, so for this purpose, batch normalization of the dataset has been done, which was achieved by the gene standardization method, a location-scale method. Gene-wise standardization modifies the values of all genes such that their means equal zero and standard deviations (SDs) equal one. This is performed by removing the mean from each gene's sample data and dividing the resulting value by its standard deviation. Batch normalized expression data was further quantile normalized to remove additional biases from the obtained expression data. Quantile normalization substitutes each attribute (row) in the data with the mean of all attributes across all samples in the same order. The following procedure was employed to normalize a raw high-throughput data collection including multiple samples: Sort the attribute values included inside each sample. (2) Calculate the mean of each attribute's rows. Replace the raw characteristic with its average value. (4) Rearrange all altered values such that they are in the same order as before they were updated.

2.2.1.3 MACHINE LEARNING MODELS IMPLEMENTATION

The training and testing sets were made from the dataset randomly in a ratio of 80 to 20. ML techniques such as SVMs, KNNs, etc., have recently gained more popularity in healthcare fields such as gene expression analysis, drug discovery, omics data analysis, imaging, etc., it was tempting to apply such ML techniques to our dataset and observe the intriguing outcomes. Because of its huge popularity, we have used the XGBoost ML classifier on our training datasets to generate prediction models, and the testing sets were then used to evaluate the performance of the prediction models. All the XGBoost ML models were validated based on their confusion matrix and the accuracy generated using the testing dataset. The XGBoost is a machine learning classifier that is based on decision trees known to boost the performance of the ML model and has been frequently reported to have beaten other ML algorithms, including random forest, decision trees, regression, etc. Despite having compatibility with several computer languages, XGBoost frameworks are most popular for Python and the associated

scikit-learn framework.

2.2.1.4 EXPLAIN THE ABILITY OF THE TRAINED MODEL

The trained XGBoost model was analyzed by the Explainable artificial intelligence (XAI) analysis with the help of the SHAP library. As XAI is concerned with the decision-making process, it helps in the identification of the features significantly impacting the model's prediction. The implementation of XAI analysis will help in identifying the significant genes, and thereafter further identification/classification of the phenotype/condition, such as test or control, will be done by trained models. A local summary plot was formed to exhibit the values indicating the features contributing to the decision confidence with the help of SHAP values. SHAP stands for Shapley Additive exPlanations. The global feature relevance from training data was shown by the SHAP summary plot, and the top 10 genes (top ranked average SHAP value) features were used to train new XGBoost models again, and the significance of 10 selected genes was validated by comparing new XGBoost models to those previously trained on 16,000 genes.

2.2.2 IDENTIFICATION OF COMMON DIFFERENCIALLY EXPRESSED GENES IN TUMOR PATIENT SAMPLES AND PBMC SAMPLES

2.2.2.1 DATA COLLECTION

Gene expression data of HNSC samples and PBMCs of HNSC Patients were collected from Gene Expression Omnibus (GEO) NCBI[71] with accession no. **GSE83519 and GSE39400** [72]**,** respectively. In GSE83519, 22 HNSC tumors and 22 paired normal samples were studied from the same patients. In GSE39400, there are 28 samples of peripheral blood cells of HNSC patients who underwent surgery by means of expression profiling with a controlled group of 11 patients who underwent surgery in the head and neck region for non-HNSC reasons. RNA was extracted from PBMCs using RNA-bee (Campro Scientific bv., Veenendaal, Netherlands). Microarrays Agilent Low RNA Input Fluorescent Linear Amplification Kit and 4x44K Whole

Human Genome Arrays were used for microarray hybridization (Agilent Technologies, Amstelveen, The Netherlands).

GSE85871 [73] contains gene expression profiles of MCF7 cells cultured in MEM/EBSS (Hyclone),10% fetal bovine serum, 1 mmol/L sodium pyruvate, and 100 mg/mL streptomycin in an incubator containing 5% CO2 at 37 °C with 102 different molecules in TCM (Traditional Chinese Medicines), vehicle control (DMSO). Concentration and duration of compound administration may influence the gene expression patterns. According to the CMAP database, the concentration of natural compounds was set to a single dosage of 10 μM for 12 hours, an internationally accepted concentration for high-throughput screening [74]. Two biological replicates for each group and the data set includes profiles for 212 samples. RNA was isolated from MCF7 cells using TRIzol after pre-treatment (Life Technologies, Carlsbad, CA, US) and analyzed with Affymetrix Human Genome U133A 2.0 (Santa Clara, California, US) for gene expression patterns.

2.2.2.2 DIFFERENTIAL GENE EXPRESSION ANALYSIS

Differential gene expression analysis of geo datasets of GSE83519 and GSE39400 were achieved by the GEO2R tool. Users can compare two or more sets of Samples in a GEO Series to find genes differentially expressed across experimental settings using GEO2R [75]. Annotated gene tables and graphs are provided to help normalize the data, remove the data error, and visualize differentially expressed genes (DEGs). GEO2R is an online microarray data analysis tool that helps compare the raw data files with the processes data files to give DEGs using Bioconductor and limma packages. Bioconductor is an R-based tool that provides different high-throughput genetic data analysis packages. GEO query parses GEO data into R data structures that other R tools can use. Linear Models for Microarray Analysis (limma) is a popular R tool for detecting differentially expressed genes. It can handle various experimental designs and data sources, fixing P-values for multiple testing. Users can execute R statistical

analysis without command-line experience by providing a simple interface. Unlike GEO's other Dataset analysis tools, GEO2R uses the original Series Matrix data file to analyze, enabling faster GEO data analysis. It is vital to understand that this tool can access and analyze practically any GEO Series, regardless of data type or quality.

2.2.2.3 COMMON DIFFERENTIAL GENE ANALYSIS

Common upregulated and downregulated differential genes were selected from the HNSC patient's tumor samples and PBMCs of HNSC patients by comparing their list of differentially expressed genes in Microsoft excel. Microsoft Excel is a platform for computation tools, graphing tools, pivot tables, and Visual Basic for Applications, a macro programming language (VBA). The set of differential genes from datasets GSE83519 and GSE39400 were compared in Microsoft excel and by using conditional formatting > Highlight cells rule > Duplicate Values.

2.2.2.4 ENRICHMENT ANALYSIS

Enrichment analysis of selected upregulated and downregulated genes were achieved by the FunRich tool [76], a standalone tool used for the functional enrichment analysis of genes. Results can be depicted in various forms like Doughnut, Venn, pie, Bar etc., and it can handle irrespective of the organism's verity of gene/protein datasets. Users can search either against the default background database or customized database for functional enrichment analysis in biological processes, pathways, etc.

2.2.3 SELECTION OF DUAL ROLE BIOMARKERS FOR TUMOR SUPPRESSION AND IMMUNE MODULATION

2.2.3.1 DATA COLLECTION

500 genes associated with HNSC were collected from string disease query database [69]. Gene expression data for HNSC was retrieved from NCBI’s GEO. Natural compounds data was collected from np_care database [70] and literature. Gene expression data for the natural

compounds were retried from NCBI's GEO.

2.2.3.2 FUNCTIONAL ENRICHMENT ANALYSIS

Five hundred genes associated with HNSC cancer were imported in Cytoscape from disease query database and functional enrichment analysis with GO Process. Gene ontology (GO) such as biological process, molecular functions, cellular processes, and protein domain analysis associated with these genes were identified. Biological processes involved in immune system were selected and further filtered for those processes involved in immune suppression. Cytoscape is a web tool containing a collection of applications for visualizing molecular biological interactions. and biological pathways and with added annotations like gene expression profiles, enrichment analysis and other state of data. Cytoscape core distribution provides a basic set of features for data integration, analysis, and visualization work can be achieved by the core distribution of cytoscape. Adj. P-value ≤ 0.05 was considered as the significantly enriched biological processes.

2.2.3.3 GENE EXPRESSION ANALYSIS AND LITERATURE EXPLORATION

Gene expression analysis was achieved by the GEO2R, which is a tool that allows users to compare two or more groups of samples to identify genes that are differentially expressed across experimental conditions. Differentially expressed genes are presented as a table ordered by p-value and adjusted P-value significance, and with graphic plots to assess data set quality and visualize differentially expressed genes with their P-value and logFC value. GEOquery and limma R packages from the Bioconductor project are used for comparisons on original submitter-supplied processed data tables. Differential genes were selected based on p-value$\leq$0.05, and |logFC value|$\geq$ 1.

Gene expression data were checked for these 53 genes which were associated with immune system processes in enrichment analysis so that we could select up-regulated genes only. 21 genes were found up-regulated. Literature was explored for these 21 genes for evidence as

tumor promoter and immune suppressor. Out of these 21 genes against 10 genes had enough evidence found for both immune suppressor as well as tumor promotor.

2.2.3.4 NETWORK ANALYSIS OF SELECTED GENES

Selected genes are input as a list in STRING which is a database of known and predicted protein-protein interactions. Physical and functional associations are both included in these interactions, they are curated from interactions aggregated from other (primary) databases, from computational prediction, from knowledge transfer between organisms. Genes with the highest degree of interaction were selected from these ten genes.

2.3 RESULTS AND DISCUSSION

2.3.1 DEVELOPMENT OF EFFICIENT MACHINE LEARNING ALGORITHM FOR IDENTIFICATION OF PBMC BASED BIOMARKERS

2.3.1.1 DATA CLASSIFICATION

The array data for PBMCs of breast cancer (BC) patients obtained from the GEO database was retrieved in normalized and calibrated form, which can be found in **Table 2.2**. Search terms like Breast Cancer and PBMCs were used to obtain the datasets. After retrieval, the datasets were merged based on the attribute "common gene symbols," About sixteen thousand such common genes were incorporated along with their values as features.

GEO Accession Number	Total Sample	Sample class in the dataset	Sample Size	Classification of samples for ML
GSE27562	162	Malignant	57(test)	Test – 252
		Benign	37(test)	
		Ectopic	37(eliminated)	
		Normal	31(control)	
GSE47862	321	Breast cancer without a family history	52(test)	Control- 194
		Normal without a family history	43(control)	
		Breast Cancer with a family history	106(test)	
		Normal with family history	120(control)	

Table 2.2: The table demonstrates the Microarray dataset obtained from the GEO database along with the familial description and the classification of samples that have further been used for ML analysis.

2.3.1.2 DATA PRE-PROCESSING

GSE27562 and GSE47862 GEO datasets were integrated to construct the final dataset and finally, 16,000 common genes were identified in both datasets Their expression profiles were merged and the batch was normalized using the gene standardization method, a location-scale method for batch normalization of data integrated from different datasets. Both datasets are already log-transformed; therefore, quantile normalization was applied to the batch-normalized data to remove further biases from the obtained expression data. Different samples were classified into a binary classification problem: test vs control. The test was the samples of BC patients, and the control was the samples from healthy women.

The normalized expression density plot was created with the help of quantile normalization, shown in **Figure 2.2**.

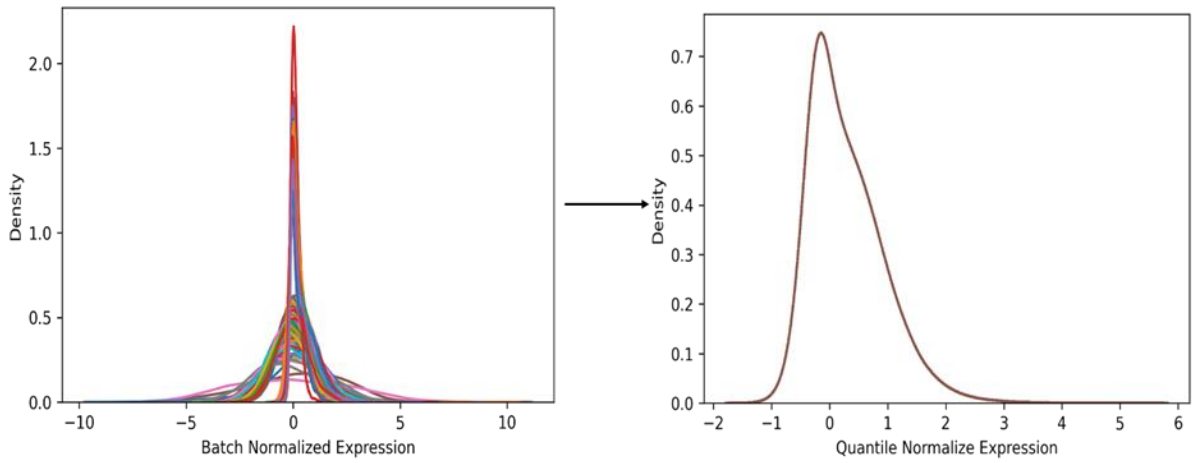


Figure 2.2: The figure demonstrates Batch normalized expression data distribution curves followed by quantile normalized expression data curves.

2.3.1.3 XGBOOST IMPLEMENTATION RESULTS

The dataset was randomly divided into a training set (80%) and a test set (20%) to apply machine learning. With the help of the scikit-learning library, the XGBoost algorithm was applied. The training dataset trained XGBoost Model for further classification on our test vs control dataset. The performance of the model was then checked using the testing sets. The confusion matrix was implemented to check the model's accuracy using the training sets, and

the model's accuracy was obtained using the test set thereafter. There were 28 true positive events, 2 false positive events, 1 false negative event, and 59 true negative events found in the confusion matrix. The accuracy here implies a prediction of the model's performance, which stands for the percentage of correct predictions the model has made. For binary classification, the accuracy was calculated in the form of positives and negatives. as described by the following equation:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

A predictive accuracy of 96.67 % for the test vs control dataset was obtained using the XGBoost ML classifier, which implies that the model did well in distinguishing the features of the test and control.

2.3.1.4 XGBOOST MODELS EXAMINATION WITH XAI

With the help of python's SHAP package, the XAI analysis was implemented on the XGBoost trained model, which is all about the model's decision-making & identifies the features that influence the model's prediction confidence to a great extent, and this analysis helped in finding out the valuable genes from which trained model can separate the corresponding dataset into test (PBMCs of Cancer Patients) and control (PBMCs of Healthy Women). The corresponding SHAP values representing the respective share of a particular attribute to the accuracy of the model's decision were displayed with the help of a local summary plot.

The global significance for every gene was found as the average absolute value of that particular over all of the given samples, with the help of a global feature importance plot that was obtained by the bar plot function where SHAP values were passed as an array. The inference obtained from this global feature importance plot points out the most significant genes in descending order, suggesting the more contribution of genes on the top towards the model's prediction. The bar plot sorts out the most important genes placed on the top. The gene of utmost significance in our machine learning model was STIV, exhibiting a high predictive

value.

With the implementation of SHAP values on the trained models, genes of the highest significance were obtained from the bar plot. The most significant genes in immune cells involved in the progression of Breast Cancer were identified by SHAP listed in **Table 2.3**.

Datasets	Significant Genes
Breast cancer patient's PBMCs vs. Healthy person PBMCs	**SVIP, BEND3, MDGA2, LEF1-AS1, PRM1, TEX14, MZB1, TMIGD2, KIT, FKBP7**, ZNF563, TC2N, LYZ, MAP3K19, GYPE, DSP, ID2, POLR2K, GFPT1, STAM, IRF8, MRPL57, CRYM, SERPIND1, DSG3, APCS, CDH16, HOXD10, TM4SF1, PMEL, COL4A6, MEGF6, HMGB3P1, LRRC20, ZNF668, CLIC3, LRP1B, STK32B, SLC16A10, TSHZ2, PDZRN4, UIMC1, SLC26A6, PIPOX, TMA7, POMGNT2, C19ORF44, CYYR1, DPP10-AS1

Table 2.3: The table shows a list of genes contributing to the model prediction obtained from the merged datasets.

2.3.1.5 EXAMINATION OF XAI OUTPUT

The authenticity of results was checked by applying ML classifier XGBoost on selected genes on the bases of their significance in model prediction. The top ten genes selected by their corresponding significant SHAP values were used to examine the reliability of the results by the ML classifier, namely XGBoost, highlighted in **Table 2.2**. The model's accuracy was 94.44% when trained with the top ten significant genes. **Table 2.4** depicts the accuracy of both the gene sets, i.e., before and after implementing XAI on binary datasets, showing the prediction model's performance in terms of accuracy. The confusion matrix of the model shows that there were 37 true positive events, 48 true negative events, 3 false positive, and 2 false negative events in the model's prediction. The confusion matrix of datasets with 16000 genes and the top 10 genes are shown in **Figure 2.3**.

Datasets	**Breast cancer patient's PBMCs vs Healthy person PBMCs**	
Number of genes in dataset	16000 genes	10 genes
Percentage Accuracy	96.67%	94.44%

Table 2.4: The table shows a comparison of accuracy between the prediction model for the 16000 genes set and 10 selected gene sets.

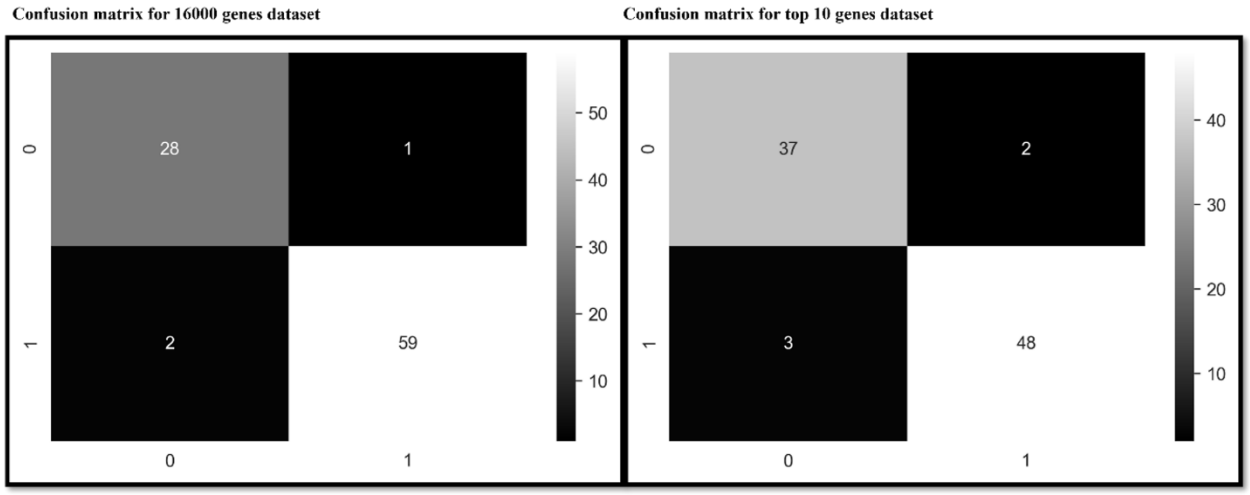


Figure 2.3: The figure shows a comparison of the confusion matrix for PBMCs obtained from Breast cancer patients vs the Healthy person dataset for all 16,000 genes and the top 10 genes. True positive, False positive, False negative, and True negative instances are indicated by a grey box, Black box, Black box, and white box respectively.

The SHAP plot of the top 10 significant genes, shown in **Figure 2.4**, indicates the contribution of the gene to the model's prediction in descending order, which shows SVIP had the highest impact, followed by BEND3, MDGA2, LEF1-AS1, PRM1, TEX14, MZB1, TMIGD2, KIT, FKBP7 respectively.

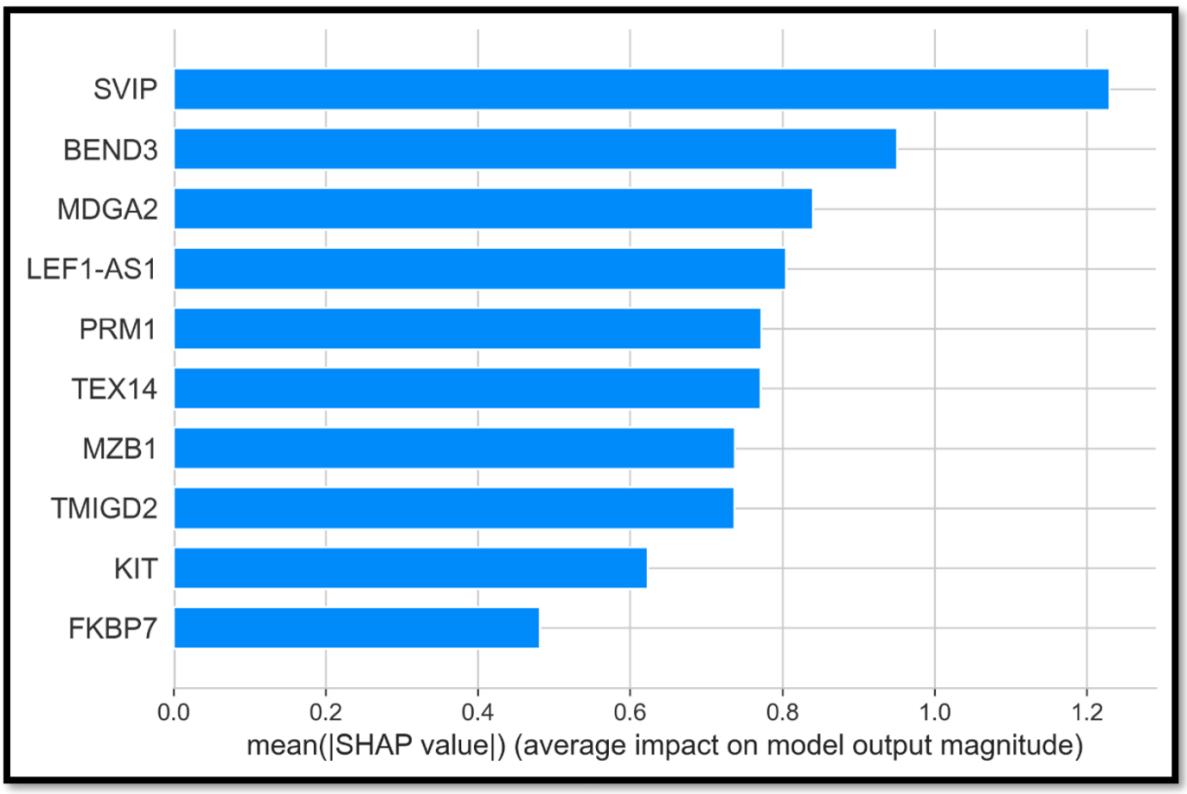


Figure 2.4: SHAP Bar plot illustrates the most significant genes and their SHAP values. The x-axis represents the average/mean absolute value for each gene across all the available data, while the y-axis represents the top 10 genes.

Furthermore, to visualize the predictor's positive & negative associations with the respective genes, the SHAP summary plot was also made, as shown in **Figure 2.5**. The inferences obtained from the SHAP summary plots are as follows: -The ranking of genes (vertically) in descending order signifies their attribute importance. The horizontal line depicts the association of the effect of an attribute on the extent of prediction. The color signifies the impact of a particular gene, maximum significance (in red color) or minimum significance (in blue color). The strong positive impact of SVIP on the SHAP Summary plot indicates the correlativity of the individual gene, where the X-axis signifies the positive impact, and the red color signifies the level (high in this case). Similarly, the inverse connection of BEND3 to the target variable can be ruled out.

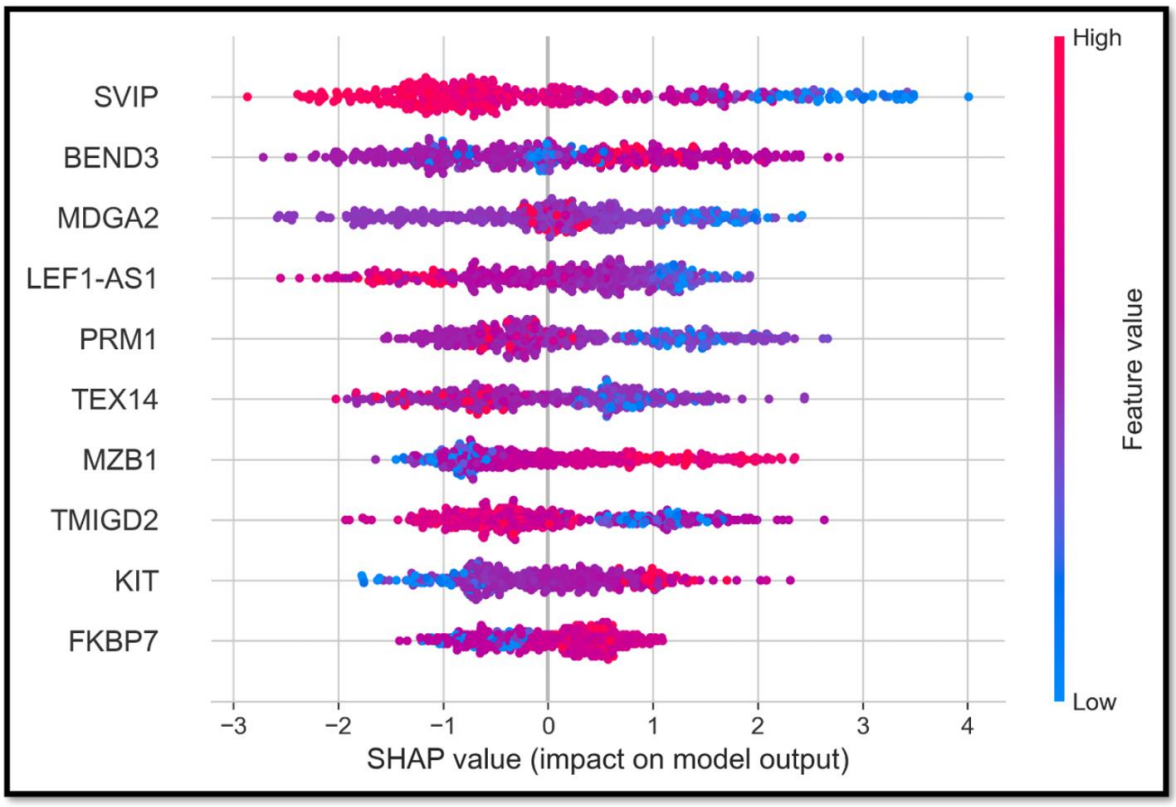


Figure 2.5: The figure illustrates the SHAP Summary diagram, which shows the highly significant genes and their influence on the dataset. On the y-axis, selected genes are sorted in descending order, based on the significance of their characteristic. On the other hand, the x-axis shows the influence of genes on the prediction, illustrating the gene's impact on the model output. The color indicates the influence of a particular gene on a prediction, whether it is statistically significant (in red) or low significance values (in blue).

2.3.1.6 SHORTLISTED GENES STATISTICAL SIGNIFICANCE

The iDEP tool was used to identify key genes differentially expressed in PBMCs during Breast cancer development. P-value ≤0.05 was the criteria for identification as statistically significant SVIP, MDGA2, TMIGD2, LEF1-AS1 and TEX14 were found to be downregulated while BEND3, FKBP7, MZB1, PRM1, and KIT were found to be upregulated in PBMCs of Breast cancer patients as shown in **Table 2.5**.

Genes	P-value	logFC
Dataset: Breast cancer patient's PBMCs vs Healthy person PBMCs		
SVIP	4.82E-14	-1.28E-01
BEND3	1.56E-02	8.31E-01
MDGA2	1.56E-04	-1.29E-01
FKBP7	3.25E-02	7.02E-02
TMIGD2	1.23E-04	-7.31E-02
LEF1-AS1	1.08E-04	-2.14E-01
MZB1	1.27E-03	1.77E-01
TEX14	4.12E-02	-6.66E-01
PRM1	5.63E-03	4.49E-01
KIT	3.67E-04	1.21E-01

Table 2.5: The table demonstrates the P-value and log FC value of the top 10 genes.

Despite the fact that tissue-specific biomarkers, such as aberrant cells, alterations in tumor gene expression, and other malignant abnormalities, may be accurate cancer biomarkers, they have several limitations [131]. It is challenging to employ tissue-specific biomarkers to assess therapy response in real-time due to the invasive nature of biopsy collection [132].TILs may be a valuable prognostic sign for identifying individuals who are most likely to respond to therapy. Biopsies and mammography, which are presently used to identify breast cancer, are painful, costly, and only effective in situations of advanced disease [133]. Mammography may not identify breast cancer immediately since its sensitivity is dependent on tumor size, ranging from 26% at 5 mm to 91% at 10 mm [134]. Breasts with thicker tissue hinder mammography's ability to detect breast cancer [135]. A high level of sensitivity and specificity is required for early cancer detection to increase patient survival rates.

When searching for symptoms of cancer, intrusive tissue collection may be dangerous and may not be the best method for old or delicate individuals [136]. Less invasive and more universally accessible techniques of acquiring biological samples, such as blood collection, may be more acceptable to patients, which might result in a quicker diagnosis [137]. A high level of sensitivity and specificity is required for early cancer detection to increase patient survival rates.

PBMCs mediate the immunological response of the host to tumor cells; hence, peripheral blood

profiling may be used to assess the host's reaction to cancer and offers the possibility of minimally invasive early cancer detection (even before the beginning of clinical symptoms). It can anticipate the prognosis and developmental trajectory of tumors and the clinical outcome. Multiple studies have attempted to identify alterations in PBMC gene expression within breast cancer to categorize subtypes. In individuals with breast cancer, the PBMC transcriptomes correlate poorly with conventional subtypes and are diverse. Using RNA sequencing, Ming et al. determined that ER, PR, and HER2 were not associated with transcriptome-wide PBMC gene expression patterns. The expression of PBMC genes indicates that blood mononuclear cells are immunologically reactive to tumor cells. Therefore, this is not entirely surprising. Similar results were seen for lung cancer patients, who showed high diversity in peripheral blood leucocyte transcriptomes regardless of histological type, with no discernible impact on the peripheral immune system. Therefore, we included PBMC samples from different types of breast cancer patients in our study concerning the stage of cancer, the patient's history of cancer, and different subtypes of breast cancer. 252 breast cancer samples were included in this study. Of them, 37 were associated with benign stage of cancer, 57 were associated with malignant stage of cancer, 106 were from the patient with a family history of breast cancer, and 52 were from patients with no breast cancer history. 194 normal PBMC samples were included in this study for comparison with the tumor PBMC samples. Healthy individuals with a family history and without a family history were also included in the healthy control category. Machine learning algorithm XGBoost was applied to the binary classified dataset for classification, which is followed by the XAI to identify significant genes based on their contribution to the model's prediction. Ten genes were identified in PBMCs of BC patients, which contribute the highest to the models' prediction. These genes were further analyzed for their biological significance and their involvement in different biological processes and their regulation.

2.3.1.7 BIOLOGICAL SIGNIFICANCE OF THE GENES

Each of the top 10 genes was further analyzed for their involvement in biological processes and their regulation to ascertain their impact on cancer progression.

SVIP has tumor suppressor properties, and its restoration is linked to enhanced ER stress and growth inhibition [138]. According to proteomic and metabolomic studies, mitochondria enzymes and oxidative respiration activity are diminished in tumor cells with SVIP epigenetic deletion [139].

$BEND3^+$ T cells generated more significant quantities of IL-6 and IL-8 than $BEND3^-$ T cells. Multiple inflammatory cells, including neutrophils, basophils, and T lymphocytes, are attracted by IL-8. Activation of $BEND3^+$ T cells, which may produce IL-6 and IL-8 in response to TCR/CD3 stimulation, may be essential for the significant and rapid initiation and development of inflammatory responses at the onset of inflammation. $BEND3^+$ T cell dysregulation may result in chronic inflammation [140]. BEND3 attaches to the promoters of differentiation-associated factors and important cell cycle regulators, such as CDKN1A, which encodes p21 and represses differentiation-associated gene expression by increasing H3K27me3 expression [141].

MDGA2 plays the role of tumor suppressor in many cancers. Hypermethylation of MDGA2 is a prognostic marker in gastric cancer [142]. MDGA2 knockdown enhances cell viability, boosts colony formation, and advances the cell cycle but reduces apoptosis. MDGA2-encoded proteins form a new subfamily of the Ig superfamily and have a distinct structural organization consisting of six immunoglobulin chains [143].

Dysregulation of PRM1 was absorbed in different tumor tissues and peripheral blood of cancer patients [144][145][146]. An abnormal expression of the CTA family gene PRM1 results in a particular humoral immune response [147]. It regulates the invasion, migration, and proliferation of cancer cells [146].

TEX14 upregulation was associated with the abundance of tumor suppressor protein REST in different cancer so it could be a potential therapeutic target [148]. It is essential for kinetochore-microtubule attachment and helps in metaphase to anaphase transition [149].

KIT auto phosphorylates on numerous Y residues that serve as docking sites for downstream effectors once activated. Several downstream mechanisms regulate cell survival and proliferation [150]. SFKs, PI3K p85, phospholipase C-gamma, and adaptors that activate MAP kinase pathways attach to phospho-Y residues on the receptor. KIT mutations are also associated with different types of cancers [151]. KIT plays an important part in the activation of different immune cells like Mast cells, dendritic cells, eosinophils, etc. [152].

TMIGD2, also known as CD28H (CD28 homolog), expressed in Homo sapiens and monkeys, while not in mice, enhanced angiogenesis when overexpressed in different cancers. It's a naïve T cell expressed stimulatory receptor. TMIGD2 is a member of the Ig superfamily and has an IgV-like domain, transmembrane region, and cytoplasmic tail. TMIGD2 has various functions depending on cell types and signaling pathways. It is a receptor of HHLA2 and hence could be a therapeutic target for various anti-cancer therapies [153].

LEF1-AS1 (long noncoding RNA) overexpression is associated with the malignant growth of various tumors, and its knockdown inhibits the progression of many cancers. LEF1-AS1 mainly regulates ERK, Akt/mTOR signaling, Wnt/β-Catenin, and Hippo signaling pathways hence playing diverse roles in tumor progression and immune regulation.

FKBP7 could be the therapeutic target for various cancer, especially in case of drug resistance, like the taxane-resistance mTOR pathway can be controlled by targeting FKBP family proteins [154][155].

MZB1 expression is associated with the progression of different cancers and patients' disease-free survival [156][157]. MZB1 is essential for plasma cell differentiation and humoral immune response independent of T-cells by plasma cells [158] and enhances the secretion of interferon

α by dendritic cells [159].

2.3.1.8 BIOLOGICAL PROCESSES REGULATED BY GENES

Enrichment analysis of the top ten selected genes was achieved by the Funrich tool [160]. Biological processes which are statistically significantly regulated by these genes were identified based on their P-value, which should be less than 0.05. It was found that these genes were mainly involved in Apoptosis, Signal transduction, regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism, and Cell communication as shown in **Figure 2.6**.

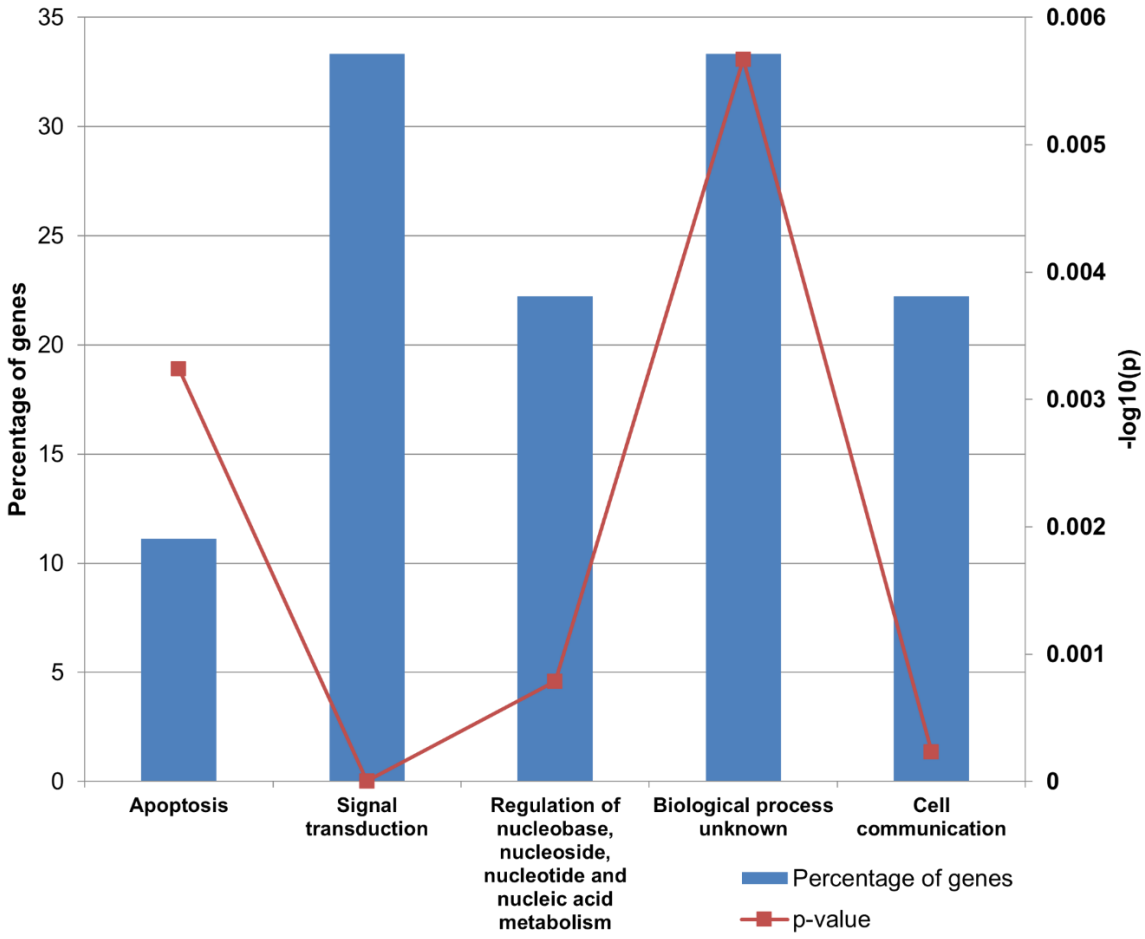


Figure 2.6: The figure demonstrates the percentage of the top 10 genes that are involved in different biological processes or their regulation (bar graphs) and the p-value for their statistical significance in each case is represented by line plots.

2.3.1.9 BIOLOGICAL PATHWAYS REGULATED BY GENES

Biological pathways regulated by these 10 genes were analyzed by the Funrich tool[160] and it was found that KIT signaling, GM-CSF signaling, NOTCH, TGFBR, interleukins signaling,

wnt signaling, cytokine signaling in the immune system, CDC42 signaling and EGF receptor signaling were the main pathways regulated by them. The significance of these pathways was analyzed statistically based on their P-value which should be less than 0.05 as shown in **Figure 2.7**.

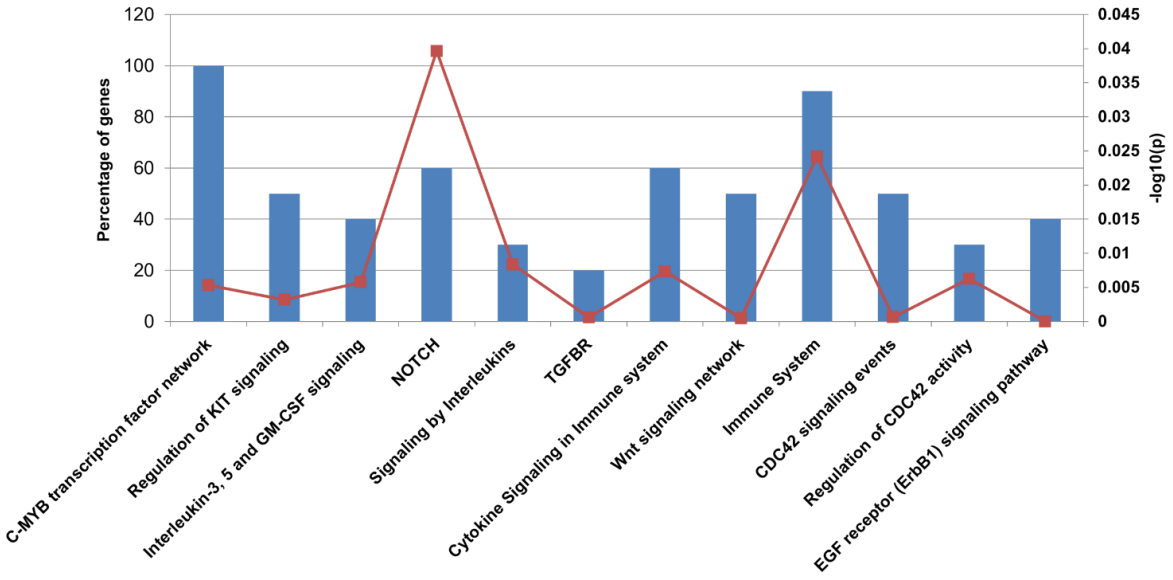


Figure 2.7: The figure demonstrates the percentage of the top 10 genes that are involved in different biological pathways or their regulation (bar graphs) and the p-value for their statistical significance in each case is represented by line plots.

The significance of these 10 genes was noticed to play a significant role in the regulation of cancer progression and regulation of the immune system that is actively involved in cancer mitigation. They were found to be related to biological processes and pathways that are very much involved in the regulation of cancer metastatic progression. Significant evidence was found in the literature proving their immunological role and contribution to cancer progression. Therefore, these genes could be the potential PBMC biomarkers of breast cancer which can help in early detection and could be the non-invasive alternative to breast cancer detection.

2.3.2 SELECTION OF DUAL ROLE BIOMARKERS FOR TUMOR SUPPRESSION AND IMMUNE MODULATION

2.3.2.1 DIFFERENTIAL GENE EXPRESSION ANALYSIS

Data Normalization and Differential gene expression analysis of the GSE83519 dataset was

achieved by the GEO2R tool, and a cut-off of adj. P-value ≤0.05, logFC Value≥1 for differentially upregulated genes and logFC Value≤-1 for differentially downregulated genes. This dataset's Normalization plot and volcano plot are shown in the **Figure 2.8.**

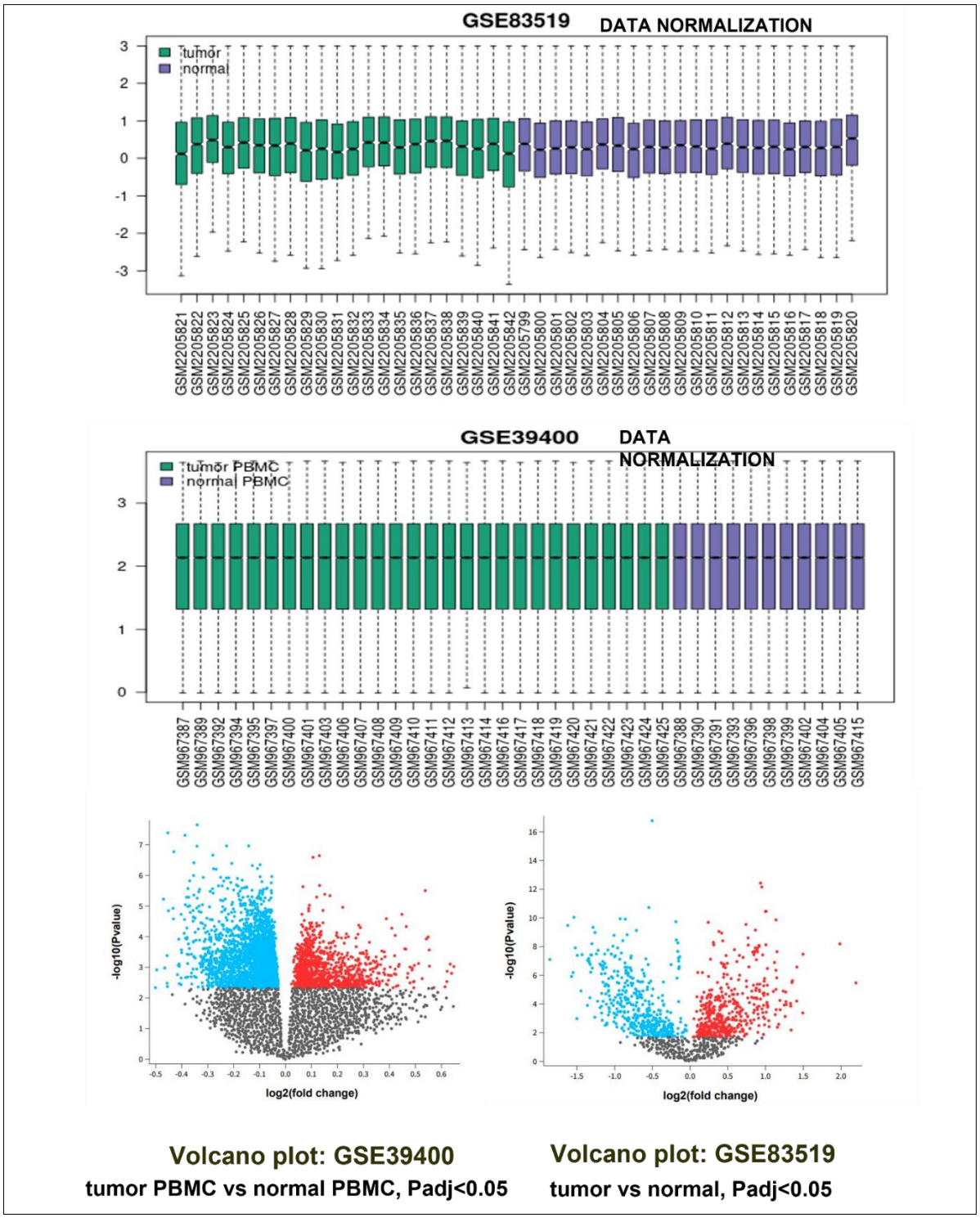


Figure 2.8: Above figure demonstrate the normalization plot and volcano plot for GEO83519 dataset between HNSC tumor vs Normal samples and GEO39400 dataset between tumor PBMC vs normal PBMC with an adjusted Pvalue <0.05 and blue dots shows upregulated genes and red dots shows downregulated genes there.

1094 genes were found to be differentially upregulated, these genes had high expression in the HNSC tumor sample as compared to the respective normal sample of the same patient, and 889 genes were found to be differentially downregulated; these genes had low expression in the HNSC tumor as compared to the respective normal sample. GSE83519 dataset contains the tumor microenvironment samples, which include not only tumor cells but also other cell mediators like immune cells, fibroblast, blood vessels etc., so when immune cell comes in contact with this tumor microenvironment, they may alter their gene expression profile; therefore the expression data of PBMCs need to be analyses individually so we can identify the differentially expressed genes in PBMCs because it is beneficial to target immune cell for the tumor regression along with only targeting tumor cells. Differential gene expression analysis of the GSE39400 dataset was also achieved by GEO2R Tool with a cut-off adj. P-value ≤ 0.05, logFC Value>0 for upregulated genes and logFC Value<0 for downregulated genes. The volcano plot of this dataset is shown in the **Figure 2.8.**

737 genes were found to be upregulated, these genes had high expression in the PBMCs which were retrieved from HNSC patients after surgery as compared to the PBMCs retrieved from patients who got head and neck surgery for a non-HNSC reason, and 1954 genes were downregulated in PBMCs of tumor patients after surgery when compared with patients who got head and neck surgery for a non-HNSC reason as shown in **Figure 2.9**. PBMCs include dendritic cells, lymphocytes (NK cells, B cells, T cells), and monocytes. Therefore, these DEGs are mainly present in dendritic cells, lymphocytes, and monocytes. They might alter these immune cells' function and help tumor cells escape the immune system. These genes can be further analyzed to screen immunological biomarkers for HNSC.

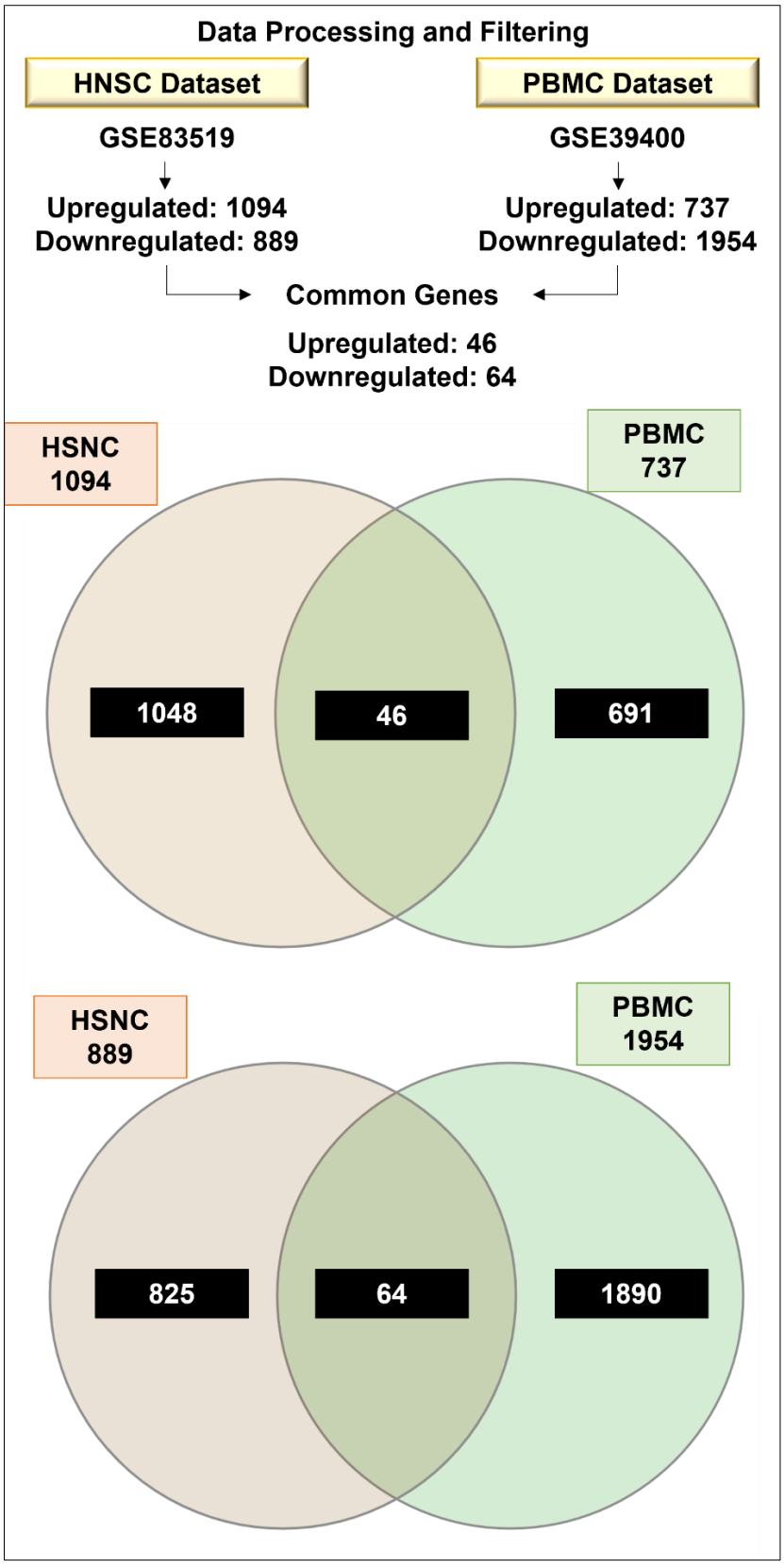


Figure 2.9: Following figure demonstrate the data processing and filtering. HNSC Dataset GSE83519: there are 1094 upregulated genes and 889 downregulated genes and PBMC dataset GSE39400: There are 737 upregulated genes and 1954 downregulated genes. Ven diagram showing that there is 46 common upregulated genes and 64 common downregulated genes in HNSC cancer tissue sample and PBMC of HNSC cancer patients.

2.3.2.2 SCREENING OF COMMON DIFFERENTIALLY EXPRESSED GENES

Such DEGs derived from immune regulators and those differentially expressed in the tumor microenvironment also need to be identified. DEGs of HNSC tumors were compared with the DEGs of PBMCs so that common DEGs could be screened. Therefore, 1094 upregulated DEGs of HNSC were compared with 737 upregulated DEGs of PBMCs and found 46 common genes. These 46 DEGs were upregulated in HNSC tumor samples and PBMCs of HNSC patients. As shown in Venn diagram in **Figure 2.9**. PBMCs may infiltrate the tumor and affect tumor progression. The influence of the tumor microenvironment alters the expression of these genes in the tumor-infiltrating lymphocytes. These genes might be involved in the alteration of gene regulation in the subset of the immune cells in the vicinity of the tumor in HNSC patients due to the complex interplay of cells in the tumor microenvironment. The list of common upregulated genes with their Adj. P-value and logFC in HNSC tumor samples and PBMCs of HNSC patients are shown in **Table 2.6**.

Gene Symbol	HNSC		PBMC	
	logFC	Adj. Pvalue	logFC	Adj. Pvalue
MITF	2.4516	2.65E-11	0.1022	0.0333
HFE	1.6367	7.44E-09	0.3233	0.0095
TRPM6	1.1952	2.03E-08	0.2776	0.0173
DOCK4	1.2461	7.59E-06	0.1514	0.0265
RABEP1	2.0842	2.45E-10	0.1074	0.0273
SLC44A1	2.8497	4.33E-12	0.089	0.0312
GK	3.9583	2.49E-10	0.0854	0.0108
NSUN7	1.7051	6.94E-04	0.1985	0.0433
VEGFA	1.0078	2.59E-05	0.1314	0.0185
RRAGD	2.3623	1.53E-14	0.081	0.0447
PCMT1	3.3705	5.23E-08	0.1044	0.0079
FCER1G	1.4801	6.60E-11	0.0648	0.0428
LATS1	1.1526	2.78E-10	0.2295	0.0455
S100A9	1.1011	1.32E-06	0.1056	0.001
LAT2	4.7458	1.79E-14	0.0338	0.0423
MAP2K6	1.6879	4.57E-05	0.3708	0.0128
MAPK14	1.9209	1.54E-03	0.0705	0.0165
PPM1A	1.8701	2.81E-06	0.0659	0.0406
IL18	1.7222	1.63E-05	0.1321	0.0452
ACTB	1.6093	3.43E-07	0.0454	0.033
TLR5	1.1144	5.67E-05	0.1595	0.0135
PANK3	3.2487	7.05E-21	0.1055	0.0135
AHR	1.4695	4.56E-07	0.0692	0.0498
LAP3	1.5307	4.07E-04	0.0877	0.0187
CYP1A2	2.8793	1.59E-05	0.1617	0.0334
SMPDL3A	1.2163	4.34E-06	0.1037	0.0197
RNASE1	1.8432	2.29E-04	0.1378	0.0327

WAC	1.004	4.78E-09	0.0464	0.0135
SERPINB2	2.0987	1.21E-07	0.4873	0.0242
RBM47	1.3479	7.22E-08	0.0566	0.0359
GPR15	3.2539	8.28E-12	0.2827	0.0444
CCDC88A	1.0265	1.35E-04	0.125	0.0171
PRNP	1.0994	4.91E-10	0.0544	0.0286
ACPP	2.6565	3.18E-09	0.1566	0.0493
ST3GAL6	2.7894	6.49E-11	0.0902	0.0188
UBOX5	2.5519	2.37E-05	0.0558	0.0207
GRM2	1.8173	5.71E-07	0.2123	0.0442
ROCK2	1.3506	7.94E-05	0.1742	0.0096
ZEB2	2.9773	2.25E-07	0.1066	0.0111
MFSD1	1.0859	3.21E-10	0.0341	0.0499
CORO1B	1.1436	1.13E-03	0.1089	0.0187
PPARG	1.1479	1.10E-10	0.2847	0.037
IPO7	1.2914	5.18E-07	0.216	0.0217
PAX9	1.9269	4.35E-12	0.4035	0.0215
ERN1	1.2908	8.75E-10	0.2779	0.0454
CLDN9	1.2167	1.29E-07	0.2484	0.0213

Table 2.6: logFC and adj. Pvalue of the 46 common differentially upregulated genes in HNSC tumor samples vs normal samples and PBMCs of HNSC cancer patient's vs PBMCs of normal person. Different tones of colors (light to dark) in the given table demonstrate the level of expression of upregulated genes.

Similarly, 889 downregulated DEGs of HNSC tumor samples were compared with 1954 downregulated DEGs of PBMCs of HNSC patients, and 64 genes were common in both. The list of common downregulated genes with their Adj. P-value and logFC in HNSC tumor samples and PBMCs of HNSC patients are shown in **Table 2.7**. As shown in **Figure 2.9**, there were 110 common DEGs in HNSC samples and PBMCs, out of which 46 common upregulated genes as shown in the first Venn diagram and 64 common downregulated genes as shown in the second Venn diagram.

Gene Symbol	HSNC		PBMC	
	logFC	Adj. Pvalue	logFC	Adj. Pvalue
ZAP70	-2.201727	1.41E-06	-0.129217	0.016018
RPL27A	-1.290623	2.47E-12	-0.044947	0.006435
POM121	-1.176432	4.38E-03	-0.204749	0.029512
PPP1R13B	-1.705423	1.32E-05	-0.225206	0.025134
CD6	-1.026282	7.84E-05	-0.100521	0.020264
ZNF764	-1.262351	5.95E-07	-0.080366	0.022201
ARHGEF5	-1.215285	1.23E-08	-0.16608	0.010322
TRAM2	-2.03586	9.60E-03	-0.108028	0.024255
RIC3	-1.504849	2.89E-04	-0.372801	0.00283
TBRG4	-2.395657	2.21E-09	-0.08776	0.01264
KAT6B	-1.269801	4.11E-09	-0.266388	0.034823
DNAJA4	-1.983303	1.91E-04	-0.086775	0.023274
SHMT2	-2.164514	9.19E-14	-0.071538	0.02023
CYP2U1	-1.410565	1.20E-10	-0.101269	0.013115
MCF2L	-1.081094	9.52E-08	-0.300434	0.017142
MLF1	-1.010867	4.47E-08	-0.266692	0.035152
SOD1	-1.302701	2.45E-03	-0.077911	0.004778

ACSF2	-1.409186	7.85E-07	-0.113655	0.027057
WDR74	-1.168891	5.26E-07	-0.092598	0.004078
PDE9A	-1.383244	1.02E-09	-0.285779	0.014934
WDHD1	-1.299895	5.06E-13	-0.135075	0.016697
MRPL58	-1.147314	5.17E-03	-0.065359	0.026208
SAFB2	-1.554142	3.05E-04	-0.071198	0.010002
CACNA2D2	-1.891351	6.07E-08	-0.194732	0.011568
GPD1L	-1.769583	2.66E-05	-0.093498	0.003471
ZNF318	-1.600488	7.17E-12	-0.194653	0.015039
SMARCC2	-2.490448	7.71E-11	-0.055608	0.017597
KDM8	-1.6758	7.66E-10	-0.255286	0.026751
BCR	-2.256436	1.17E-10	-0.071938	0.013975
SRSF8	-1.439907	1.12E-06	-0.08949	0.032799
SNRNP40	-1.285771	2.58E-04	-0.090312	0.038549
SLC4A10	-1.323454	4.61E-06	-0.415896	0.029181
DHX30	-1.057061	6.04E-05	-0.082518	0.008847
PCNX2	-1.664282	1.61E-04	-0.209404	0.012008
RPAIN	-1.555945	1.14E-04	-0.061706	0.048962
TRADD	-1.501988	8.72E-06	-0.048286	0.034225
PTGDR	-1.063566	1.51E-05	-0.123149	0.040107
ZBTB20	-1.348264	5.41E-09	-0.115621	0.00839
PIP4K2B	-1.329684	5.76E-10	-0.052544	0.042676
PPARD	-1.157839	1.10E-03	-0.051296	0.023755
NHP2	-1.237519	2.75E-04	-0.055893	0.04521
ACACA	-1.138407	6.28E-06	-0.109746	0.014158
AMMECR1	-2.10191	1.09E-08	-0.141336	0.027549
FANCI	-1.100655	1.35E-07	-0.105654	0.01507
CSNK1E	-1.008658	4.98E-03	-0.058505	0.012541
CKS1B	-1.360469	2.10E-14	-0.063722	0.030507
NF2	-1.232306	4.07E-03	-0.165997	0.017791
ZNF205-AS1	-1.667942	1.77E-04	-0.26602	0.019041
CD22	-2.681652	5.46E-12	-0.195684	0.01068
PTCD2	-1.199056	3.69E-03	-0.111281	0.013654
DOLPP1	-1.282885	1.51E-12	-0.063006	0.022189
SERGEF	-1.053468	3.30E-07	-0.107393	0.003703
RFC4	-2.605413	6.72E-11	-0.086207	0.010768
MMP11	-1.187203	1.98E-06	-0.333011	0.003549
GPX7	-1.874631	1.61E-05	-0.100866	0.021288
PROCR	-1.335837	4.47E-10	-0.16094	0.020739
UBFD1	-1.017601	8.20E-03	-0.071952	0.027642
RPS17	-1.140295	4.07E-08	-0.053202	0.004057
FEZ1	-1.945739	2.64E-06	-0.200686	0.018734
GPALPP1	-2.010168	1.67E-08	-0.083729	0.045393
NCR3	-1.262607	3.52E-07	-0.188793	0.00238
SKI	-1.105797	7.90E-08	-0.22183	0.0472
RARRES3	-1.54742	1.83E-10	-0.068442	0.021288
DENND2D	-1.056931	1.03E-04	-0.14677	0.047475

Table 2.7: logFC and adj. Pvalue of the 46 common differentially downregulated genes in HNSC tumor samples vs normal samples and PBMCs of HNSC cancer patient's vs PBMCs of normal person. Different tones of colors (light to dark) in the given table demonstrate the level of expression of downregulated genes.

Enrichment analysis of these 110 DEGs was achieved and found that 46 common upregulated genes were mainly involved in biological processes like signal transduction, cell migration, RNA metabolism, Anti-apoptosis, regulation of cell cycle, regulation of gene expression, cell

communication, energy pathways, transport, protein metabolism, immune response, cell growth and/or maintenance. These biological processes might help in tumor progression because processes like suppression of apoptosis, cell migration, cell cycle regulation, cell growth and/or maintenance directly support tumor growth. Biological processes like immune response, signal transduction, cell communication, etc., could play an essential role in the tumor microenvironment for tumor progression. Therefore, the overexpression of these genes enhances these biological processes in the tumor microenvironment, which could help in tumor progression. 64 common downregulated genes were mainly involved in biological processes like CGMP-mediated signaling, ribosome biogenesis and assembly, immune response, regulation of signal transduction, RNA metabolism, Transcription, DNA repair, signal transduction, cell communication, transport, protein metabolism, energy pathways, metabolism, apoptosis. These biological processes are also linked with tumor progression or regression; therefore, downregulation of these genes could help tumor progression. So those drugs should be screened, which alter the expression of these genes to restore normal expression levels such that normal biological processes are restored. Enrichment analysis of 46 upregulated and 64 downregulated genes are shown separately in **Figure 2.10**.

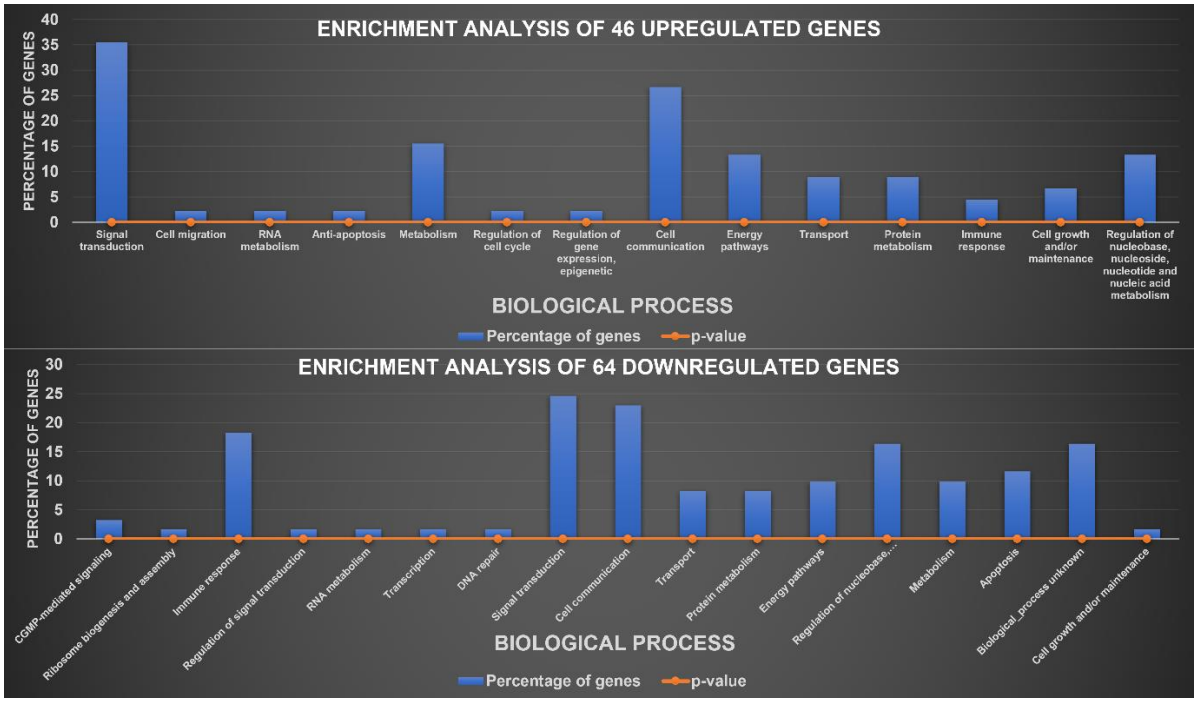


Figure 2.10: Following figure demonstrate the enrichment analysis of 46 common upregulated genes and 64 common downregulated genes with respect to biological processes.

2.3.3 IDENTIFICATION OF COMMON DIFFERENCIALLY EXPRESSED GENES IN TUMOR PATIENT SAMPLES AND PBMC SAMPLES

2.3.3.1 ENRICHMENT ANALYSIS

An enrichment analysis of 500 genes that are known through experimental validation to be ones that are most explicitly associated with HNSC cancer, were conducted and immune-associated genes were selected. 256 genes related to the immune system were found out of which 53 genes were found to be associated with the negative regulation of immune system associated processes, such as negative regulation of T cell activation, negative regulation of B cell activation, negative regulation of B cell proliferation etc. as shown in **Figure 2.11**.

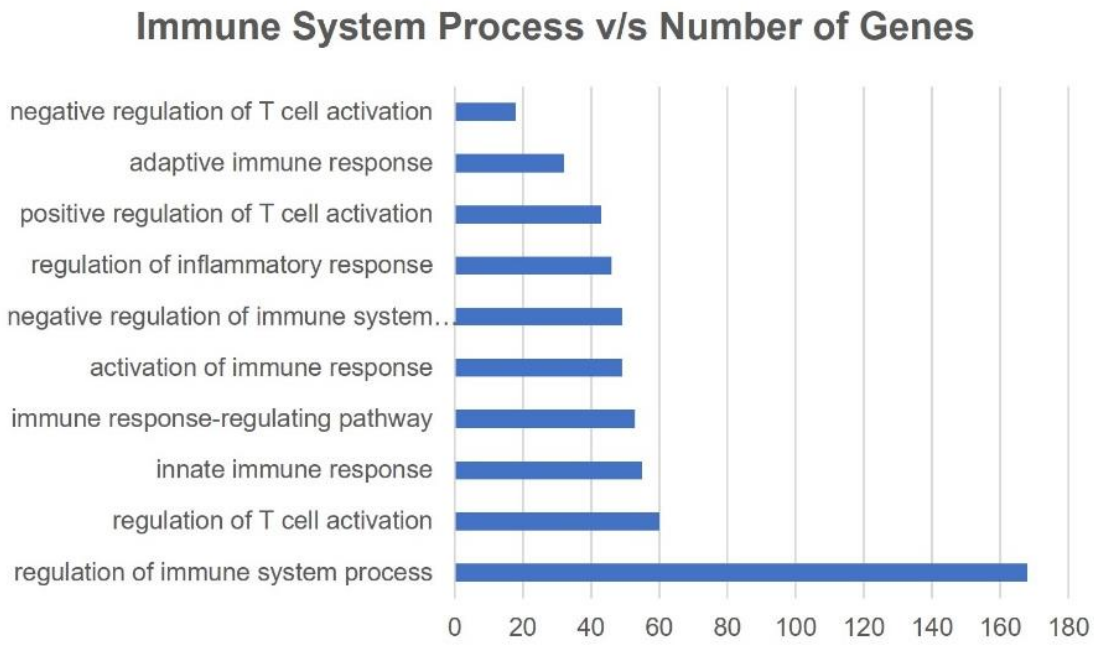


Figure 2.11: The following figure demonstrate enrichment analysis data by GO Process involved in immune system related processes which shows number of genes from the HNSC associated involved in different immune system related processes by enrichment analysis GO Process.

2.3.3.2 GENE EXPRESSION ANALYSIS

Gene expression data was analyzed for the 53 genes associated with the negative regulation of the immune system, out of which 21 genes were identified as having a LogFc value greater than or equal to 1. These genes could be associated with both immune suppression as well as tumor progression. Complete gene expression data is provided in the supplementary table1.

These 21 genes were functionally validated by annotation from literature for their association with tumor progression related process like cell proliferation, metastasis etc. and immune suppression related process like T cell inactivation, development of tumor-associated macrophage etc.

It was found that 10 genes were associated with both the above-mentioned processes and 7 genes were associated with immune suppression only and 2 genes were associated with tumor progression only and remaining 2 genes were associated with HNSC cancer due to alteration of their function by mutations as shown in **Table 2.8**.

S.No.	Genes	Tumor Progression Role	Immunosuppression Role	Targeting Drugs (FDA approved / in Clinical Trials)
1	TGFB1	Epithelial-Mesenchymal Transition [77][78] Metastasis initiation [80]	Inhibits CD8+ T-cell and NK-cell mediated anti-tumor immune responses. Inhibits activation of neutrophils [79].	lerdelimumab and metelimumab (in clinical trial)
2	IRF1	Upregulate PD-L1 in the tumor cell [81].		
3	TWSG1	Enhancing tumor growth and malignant cell behavior and stimulating tumor-associated angiogenesis [82].		
4	CDK6	Regulates the progression of the cell cycle. Transcriptional role in tumor angiogenesis [84].	CDK6 inhibition triggers antitumor immunity [83].	Palbociclib (FDA approved)
5	AXL	Tumor proliferation, survival, metastasis, and resistance to cancer therapy [85].	Small-Molecule Inhibition of Axl Targets Tumor Immune Suppression [86].	
6	FADD	Cell cycle progression and cell proliferation [87].	A negative regulator of T-cell receptor–mediated necroptosis [88].	
7	HAVCR2	Induce epithelial-mesenchymal transition by JAK-STAT3 signaling pathway [89].	Over expression of HAVCR2 observed in tumor-infiltrating lymphocytes which is associated with adaptive resistance to immunotherapy[90].	BMS-986258(in clinical trial)
8	PRKDC	Promotes tumor cell growth via p38 MAPK signalling [91]**.**	PRKDC is not only a predictive biomarker but also a drug target for immune checkpoint inhibitors [92].	NU7026 (FDA Approved)
9	IL10	An association exists between IL-10 expression and tumor-related markers such as Bcl-2 [93].	Inhibited T-cell proliferation and function [94]. It seems that TAMs cause drug resistance via the IL- 10/Stat3/Bcl-1/BCL2 signaling pathway [93][95].	GIT 27 (in clinical trials)
10	SOCS1	SOCS1 downregulation inhibits cell proliferation via cell cycle progression, resulting in accumulation of G0/G1 phase and reduction of S phase [96].	SOCS1 involved in inactivation of CD8+ T cells against tumor cells [97].	
11	MICA		Anti-MICA antibodies can promote the anti-tumor immunity through the induction of direct anti-tumor effects (antibody-dependent cell-mediated cytotoxicity, ADCC) [98].	
12	TIGIT		It suppresses the Function of NK cells and CD8+ T Cells [99] [100].	
13	IDO1	IDO1 expression associated with the progression of tumor [101] by drug resistance mechanism against response of different drugs [102].	IDO overexpression associated with escape from anti-tumor immunity by suppression of mTOR in T cells [103] [104].	Indoximod (in clinical trials)
14	CDKN2A	Mutated	Mutated	
15	LAG3		Inactivates the CD4+ T cells. Reduces the effector function of CD8+	BI 754111(in Clinical Trials)

			T cells. Promotes the suppressor activity of Tregs [105].	
16	CTLA4		Inhibits the activation and proliferation of T cell [106].	Ipilimumab (FDA Approved)
17	CD274	Promotes tumor cell growth, migration and invasion via WIP and β-catenin signalling [107].	CD274 overexpression negatively regulate the T-cell mediated immune response in peripheral tissues[39].	Nivolumab (FDA Approved)
18	HLA-A	Highly polymorphic	Highly polymorphic	
19	PDCD1LG2		PDCD1LG2 overexpression suppressed the tumor antigen specific CD8+ T cells [108].	Atezolizumab (FDA approved)
20	FOXP3	FOXP3 overexpression promotes cell proliferation, migration, and invasion [109].	FoxP3 plays a key role in the development of Treg cells [110].	RPG (FDA approved)
21	PDCD1		PDCD1 overexpression suppress the immune response against tumor [111].	Avelumab (FDA approved)

Table 2.8: List of genes with their tumor progression and immunosuppression roles

The specific immunomodulatory roles of these genes in tumor progression and immune suppression were explored such that single natural inhibitor could impact multiple immune cells and would therefore be crucial in augmenting a concerted immune response in tumor microenvironment.

TGFB1 induced an epithelial-mesenchymal transition (EMT) to increase the invasion of cancer cells. It induces genes that help in metastatic colonization at secondary organ sites, so that TGFB1 works as a promoter of metastases [112]. TGFB1 also functions as an immune suppressor by influencing the development, differentiation, tolerance and homeostasis of immune cells [113]. TGFB1 may promote the development of Treg cells by inducing Foxp3 expression [114]. Other studies have shown that the deletion of TGFB1 in T cells alone, not in tumor cells, has suppressed tumor development in different tumor models [115]. So TGFB1 signaling is crucial for tumor progression as well as immune suppression. Hence, TGFB1 Downregulation is an effective therapeutic strategy.

CDK6 overexpression associated with many cancer progression, it is involved in cell cycle progression, up-regulation of CDK6 also result in increased tumor angiogenesiss [116] it also shows its role in immune suppression as shown in our gene enrichment analysis we also get

evidence from literature such that CDK6 inhibition enhance immunity against cancer by activating T cell, reductio n in immunosuppressive regulatory T cells, increase in antigen presentation [117].

In tumor cell AXL promote proliferation, EMT, metastasis, and resistance to apoptosis [118]. AXL involved in immune regulation by inhibition of T cell activation by NK cells and dendritic cell, TLR signaling, specific tumor killing by NK cells [119].

Overexpression of HAVCR2 helps in tumor progression as well as immune suppression. It might induces EMT hence helps in metastasis [89]. HAVCR2 is over-expressed on tumor infiltrating dendritic cells and can compete with nucleic acid binding to HGMB1, therefore inhibiting anti-tumor immunity [120]. HAVCR2 can form a heterodimer with ceacam-1 which acts as a negative regulator of t cytotoxic cell response [90].

Overexpression of PRKDC promote tumor cell growth and proliferation via p38 MAPK signaling [91].

IL-10 overexpression associated with tumor aggressiveness through CIP2A over-expression via PI3K signaling pathway [121]. IL-10 overexpression inhibited proliferation, cytokine production and migration capacities of effector T cells [122]. Expression of IL-10 on dendritic cells generates Treg, which were produced IL-10 too [123]. IL-10 expression also influence the expression of Foxp3[94], TGF-beta [124] and TGFBR2 [125]. Inhibition reduce the cell proliferation by inhibiting cell cycle progression, resulting in accumulation of G0/G1 phase and reduction of S phage [96]. Overexpression of SOCS1 also involved in the reduction of CD8+ T cells activity against tumor cells [97].

IDO1 expression associated with the progression of tumor [126] by drug resistance mechanism against response of different drugs [127]. IDO1 expression in tumor cell mediate the catabolism of tryptophan is a critical factor of immune escape by suppression of anti-tumor immunity[128]. IDO1 overexpression increases the proliferation of Tregs [129].

CD274 regulates tumor growth, proliferation, migration and invasion by targeting WIP and beta-catenin signaling [107]. CD274 is a well know immune checkpoint. It helps in tumor cell survival by PD-1/PD-L1 interaction which inhibit T cell activation [130].

Foxp3 regulates cell proliferation, migration and invasion of tumor cells [109]. Foxp3 regulates the development and proliferation of immunosuppressor Treg cells [110].

2.3.3.3 NETWORK ANALYSIS OF SELECTED GENES

10 genes were selected which play a crucial role in immune suppression as well as tumor progression. These are Foxp3, CD274, IDO1, IL-10, SOCS1, PRKDC, AXL, CDK6, TGFB1, FADD. Network analysis in string revealed that CD274 has seven degrees of interaction, IDO1 has five degrees of interactions as shown in **Figure 2.12**. Hence these two genes can serve as the preferred target for modulating immune regulation and will impact multiple immune cells and determine tumor prognosis. CD274 gene encodes protein PD-L1 which is an immune suppressor ligand. It is expressed in different tissues but is mainly expressed in activated T cells and B cells, monocytes, dendritic cells and different tumor cells. The interaction of this ligand with PD1 results in immune escape by the tumor cells, by inhibiting T-cell activation and cytokine production. High expression of this gene is a prognostic marker in many cancers. Indoleamine 2,3-dioxygenase 1 (IDO1) is a heme enzyme that catalyzes the first and rate limiting step in the catabolism of tryptophan which changes the behavior of T-cells. This enzyme plays a role in a variety of pathophysiological processes such as antioxidant activity, antimicrobial defense, neuropathology, immunoregulation, antitumor defense. Overexpression of IDO1 is found in different cancers, which is associated with poor prognosis. IDO1 can be inhibited by the cancer-suppression gene bridging integrator 1 (Bin1) and up-regulated by some immune checkpoint molecules and cytokines such as IFN-γ, pathogen-associated molecular patterns (PAMPs, such as Toll-like receptor (TLR) 3, TLR4, TLR7, TLR8, and TLR9), IL-6, prostaglandin E2 (PGE2), damage-associated molecular patterns (DAMPs), immune

checkpoint (including PD-1, glucocorticoid-induced TNF receptor-related protein (GITRL), CTLA-4), and TNF-α, TGF-β to establish an immunosuppressive environment.

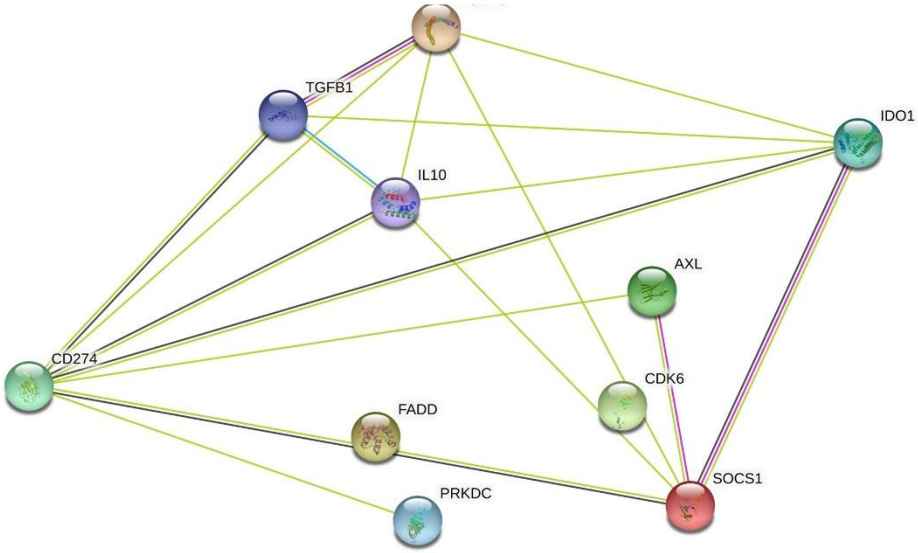


Figure 2.12: Interaction network of ten genes which are immune suppressor as well as tumor progressor for checking degree of interaction. As we can see here CD274 have highest degree of interaction 7 and IDO1 have second highest degree of interaction 5.

CHAPTER III

Objective 2

- Mitigation of side effects of chemotherapeutic drugs using natural compounds.

CHAPTER III: OBJECTIVE 2

3.1. RATIONALE OF THE STUDY

Toxicity of chemotherapeutic agents to normal cells is often due to additional binding of therapeutic drug to off target receptors leading to unpleasant side effects. Hence the efficacy of chemotherapy lies in the fact that chemotherapeutic agent has a significant selectivity for cancer cells over normal host cells. Side effect of FDA approved chemotherapeutic drug were analyzed by intside database (https://intside.irbbarcelona.org). It was found that most of chemotherapeutic drug have their own specific side effects as shown in **Table 3.1**. For example:

Doxorubicin (Adriamycin): Hair loss, nausea and vomiting, fatigue, increased risk of infections, anemia, mouth sores, heart damage (rare but serious), darkening of the nails.

Cyclophosphamide (Cytoxan): Nausea and vomiting, hair loss, increased risk of infections, anemia, bladder irritation, increased risk of secondary cancers.

Paclitaxel (Taxol): Hair loss, joint and muscle pain, peripheral neuropathy, low blood cell counts, nausea and vomiting, allergic reactions, changes in nail color, fluid retention.

Cisplatin: Nausea and vomiting, kidney damage, hearing loss, peripheral neuropathy, low blood cell counts, electrolyte imbalances, allergic reactions.

Methotrexate: Nausea and vomiting, mouth sores, low blood cell counts, liver toxicity, skin reactions, increased risk of infections, kidney damage.

Fluorouracil (5-FU): Nausea and vomiting, diarrhea, mouth sores, low blood cell counts, hand-foot syndrome (redness, pain, and peeling of the hands and feet), increased sensitivity to sunlight.

Vincristine: Peripheral neuropathy, constipation, hair loss, jaw pain, difficulty walking, low blood cell counts, nerve damage, increased risk of infections.

Bleomycin: Lung toxicity (causing cough, shortness of breath), skin reactions, mouth sores, low blood cell counts, increased risk of infections.

Etoposide: Nausea and vomiting, low blood cell counts, hair loss, mouth sores, increased risk of infections, allergic reactions.

Vinblastine: Hair loss, constipation, low blood cell counts, peripheral neuropathy, mouth sores,

increased risk of infections.

Drug	Side Effects	Drug	Side Effects	Drug	Side Effects
Abemaciclib	Nausea, fatigue, diarrhea, decreased appetite	Everolimus	Mouth sores, diarrhea, nausea, fatigue	Palbociclib	Nausea, vomiting, hair loss, fatigue, anemia
Abraxane	Bone marrow suppression, hair loss, neuropathy	Exemestane	Hot flashes, fatigue, joint pain	Pembrolizumab	Fatigue, rash, joint pain, diarrhea, decreased appetite
Ado-Trastuzumab Emtansine	Nausea, fatigue, decreased appetite	Fluorouracil Injection	Nausea, vomiting, diarrhea, fatigue	Pertuzumab	Diarrhea, nausea, vomiting, fatigue, fever
Afinitor	Fatigue, nausea, diarrhea, hyperglycemia	Fam-Trastuzumab Deruxtecan-nxki	Nausea, vomiting, fatigue, hair loss, anemia	Pertuzumab, Trastuzumab, and Hyaluronidase-zzxf	Diarrhea, nausea, vomiting, fatigue, fever
Afinitor Disperz	Fatigue, nausea, diarrhea, hyperglycemia	Fareston	Hot flashes, vaginal discharge, fatigue	Piqray (Alpelisib)	Nausea, vomiting, diarrhea, rash, hyperglycemia
Alpelisib	Hyperglycemia, nausea, fatigue	Faslodex	Injection site reactions, hot flashes, fatigue	Ribociclib	Nausea, vomiting, hair loss, fatigue, anemia
Anastrozole	Hot flashes, joint pain, vaginal dryness	Femara	Hot flashes, fatigue, joint pain	Sacituzumab Govitecan-hziy	Nausea, vomiting, diarrhea, fatigue, anemia
Aredia	Flu-like symptoms, bone pain	Fulvestrant	Injection site reactions, hot flashes, fatigue	Soltamox	Hot flashes, vaginal discharge, mood changes
Arimidex	Hot flashes, joint pain, vaginal dryness	Gemcitabine Hydrochloride	Bone marrow suppression, fatigue, nausea, vomiting, rash, fever, cough, peripheral neuropathy	Talazoparib Tosylate	Fatigue, nausea, anemia, myelosuppression
Aromasin	Hot flashes, joint pain, vaginal dryness	Goserelin Acetate	Hot flashes, vaginal dryness, loss of libido, decreased bone density	Talzenna	Fatigue, anemia, nausea, vomiting, myelodysplastic syndrome
Capecitabine	Hand-foot syndrome, nausea, diarrhea	Lapatinib Ditosylate	Diarrhea, nausea, rash, hand-foot syndrome, fatigue, decreased left ventricular e	Tamoxifen Citrate	Hot flashes, vaginal discharge, irregular menstruation
Cyclophosphamide	Nausea, hair loss, bone marrow suppression	Lapatinib Ditosylate	Diarrhea, rash, nausea, fatigue, hepatotoxicity	Taxotere	Hair loss, nausea, vomiting, anemia, fatigue, peripheral neuropathy
Docetaxel	Bone marrow suppression, hair loss, neuropathy	Letrozole	Hot flashes, headache, dizziness, fatigue, nausea	Tecentriq	Fatigue, decreased appetite, cough, fever
Doxorubicin Hydrochloride	Cardiotoxicity, hair loss, nausea	Margetuximab-cmkb	Fatigue, nausea, vomiting, diarrhea, headache	Tepadina	Nausea, vomiting, hair loss, decreased white blood cell count
Elacestrant Dihydrochloride	Fatigue, nausea, decreased appetite	Megestrol Acetate	Weight gain, hot flashes, fluid retention	Thiotepa	Nausea, vomiting, hair loss, decreased white blood cell count
Ellence	Nausea, vomiting, hair loss, fatigue,	Methotrexate Sodium	Nausea, vomiting, diarrhea, myelosuppression	Toremifene	Hot flashes, vaginal discharge, irregular menstruation

	anemia				
Enhertu	Nausea, vomiting, fatigue, hair loss, anemia	Neratinib Maleate	Diarrhea, nausea, vomiting, fatigue, hepatotoxicity	Trastuzumab	Fatigue, nausea, vomiting, headache, fever
Epirubicin Hydrochloride	Nausea, vomiting, hair loss, fatigue, anemia	Olaparib	Nausea, vomiting, diarrhea, fatigue, myelosuppression	Trastuzumab and Hyaluronidase-oysk	Fatigue, nausea, vomiting, headache, fever
Eribulin Mesylate	Fatigue, hair loss, nausea, constipation	Paclitaxel	Nausea, vomiting, hair loss, muscle/joint pain	Trexall (Methotrexate Sodium)	Nausea, vomiting, fatigue, hair loss, liver damage
Tucatinib	Diarrhea, nausea, vomiting, fatigue, liver damage	Paclitaxel Albumin-stabilized Nanoparticle Formulation	Nausea, vomiting, hair loss, muscle/joint pain	Trodelvy (Sacituzumab Govitecan-hziy)	Nausea, vomiting, diarrhea, fatigue, hair loss

Table 3.1: List of FDA approved drugs with their side effects.

Therefore, alternative compounds need to be identified to mitigate these side effects. As we know Natural compounds are less toxic as compared to synthetic drugs often due to the fact that they are multi targeting. Natural compounds are often perceived to be less toxic than synthetic chemotherapy drugs due to several reasons:

Evolutionary Compatibility: Natural compounds have been present in the environment for thousands of years and have evolved alongside living organisms. As a result, many organisms have developed mechanisms to interact with and metabolize these compounds. This evolutionary compatibility often leads to a reduced toxicity in comparison to synthetic drugs, which may be entirely new to biological systems and lack specific metabolic pathways for their efficient breakdown.

Complexity and Diversity: Natural compounds found in plants, fungi, and other organisms often possess complex structures and are composed of multiple components. This complexity can make it difficult for these compounds to interact with specific targets in the body, reducing the likelihood of causing toxic effects. In contrast, synthetic drugs are designed to have specific molecular interactions with targets in the body, which can lead to more potent and targeted effects, but also potentially higher toxicity.

Co-Evolved Mechanisms: Natural compounds frequently have evolved in plants and other organisms to serve specific biological functions, such as defense against predators or pathogens. These compounds often have mechanisms in place that mitigate their potential toxicity to the host organism or ensure that their effects are localized. On the other hand, synthetic chemotherapy drugs are primarily designed to

target cancer cells or disrupt cancer-related processes, which may not have the same level of specificity or built-in safety mechanisms.

Regulatory Factors: Natural compounds are often subject to regulatory mechanisms within living organisms that help maintain homeostasis and prevent excessive toxicity. For example, enzymes and transporters in the body may contribute to the metabolism and elimination of natural compounds, reducing their potential toxic effects. Synthetic drugs, especially those used in chemotherapy, are typically designed to be more resistant to degradation and elimination, allowing them to persist in the body for longer periods and potentially increasing their toxicity.

Phytochemicals, which are natural compounds found in plants, have been studied for their potential to mitigate the side effects of chemotherapeutic drugs. These compounds possess various biological activities and can offer supportive effects during cancer treatment. Here are some examples of phytochemicals that have shown promise in mitigating the side effects of chemotherapeutic drugs:

Quercetin: Quercetin is a flavonoid found in various fruits, vegetables, and herbs. It exhibits antioxidant, anti-inflammatory, and anticancer properties. Quercetin has been studied for its potential to reduce chemotherapy-induced oxidative stress, inflammation, and DNA damage.

Resveratrol: Resveratrol is a polyphenol found in grapes, berries, and peanuts. It has antioxidant and anti-inflammatory properties and has been shown to protect against chemotherapy-induced organ toxicity, such as heart and kidney damage.

***Epigallocatechin Gallate (EGCG)*:** EGCG is a catechin found in green tea. It possesses antioxidant, anti-inflammatory, and anticancer effects. EGCG has been investigated for its potential to mitigate the side effects of chemotherapy, including cardiotoxicity and gastrointestinal toxicity.

Curcumin: Curcumin is the active component of turmeric and has been extensively studied for its health benefits. It has antioxidant, anti-inflammatory, and anticancer properties. Curcumin has shown potential in reducing chemotherapy-induced side effects such as nausea, vomiting, and inflammation.

Sulforaphane: Sulforaphane is a compound found in cruciferous vegetables like broccoli, cabbage, and Brussels sprouts. It has antioxidant and anti-inflammatory effects and may help protect against chemotherapy-induced toxicity and enhance the efficacy of certain chemotherapeutic drugs.

Gingerols: Gingerols are bioactive compounds found in ginger. They possess anti-inflammatory and

antioxidant properties. Gingerols have been studied for their ability to alleviate chemotherapy-induced nausea and vomiting.

Lycopene: Lycopene is a carotenoid pigment found in tomatoes, watermelon, and other red fruits and vegetables. It has antioxidant properties and has been investigated for its potential to protect against chemotherapy-induced toxicity, particularly in the context of prostate cancer treatment.

3.2. METHODOLOGY AND MATERIALS REQUIRED

3.2.1. GENE EXPRESSION DATA OF CANCER CELL LINES AFTER TREATMENT WITH NATURAL COMPOUNDS

GSE85871, GSE24743, GSE158788 datasets were retrieved from GEO Database. GSE85871 contain the gene expression data of cancer cell lines after treatment with 102 natural compounds individually. Two replicates for each compound were considered a test set and compared with the vehicle control set in which only DMSO was present one by one, and the differential expression profile for all genes was collected. GSE24743, GSE158788 datasets contain the gene expression data of cancer cell line after treatment with Shikonin and Gallic acid respectively.

3.2.2. IDENTIFICATION OF NATURAL COMPOUNDS TARGETING SELECTED GENES

Np care database [70] and IMPACT databases were explored for the selection of natural compounds targeting these selected genes. We further explored the literature for those genes which are not found in the Np care database.

3.2.3. GENE EXPRESSION ANALYSIS OF NATURAL COMPOUNDS TREATMENT EFFECTS

Selected datasets were analysed by GEO2R and iDEP (integrated Differential Expression and Pathway analysis) Tool for the identification of differentially expressed genes. iDEP is a user-friendly interface for bioinformatic analysis of gene-level data for differential expression analysis and pathway analysis. User can generate reports to analyse RNA-seq datasets from NCBI's GEO which contains differential expression and enrichment analyses etc.

3.2.4. IDENTIFICATION OF NUMBER OF GENES REGULATED BY EACH NATURAL COMPOUND FROM THE SELECTED GENES

The expression of common upregulated genes was matched with the expression profile of natural compounds in Microsoft excel and selected compounds that alter the maximum number of gene expressions. Similarly, it was achieved for common downregulated genes.

3.3. RESULTS AND DISCUSSION

3.3.1. NATURAL COMPOUNDS SELECTION FOR TARGETING CD274 AND IDO1

We selected natural compound against the two selected genes (**Table 3.2**). Gallic Acid (3,4,5-trihydroxybenzoic acid) was inhitory against CD274 while three compounds, dihydrotanshinone I, shikonin, 9-O-demethyltrigonostemone were inhibitory against IDO1. Gallic acid is a phenolic acid which is found in sumac, gallnuts, tea leaves, oak bark, witch hazel and other plants. Dihydrotanshinone I (DI) is a natural compound found in the salvia miltiorrhiza which is a Chinese medicinal plant. It has been reported to have cytotoxicity to a variety of tumors. Shikonin is a naphthoquinone compound which is found in the roots of shikonin plant (Lithospermum erythrorhizon) and it is used as a traditional Chinese medicine. 9-O-demethyltrigonostemone is a natural compound found in the roots of Strophioblachia fimbricalyx which shows cytotoxic activity in different tumors.

Gene	Natural Compounds	Plant Origin
CD274	Cosmosiin	Teucrium gnaphalodes
	Fisetin	Strawberries, apples, persimmons, onions and cucumbers
	Gallic Acid	Banana, walnut, hazelnut, green tea, avocado, guava, mango, mulberry
	Kaempferol	Kale, beans, tea, spinach, and broccoli
IDO1	Dihydrotanshinone I	Salvia miltiorrhiza
	Shikonin	Lithospermum erythrorhizon, Alkanna, Arnebia, Onosma, Onosma sericeum Willd
	9-O-demethyltrigonostemone	Strophioblachia fmbricalyx

Table 3.2: Following table demonstrate the natural compounds targeting CD274, IDO1 and their natural sources.

We searched NCBI's GEO for the selected four compounds and found experimental data corroborating functional inhibitory characteristics of the two compounds which are Gallic Acid and shikonin. Dataset accession number for Gallic Acid is GSE158788. Gene Expression Profile Analysis of Gallic Acid-induced Cell Death Process using Hela cells treated with gallic acid (50 μg/ml) for 0 hour (GA0hr), 2

hours (GA2hr), 4 hours (GA4hr), 6 hours (GA6hr), and 9 hours (GA9hr) were studied. Dataset accession number for shikonin is GSE24743 and its effect on the gene expression of human lymphoma U937 cells was studied[161]. In this dataset U937 cells were treated with 100 Nm shikonin and followed by incubation for 3h at 37°C. The cells treated with dimethyl sulfoxide served as control. The microarray dataset was analyed with Geo2R tool and their volcano plot and heat map showed there are numerous genes whose expression altered by treatment as shown in **Figure 3.1** and significantly differentially expressed genes were identified.

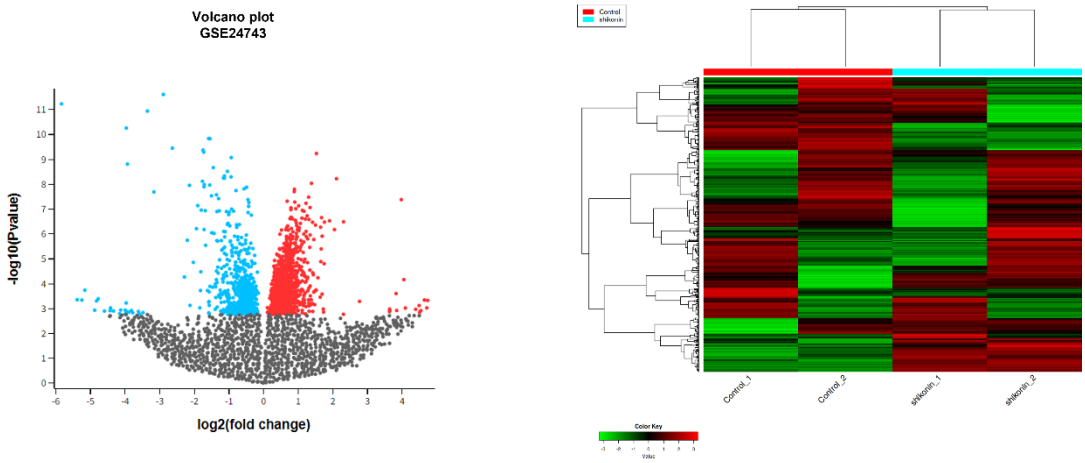


Figure 3.1: Shows the volcano plot of GEO dataset: GSE24743 Effects of Shikonin on the gene expression, red dots denote the genes which are differentially up-regulated and blue dots denotes the genes which are differentially down-regulated with an adjusted P-value less than 0.05 (in left). Right shows the Heatmap heat map of different genes in different control and test samples.

We analysed GSE158788 data with iDEP tool for differential gene expression and results are shown in the **Figure 3.2.** Their volcano plot and heat map showed there are numerous genes whose expression altered by treatment and significantly differentially expressed genes were identified.

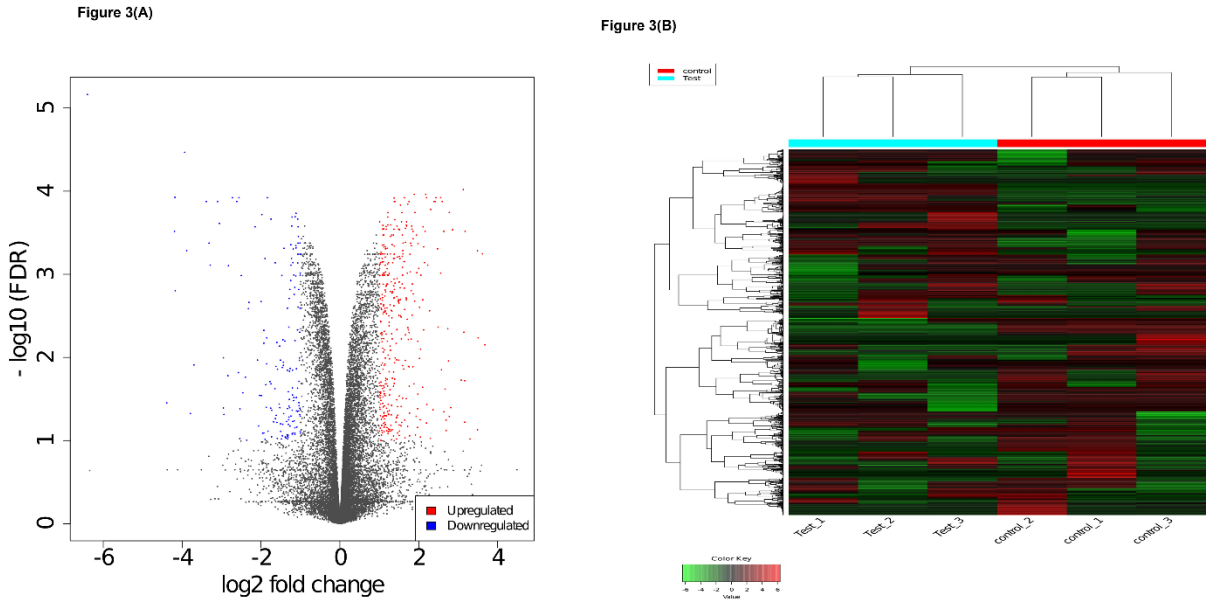


Figure 3.2: The volcano plot of GEO dataset: Effects of Gallic acid on the gene expression, blue dots denote the genes which are differentially up-regulated and red dots denotes the genes which are differentially down-regulated with an adjusted P-value less than 0.05. Right side shows the heatmap of different genes in different control and test samples.

3.3.2. IDENTIFICATION OF NUMBER OF GENES REGULATED BY EACH NATURAL COMPOUND FROM THE SELECTED GENES

The expression profiles of 102 natural compounds already known for their antitumor activity were analyzed with GEO2R, and Differentially expressed genes were studied. These differentially expressed genes were those whose expression was altered after the treatment with a particular natural compound. So, the list of differential genes was made of all the 102 Natural compounds with their logFC Values. Those natural compounds that could downregulate the maximum number of genes from the 46 common upregulated genes and upregulate the maximum number of genes from the 64 common downregulated genes in cancer patients were further identified. The list of 110 common DEGs (46 upregulated genes + 64 downregulated genes) was compared with the natural compounds' DEGs. The natural compounds were sorted based on their ability to reverse the expression of the maximum number of genes. The list of natural compounds with respective number of genes whose expression was reversed by that natural compound is shown in **Table 3.3.** Natural compounds are multitargeting; therefore, they can target multiple pathways for tumor suppression. Tumor cells change their gene morphology very fast when exposed to single target drugs to find a way to escape the drug effect. So, the problem of drug resistance can be minimized by using a combination of natural compounds because they target multiple pathways,

and tumor progression can be effectively reversed.

Natural Compound	Genes	Natural Compound	Genes	Natural Compound	Genes
Salidroside	66	Phillyrin	57	Nitidine chloride	52
Schisantherin A	65	Resibufogenin	57	Bruceine	51
Oxymatrine	64	Alantolactone	56	Ginsenoside Rb3	51
Silybin	64	Ginkgolide B	56	Macrozamin	51
Daidzin	63	Hyodeoxycholic acid	56	Tanshinone IIA	51
Scutellarein	62	Matrine	56	Bufalin	50
Gastrodin	60	Osthole	56	Cinnamic acid	50
Ginseooside Rd	60	stachydrine hydrochloride	56	Honokiol	50
Glycyrrhizic acid (2)	60	Ursodeoxycholic acid	56	Hyperoside	50
Schizandrin	60	ethyl caffeate	55	Puerarin	50
Astragaloside IV	59	Ferulic acid	55	Saikosaponin A	50
Benzoylaconitine	59	Isoalantolactone	55	Sanguinarine	50
Ginsenoside Rc	59	Tetrahydropalmatine	55	Sennoside A	50
Imperatorin	59	1bita-hydroxyalantolactone	54	Acteoside	49
L-scopolamine	59	Benzyl benzoate	54	Hesperidin	49
Saikosaponin D	59	Gallic acid	54	Hydroxysafflor yellow	49
Aconitin	58	Ginsenoside Rb1	54	Magnolol	49
Arenobufagin	58	Isoborneol	54	Resveratrol	49
Chlorogenic acid	58	Protocatechuic aldehyde	54	6-gingerol	48
Cinnamaldehyde	58	Salvianic acid A sodium	54	Andrographolide	48
Cinobufotain	58	Bacopaside l	53	Cholic acid	48
Deoxycholic acid	58	Benzoylhypaconitine	53	Ginsenoside Re	48
Gamabufatalin	58	bita-ecdysterone	53	Japonicone A	48
Uridonin	58	Chenodeoxycholic acid	53	Narciclasine	48
Anhydroicaritin	57	Ginsenoside Rb2	53	Notoginsenoside R1	48
Borneol	57	Muscone	53	Hypaconitine	47
Chelerythrine	57	Paeoniflorin	53	Liquiritin	47
Curcullgoside	57	Salvianolic acid B	53	Berberine hydrochloride	46
Dioscin	57	Telocinobufagin	53	Bufotaline	46
Emodin	57	Bilobalide	52	Ainsliadimer A	45
Gentiopicroside	57	Britanin	52	Geniposide	45
Ginsenoside Rg1	57	Cinobufogenin	52	Santonin	44
Lobetyolin	57	Ephedrine hydrochloride	52	Strychnine	44
Oleanic acid	57	Glycyrrhizic acid	52	Artemisinin	36

Table 3.3: Following table demonstrate the list of natural compounds with the number of genes regulated.

CHAPTER IV

Objective 3

- Combinatorial potential of natural compounds and their validation via *in vitro* experiments

CHAPTER IV: OBJECTIVE 3

4.1. RATIONALE OF THE STUDY

There are several drawbacks associated with using a single drug in tumor treatment. Here are some key limitations:

Development of Resistance: Tumors can develop resistance to a single drug over time, rendering it less effective or completely ineffective. This resistance can occur due to various mechanisms, including genetic mutations or alterations in the tumor cells. When resistance develops, the tumor can continue to grow and spread despite treatment.

Incomplete Tumor Targeting: Tumors are often heterogeneous, meaning they consist of different cell populations with varying characteristics. A single drug may only target a specific subset of tumor cells, leaving other cells unaffected. This can result in incomplete tumor eradication and potential regrowth of the tumor.

Limited Efficacy: Tumors can have multiple pathways or molecular targets involved in their growth and survival. A single drug may only target one specific pathway or target, limiting its overall efficacy.

Side Effects and Toxicity: Some drugs used in tumor treatment can have significant side effects or toxicity, especially at higher doses required for optimal efficacy. Using a single drug may necessitate higher doses, increasing the risk and severity of adverse effects. Combining multiple drugs can allow for the use of lower doses of each drug, potentially reducing side effects while still achieving therapeutic benefits.

Lack of Personalized Treatment: Tumors can have diverse characteristics and response patterns among different patients. Using a single drug may not adequately address the individual characteristics of a patient's tumor. Personalized treatment approaches, including the use of combination therapy, can help target specific tumor characteristics and optimize treatment outcomes.

4.1.1. BENEFITS OF COMBINATORIAL THERAPY

It refers to the use of multiple drugs in the treatment of a particular condition. Here are some reasons why combination therapy may be preferred:

Increased Efficacy: Combining drugs with different mechanisms of action can target multiple pathways or targets involved in the disease process. This approach can lead to synergistic effects, where the combined action of the drugs is more effective than each drug alone. Combination therapy is commonly used in the treatment of complex diseases like cancer, HIV, tuberculosis, and hepatitis.

Reduced Drug Resistance: The use of a single drug over time can lead to the development of drug resistance in certain pathogens or tumors. Combining multiple drugs with different mechanisms of action can help prevent or delay the development of resistance by attacking the disease from different angles.

Improved Tolerability: Some drugs may have side effects or toxicity at higher doses required for optimal efficacy. Combination therapy allows for the use of lower doses of each drug, reducing the risk and severity of side effects while still achieving the desired therapeutic effect.

Targeting Different Disease Stages: Certain diseases have multiple stages or phases, and each stage may require a different treatment approach. Combination therapy can be used to target different stages of the disease, maximizing the effectiveness of treatment throughout the course of the illness.

Individualized Treatment: Combination therapy can be tailored to individual patients based on their specific characteristics, such as disease severity, genetic factors, and response to treatment. By combining drugs with different modes of action, physicians can customize treatment regimens to better meet the unique needs of each patient.

4.2. METHODOLOGY AND MATERIALS REQUIRED

4.2.1. SCREENING OF NATURAL COMPOUNDS

Natural compounds were screened on the bases of their regulation of common DEGs. The compound which regulates the maximum number of genes from common DEGs was selected first. Then the second compound was selected based on who can alter the expression of the maximum number of genes from the remaining ones (which remain unaltered by the first compound). Similarly, the remaining compounds were screened.

4.2.2. ANALYSIS OF SELECTED COMBINATION OF NATURAL COMPOUNDS

The selected combination of natural compounds was analyzed with the help of a Venn diagram of genes regulated by selected compounds and Enrichment analysis of the shared genes by the FunRich tool. Different biological processes and pathways regulated by these compounds were also analyzed from the literature.

4.2.3. COMPARISON OF GENE EXPRESSION DATA OF TUMOR AND NATURAL COMPOUND TREATED SAMPLE

Differentially expressed gene (DEGs) of HNSC cancer dataset was compared with the Differential gene expression data of both the selected compounds for conforming that these compounds reverse the expression of differentially expressed genes. This conformation was achieved in the Microsoft-excel by comparing their logFC values in different samples.

4.2.4. BIOLOGICAL PATHWAY ANALYSIS OF THE INVOLVED GENES

Pathway analysis of the genes whose expression was altered by gallic acid and shikonin was achieved by Funrich tool. We devided genes into four categories for pathway analysis these are HNSC upregulated genes down regulated by Gallic acid, HNSC upregulated genes down regulated by shikonin, HNSC downregulated genes upregulated by Gallic acid, HNSC downregulated genes upregulated by shikonin. Then gene involved in these respective categories are analyzed for the pathways in which they are involved in Funrich tool individually. Funrich is a standalone tool used for enrichment analysis which gives enrichment

analysis on the basis of biological pathways, biological processes, functional, transcription factors etc.

4.2.5. MATERIAL AND METHODOLOGY FOR *IN-VITRO* EXPEREMENTS

4.2.5.1.MATERIALS

MDA-MB-231 cell lines were obtained from the National Centre for Cell Sciences (NCCS) in Pune, India. Gallic acid and Shikonin, the natural compounds of interest, were procured in powdered form from Sigma-Aldrich and stored at 4 degrees Celsius until utilized. For cell culturing, RPMI medium was employed, purchased from Sigma-Aldrich, and supplemented with antibiotics Penicillin and streptomycin along with Fetal Bovine Serum (FBS) obtained from Gibco, Thermo Fischer, along with antibiotics. The Trypsin-EDTA solution 1X, essential for cell detachment, was purchased from HIMEDIA. For the MTT assay, the MTT reagent came from BioAssay Systems, and it was supplied within the MTT assay kit alongside a solubilizer to process the assay. Dimethyl sulfoxide (DMSO) from Merck was utilized as the vehicle control in the experiments.

4.2.5.2.METHODOLOGY

The cells were cultured in RPMI at 37 degrees celsius and 5% CO2 till they reached >90% confluency, following standard cell culture protocols and there was enough number of cells for performing the cell viability assay. The cells were passaged and centrifuged followed by counting using Hemocytometer (Neubarr's chamber) and trypan blue dye. The cells were suspended in RPMI so that the final concentration is 625000 cells/mL (approximately). Cells were seeded in the 96-well tissue culture plate at the density 5000 cells per well in 80 µl culture media. The natural compounds were dissolved in DMSO and drug solutions of different concentration were prepared with desired volume of culture media. After the cells adhered to the surface of wells (22 hours) the medium in each well was replaced with DMSO solution that contained the natural compounds gallic acid and shikonin at different concentrations, as well

as their mixture. DMSO without the natural compounds was added in 3 wells and these were considered as the control wells. The cells were incubated with the drugs for 48 hours at 37 degrees Celsius. Then media was removed carefully and different concentration of 100 µl media solution of Gallic acid and Shikonin were added in different wells (Three replicates for each dose) in different concentration. After 12, 24, 48 hours of incubation 15 µl of the MTT reagent was added to each well followed by 4 hours of incubation for the formation of formazan crystals by viable cells. 100 µl of solublizer was added to each well and mixed on orbital shaker for 1 hours on CO2 incubator. After treating the cells with the different natural compound combination, and the control (DMSO), the absorbance of each sample was measured using a microplate reader. Absorbance was calculated using the Microplate reader with absorbance measurement capability at 570 nm because maximum observance of formazan dye lies between 560 nm and 590 nm. The absorbance data obtained from each sample was used to calculate the percentage cell viability. This calculation involves comparing the absorbance values of the treated samples to that of the control (DMSO). The formula for percentage viability is typically:

Percentage Viability = (Absorbance of treated sample- Absorbance of blank / Absorbance of DMSO control- Absorbance of blank) x 100

To understand the dose-dependent effect of the drugs, the percentage viability data was plotted against the dose of each drug. Dose-response curve allows to observe how cell viability changes as the drug concentration varies. Bar plot was used to analyse the percentage viability of cells after treatment with different dose combination of Shikonin and Gallic acid.

4.3.RESULTS AND DISCUSSION

4.3.1. COMBINATION OF NATURAL COMPOUNDS

A combination of natural compounds is selected that alter the expression of the maximum number of genes out of 110 common DEGs. Therefore, we first selected that natural compound, which altered the expression of the maximum number of genes, Salidroside. Salidroside altered

the expression of 66 genes from 110 common DEGs. Then the remaining 44 commonly altered genes were studied for their susceptibility to restorative regulation by other such natural compounds that impacted the expression of the maximum number of the gene. Therefore, Ginsenoside Rd was found to regulate the expression of 20 genes. Oridonin was found to regulate12 genes out of the 24 common DEGs. Britanin was found to regulate 6 genes, and Scutellarein regulated 4 genes. Therefore, these five compounds together resulted in the regulation of 108 genes out of 110 common DEGs, restoring the gene expression to that in normal matched tissues. Two genes, GPR15 and CYP2U1, were not suitably regulated by our combination of natural compounds. **Figure 4.1** shows the screening process of natural compounds for targeting 110 common DEGs.

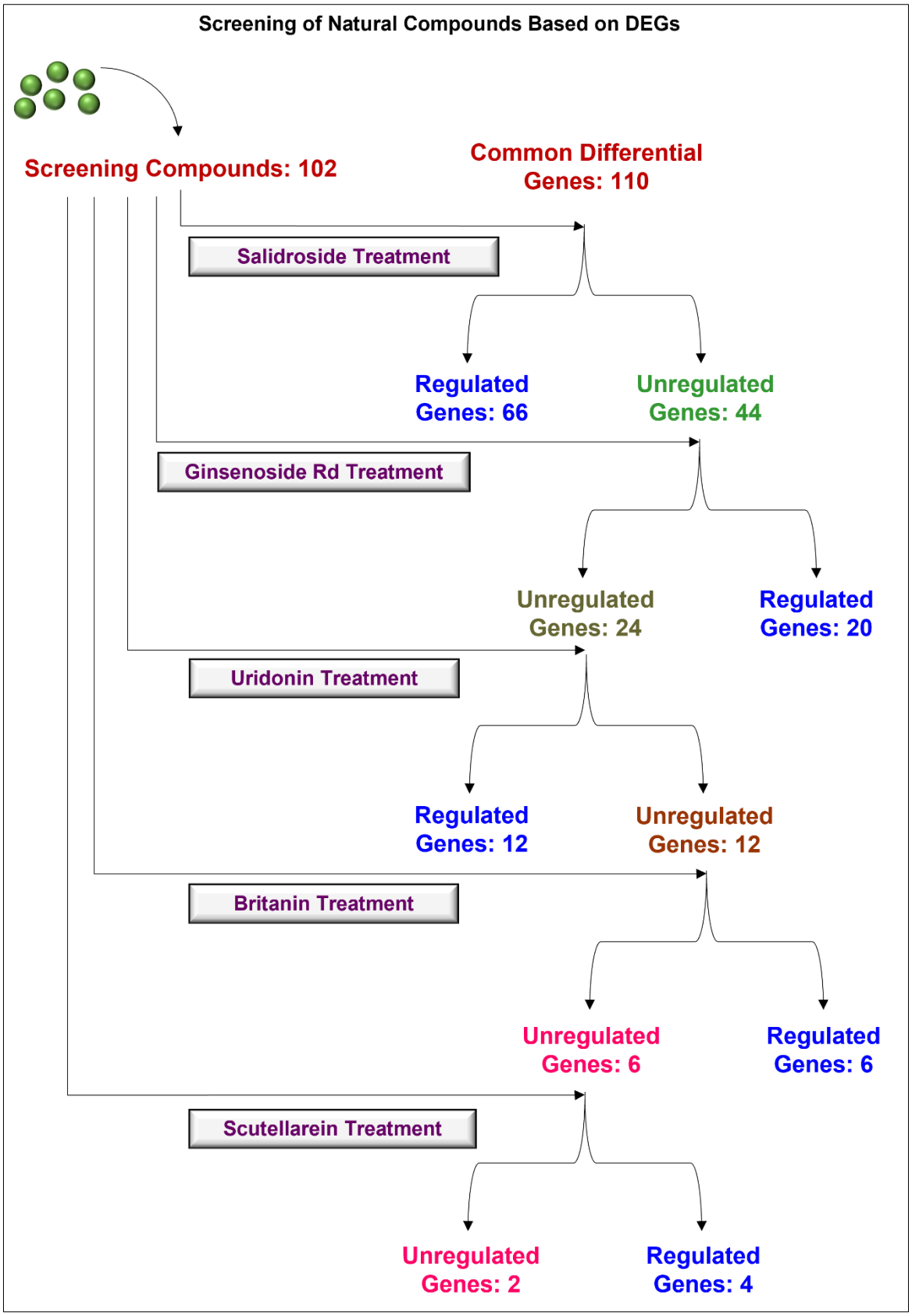


Figure 4.1: Following figure demonstrate the screening process of natural compounds against 110 common DEGs. 66 genes were regulated by salidroside therefore remaining genes were analysed for other targeting compounds and found that Ginsenoside Rd, Uridonin, Britanin, Scutellarein were regulated 20, 12, 6,4 genes respectively.

A combination of five compounds was selected and further checked for their combined effect on these common 110 DEGs. Therefore, the expression of these 110 DEGs was compared with the expression of the individual compounds. Salidroside, Ginsenoside Rd, Oridonin, Britanin, and Scutellarein individually regulated 66, 60, 58, 52, 62 genes. The gene expression regulation of common upregulated genes by different selected natural compounds separately is listed in **Table 4.1.** The gene expression regulation of common downregulated genes by different selected natural compounds separately is listed in **Table 4.2.**

Common Downregulated Genes	Salidroside	Ginsenoside Rd	Uridonin	Britanin	Scutellarein
ZBTB20	NA	0.112	NA	0.032402	0.010433
CD6	0.128491	NA	0.101702	0.065331	0.011821
SHMT2	0.011107	0.0606	0.003524	0.005039	NA
SERGEF	0.123762	0.0756	NA	0.07678	0.060725
ZNF205-AS1	0.022565	0.156	0.018661	NA	0.064165
NCR3	NA	0.244	0.31222	NA	NA
SRSF8	0.002687	0.0293	NA	NA	0.009713
TRADD	0.020171	1.06	NA	NA	NA
SKI	0.076444	0.109	NA	NA	0.105307
ACACA	NA	0.00176	0.005808	0.040862	NA
ZAP70	1.464565	1.03	0.85929	1.166922	1.143146
UBFD1	NA	NA	0.003933	NA	NA
CKS1B	NA	0.00666	0.006938	NA	NA
D NAJA4	NA	NA	NA	0.007385	NA
PIP4K2B	NA	NA	NA	NA	0.018495
RIC3	NA	0.154	NA	NA	NA
PDE9A	0.054939	0.0349	NA	NA	0.018845
GPX7	NA	NA	0.204458	NA	0.08836
RFC4	NA	NA	0.015399	NA	NA
RPS17	NA	0.00944	NA	5.08E-05	NA
PTGDR	NA	0.0797	NA	NA	NA
PPARD	0.004411	NA	NA	NA	NA
BCR	0.011611	NA	0.017969	0.002361	7.46E-05
WDR74	0.006102	NA	0.014646	NA	NA
TBRG4	0.10012	0.11	7.75E-05	NA	0.175699
DHX30	6.69E-05	NA	0.006934	NA	NA
MCF2L	0.040489	NA	NA	NA	0.004955
ARHGEF5	0.019291	0.0202	0.006395	0.06986	0.00123
ACSF2	0.012821	0.0474	0.002574	NA	0.046479
MMP11	NA	NA	0.027858	NA	NA
FEZ1	NA	0.00501	NA	NA	0.018955
ZNF764	0.000229	0.0423	NA	NA	0.016068

SLC4A10	0.050089	0.225	0.154503	0.306708	0.092965
SNRNP40	0.020221	NA	NA	NA	0.00371
GPD1L	0.027114	0.0283	NA	NA	0.003768
TRAM2	0.025506	0.0292	0.00835	0.094109	0.009603
KDM8	0.075093	NA	0.067595	NA	0.064676
SAFB2	0.009386	0.582	0.013734	0.012861	0.002413
GPALPP1	0.008237	NA	0.083655	NA	0.003773
RARRES3	0.079535	NA	0.152624	NA	0.260278
PTCD2	0.189298	NA	0.132122	0.129328	0.074555
CD22	NA	0.159	NA	NA	NA
CAC NA2D2	0.203298	0.291	NA	0.089563	0.079899
ZNF318	NA	NA	0.027326	NA	0.009677
POM121	NA	0.0757	0.132027	0.037441	0.04488
RPL27A	0.013419	0.0731	NA	0.143882	0.001093
KAT6B	0.014566	NA	NA	0.016648	0.007702
DOLPP1	0.014736	0.0123	0.013935	0.093057	NA
SOD1	NA	NA	0.014987	0.026234	NA
NHP2	0.007297	0.000815	NA	NA	NA
PPP1R13B	0.019158	NA	0.025801	0.102103	0.025379
RPAIN	0.067022	0.0942	0.025175	0.04728	0.026807
PROCR	NA	NA	NA	NA	0.014211
WDHD1	0.010003	NA	9.77E-06	0.006522	1.13E-05
MLF1	0.021094	0.23	NA	0.072687	0.026039
MRPL58	NA	NA	0.011365	0.01975	NA
NF2	NA	0.0178	0.003165	0.037192	NA
CYP2U1	NA	NA	NA	NA	NA
DENND2D	NA	NA	NA	0.008037	NA
PCNX2	0.077608	0.0228	0.07015	0.087913	0.124665
AMMECR1	0.016596	0.00637	NA	NA	0.015094
SMARCC2	NA	NA	0.019569	NA	NA
CSNK1E	NA	0.08	NA	NA	0.004997
FANCI	NA	0.0162	NA	NA	NA

Table 4.1: Different tones of colors (light to dark) in the given table demonstrate the level of expression of downregulated genes altered by different compounds individually. Light tones indicate genes which are altered on a small level, while dark tones indicate genes which are highly altered and moderate tones indicating moderately altered genes by different compounds.

Common Upregulated Genes	Salidroside	Ginsenoside Rd	Uridonin	Britanin	Scutellarein
PPARG	-0.01246	NA	NA	NA	NA
CYP1A2	-0.06406	-0.0523	-0.51742	NA	-0.06783
RABEP1	-0.01659	-0.125	NA	-0.13887	-0.00356
AHR	-0.03835	NA	NA	NA	-0.0275
MITF	-0.00938	NA	-0.02652	-0.14417	NA
FCER1G	-0.67839	NA	-0.11665	-0.09313	-0.01787
NSUN7	-0.20027	-0.388	-0.14311	-0.14516	NA
ACTB	-0.00096	-0.0031	NA	NA	-0.00317

MAPK14	-0.00098	-2.79	NA	-0.00988	NA
RBM47	-0.00758	NA	-0.04291	NA	NA
PAX9	-0.12917	-0.0784	-0.19931	-0.26224	-0.04412
PRNP	-0.01957	NA	NA	-0.07356	-0.03223
R NASE1	-1.01883	-0.000441	NA	NA	NA
MAP2K6	NA	NA	NA	NA	-0.96137
ERN1	-0.17315	-0.258	-0.06998	-0.0927	-0.05894
WAC	NA	-0.00646	-0.00046	-0.04323	-0.01332
IPO7	NA	NA	-0.00512	-0.02105	NA
CLDN9	-0.01543	-0.000966	-0.03241	NA	-0.01317
PCMT1	NA	-0.0169	-0.00551	-0.02151	-0.00263
GPR15	NA	NA	NA	NA	NA
ST3GAL6	NA	NA	NA	-0.961	NA
VEGFA	-0.01287	-0.896	NA	NA	NA
LAP3	-0.01316	NA	NA	NA	NA
UBOX5	-0.07102	-0.0271	-0.02715	NA	-0.06819
GK	NA	NA	-0.08025	-0.01122	NA
PANK3	-0.03335	NA	-0.04324	NA	-0.01208
CORO1B	NA	-0.0291	NA	-0.1771	-0.00315
ACPP	NA	NA	NA	-0.37689	NA
LATS1	NA	NA	NA	NA	-0.02555
CCDC88A	-0.74399	-1.6	-0.86328	NA	NA
IL18	-0.07134	-0.254	NA	NA	-0.01199
LAT2	-0.1308	NA	-0.16805	-0.062	-0.0292
SMPDL3A	-0.00835	-0.0179	-0.03109	-0.21941	NA
SLC44A1	-0.20994	NA	-0.38872	NA	-0.09192
S100A9	NA	-0.0255	NA	NA	NA
SERPINB2	-0.15888	NA	NA	NA	NA
PPM1A	NA	NA	NA	-0.09572	NA
DOCK4	NA	NA	NA	-0.1361	NA
ZEB2	-0.54734	NA	-1.72513	NA	-0.15526
TRPM6	-0.66426	-0.162	-1.29945	-1.1726	-2.06793
TLR5	NA	-0.0487	NA	NA	NA
RRAGD	NA	NA	-0.00748	-0.01268	NA
HFE	-0.0863	-0.0611	-0.01804	-0.03487	-0.04517
ROCK2	NA	-0.0406	NA	-0.2765	NA
GRM2	NA	-0.259	NA	-0.33371	-0.09282
MFSD1	NA	NA	-0.0043	NA	NA

Table 4.2: Different tones of colors (light to dark) in the given table demonstrate the level of expression of upregulated genes altered by different compounds individually. Light tones indicate genes which are altered on a small level, while dark tones indicate genes which are highly altered individually by different compounds.

This data analysis showed that many genes were regulated more efficiently in combination rather than isolation; therefore, these compounds might show synergistic effects. The alteration of expression of genes by more than one compound is also beneficial to preventing drug resistance and toxic side effects

due to alternative engagement of redundant pathways. The expression of different genes in PBMCs and regulated by the different compounds are shown in the **Figure 4.2.**

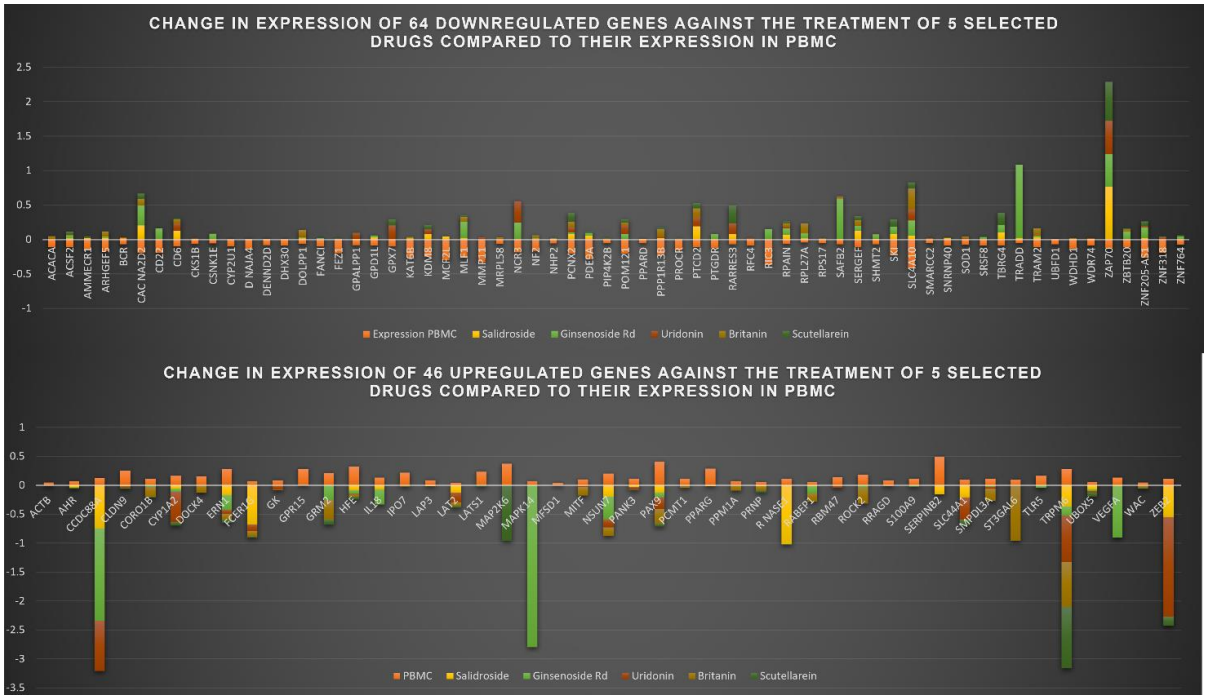


Figure 4.2: Following figure demonstrate the change in expression of 64 upregulated genes against the treatment of 5 selected compounds Salidroside, Ginsenoside Rd, Uridonin, Britanin, Scutellarein and their expression in PBMCs. Different color indicates the effect of different compounds as mentioned in the figure.

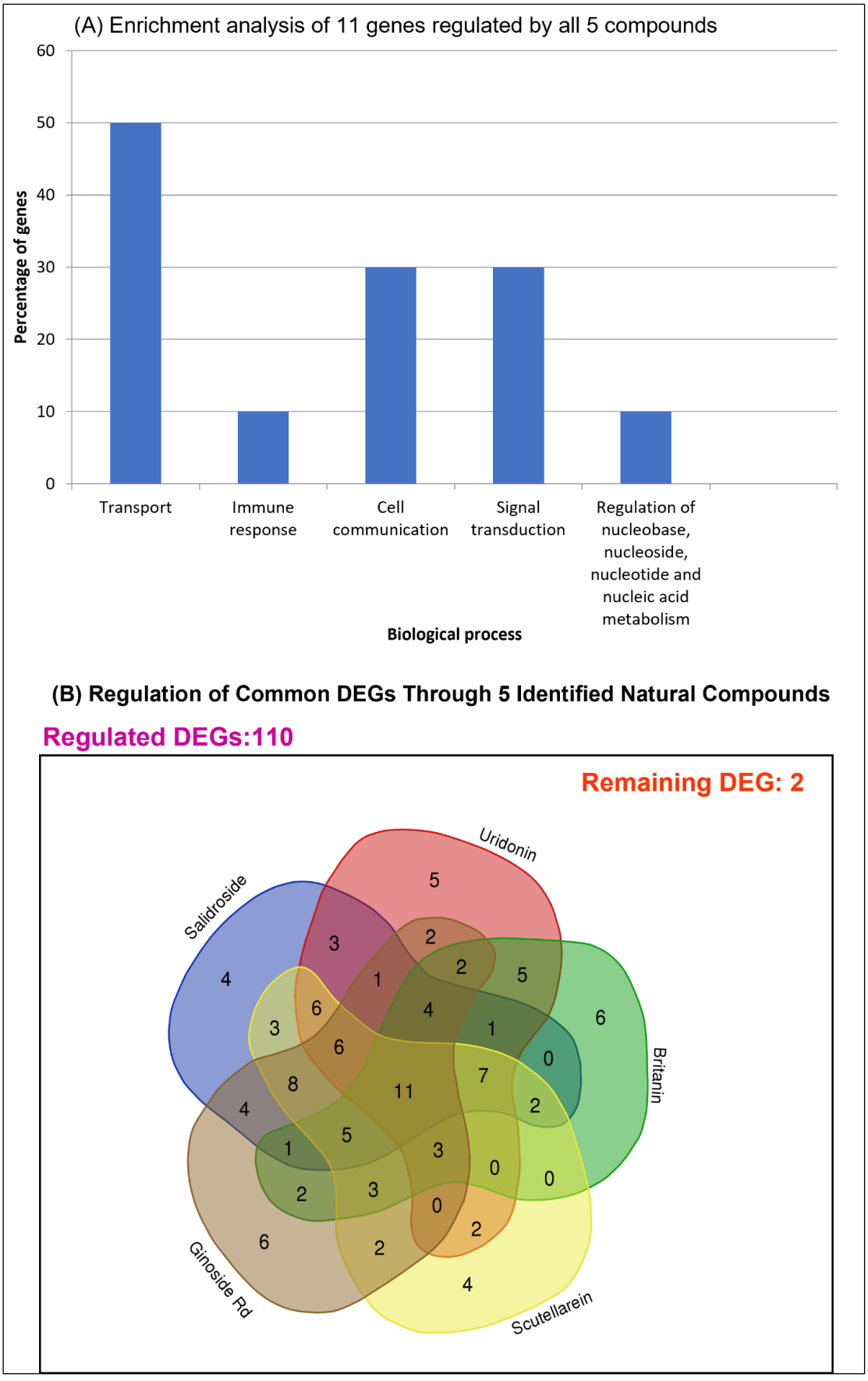


Figure 4.3: - (A) following figure demonstrate the enrichment analysis of 11 genes which were regulated by all five natural compounds. (B) Venn diagram shows the gene regulation of 110 common DEGs by

different natural compounds 11 genes were regulated by all five compounds.

Figure 4.3 Venn diagram shows that all the five compounds regulated 11 genes, i.e., ZAP70, HFE, TRPM6, RPAIN, PAX9, PCNX2, TRAM2, ARHGEF5, ERN1, SAFB2, and SLC4A10, 25 genes were regulated by any four compounds, 24 genes were regulated by any three compounds, 23 genes were regulated by any two compounds, 25 were regulated only one compound out of these five compounds.

Enrichment analysis of these 11 compounds was achieved and found that these compounds were involved in biological processes like transport, immune response, signal transduction, and regulation of nucleic acid metabolism. These biological processes are related to cancer progression and the immune system. Therefore, synergistic targeting of these genes would be beneficial for efficient combating HNSC tumors. All the selected compounds were efficient at targeting the expression of these eleven genes that effected the above-mentioned biological processes.

Natural compounds are multitargeting compounds, so targeting genes that positively impact cancer regression may result in undesirable side effects. Hence, to test the synergistic potential for targeting of the selected natural compounds, their mechanisms of action were retrieved from the literature and congruence analysis was done. The drug combination of natural compounds must ideally target diverse pathways that converge to result in effective tumor regression. Salidroside induced autophagy via PI3K/Akt/mTOR signaling. mTOR is highly upregulated in tumor cells, hence inhibiting autophagy. PI3K/Akt plays an essential role in the regulation of mTOR. Salidroside regulates the PI3K/AKT pathway, decreasing anti-apoptotic factors and increasing pro-apoptotic factors, thus inducing caspase-dependent and mitochondria-mediated apoptotic cell death [163]. Salidroside inhibits proliferation, migration, and invasion of tumor cells by inhibiting ROS, which regulates Src, and downregulates HSP70 via Akt/ERK signaling [164]. Salidroside reduces the pro-inflammatory cytokine secretion via activating IkBα/NF-kB pathway and induces apoptosis via p53 and caspase signaling [165]. Oridonin inhibits angiogenesis via the HIF-1α/VEGF pathway and shows anti-migratory, anti-invasive and anti-adhesive properties [166]. Oridonin inhibits the proliferation and migration of tumor cells via targeting TRPM7 through the inactivation of ERK/AKT signaling [166]. Oridonin induces phagocytosis via activating ERK, which activates NFκB [167]. Ginsenoside Rd reduces metastasis via miR-18a-mediated downregulation of SMAD2 [168]. Ginsenoside Rd increases the expression of miR-144-5p,

which inhibits the expression of TLR2 hence reducing the proliferation and metastasis of tumor cells [169]. Ginsenoside Rd inhibits VEGF-induced migration, tube formation, and proliferation and suppresses VEGF-induced regulation of Akt/mTOR signaling pathways, inducing apoptosis and inhibiting cell proliferation [170]. Ginsenoside Rd inhibits proliferation, and metastasis mainly reverses EMT via STAT3/JAK2 signaling and STAT3 is the direct target of Ginsenoside Rd [171]. Britanin inhibits NF-kB via downregulation of IKK1/1KK2, controlling tumor cell proliferation and angiogenesis [172]. Britanin shows an anti-inflammatory response via inhibiting NF-κB signaling [173]. Britanin downregulates cMyc and HIF1α via upstream effectors like mTOR, reducing the expression of specific proteins, including PD-L1, leading to the inhibition of angiogenesis and cell proliferation [174]. Britanin induces apoptosis and autophagy via activating AMPK signaling regulated by ROS [175]. Scutellarin inhibits ALS and AST hence regulating the immune system against tumor cells [176][177]. Scutellarin inhibits MCP1, thus inhibiting cell migration and reducing inflammation [178]. Scutellarin downregulates ICAM-1 and inhibits the activation of NF-κB hence inhibiting adhesion and showing an anti-inflammatory effect [179]. Scutellarin induces vasodilation via eNOS/NO/PKG pathways [180].

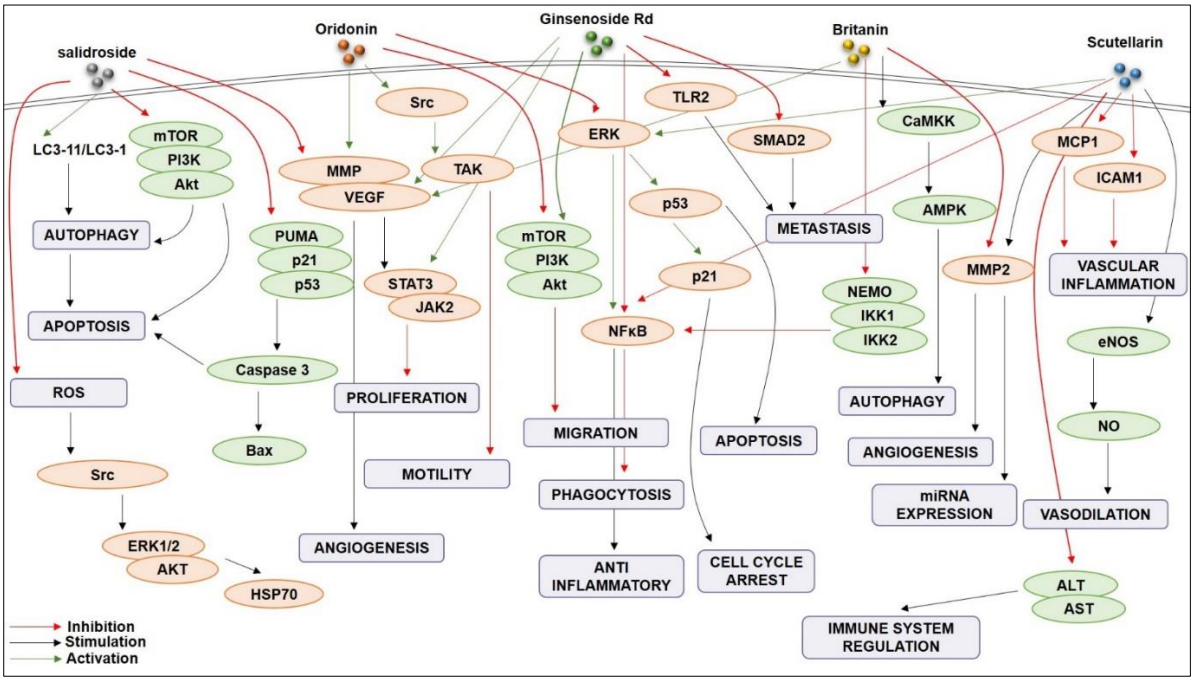


Figure 4.4: Following figure demonstrate the different biological processes and genes regulated by different natural compounds where different signs were used for inhibition, stimulation and activation of genes and biological processes.

As shown in **Figure 4.4,** These drugs regulate many different pathways and can target more than one process. So, our drug combination offers a highly potent multifaceted antitumor and immunomodulatory role and helps in the regression of HNSC cancer. The biological network of the following compounds is shown in **Figure 4.5**, which shows the key genes regulated by these compounds.

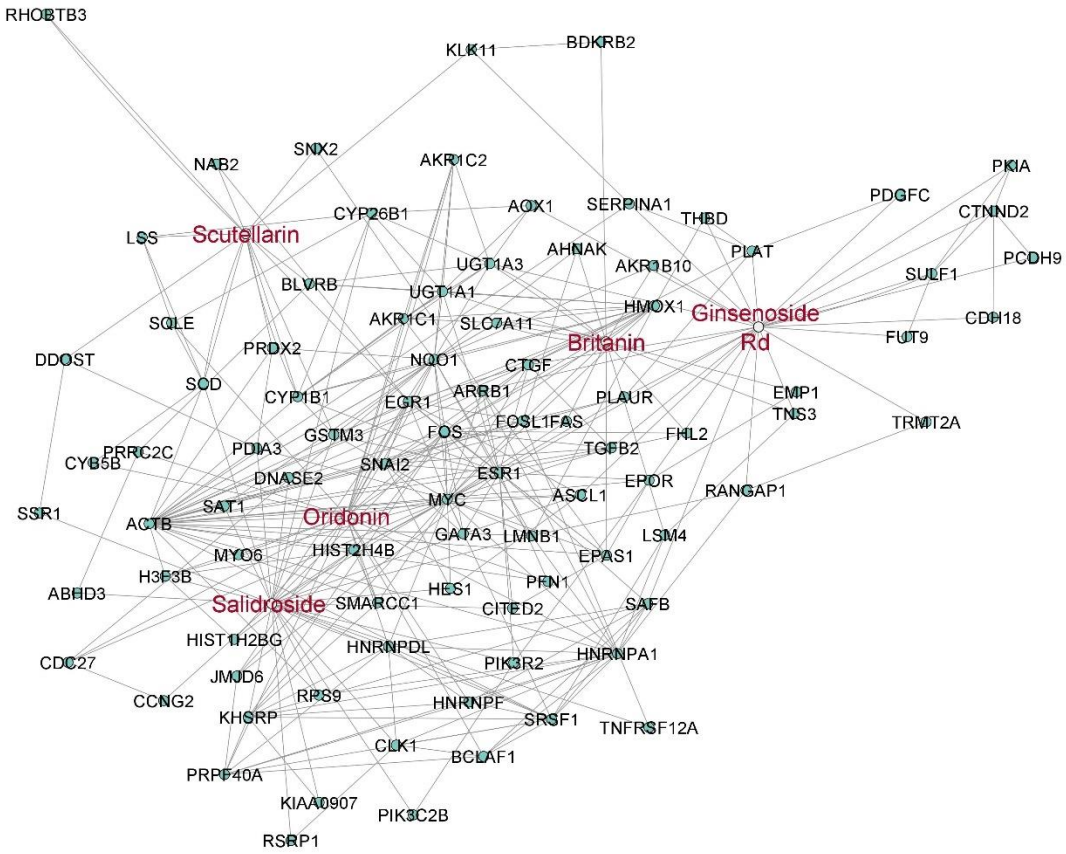


Figure 4.5: The above figure demonstrates the biological network of the selected compounds with their regulating genes.

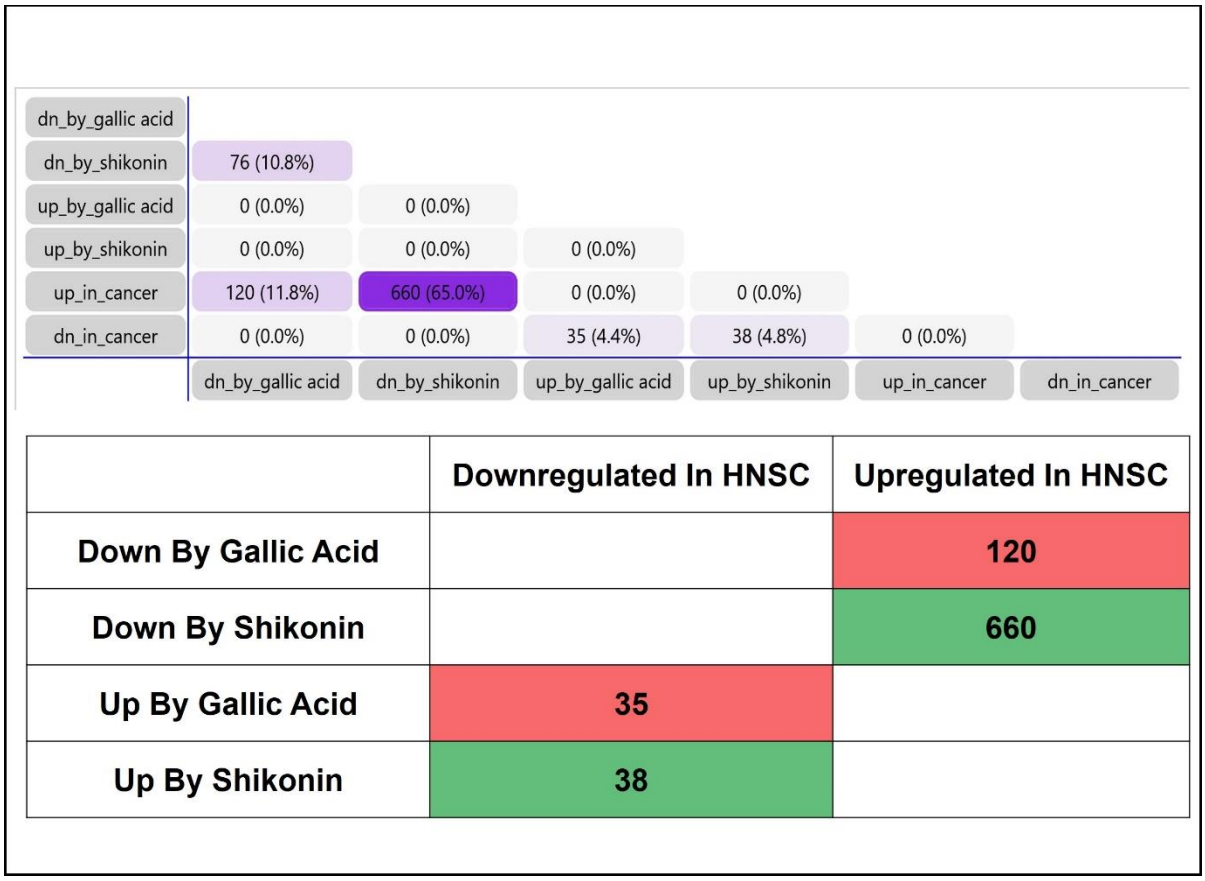
A combination of Salidroside, Ginsenoside Rd, Oridonin, Britanin, and Scutellarein was chosen such that they can alter the expression of 108 genes out of the selected 110 genes. Salidroside is widely found in Rhodiola plants. Rhodiola sacinehalnsis, Rhodiola rosea, Rhodibetic tibetica and large Rhodiola. Ligustrum lucidum, in the leaves of Salix triandra L. and Willow bark, Vaccinium vitisidaea L leaves of Oleaceae, Veroniceae of Veronica minor. Salidroside was found to induce autophagy, inducing caspase-dependent and mitochondria-mediated apoptotic cell death, and inhibiting proliferation, migration, and invasion of tumor cells via PI3K/Akt/mTOR signaling, IkBα/NF-κb signaling [163][164][165][163]. Salidroside is generally deemed safe and effective. In the experimental

conditions, salidroside at doses of 0.5, 0.25, and 0.125 g/kg in SD rats did not cause maternal or embryonic toxicity, nor did it have teratogenic consequences [181]. Genotoxicity testing is critical in drug risk assessment. Salidroside is not genotoxic at a clinical dose (150 mg/60 kg/day) in humans, according to the Ames test, reverse mutation, chromosomal abnormalities, and mice micronucleus studies [182]. Another study of 60 breast cancer patients found no clinical adverse effects when an effective dose of salidroside (600 mg/kg/day) was given throughout the therapy procedure [183]. The lack of negative effects in pre-clinical and clinical trials suggests salidroside is a safe common clinical medication. Ginsenoside Rd is mainly found in plants like P. ginseng, Panax notoginseng, P. quinquefolius, Panax japonicas etc. Ginsenoside Rd reduces metastasis, proliferation, migration, inducing apoptosis, reverses EMT via different signaling pathways like Akt/mTOR signaling, STAT3/JAK2 signaling, miR-18a-mediated downregulation of SMAD2 [168][169][170][171]. Numerous studies show that Ginsenoside Rd has no significant side effects [184][185]. Oridonin is primarily found in plants like Rabdosia rubescens, Isodon japonicus Hara, Isodon trichocarpus, Isodon enanderianus, and I. lophanthoides. Oridonin inhibits angiogenesis, migration, invasion and adhesion, proliferation, and phagocytosis properties via HIF-1α/VEGF, ERK/AKT, ERK/NFκB signaling [166][166][167]. Oridonin reduces the side effects of various other cancer drugs when used in combination [186]. Oridonin shows anticancer properties with very low side effects [187]. Britanin is mainly found in plants like Inula lineariifolia Turcz. (Asteraceae), Inula japonica, Inula Britannica. Britanin induces apoptosis and autophagy and inhibits cell proliferation and angiogenesis via regulating different pathways like IKK1/1KK2, NF-κB, and AMPK signaling [172][173][175]. Britanin shows tolerable side effects at low dose administration *in vivo* [188]. Scutellarein is found primarily in plants like Scutellaria lateriflora, Asplenium belangeri, Mexican oregano, sweet orange, Scutellaria barbata. Scutellarein inhibits cell migration, adhesion, reducing inflammation, induces vasodilation via regulating different pathways like eNOS/NO/PKG, NF-κb [176][177][180]. No side effects were absorbed when treated with Scutellarein in various studies [189].

4.3.2. ANALYSIS OF IMPACT OF GALLIC ACID AND SHIKONIN ON DEG OF HNSC.

Expression data of HNSC cancer and both the natural compounds merged in Microsoft-excel were studied and filtered expression of HNSC cancer with a cut of greater than one. 1016 genes were found

differentially over-expressed in HNSC cancer and were compared this with the expression data of Gallic acid and Shikonin. Gallic acid results in downregulation of 120 of these differentially over-expressed genes and shikonin down-regulates 660 genes from these 1016 over-expressed genes. Again, for down regulated gene, Expression data of HNSC cancer filtered with a cut off less than or equals to -1. 795 genes were found differentially down expressed in the HNSC cancer and were compared with the expression data of Gallic acid and Shikonin. Gallic acid results in upregulation of 35 genes and Shikonin up-regulates 38 genes from these 795 down-regulated genes. So, this combination of gallic acid and shikonin could be effective for the HNSC cancer treatment. **(Figure 4.6)**



	Downregulated In HNSC	Upregulated In HNSC
Down By Gallic Acid		120
Down By Shikonin		660
Up By Gallic Acid	35	
Up By Shikonin	38	

Figure 4.6: Figure demonstrate that expression of no. of genes altered by gallic acid and Shikonin from the differentially expressed genes. Gallic acid and Shikonin downregulates 120 genes and 660 genes, respectively that are upregulated in HNSC, whereas gallic acid and Shikonin upregulates 35 and 38 genes, respectively that are downregulates in HNSC.

Gallic acid shows anti-cancer activity by its selective cell death effect in various cancer cells but not in normal cells[161]. The molecular targets and function of gallic acid are activation of NF-B inhibition, ATM kinase, UDP-glucose dehydrogenase inhibition, Apoptosis induction, Ribonucleotide reductase inhibition, Cyclooxygenase inhibition, GSH depletion, Invasion inhibition[162]. Shikonin deregulates

the cellular Ca^{2+} and ROS levels in the mitochondria which leads to breakdown of mitochondrial membrane potential, dysfunction of microtubules, cell-cycle arrest, and ultimately induction of apoptosis. The structure and metabolism of mitochondria is very different in cancer as compared to normal cells, hence Shikonin is a promising candidate for next generation of chemotherapy as shown in **Figure 4.7.**

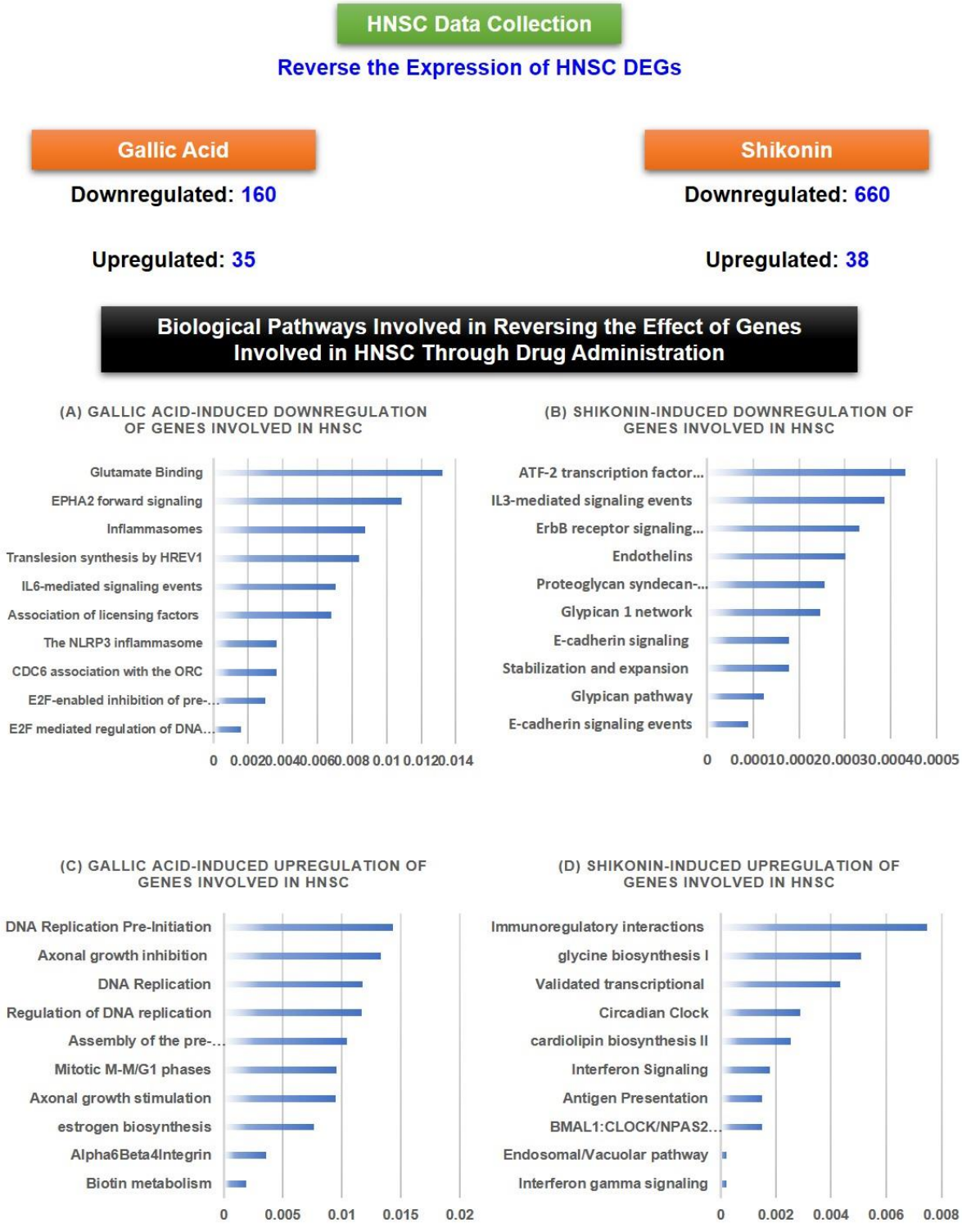


Figure 4.7: Following figure demonstrate the pathway analysis of differential genes involved in HNSC and whose expression are reverse by the action of gallic acid and Shikonin, respectively.

Funrich tool was used for the pathway analysis of regulated genes by Gallic acid and Shikonin from the differentially expressed genes results show genes getting downregulated by Gallic acid involved in Pathways like Glutamate Binding, inflammasomes, translation synthesis by HREV1, IL6-mediated signalling events, association of licensing factor, The NLRP3 inflammasome, CDC6 association with the ORC Pathways, EPHA2 forward signalling. Genes upregulated by Gallic acid involved in pathways like axonal growth inhibition, DNA Replication Pre-Initiation, axonal growth stimulation, estrogen biosynthesis, biotin metabolism etc. Genes downregulated by Shikonin involved in pathways like ATF-2 transcription factor, IL3-mediated signalling events, ErbB receptor signalling, endothelin, E-cadherin signalling, stabilization and expansion, glypican pathway, E-cadherin signalling events etc. Genes upregulated by Shikonin were involved in pathways like immunoregulatory interactions, glycine biosynthesis, validated transcriptional, circadian Clock, interferon Signalling, antigen presentation etc. **Figure 4.8.**

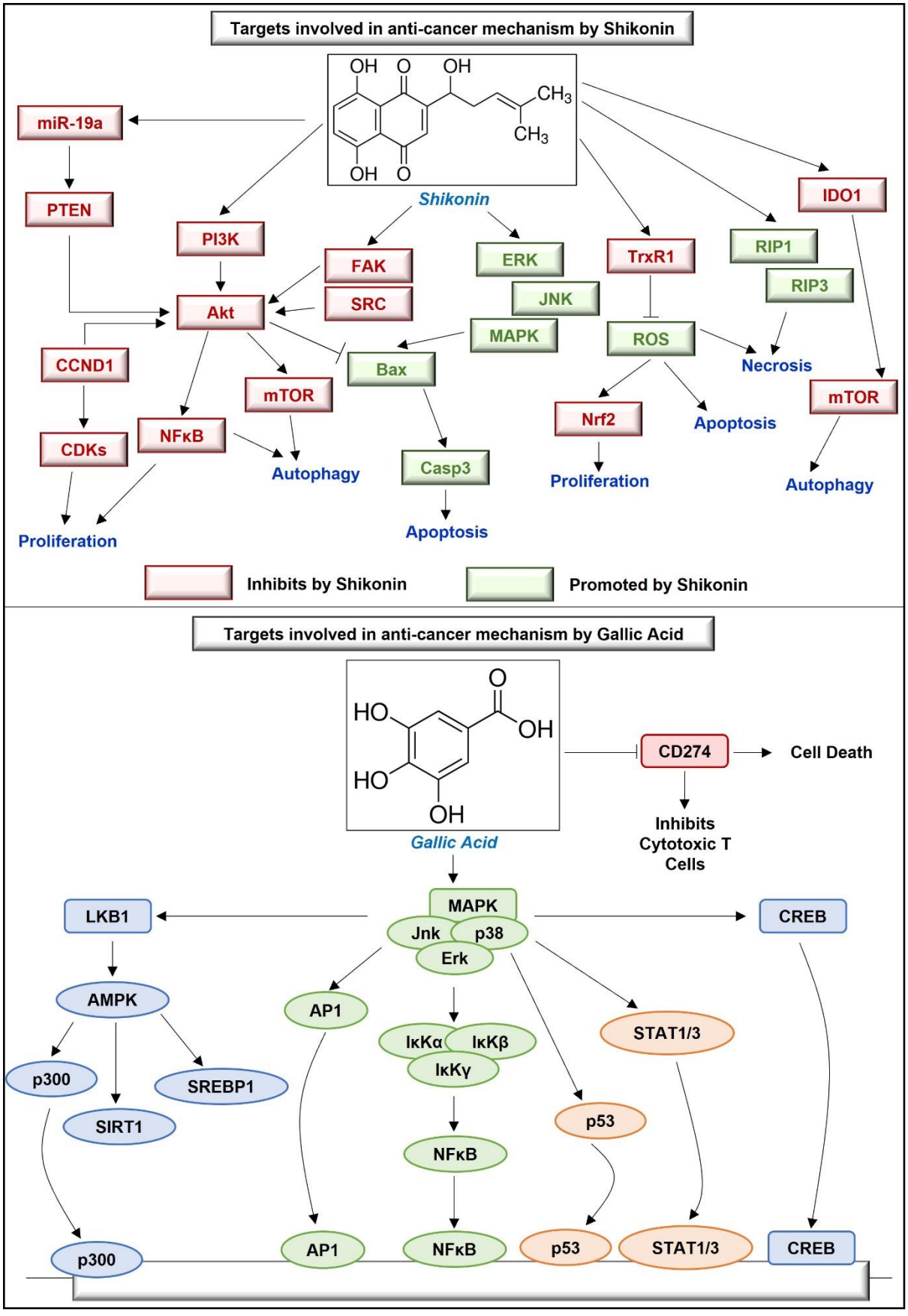


Figure 4.8: Signaling molecules regulated by two phytochemicals, namely Shikonin and Gallic acid.

For example, Shikonin Inhibits PI3K, mir19a, FAK, TrxR1, IDO1 that regulates cell proliferation, autophagy, and apoptosis. In addition, Shikonin upregulates the activity of Erk, Casp3, and RIP1 that are involved in the regulation of autophagy, apoptosis, proliferation, and necrosis. Moreover, Gallic acid targets MAPKs, namely Jnk, p38, and Erk, which regulates the transcriptional status of various signaling molecules, namely CREB, STAT1/3, LKB1, AP1, and p53.

Hence, the genes involved in HNSC which contributed to cancer prognosis were shown to be regulated by the natural compounds that can potentially impact cancer progression and immunity related pathways.

4.3.3. VALIDATION OF NATURAL COMPOUND COMBINATION IN IN-VITRO CONDITIONS

OD observations and percentage viability of cells after treatment with Shikonin at different doses shown in **Table 4.3**. Dose response curve shown in **Figure 4.9(a).** The IC50 values of Shikonin were shown to be 13.86, 11.95, and 10.89 at 12h, 24h, and 48h treatment, respectively. Cell viability remained relatively high (above 90%) at low doses (1 µl and 2 µl) across all time points, implying that these doses had a minor inhibitory effect on cell growth. Cell viability was progressively decreased at all time points as the dose was increased (4 to 20). A greater inhibitory effect on cell viability was observed with increased doses of Shikonin. At higher doses (10 to 20), cell viability significantly dropped, with percentages as low as 14.7% at 48 hours for 20µl, suggesting that higher Shikonin doses negatively impacted cell viability and growth, resulting in a significant reduction in viable cells. The IC_{50} values (13.86 after 12 hours, 11.95 after 24 hours, and 10.89 after 48 hours) were represented as the drug concentration required to inhibit 50% of cell viability. A lower IC50 value indicated that the drug was more effective at inhibiting cell viability at low concentrations.

Serial no.	Dosage conc.	OD at 12h	OD at 24h	OD at 48h	% cell viability at 12 h	% cell viability at 24 h	% cell viability at 48 h
1	0	0.711	0.739	0.768	100	100	100
2	1	0.68	0.69	0.719	95.33133	92.91908	93.20388
3	2	0.666	0.69	0.705	93.22289	92.91908	91.26214
4	4	0.657	0.672	0.681	91.86747	90.31792	87.93343
5	6	0.641	0.641	0.656	89.45783	85.83815	84.46602
6	8	0.619	0.638	0.579	86.14458	85.40462	73.78641
7	10	0.556	0.581	0.452	76.65663	77.16763	56.17198
8	12	0.442	0.359	0.327	59.48795	45.08671	38.83495
9	16	0.297	0.243	0.194	37.6506	28.3237	20.38835
10	20	0.224	0.173	0.153	26.65663	18.20809	14.7018

Table 4.3: Represents the effect of different doses of Shikonin on MDA-MB-231 cell lines for 12h, 24h, 48h in term of OD value and percentage viability.

OD observations and % viability of cells after treatment with Gallic acid at different doses shown in **Table 4.4**. Dose response curve shown in **Figure 4.9(b).** The IC50 values of Gallic acid were shown to be 46.87, 59.37, and 93.75 at 12h, 24h, and 48h treatment, respectively. Cell viability remained relatively high (above 90%) across all time points at low doses (10 µl and 20 µl), implying that these doses had a minor inhibitory effect on cell growth. Cell viability was progressively decreased at all time points as the dose was increased (50 µl to 100 µl). With increasing doses of Gallic acid, the inhibitory effect on cell growth and viability became more pronounced. At higher doses (60 µl to 100 µl), cell viability dropped significantly, with percentages as low as 15.10% at 48 hours for 100 µl dose.

Serial no.	Dosage conc.	OD at 12h	OD at 24h	OD at 48h	% cell viability at 12 h	% cell viability at 24 h	% cell viability at 48 h
1	0	0.711	0.739	0.768	100	100	100
2	10	0.705	0.736	0.755	99.09639	99.56647	92.1875
3	20	0.698	0.714	0.744	98.04217	96.38728	90.75521
4	30	0.686	0.681	0.698	96.23494	91.6185	84.76563
5	40	0.678	0.643	0.536	95.03012	86.12717	63.67188
6	50	0.64	0.538	0.38	89.30723	70.95376	43.35938
7	60	0.607	0.387	0.253	84.33735	49.13295	26.82292
8	70	0.481	0.365	0.22	65.36145	45.95376	22.52604
9	80	0.406	0.28	0.194	54.06627	33.67052	19.14063
10	90	0.395	0.222	0.167	52.40964	25.28902	15.625
11	100	0.352	0.197	0.163	45.93373	21.6763	15.10417

Table 4.4: Represents the effect of different doses of Gallic acid on MDA-MB-231 cell lines for 12h, 24h, 48h in term of OD value and percentage viability.

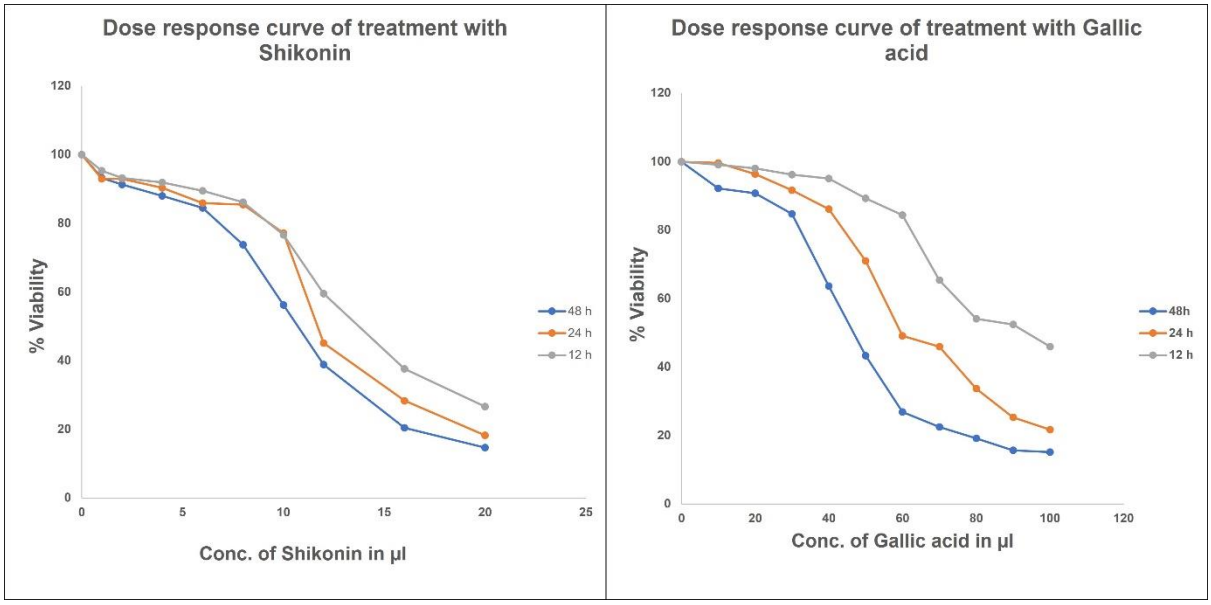


Figure 4.9: Figure showed the percentage cell viability at different concentration of Shikonin (left) and gallic acid (right).

OD observations and % viability of cells after treatment with different dose combination of Shikonin and Gallic acid are shown in **Table 4.5**. The bar plot in **Figure 4.10** showed the percentage cell viability after treatment with combinations of shikonin and gallic acid in the ratios as mentioned in the figure exposed for different time periods. Cell viability percentages after 12 hours, 24 hours, and 48 hours for dose combination G40 µl +S12 µl were 47.59%, 38.52%, and 31.54%, respectively. Cell viability was further shown to decrease at all time points as the dosage concentration of both Gallic acid and Shikonin was increased in combination treatments. At 48 hours, the lowest cell viability percentage of 14.46% was achieved with the drug combination dose G80µl+S16µl.

Serial no.	Dosage conc.	OD at 12h	OD at 24h	OD at 48h	% cell viability at 12 h	% cell viability at 24 h	% cell viability at 48 h
1	$G_{40µl}+S_{12µl}$	0.363	0.317	0.276	47.59036	38.51641	31.5427
2	$G_{40µl}+S_{16µl}$	0.285	0.227	0.186	35.84337	25.6776	19.14601
3	$G_{40µl}+S_{20µl}$	0.212	0.191	0.182	24.8494	20.54208	18.59504
4	$G_{60µl}+S_{12µl}$	0.329	0.302	0.257	42.46988	36.3766	28.92562
5	$G_{60µl}+S_{16µl}$	0.269	0.205	0.174	33.43373	22.53923	17.49311
6	$G_{60µl}+S_{20µl}$	0.271	0.247	0.201	33.73494	28.53067	21.21212
7	$G_{80µl}+S_{12µl}$	0.342	0.296	0.231	44.42771	35.52068	25.34435
8	$G_{80µl}+S_{16µ}l$	0.246	0.181	0.152	29.96988	19.11555	14.46281
9	$G_{80µl}+S_{20µl}$	0.217	0.189	0.168	25.60241	20.25678	16.66667

Table 4.5: Represents the effect of combination of gallic acid and shikonin on MDA-MB-231 cell lines in a dose dependent manner.

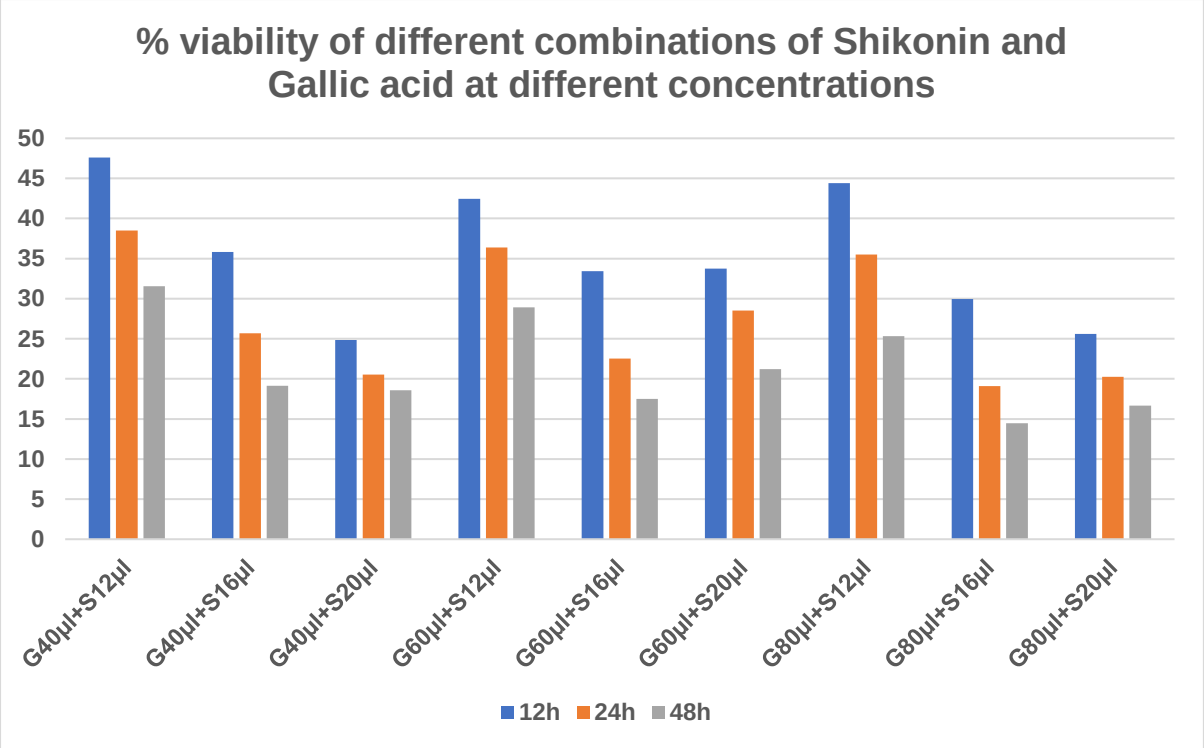


Figure 4.10: Represents the effect of combination of gallic acid and shikonin on MDA-MB-231 cell lines at different conc. And time interval.

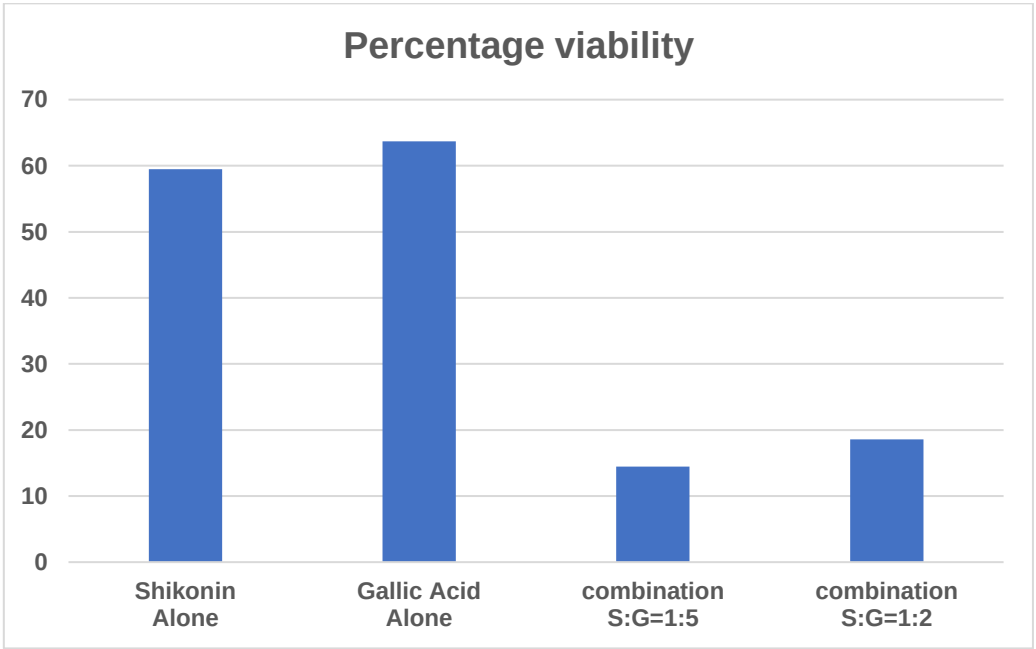


Figure 4.11: Comperition of percentage viability of cells after treatment with shikonin and gallic acid individually and their combinations. (X-axis: percentage of cell viability, Y-axis: individual effect of single natural compound and combination of natural compound)

Cell viability after Shikonin and Gallic acid treatment (individually) were compared to the combination treatments in Figure 4.11, it was observed that the drug combination had a stronger inhibitory effect on cell viability. The data suggested that Gallic acid and Shikonin had a synergistic effect on MDA-MB-

231 cell viability. The combination treatment was found to inhibit cell viability more effectively than the individual drugs alone.

Therefore, it can be stated that the combination of gallic acid and shikonin could be benefcial for the combinatorial treatment of HNSC cancer. Many plants, including Lithospermum erythrorhizon, Alkanna, Arnebia, Onosma, Onosma sericeum Willd, and Echium generate shikonin and research have previously shown that shikonin regulates various functions in these plants, including transgene expression. Shikonin has been used as a red dye for centuries and is reported to possess medicinal properties. It was evaluated as a multi-functional antibacterial and UV protective agent on a silk fabric, exhibits insulin-like activities by inhibiting phosphatase and tensin homologue deleted on Chromosome 10 (PTEN). Further, the drug has shown various properties, such as anti-viral, anti-tumor, cardiotonic and contraceptive properties. Similarly, gallic acid is found in many food sources like banana, walnut, hazelnut, green tea, avocado, guava, mango, mulberry, pomegranate, blackcurrant, cashew, red wine, strawberry, blueberry, apple, grape etc.

Gallic acid is a typical antioxidant tea formulation, and thus considered as potential natural antioxidant. Moreover, Gallic acid in addition to its phytochemical activity is also utilised in tanning, ink colours, and paper manufacturing. Gallic acid, commonly known as 3,4,5-trihydroxybenzoic acid, is a phenolic chemical, which can be found both in its free form and as a component of tannins, specifcally gallotannin. Additionally, gallic acid and its derivatives can be found in almost all parts of the plant, including the bark, wood, leaf, fruit, root, and seed.

CHAPTER V

Conclusion

CHAPTER V: CONCLUSION

Cancer remains a complex disease with diverse underlying causes and mechanisms. Addressing the burden of cancer requires a comprehensive approach involving prevention, early detection, and effective treatment strategies. By fostering collaboration between researchers, healthcare professionals, policymakers, and the public, we can strive towards reducing the global impact of cancer and improving the lives of affected individuals.

Early detection of cancer using tumor-derived biomarkers for cancer has several lacunae and has been discussed extensively. Since PBMCs are immune cells in the blood that help the host immune system respond to tumor cells, peripheral blood profiling can be used for early detection of cancer based on immune marker profiling that alters due to the host immune system's reaction to cancer. Cancer patients often exhibit a high frequency of heterogenicity in tumor expressed biomarkers. This heterogenicity results in the success of certain therapy in some patients while others do not respond. Hence there is variable efficacy of different drugs in different populations. However, biomarkers that are differentially expressed as result of immunological response to tumor show significantly lower frequency of heterogenicity. Our study that focuses on the immunomodulatory biomarkers would therefore be effective in developing therapy that would show universal responsiveness in diverse patients. It also offers the possibility of early cancer detection with minimally invasive methods (even before clinical symptoms appear). It can also be useful for predicting how a tumor will grow and how a patient will fair and the prognosis of clinical progression. However, since these results are based on computational biology, in vivo studies are necessary to validate them. This study promotes the application of XAI on ML models for quantifying & comprehensively examining the predicted findings, particularly in biology, for the development of biomarkers of predictive and prognostic significance. Traditional biomarkers of tumor cells showed polymorphism in patients therefore immunological biomarkers might be a better alternative as a therapeutic

target.

Synthetic drugs even FDA approved have side effect on normal cells which is often due to additional binding of therapeutic drug to off target receptors leading to unpleasant side effects. Natural compounds are multitargeting; therefore, they can simultaneously target multiple pathways and many biological processes, helping in tumor regression. Combination of natural compounds was developed that can help in HNSC tumor regression and immune modulation with minimal side effects. As this combination was further analyzed, it was found that many biological processes were regulated by more than one compound via different pathways; therefore, it might not be easy for tumor cells to escape this regression mechanism. Further, tumor cells cannot gain drug resistance easily against them. Immunotherapy is typically associated with side effects that often deter the use of such treatment strategies. Our combination of natural compounds holds a better immunotherapeutic potential without the commonly associated side effects typically seen with chemical immunomodulatory drugs. Our study has opened a new dimension for developing a combinatorial natural compound cocktail as a potential immunomodulatory drug alternative. Thus, we propose that such a combination be further analyzed in *in vivo* studies to develop better treatment for tumor patients.

REFERENCES

[1] R. Fisher, L. Pusztai, and C. Swanton, "Cancer heterogeneity: implications for targeted therapeutics," *Br. J. Cancer*, vol. 108, no. 3, p. 479, Feb. 2013, doi: 10.1038/BJC.2012.581.

[2] A. Wadhawan, M. Chatterjee, and G. Singh, "Present Scenario of Bioconjugates in Cancer Therapy: A Review," *Int. J. Mol. Sci. 2019, Vol. 20, Page 5243*, vol. 20, no. 21, p. 5243, Oct. 2019, doi: 10.3390/IJMS20215243.

[3] K. D. Miller *et al.*, "Cancer treatment and survivorship statistics, 2016," *CA. Cancer J. Clin.*, vol. 66, no. 4, pp. 271–289, Jul. 2016, doi: 10.3322/CAAC.21349.

[4] L. Zhong *et al.*, "Small molecules in targeted cancer therapy: advances, challenges, and future perspectives," *Signal Transduct. Target. Ther. 2021 61*, vol. 6, no. 1, pp. 1–48, May 2021, doi: 10.1038/s41392-021-00572-w.

[5] P. Economopoulou, R. de Bree, I. Kotsantis, and A. Psyrri, "Diagnostic tumor markers in head and neck squamous cell carcinoma (HNSCC) in the clinical setting," *Front. Oncol.*, vol. 9, no. AUG, p. 827, 2019, doi: 10.3389/FONC.2019.00827/BIBTEX.

[6] M. R. I. Young, "Protective mechanisms of head and neck squamous cell carcinomas from immune assault," *Head Neck*, vol. 28, no. 5, pp. 462–470, May 2006, doi: 10.1002/HED.20331.

[7] G. L. Beatty and W. L. Gladney, "Immune escape mechanisms as a guide for cancer immunotherapy," *Clin. Cancer Res.*, vol. 21, no. 4, p. 687, Feb. 2015, doi: 10.1158/1078-0432.CCR-14-1860.

[8] S. M. Y. Chen, A. L. Krinsky, R. A. Woolaver, X. Wang, Z. Chen, and J. H. Wang, "Tumor Immune Microenvironment in Head and Neck Cancers," *Mol. Carcinog.*, vol. 59, no. 7, p. 766, Jul. 2020, doi: 10.1002/MC.23162.

[9] R. Baghban *et al.*, "Tumor microenvironment complexity and therapeutic implications

at a glance," *Cell Commun. Signal. 2020 181*, vol. 18, no. 1, pp. 1–19, Apr. 2020, doi: 10.1186/S12964-020-0530-4.

[10] D. S. Vinay *et al.*, "Immune evasion in cancer: Mechanistic basis and therapeutic strategies," *Semin. Cancer Biol.*, vol. 35, pp. S185–S198, Dec. 2015, doi: 10.1016/J.SEMCANCER.2015.03.004.

[11] A. D. Waldman, J. M. Fritz, and M. J. Lenardo, "A guide to cancer immunotherapy: from T cell basic science to clinical practice," *Nat. Rev. Immunol. 2020 2011*, vol. 20, no. 11, pp. 651–668, May 2020, doi: 10.1038/s41577-020-0306-5.

[12] P. L. Bedard, A. R. Hansen, M. J. Ratain, and L. L. Siu, "Tumour heterogeneity in the clinic," *Nature*, vol. 501, no. 7467, pp. 355–364, Sep. 2013, doi: 10.1038/nature12627.

[13] A. Tadimety, A. Closson, C. Li, S. Yi, T. Shen, and J. X. J. Zhang, "Advances in liquid biopsy on-chip for cancer management: Technologies, biomarkers, and clinical analysis," *Crit. Rev. Clin. Lab. Sci.*, vol. 55, no. 3, pp. 140–162, Apr. 2018, doi: 10.1080/10408363.2018.1425976.

[14] G. Yang *et al.*, "A Systematic Review of Oral Biopsies, Sample Types, and Detection Techniques Applied in Relation to Oral Cancer Detection," *BioTech*, vol. 11, no. 1, p. 5, Mar. 2022, doi: 10.3390/biotech11010005.

[15] M. Akram, M. Iqbal, M. Daniyal, and A. U. Khan, "Awareness and current knowledge of breast cancer," *Biol. Res.*, vol. 50, no. 1, p. 33, Dec. 2017, doi: 10.1186/s40659-017-0140-9.

[16] I. Jatoi and P. F. Pinsky, "Breast Cancer Screening Trials: Endpoints and Overdiagnosis," *JNCI J. Natl. Cancer Inst.*, vol. 113, no. 9, pp. 1131–1135, Sep. 2021, doi: 10.1093/jnci/djaa140.

[17] J. Li *et al.*, "Non-Invasive Biomarkers for Early Detection of Breast Cancer," *Cancers (Basel).*, vol. 12, no. 10, p. 2767, Sep. 2020, doi: 10.3390/cancers12102767.

[18] S. Kure *et al.*, "Breast Cancer Detection from a Urine Sample by Dog Sniffing: A Preliminary Study for the Development of a New Screening Device, and a Literature Review," *Biology (Basel).*, vol. 10, no. 6, p. 517, Jun. 2021, doi: 10.3390/biology10060517.

[19] A. B. Nixon, K. A. Schalper, I. Jacobs, S. Potluri, I. M. Wang, and C. Fleener, "Peripheral immune-based biomarkers in cancer immunotherapy: can we realize their predictive potential?," *J. Immunother. Cancer*, vol. 7, no. 1, pp. 1–14, Nov. 2019, doi: 10.1186/S40425-019-0799-2/TABLES/2.

[20] H. Hou *et al.*, "Peripheral blood transcriptome identifies high-risk benign and malignant breast lesions," 2020, doi: 10.1371/journal.pone.0233713.

[21] J. Marrugo-Ramírez, M. Mir, and J. Samitier, "Blood-Based Cancer Biomarkers in Liquid Biopsy: A Promising Non-Invasive Alternative to Tissue Biopsy," *Int. J. Mol. Sci.*, vol. 19, no. 10, p. 2877, Sep. 2018, doi: 10.3390/ijms19102877.

[22] D. Lucchetti, C. R. Tenore, F. Colella, and A. Sgambato, "Extracellular Vesicles and Cancer: A Focus on Metabolism, Cytokines, and Immunity," *Cancers (Basel).*, vol. 12, no. 1, Jan. 2020, doi: 10.3390/CANCERS12010171.

[23] F. Mattei *et al.*, "The Tumor Microenvironment: A Milieu Hindering and Obstructing Antitumor Immune Responses," *Front. Immunol.* | *www.frontiersin.org*, vol. 1, p. 940, 2020, doi: 10.3389/fimmu.2020.00940.

[24] G. Wen, T. Zhou, and W. Gu, "The potential of using blood circular RNA as liquid biopsy biomarker for human diseases," *Protein Cell*, vol. 12, no. 12, pp. 911–946, Dec. 2021, doi: 10.1007/S13238-020-00799-3.

[25] P. Sharma *et al.*, "Early detection of breast cancer based on gene-expression patterns in peripheral blood cells," *Breast Cancer Res.*, vol. 7, no. 5, Jun. 2005, doi: 10.1186/BCR1203.

[26] H. Čelešnik and U. Potočnik, "Peripheral Blood Transcriptome in Breast Cancer Patients as a Source of Less Invasive Immune Biomarkers for Personalized Medicine, and Implications for Triple Negative Breast Cancer," *Cancers (Basel).*, vol. 14, no. 3, Feb. 2022, doi: 10.3390/CANCERS14030591.

[27] P. L. Crispen and S. Kusmartsev, "Mechanisms of immune evasion in bladder cancer.," *Cancer Immunol. Immunother.*, vol. 69, no. 1, pp. 3–14, Jan. 2020, doi: 10.1007/s00262-019-02443-4.

[28] G. J. Veal *et al.*, "Pharmacodynamic Therapeutic Drug Monitoring for Cancer: Challenges, Advances, and Future Opportunities," *Ther. Drug Monit.*, vol. 41, no. 2, pp. 142–159, Apr. 2019, doi: 10.1097/FTD.0000000000000606.

[29] N. Vasan, J. Baselga, and D. M. Hyman, "A view on drug resistance in cancer," *Nature*, vol. 575, no. 7782. Nature Publishing Group, pp. 299–309, Nov. 14, 2019, doi: 10.1038/s41586-019-1730-1.

[30] I. Bozic and M. A. Nowak, "Timing and heterogeneity of mutations associated with drug resistance in metastatic cancers," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 45, pp. 15964–15968, Nov. 2014, doi: 10.1073/pnas.1412075111.

[31] E. A. Zaal and C. R. Berkers, "The influence of metabolism on drug response in cancer," *Frontiers in Oncology*, vol. 8, no. NOV. Frontiers Media S.A., 2018, doi: 10.3389/fonc.2018.00500.

[32] L. Chen, Y. Zeng, and S.-F. Zhou, "Role of Apoptosis in Cancer Resistance to Chemotherapy," in *Current Understanding of Apoptosis - Programmed Cell Death*, InTech, 2018.

[33] M. R. Salehan and H. R. Morse, "DNA damage repair and tolerance: A role in chemotherapeutic drug resistance," *British Journal of Biomedical Science*, vol. 70, no. 1. Step Publishing Ltd., pp. 31–40, 2013, doi: 10.1080/09674845.2013.11669927.

[34] X. Xue and X. J. Liang, “Overcoming drug efflux-based multidrug resistance in cancer with nanotechnology,” *Chinese Journal of Cancer*, vol. 31, no. 2. BioMed Central, pp. 100–109, 2012, doi: 10.5732/cjc.011.10326.

[35] K. O. Alfarouk *et al.*, “Resistance to cancer chemotherapy: Failure in drug response from ADME to P-gp,” *Cancer Cell International*, vol. 15, no. 1. BioMed Central Ltd., p. 71, Jul. 15, 2015, doi: 10.1186/s12935-015-0221-1.

[36] J. Deng *et al.*, “CDK4/6 inhibition augments antitumor immunity by enhancing T-cell activation,” *Cancer Discov.*, vol. 8, no. 2, pp. 216–233, Feb. 2018, doi: 10.1158/2159-8290.CD-17-0915.

[37] R. L. Ferris, “Immunology and Immunotherapy of Head and Neck Cancer,” *J. Clin. Oncol.*, vol. 33, no. 29, pp. 3293–3304, Oct. 2015, doi: 10.1200/JCO.2015.61.1509.

[38] M. E. Keir, M. J. Butte, G. J. Freeman, and A. H. Sharpe, “PD-1 and Its Ligands in Tolerance and Immunity,” *Annu. Rev. Immunol.*, vol. 26, no. 1, pp. 677–704, Apr. 2008, doi: 10.1146/annurev.immunol.26.021607.090331.

[39] A. Akinleye and Z. Rasool, “Immune checkpoint inhibitors of PD-L1 as cancer therapeutics,” *Journal of Hematology and Oncology*, vol. 12, no. 1. BioMed Central Ltd., p. 92, Sep. 05, 2019, doi: 10.1186/s13045-019-0779-5.

[40] S. Sakhnevych, “Tim-3-galectin-9 immunosuppressive pathway in human acute myeloid leukaemia and solid tumour cells and biochemical functions of its crucial components,” 2019.

[41] “TIGIT & CD155 Immune Checkpoint Pathway.” https://www.sinobiological.com/research/immune-checkpoint/tigit-pathway (accessed May 22, 2020).

[42] S. Chikuma, “CTLA-4, an essential immune-checkpoint for T-Cell activation,” in *Current Topics in Microbiology and Immunology*, vol. 410, Springer Verlag, 2017, pp.

99–126.

[43] L. P. Andrews, A. E. Marciscano, C. G. Drake, and D. A. A. Vignali, “LAG3 (CD223) as a cancer immunotherapy target,” *Immunological Reviews*, vol. 276, no. 1. Blackwell Publishing Ltd, pp. 80–96, Mar. 01, 2017, doi: 10.1111/imr.12519.

[44] S. Gilfillan *et al.*, “DNAM-1 promotes activation of cytotoxic lymphocytes by nonprofessional antigen-presenting cells and tumors,” *J. Exp. Med.*, vol. 205, no. 13, pp. 2965–2973, Dec. 2008, doi: 10.1084/jem.20081752.

[45] C. J. Kearney, K. M. Ramsbottom, I. Voskoboinik, P. K. Darcy, and J. Oliaro, “Loss of DNAM-1 ligand expression by acute myeloid leukemia cells renders them resistant to NK cell killing,” *Oncoimmunology*, vol. 5, no. 8, Aug. 2016, doi: 10.1080/2162402X.2016.1196308.

[46] O. Marinelli, M. Nabissi, M. B. Morelli, L. Torquati, C. Amantini, and G. Santoni, “ICOS-L as a Potential Therapeutic Target for Cancer Immunotherapy,” *Curr. Protein Pept. Sci.*, vol. 19, no. 11, pp. 1107–1113, Jul. 2018, doi: 10.2174/1389203719666180608093913.

[47] D. A. Knee, B. Hewes, and J. L. Brogdon, “Rationale for anti-GITR cancer immunotherapy,” *European Journal of Cancer*, vol. 67. Elsevier Ltd, pp. 1–10, Nov. 01, 2016, doi: 10.1016/j.ejca.2016.06.028.

[48] A. Zheng *et al.*, “PD-L1 promotes head and neck squamous cell carcinoma cell growth through mTOR signaling,” *Oncol. Rep.*, vol. 41, no. 5, pp. 2833–2843, May 2019, doi: 10.3892/or.2019.7053.

[49] M. Matsushita and M. Kawaguchi, “Immunomodulatory Effects of Drugs for Effective Cancer Immunotherapy,” *J. Oncol.*, vol. 2018, 2018, doi: 10.1155/2018/8653489.

[50] P. Desai *et al.*, “An analysis of the effect of statins on the risk of Non-Hodgkin’s Lymphoma in the Women’s Health Initiative cohort,” *Cancer Med.*, vol. 7, no. 5, pp.

2121–2130, May 2018, doi: 10.1002/cam4.1368.

[51] F. V. Pereira *et al.*, "Metformin exerts antitumor activity via induction of multiple death pathways in tumor cells and activation of a protective immune response," *Oncotarget*, vol. 9, no. 40, pp. 25808–25825, 2018, doi: 10.18632/oncotarget.25380.

[52] T. Liu, F. Guo, X. Zhu, X. He, and L. Xie, "Thalidomide and its analogues: A review of the potential for immunomodulation of fibrosis diseases and opthalmopathy," *Experimental and Therapeutic Medicine*, vol. 14, no. 6. Spandidos Publications, pp. 5251–5257, Dec. 01, 2017, doi: 10.3892/etm.2017.5209.

[53] F. Majolo, L. K. de Oliveira Becker Delwing, D. J. Marmitt, I. C. Bustamante-Filho, and M. I. Goettert, "Medicinal plants and bioactive natural compounds for cancer treatment: Important advances for drug discovery," *Phytochem. Lett.*, vol. 31, pp. 196–207, Jun. 2019, doi: 10.1016/j.phytol.2019.04.003.

[54] T. N. Aung, Z. Qu, R. D. Kortschak, and D. L. Adelson, "Understanding the effectiveness of natural compound mixtures in cancer through their molecular mode of action," *International Journal of Molecular Sciences*, vol. 18, no. 3. MDPI AG, Mar. 17, 2017, doi: 10.3390/ijms18030656.

[55] C. A. Buckner, R. M. Lafrenie, J. A. Dénommée, J. M. Caswell, and D. A. Want, "Complementary and alternative medicine use in patients before and after a cancer diagnosis," *Curr. Oncol.*, vol. 25, no. 4, p. e275, Aug. 2018, doi: 10.3747/CO.25.3884.

[56] A. G. Atanasov *et al.*, "Natural products in drug discovery: advances and opportunities," *Nat. Rev. Drug Discov. 2021 203*, vol. 20, no. 3, pp. 200–216, Jan. 2021, doi: 10.1038/s41573-020-00114-z.

[57] J. Sharifi-Rad *et al.*, "Natural Products and Synthetic Analogs as a Source of Antitumor Drugs," *Biomolecules*, vol. 9, no. 11, Nov. 2019, doi: 10.3390/BIOM9110679.

[58] R. R. Ramsay, M. R. Popovic-Nikolic, K. Nikolic, E. Uliassi, and M. L. Bolognesi, "A

perspective on multi-target drug discovery and design for complex diseases," *Clin. Transl. Med.*, vol. 7, no. 1, pp. 1–14, Dec. 2018, doi: 10.1186/S40169-017-0181-2/FIGURES/4.

[59] A. Karimi, M. Majlesi, and M. Rafieian-Kopaei, "Herbal versus synthetic drugs; beliefs and facts," *J. Nephropharmacology*, vol. 4, no. 1, p. 27, 2015, Accessed: Apr. 11, 2022. [Online]. Available: /pmc/articles/PMC5297475/.

[60] M. Huang, J. J. Lu, and J. Ding, "Natural Products in Cancer Therapy: Past, Present and Future," *Nat. Products Bioprospect.*, vol. 11, no. 1, pp. 5–13, Feb. 2021, doi: 10.1007/S13659-020-00293-7/FIGURES/4.

[61] W. P. Steward and K. Brown, "Cancer chemoprevention: a rapidly evolving field," *Br. J. Cancer*, vol. 109, no. 1, p. 1, Jul. 2013, doi: 10.1038/BJC.2013.280.

[62] G. Mohan Shankar, M. Swetha, C. K. Keerthana, T. P. Rayginia, and R. J. Anto, "Cancer Chemoprevention: A Strategic Approach Using Phytochemicals," *Front. Pharmacol.*, vol. 12, p. 4044, Jan. 2022, doi: 10.3389/FPHAR.2021.809308/BIBTEX.

[63] R. Kotecha, A. Takami, and J. L. Espinoza, "Dietary phytochemicals and cancer chemoprevention: a review of the clinical evidence," *Oncotarget*, vol. 7, no. 32, p. 52517, Aug. 2016, doi: 10.18632/ONCOTARGET.9593.

[64] R. B. Mokhtari *et al.*, "Combination therapy in combating cancer," *Oncotarget*, vol. 8, no. 23, pp. 38022–38043, 2017, doi: 10.18632/ONCOTARGET.16723.

[65] N. M. Ayoub, "Editorial: Novel Combination Therapies for the Treatment of Solid Cancers," *Front. Oncol.*, vol. 11, p. 2377, Jun. 2021, doi: 10.3389/FONC.2021.708943/BIBTEX.

[66] Y. Li, Z. Wang, J. A. Ajani, and S. Song, "Drug resistance and Cancer stem cells," *Cell Commun. Signal. 2021 191*, vol. 19, no. 1, pp. 1–11, Feb. 2021, doi: 10.1186/S12964-020-00627-5.

[67] S. A. Eccles and D. R. Welch, "Metastasis: recent discoveries and novel treatment strategies," *Lancet*, vol. 369, no. 9574, p. 1742, May 2007, doi: 10.1016/S0140-6736(07)60781-8.

[68] L. Yang *et al.*, "Targeting cancer stem cell pathways for cancer therapy," *Signal Transduct. Target. Ther. 2020 51*, vol. 5, no. 1, pp. 1–35, Feb. 2020, doi: 10.1038/s41392-020-0110-5.

[69] D. Szklarczyk *et al.*, "The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible," *Nucleic Acids Res.*, vol. 45, no. D1, pp. D362–D368, 2017, doi: 10.1093/nar/gkw937.

[70] H. Choi *et al.*, "NPCARE: database of natural products and fractional extracts for cancer regulation," *J. Cheminform.*, vol. 9, no. 1, p. 2, Dec. 2017, doi: 10.1186/s13321-016-0188-5.

[71] T. Barrett *et al.*, "NCBI GEO: archive for functional genomics data sets—update," *Nucleic Acids Res.*, vol. 41, no. D1, pp. D991–D995, Nov. 2012, doi: 10.1093/nar/gks1193.

[72] B. Braakhuis *et al.*, "Expression signature in peripheral blood cells for molecular diagnosis of head and neck squamous cell carcinoma," *Oral Dis.*, vol. 19, no. 5, pp. 452–455, Jul. 2013, doi: 10.1111/odi.12019.

[73] C. Lv *et al.*, "The gene expression profiles in response to 102 traditional Chinese medicine (TCM) components: a general template for research on TCMs," *Sci. Rep.*, vol. 7, no. 1, p. 352, Dec. 2017, doi: 10.1038/s41598-017-00535-8.

[74] X. A. Qu and D. K. Rajpal, "Applications of Connectivity Map in drug discovery and development," *Drug Discov. Today*, vol. 17, no. 23–24, pp. 1289–1298, Dec. 2012, doi: 10.1016/j.drudis.2012.07.017.

[75] G. K. Smyth, "Linear Models and Empirical Bayes Methods for Assessing Differential

Expression in Microarray Experiments," *Stat. Appl. Genet. Mol. Biol.*, vol. 3, no. 1, pp. 1–25, Jan. 2004, doi: 10.2202/1544-6115.1027.

[76] M. Pathan *et al.*, "FunRich: An open access standalone functional enrichment and interaction network analysis tool," *Proteomics*, vol. 15, no. 15, pp. 2597–2601, Aug. 2015, doi: 10.1002/pmic.201400515.

[77] S. H. A. Mahdi, H. Cheng, J. Li, and R. Feng, "The effect of TGF-beta-induced epithelial-mesenchymal transition on the expression of intracellular calcium-handling proteins in T47D and MCF-7 human breast cancer cells," *Arch. Biochem. Biophys.*, vol. 583, pp. 18–26, Aug. 2015, doi: 10.1016/j.abb.2015.07.008.

[78] R. L. Furler, D. F. Nixon, C. A. Brantner, A. Popratiloff, and C. H. Uittenbogaart, "TGF-β sustains tumor progression through biochemical and mechanical signal transduction," *Cancers*, vol. 10, no. 6. MDPI AG, Jun. 14, 2018, doi: 10.3390/cancers10060199.

[79] L. Yang, Y. Pang, and H. L. Moses, "TGF-β and immune cells: an important regulatory axis in the tumor microenvironment and progression," *Trends in Immunology*, vol. 31, no. 6. NIH Public Access, pp. 220–227, Jun. 2010, doi: 10.1016/j.it.2010.04.002.

[80] A. Calon *et al.*, "Dependency of Colorectal Cancer on a TGF-β-Driven Program in Stromal Cells for Metastasis Initiation," *Cancer Cell*, vol. 22, no. 5, pp. 571–584, Nov. 2012, doi: 10.1016/j.ccr.2012.08.013.

[81] L. Shao *et al.*, "IRF1 inhibits antitumor immunity through the upregulation of PD-L1 in the tumor cell," *Cancer Immunol. Res.*, vol. 7, no. 8, pp. 1258–1266, 2019, doi: 10.1158/2326-6066.CIR-18-0711.

[82] S. Xia, R. Ji, Y. Xu, X. Ni, Y. Dong, and W. Zhan, "Twisted gastrulation BMP signaling modulator 1 regulates papillary thyroid cancer cell motility and proliferation," *J. Cancer*, vol. 8, no. 14, pp. 2816–2827, 2017, doi: 10.7150/jca.18482.

[83] S. Goel *et al.*, "CDK4/6 inhibition triggers anti-tumour immunity," *Nature*, vol. 548, no.

7668, pp. 471–475, Aug. 2017, doi: 10.1038/nature23465.

[84] S. Tadesse, M. Yu, M. Kumarasiri, B. T. Le, and S. Wang, "Targeting CDK6 in cancer: State of the art and new insights," *Cell Cycle*, vol. 14, no. 20. Taylor and Francis Inc., pp. 3220–3230, 2015, doi: 10.1080/15384101.2015.1084445.

[85] E. B. Rankin and A. J. Giaccia, "The receptor tyrosine kinase AXL in cancer progression," *Cancers*, vol. 8, no. 11. MDPI AG, Nov. 01, 2016, doi: 10.3390/cancers8110103.

[86] K. F. Ludwig *et al.*, "Small-molecule inhibition of Axl targets tumor immune suppression and enhances chemotherapy in pancreatic cancer," *Cancer Res.*, vol. 78, no. 1, pp. 246–255, Jan. 2018, doi: 10.1158/0008-5472.CAN-17-1973.

[87] G. Papoff, N. Trivieri, R. Crielesi, F. Ruberti, S. Marsilio, and G. Ruberti, "FADD-calmodulin interaction: A novel player in cell cycle regulation," *Biochim. Biophys. Acta - Mol. Cell Res.*, vol. 1803, no. 8, pp. 898–911, Aug. 2010, doi: 10.1016/j.bbamcr.2010.04.006.

[88] S. L. Osborn *et al.*, "Fas-associated death domain (FADD) is a negative regulator of T-cell receptor-mediated necroptosis," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 29, pp. 13034–13039, Jul. 2010, doi: 10.1073/pnas.1005997107.

[89] X. Hou, G. Zhou, Y. Fan, Q. Zhang, and S. Yao, "The High Expression of CD276/HAVCR2 and CD163 is an Adverse Immune Subtype of Glioblastoma and is Closely Related to Epithelial-Mesenchymal Transition," doi: 10.21203/rs.3.rs-31174/v2.

[90] Y. H. Huang *et al.*, "CEACAM1 regulates TIM-3-mediated tolerance and exhaustion," *Nature*, vol. 517, no. 7534, pp. 386–390, Jan. 2015, doi: 10.1038/nature13848.

[91] Y. Zhang *et al.*, "High expression of PRKDC promotes breast cancer cell growth via p38 MAPK signaling and is associated with poor survival," *Mol. Genet. Genomic Med.*,

vol. 7, no. 11, Nov. 2019, doi: 10.1002/mgg3.908.

[92] K. T. Tan *et al.*, "PRKDC: New biomarker and drug target for checkpoint blockade immunotherapy," *J. Immunother. Cancer*, vol. 8, no. 1, Apr. 2020, doi: 10.1136/jitc-2019-000485.

[93] E. Sheikhpour, P. Noorbakhsh, E. Foroughi, S. Farahnak, R. Nasiri, and H. Neamatzadeh, "A Survey on the Role of Interleukin-10 in Breast Cancer: A Narrative.," *Reports Biochem. Mol. Biol.*, vol. 7, no. 1, pp. 30–37, Oct. 2018, Accessed: Aug. 26, 2020. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/30324115.

[94] O. Turovskaya, G. Kim, H. Cheroutre, M. Kronenberg, and R. Madan, "Interleukin 10 acts on regulatory t cells to maintain expression of the transcription factor foxp3 and suppressive function in mice with colitis," *Nat. Immunol.*, vol. 10, no. 11, pp. 1178–1184, 2009, doi: 10.1038/ni.1791.

[95] L. Llanes-Fernández *et al.*, "Relationship between IL-10 and tumor markers in breast cancer patients," *Breast*, vol. 15, no. 4, pp. 482–489, 2006, doi: 10.1016/j.breast.2005.09.012.

[96] J. Zhang, H. Li, J.-P. Yu, S. E. Wang, and X.-B. Ren, "Role of SOCS1 in tumor progression and therapeutic application," *Int. J. Cancer*, vol. 130, no. 9, pp. 1971–1980, May 2012, doi: 10.1002/ijc.27318.

[97] S. Chikuma, M. Kanamori, S. Mise-Omata, and A. Yoshimura, "Suppressors of cytokine signaling: Potential immune checkpoint molecules for cancer immunotherapy," *Cancer Science*, vol. 108, no. 4. Blackwell Publishing Ltd, pp. 574–580, Apr. 01, 2017, doi: 10.1111/cas.13194.

[98] N. Torres *et al.*, "Restoration of antitumor immunity through anti-MICA antibodies elicited with a chimeric protein," *J. Immunother. Cancer*, vol. 8, no. 1, p. 233, Jun. 2020, doi: 10.1136/jitc-2019-000233.

[99] X. M. Zhou *et al.*, “Intrinsic expression of immune checkpoint molecule TIGIT could help tumor growth in vivoby suppressing the function of NK and CD8+T Cells,” *Front. Immunol.*, vol. 9, no. NOV, p. 2821, Nov. 2018, doi: 10.3389/fimmu.2018.02821.

[100] H. Harjunpää and C. Guillerey, “TIGIT as an emerging immune checkpoint,” *Clin. Exp. Immunol.*, vol. 200, no. 2, pp. 108–119, May 2020, doi: 10.1111/cei.13407.

[101] L. Hornyák *et al.*, “The role of indoleamine-2,3-dioxygenase in cancer development, diagnostics, and therapy,” *Frontiers in Immunology*, vol. 9, no. JAN. Frontiers Media S.A., p. 31, Jan. 31, 2018, doi: 10.3389/fimmu.2018.00151.

[102] A. I. Thaker *et al.*, “IDO1 Metabolites Activate β-catenin Signaling to Promote Cancer Cell Proliferation and Colon Tumorigenesis in Mice,” *Gastroenterology*, vol. 145, no. 2, 2013, doi: 10.1053/j.gastro.2013.05.002.

[103] G. C. Prendergast *et al.*, “Indoleamine 2,3-dioxygenase pathways of pathogenic inflammation and immune escape in cancer,” *Cancer Immunology, Immunotherapy*, vol. 63, no. 7. Springer Science and Business Media Deutschland GmbH, pp. 721–735, 2014, doi: 10.1007/s00262-014-1549-4.

[104] “IDO: A Target for Cancer Treatment | Personalized Medicine in Oncology.” http://www.personalizedmedonc.com/publications/ito/june-2014-part-2/ido-a-target-for-cancer-treatment/ (accessed Aug. 26, 2020).

[105] L. Long *et al.*, “The promising immune checkpoint LAG-3: From tumor microenvironment to cancer immunotherapy,” *Genes and Cancer*, vol. 9, no. 5–6, pp. 176–189, May 2018, doi: 10.18632/genesandcancer.180.

[106] Y. Zhao, W. Yang, Y. Huang, R. Cui, X. Li, and B. Li, “Evolving Roles for Targeting CTLA-4 in Cancer Immunotherapy,” *Cell. Physiol. Biochem.*, vol. 47, no. 2, pp. 721–734, Jun. 2018, doi: 10.1159/000490025.

[107] W. Yu *et al.*, “PD-L1 promotes tumor growth and progression by activating WIP and β-

catenin signaling pathways and predicts poor prognosis in lung cancer," *Cell Death Dis.*, vol. 11, no. 7, pp. 1–16, Jul. 2020, doi: 10.1038/s41419-020-2701-z.

[108] T. Tanegashima *et al.*, "Immune suppression by PD-L2 against spontaneous and treatment-related antitumor immunity," *Clin. Cancer Res.*, vol. 25, no. 15, pp. 4808–4819, Aug. 2019, doi: 10.1158/1078-0432.CCR-18-3991.

[109] S. Yang *et al.*, "FOXP3 promotes tumor growth and metastasis by activating Wnt/β-catenin signaling pathway and EMT in non-small cell lung cancer," *Mol. Cancer*, vol. 16, no. 1, pp. 1–12, Jul. 2017, doi: 10.1186/s12943-017-0700-1.

[110] F. Mercer and D. Unutmaz, "The biology of FoxP3: AKey player in immune suppression during infections, autoimmune diseases and cancer," *Adv. Exp. Med. Biol.*, vol. 665, pp. 47–59, 2009, doi: 10.1007/978-1-4419-1599-3_4.

[111] Y. Han, D. Liu, and L. Li, "PD-1/PD-L1 pathway: current researches in cancer.," *Am. J. Cancer Res.*, vol. 10, no. 3, pp. 727–742, 2020, Accessed: Aug. 29, 2020. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/32266087.

[112] N. Sun, A. Taguchi, and S. Hanash, "Switching Roles of TGF-β in Cancer Development: Implications for Therapeutic Target and Biomarker Studies," *J. Clin. Med.*, vol. 5, no. 12, p. 109, Nov. 2016, doi: 10.3390/jcm5120109.

[113] J. Sheng, W. Chen, and H. J. Zhu, "The immune suppressive function of transforming growth factor-β (TGF-β) in human diseases," *Growth Factors*, vol. 33, no. 2, pp. 92–101, Apr. 2015, doi: 10.3109/08977194.2015.1010645.

[114] Y. Liu, P. Zhang, J. Li, A. B. Kulkarni, S. Perruche, and W. J. Chen, "A critical function for TGF-β signaling in the development of natural CD4+CD25+Foxp3+ regulatory T cells," *Nat. Immunol.*, vol. 9, no. 6, pp. 632–640, Jun. 2008, doi: 10.1038/ni.1607.

[115] L. D. S. Johnson and S. C. Jameson, "TGF-β sensitivity restrains CD8+ T cell homeostatic proliferation by enforcing sensitivity to IL-7 and IL-15," *PLoS One*, vol. 7,

no. 8, Aug. 2012, doi: 10.1371/journal.pone.0042268.

[116] C. J. Sherr, D. Beach, and G. I. Shapiro, “Targeting CDK4 and CDK6: From discovery to therapy,” *Cancer Discovery*, vol. 6, no. 4. American Association for Cancer Research Inc., pp. 353–367, Apr. 01, 2016, doi: 10.1158/2159-8290.CD-15-0894.

[117] A. C. Chaikovsky and J. Sage, “Beyond the cell cycle: Enhancing the immune surveillance of tumors via CDK4/6 inhibition,” *Molecular Cancer Research*, vol. 16, no. 10. American Association for Cancer Research Inc., pp. 1454–1457, Oct. 01, 2018, doi: 10.1158/1541-7786.MCR-18-0201.

[118] E. B. Rankin and A. J. Giaccia, “The receptor tyrosine kinase AXL in cancer progression,” *Cancers*, vol. 8, no. 11. MDPI AG, Nov. 01, 2016, doi: 10.3390/cancers8110103.

[119] K. De Veirman *et al.*, “Receptor Tyrosine Kinase AXL: A Potential Strategy to Counter Immune Suppression and Dormancy in Multiple Myeloma,” *Blood*, vol. 134, no. Supplement_1, pp. 4335–4335, Nov. 2019, doi: 10.1182/blood-2019-126834.

[120] S. Chiba *et al.*, “Tumor-infiltrating DCs suppress nucleic acid-mediated innate immune responses through interactions between the receptor TIM-3 and the alarmin HMGB1,” *Nat. Immunol.*, vol. 13, no. 9, pp. 832–842, Sep. 2012, doi: 10.1038/ni.2376.

[121] W. W. Sung *et al.*, “IL-10 promotes tumor aggressiveness via upregulation of CIP2A transcription in lung adenocarcinoma,” *Clin. Cancer Res.*, vol. 19, no. 15, pp. 4092–4103, Aug. 2013, doi: 10.1158/1078-0432.CCR-12-3439.

[122] K. W. Moore, R. De Waal Malefyt, R. L. Coffman, and A. O’Garra, “Interleukin-10 and the interleukin-10 receptor,” *Annual Review of Immunology*, vol. 19. Annu Rev Immunol, pp. 683–765, 2001, doi: 10.1146/annurev.immunol.19.1.683.

[123] R. Lang, D. Patel, J. J. Morris, R. L. Rutschman, and P. J. Murray, “Shaping Gene Expression in Activated and Resting Primary Macrophages by IL-10,” *J. Immunol.*, vol.

169, no. 5, pp. 2253–2263, Sep. 2002, doi: 10.4049/jimmunol.169.5.2253.

[124] M. K. Levings, R. Bacchetta, U. Schulz, and M. G. Roncarolo, "The role of IL-10 and TGF-beta in the differentiation and effector function of T regulatory cells.," *International archives of allergy and immunology*, vol. 129, no. 4. Int Arch Allergy Immunol, pp. 263–276, 2002, doi: 10.1159/000067596.

[125] R. A. Maldonado and U. H. von Andrian, "How Tolerogenic Dendritic Cells Induce Regulatory T Cells," in *Advances in Immunology*, vol. 108, no. C, Academic Press Inc., 2010, pp. 111–165.

[126] C. Bilir and C. Sarisozen, "Indoleamine 2,3-dioxygenase (IDO): Only an enzyme or a checkpoint controller?," *J. Oncol. Sci.*, vol. 3, no. 2, pp. 52–56, Jul. 2017, doi: 10.1016/j.jons.2017.04.001.

[127] A. Okamoto *et al.*, "Indoleamine 2,3-dioxygenase serves as a marker of poor prognosis in gene expression profiles of serous ovarian cancer cells," *Clin. Cancer Res.*, vol. 11, no. 16, pp. 6030–6039, Aug. 2005, doi: 10.1158/1078-0432.CCR-04-2671.

[128] M. Platten, W. Wick, and B. J. Van Den Eynde, "Tryptophan catabolism in cancer: Beyond IDO and tryptophan depletion," *Cancer Research*, vol. 72, no. 21. Cancer Res, pp. 5435–5440, Nov. 01, 2012, doi: 10.1158/0008-5472.CAN-12-0569.

[129] B. Baban *et al.*, "IDO Activates Regulatory T Cells and Blocks Their Conversion into Th17-Like T Cells," *J. Immunol.*, vol. 183, no. 4, pp. 2475–2483, Aug. 2009, doi: 10.4049/jimmunol.0900986.

[130] Y. Wu, W. Chen, Z. P. Xu, and W. Gu, "PD-L1 distribution and perspective for cancer immunotherapy— blockade, knockdown, or inhibition," *Frontiers in Immunology*, vol. 10, no. AUG. Frontiers Media S.A., 2019, doi: 10.3389/fimmu.2019.02022.

[131] U. Czerwinska, "Interpretability of Machine Learning Models," 2022, pp. 275–303.

[132] R. Ramos-Medina, S. López-Tarruella, M. Del Monte-Millán, T. Massarrah, and M.

Martín, “Technical Challenges for CTC Implementation in Breast Cancer.,” *Cancers (Basel).*, vol. 13, no. 18, Sep. 2021, doi: 10.3390/cancers13184619.

[133] M. Mutebi *et al.*, “Breast cancer treatment: A phased approach to implementation,” *Cancer*, vol. 126 Suppl 10, no. S10, pp. 2365–2378, 2020, doi: 10.1002/CNCR.32910.

[134] R. M. Mann, N. Cho, and L. Moy, “Breast MRI: State of the Art,” *Radiology*, vol. 292, no. 3, pp. 520–536, 2019, doi: 10.1148/RADIOL.2019182947.

[135] N. R. Kressin *et al.*, “Women’s Understandings and Misunderstandings of Breast Density and Related Concepts: A Mixed Methods Study,” *J. Womens. Health (Larchmt).*, Feb. 2022, doi: 10.1089/JWH.2021.0343.

[136] S. Srivastava *et al.*, “Cancer overdiagnosis: a biological challenge and clinical dilemma,” *Nat. Rev. Cancer*, vol. 19, no. 6, pp. 349–358, Jun. 2019, doi: 10.1038/S41568-019-0142-8.

[137] T. Hu, J. Wolfram, and S. Srivastava, “Extracellular Vesicles in Cancer Detection: Hopes and Hypes,” *Trends in cancer*, vol. 7, no. 2, pp. 122–133, Feb. 2021, doi: 10.1016/J.TRECAN.2020.09.003.

[138] P. Llinàs-Arias *et al.*, “Epigenetic loss of the endoplasmic reticulum–associated degradation inhibitor SVIP induces cancer cell metabolic reprogramming,” *JCI Insight*, vol. 4, no. 8, Apr. 2019, doi: 10.1172/jci.insight.125888.

[139] C. Li, “Unfolded Protein Response and Crohn’s Diseases: A Molecular Mechanism of Wound Healing in the Gut,” *Gastrointest. Disord. 2021, Vol. 3, Pages 31-43*, vol. 3, no. 1, pp. 31–43, Feb. 2021, doi: 10.3390/GIDISORD3010004.

[140] H. Shiheido *et al.*, “Human T cells expressing BEND3 on their surface represent a novel subpopulation that preferentially produces IL-6 and IL-8,” *Immunity, Inflamm. Dis.*, vol. 2, no. 1, pp. 35–43, Jun. 2014, doi: 10.1002/iid3.17.

[141] F. Kurniawan *et al.*, “BEND3 safeguards pluripotency by repressing differentiation-

associated genes," *Proc. Natl. Acad. Sci.*, vol. 119, no. 9, Mar. 2022, doi: 10.1073/pnas.2107406119.

[142] K. Wang *et al.*, "MDGA2 is a novel tumour suppressor cooperating with DMAP1 in gastric cancer and is associated with disease outcome," *Gut*, vol. 65, no. 10, pp. 1619–1631, Oct. 2016, doi: 10.1136/gutjnl-2015-309276.

[143] E. D. Litwack, R. Babey, R. Buser, M. Gesemann, and D. D. . O'Leary, "Identification and characterization of two novel brain-derived immunoglobulin superfamily members with a unique structural organization," *Mol. Cell. Neurosci.*, vol. 25, no. 2, pp. 263–274, Feb. 2004, doi: 10.1016/j.mcn.2003.10.016.

[144] S. Ren *et al.*, "The expression, function, and utilization of Protamine1: a literature review," *Transl. Cancer Res.*, vol. 10, no. 11, pp. 4947–4957, Nov. 2021, doi: 10.21037/TCR-21-1582.

[145] F. Meklat *et al.*, "Identification of protamine 1 as a novel cancer-testis antigen in early chronic lymphocytic leukaemia," *Br. J. Haematol.*, vol. 144, no. 5, pp. 660–666, Mar. 2009, doi: 10.1111/J.1365-2141.2008.07502.X.

[146] Z. Chen, C. Shi, S. Gao, D. Song, and Y. Feng, "Impact of protamine I on colon cancer proliferation, invasion, migration, diagnosis and prognosis," *Biol. Chem.*, vol. 399, no. 3, pp. 265–275, Feb. 2018, doi: 10.1515/hsz-2017-0222.

[147] S. Ren *et al.*, "The expression, function, and utilization of Protamine1: a literature review," *Transl. Cancer Res.*, vol. 10, no. 11, pp. 4947–4957, Nov. 2021, doi: 10.21037/tcr-21-1582.

[148] K. L. Karlin *et al.*, "The Oncogenic STP Axis Promotes Triple-Negative Breast Cancer via Degradation of the REST Tumor Suppressor," *Cell Rep.*, vol. 9, no. 4, pp. 1318–1332, Nov. 2014, doi: 10.1016/J.CELREP.2014.10.011.

[149] G. Mondal, A. Ohashi, L. Yang, M. Rowley, and F. J. Couch, "Tex14, a Plk1-Regulated

Protein, Is Required for Kinetochore-Microtubule Attachment and Regulation of the Spindle Assembly Checkpoint," *Mol. Cell*, vol. 45, no. 5, pp. 680–695, Mar. 2012, doi: 10.1016/j.molcel.2012.01.013.

[150] L. R�nnstrand, "Signal transduction via the stem cell factor receptor/c-Kit," *Cell. Mol. Life Sci.*, vol. 61, no. 19–20, pp. 2535–2548, Oct. 2004, doi: 10.1007/s00018-004-4189-6.

[151] L. K. Ashman and R. Griffith, "Therapeutic targeting of c-KIT in cancer," *Expert Opin. Investig. Drugs*, vol. 22, no. 1, pp. 103–115, Jan. 2013, doi: 10.1517/13543784.2013.740010.

[152] T. B. Oriss, N. Krishnamoorthy, P. Ray, and A. Ray, "Dendritic cell c-kit signaling and adaptive immunity," *Curr. Opin. Allergy Clin. Immunol.*, vol. 14, no. 1, pp. 7–12, Feb. 2014, doi: 10.1097/ACI.0000000000000019.

[153] M. Janakiram, J. M. Chinai, A. Zhao, J. A. Sparano, and X. Zang, "HHLA2 and TMIGD2: new immunotherapeutic targets of the B7 and CD28 families," *Oncoimmunology*, vol. 4, no. 8, p. e1026534, Aug. 2015, doi: 10.1080/2162402X.2015.1026534.

[154] J. M. Kolos, A. M. Voll, M. Bauder, and F. Hausch, "FKBP Ligands—Where We Are and Where to Go?," *Front. Pharmacol.*, vol. 9, Dec. 2018, doi: 10.3389/fphar.2018.01425.

[155] M. F. Garrido *et al.*, "Regulation of eIF4F Translation Initiation Complex by the Peptidyl Prolyl Isomerase FKBP7 in Taxane-resistant Prostate Cancer," *Clin. Cancer Res.*, vol. 25, no. 2, pp. 710–723, Jan. 2019, doi: 10.1158/1078-0432.CCR-18-0704/87619/AM/REGULATION-OF-EIF4F-TRANSLATION-INITIATION-COMPLEX.

[156] M. Watanabe *et al.*, "MZB1 expression indicates poor prognosis in estrogen

receptor-positive breast cancer," *Oncol. Lett.*, vol. 20, no. 5, pp. 1–1, Sep. 2020, doi: 10.3892/ol.2020.12059.

[157] M. Kanda *et al.*, "Epigenetic suppression of the immunoregulator MZB1 is associated with the malignant phenotype of gastric cancer," *Int. J. Cancer*, vol. 139, no. 10, pp. 2290–2298, Nov. 2016, doi: 10.1002/IJC.30286.

[158] V. Andreani *et al.*, "Cochaperone Mzb1 is a key effector of Blimp1 in plasma cell differentiation and β1-integrin function," doi: 10.1073/pnas.1809739115.

[159] T. Kapoor, M. Corrado, E. L. Pearce, E. J. Pearce, and R. Grosschedl, "MZB1 enables efficient interferon α secretion in stimulated plasmacytoid dendritic cells," *Sci. Rep.*, vol. 10, no. 1, p. 21626, Dec. 2020, doi: 10.1038/s41598-020-78293-3.

[160] M. Pathan *et al.*, "FunRich: An open access standalone functional enrichment and interaction network analysis tool," *Proteomics*, 2015, doi: 10.1002/pmic.201400515.

[161] C. A. Gomes, T. Girão da Cruz, J. L. Andrade, N. Milhazes, F. Borges, and M. P. M. Marques, "Anticancer Activity of Phenolic Acids of Natural or Synthetic Origin: A Structure−Activity Study," *J. Med. Chem.*, vol. 46, no. 25, pp. 5395–5401, Dec. 2003, doi: 10.1021/jm030956v.

[162] S. Verma, A. Singh, and A. Mishra, "Gallic acid: Molecular rival of cancer," *Environ. Toxicol. Pharmacol.*, vol. 35, no. 3, pp. 473–485, May 2013, doi: 10.1016/j.etap.2013.02.011.

[163] L. Rong *et al.*, "Salidroside induces apoptosis and protective autophagy in human gastric cancer AGS cells through the PI3K/Akt/mTOR pathway," *Biomed. Pharmacother.*, vol. 122, p. 109726, Feb. 2020, doi: 10.1016/J.BIOPHA.2019.109726.

[164] Z. Qi *et al.*, "Salidroside inhibits the proliferation and migration of gastric cancer cells via suppression of Src-associated signaling pathway activation and heat shock protein 70 expression," *Mol. Med. Rep.*, vol. 18, no. 1, p. 147, Jul. 2018, doi:

10.3892/MMR.2018.8958.

[165] Y. H. Kong and S. P. Xu, “Salidroside prevents skin carcinogenesis induced by DMBA/TPA in a mouse model through suppression of inflammation and promotion of apoptosis,” *Oncol. Rep.*, vol. 39, no. 6, pp. 2513–2526, Jun. 2018, doi: 10.3892/OR.2018.6381.

[166] C. Li, Q. Wang, S. Shen, X. Wei, and G. Li, “Oridonin inhibits vegf-a-associated angiogenesis and epithelial-mesenchymal transition of breast cancer in vitro and in vivo,” *Oncol. Lett.*, vol. 16, no. 2, pp. 2289–2298, Aug. 2018, doi: 10.3892/OL.2018.8943/HTML.

[167] Y. Q. Liu, S. You, S. I. Tashiro, S. Onodera, and T. Ikejima, “Activation of phosphoinositide 3-kinase, protein kinase C, and extracellular signal-regulated kinase is required for oridonin-enhanced phagocytosis of apoptotic bodies in human macrophage-like U937 cells,” *J. Pharmacol. Sci.*, vol. 98, no. 4, pp. 361–371, 2005, doi: 10.1254/JPHS.FPJ05005X.

[168] P. Wang *et al.*, “Ginsenoside Rd attenuates breast cancer metastasis implicating derepressing microRNA-18a-regulated Smad2 expression,” *Sci. Reports 2016 61*, vol. 6, no. 1, pp. 1–14, Sep. 2016, doi: 10.1038/srep33709.

[169] G. M. Liu, T. C. Lu, M. L. Sun, W. Y. Jia, X. Ji, and Y. G. Luo, “Ginsenoside Rd Inhibits Glioblastoma Cell Proliferation by Up-Regulating the Expression of miR-144-5p,” *Biol. Pharm. Bull.*, vol. 43, no. 10, pp. 1534–1541, Oct. 2020, doi: 10.1248/BPB.B20-00338.

[170] E. Zhang, H. Shi, L. Yang, X. Wu, and Z. Wang, “Ginsenoside Rd regulates the Akt/mTOR/p70S6K signaling cascade and suppresses angiogenesis and breast tumor growth,” *Oncol. Rep.*, vol. 38, no. 1, pp. 359–367, Jul. 2017, doi: 10.3892/OR.2017.5652/HTML.

[171] L. Yang *et al.*, “Protopanaxadiol inhibits epithelial–mesenchymal transition of

hepatocellular carcinoma by targeting STAT3 pathway," *Cell Death Dis. 2019 109*, vol. 10, no. 9, pp. 1–13, Aug. 2019, doi: 10.1038/s41419-019-1733-8.

[172] L. Xia *et al.*, "Role of the NFκB-signaling pathway in cancer," *Onco. Targets. Ther.*, vol. 11, p. 2063, Apr. 2018, doi: 10.2147/OTT.S161109.

[173] H. H. Park *et al.*, "Suppressive Effects of Britanin, a Sesquiterpene Compound Isolated from Inulae Flos, on Mast Cell-Mediated Inflammatory Responses," *http://dx.doi.org/10.1142/S0192415X14500591*, vol. 42, no. 4, pp. 935–947, Jul. 2014, doi: 10.1142/S0192415X14500591.

[174] C. Bailly, "Anticancer Targets and Signaling Pathways Activated by Britannin and Related Pseudoguaianolide Sesquiterpene Lactones," *Biomedicines*, vol. 9, no. 10, Oct. 2021, doi: 10.3390/BIOMEDICINES9101325.

[175] Y. Q. Cui, Y. J. Liu, and F. Zhang, "The suppressive effects of Britannin (Bri) on human liver cancer through inducing apoptosis and autophagy via AMPK activation regulated by ROS," *Biochem. Biophys. Res. Commun.*, vol. 497, no. 3, pp. 916–923, Mar. 2018, doi: 10.1016/J.BBRC.2017.12.144.

[176] L. Y. Zhao *et al.*, "The Prognostic Value of aspartate aminotransferase to lymphocyte ratio and systemic immune-inflammation index for Overall Survival of Hepatocellular Carcinoma Patients Treated with palliative Treatments," *J. Cancer*, vol. 10, no. 10, p. 2299, 2019, doi: 10.7150/JCA.30663.

[177] L. L. Xiong *et al.*, "Anti-colorectal cancer effects of scutellarin revealed by genomic and proteomic analysis," *Chinese Med. (United Kingdom)*, vol. 15, no. 1, pp. 1–15, Mar. 2020, doi: 10.1186/S13020-020-00307-Z/FIGURES/7.

[178] Y. Yuan, P. Rangarajan, E. M. Kan, Y. Wu, C. Wu, and E. A. Ling, "Scutellarin regulates the Notch pathway and affects the migration and morphological transformation of activated microglia in experimentally induced cerebral ischemia in rats and in

activated BV-2 microglia," *J. Neuroinflammation*, vol. 12, no. 1, pp. 1–21, Jan. 2015, doi: 10.1186/S12974-014-0226-Z/FIGURES/12.

[179] P. Luo, Z. H. Tan, Z. F. Zhang, H. Zhang, X. F. Liu, and Z. J. Mo, "Scutellarin isolated from Erigeron multiradiatus inhibits high glucose-mediated vascular inflammation," *Yakugaku Zasshi*, vol. 128, no. 9, pp. 1293–1299, Sep. 2008, doi: 10.1248/YAKUSHI.128.1293.

[180] Y. J. Chen *et al.*, "Scutellarin Reduces Cerebral Ischemia Reperfusion Injury Involving in Vascular Endothelium Protection and PKG Signal," *Nat. Products Bioprospect.*, vol. 11, no. 6, p. 659, Dec. 2021, doi: 10.1007/S13659-021-00322-Z.

[181] Z. Zhong, J. Han, J. Zhang, Q. Xiao, J. Hu, and L. Chen, "Pharmacological activities, mechanisms of action, and safety of salidroside in the central nervous system," *Drug Des. Devel. Ther.*, vol. 12, p. 1479, May 2018, doi: 10.2147/DDDT.S160776.

[182] J. Zhu, X. Wan, Y. Zhu, X. Ma, Y. Zheng, and T. Zhang, "Evaluation of salidroside in vitro and in vivo genotoxicity," *Drug Chem. Toxicol.*, vol. 33, no. 2, pp. 220–226, Apr. 2010, doi: 10.3109/01480540903373654.

[183] H. Zhang, W. S. Shen, C. H. Gao, L. C. Deng, and D. Shen, "Protective effects of salidroside on epirubicin-induced early left ventricular regional systolic dysfunction in patients with breast cancer," *Drugs R. D.*, vol. 12, no. 2, pp. 101–106, 2012, doi: 10.2165/11632530-000000000-00000.

[184] X. Zeng *et al.*, "Pharmacokinetics and safety of ginsenoside Rd following a single or multiple intravenous dose in healthy Chinese volunteers," *J. Clin. Pharmacol.*, vol. 50, no. 3, pp. 285–292, Mar. 2010, doi: 10.1177/0091270009344334.

[185] G. Zhang *et al.*, "Ginsenoside Rd Is Efficacious Against Acute Ischemic Stroke by Suppressing Microglial Proteasome-Mediated Inflammation," *Mol. Neurobiol.*, vol. 53, no. 4, pp. 2529–2540, May 2016, doi: 10.1007/S12035-015-9261-8/FIGURES/9.

[186] L. Kazantseva, J. Becerra, and L. Santos-Ruiz, “Oridonin enhances antitumor effects of doxorubicin in human osteosarcoma cells,” *Pharmacol. Reports*, vol. 74, no. 1, pp. 248–256, Feb. 2022, doi: 10.1007/S43440-021-00324-1/FIGURES/6.

[187] G. B. Zhou *et al.*, “Oridonin, a diterpenoid extracted from medicinal herbs, targets AML1-ETO fusion protein and shows potent antitumor activity with low adverse effects on t(8;21) leukemia in vitro and in vivo,” *Blood*, vol. 109, no. 8, pp. 3441–3450, Apr. 2007, doi: 10.1182/BLOOD-2006-06-032250.

[188] K. Li, Y. Zhou, Y. Chen, L. Zhou, and J. Liang, “A novel natural product, britanin, inhibits tumor growth of pancreatic cancer by suppressing nuclear factor-κB activation,” *Cancer Chemother. Pharmacol.*, vol. 85, no. 4, pp. 699–709, Apr. 2020, doi: 10.1007/S00280-020-04052-W.

[189] L. Wang and Q. Ma, “Clinical benefits and pharmacology of scutellarin: A comprehensive review,” *Pharmacol. Ther.*, vol. 190, pp. 105–127, Oct. 2018, doi: 10.1016/J.PHARMTHERA.2018.05.006.

LIST OF PUBLICATIONS

JOURNAL PUBLICATION

- ✓ Sunil Kumar, Asmita Das **A cocktail of natural compounds holds promise for new immunotherapeutic potential in head and neck cancer.** Chinese Journal of Integrative Medicine September 2022. DOI: 10.1007/s11655-023-3694-0
- ✓ Sunil Kumar, Asmita Das **Peripheral Blood Mononuclear Cell derived Biomarker detection using eXplainable Artificial Intelligence (XAI) provides better diagnosis of Breast Cancer.** Computational Biology and Chemistry. DOI: 10.1016/j.compbiolchem.2023.107867
- ✓ Saksham Garg, **Sunil Kumar**, Ashutosh Anand, Tarunya Menon, Nikita Sharma, Japneet Singh, Siddharth Chawla, Asmita Das **Plant-derived natural compounds aiding SOCS1 mediated JAK1 inhibition, a novel mechanism of combinatorial cancer chemotherapy.** Vegetos journal February 2022 DOI: https://doi.org/10.1007/s42535-021-00329-4.
- ✓ Sunil Kumar, Asmita Das **Elucidation of natural compounds Gallic acid and Shikonin for the treatment of HNSC cancer by targeting immune suppressor and tumour progressor genes.** Vegetos journal March 2022 DOI: https://doi.org/10.1007/s42535-022-00363-w.

CONFERENCES AND PRESENTATIONS

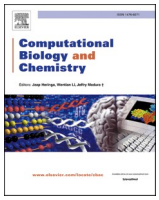
- ✓ **A combinatorial Therapy of Natural Compounds Targeting Tumor Progression and Immune Suppression for Breast Cancer** accepted for publication in international conference on innovations in Biotechnology and Life Science (ICIBLS 2020).
- ✓ **A Natural Compounds Cocktail for Breast Cancer** presented on 35th conference on Preventive Oncology and Diagnostic cancer (April 2021).

Computational Biology and Chemistry 104 (2023) 107867

Contents lists available at ScienceDirect

Computational Biology and Chemistry

journal homepage: www.elsevier.com/locate/cbac



Peripheral blood mononuclear cell derived biomarker detection using eXplainable Artificial Intelligence (XAI) provides better diagnosis of breast cancer

Sunil Kumar, Asmita Das *

Department of Biotechnology, Delhi Technological University, India

ARTICLE INFO

Keywords:
Breast cancer
Peripheral blood mononuclear cells
XGBoost
EXplainable Artificial Intelligence
Biomarker

ABSTRACT

The incidence and mortality rate of breast cancer increases yearly by an average of 1.44 % and 0.23 %, respectively. Till 2021, there were 7.8 million women who had been diagnosed with breast cancer within 5 years. Biopsies of tumors are often expensive and invasive and raise the risk of serious complications like infection, excessive bleeding, and puncture damage to nearby tissues and organs. Early detection biomarkers are often variably expressed in different patients and may even be below the detection level at an early stage. Hence PBMC that shows alteration in gene profile as a result of interaction with tumor antigens may serve as a better early detection biomarker. Also, such alterations in immune gene profile in PBMCs are more prone to detection despite variability in different breast cancer mutants.This study aimed to identify potential diagnostic biomarkers for breast cancer using eXplainable Artificial Intelligence (XAI) on XGBoost machine learning (ML) models trained on a binary classification dataset containing the expression data of PBMCs from 252 breast cancer patients and 194 healthy women.After effectively adding SHAP values further into the XGBoost model, ten important genes related to breast cancer development were discovered to be effective potential biomarkers. Our studies showed that SVIP, BEND3, MDGA2, LEF1-AS1, PRM1, TEX14, MZB1, TMIGD2, KIT, and FKBP7 are key genes that impact model prediction. These genes may serve as early, non-invasive diagnostic and prognostic biomarkers for breast cancer patients.

1. Introduction

The incidence and mortality rate of breast cancer increases yearly by 1.44 % and 0.23 %, respectively (Lima et al., 2021). It is the most common cancer in women worldwide (Sung et al., 2021). Breast cancer is a disorder in which the breast cells proliferate uncontrollably. Based on which breast cells become cancerous, we can classify the different categories of breast cancer that may be found in the breast (Feng et al., 2018). Breast cancer affects women of all ages after puberty; however, the incidence percentage rises with age (Benz, 2008). Till 2021, there were 7.8 million women who had been diagnosed with breast cancer within 5 years (Siegel et al., 2022). Breast cancer is responsible for more DALYs (disability-adjusted life years) lost in women worldwide than any other kind of cancer (Kocarnik et al., 2022).

Even though tumor sampling is frequently used to identify biomarkers, collecting tissue is difficult because of restricted accessibility, many lesions and heterogeneity of the biopsy site, and patient conditions (Bedard et al., 2013). Biopsies of tumors are often expensive, invasive, and time-consuming, and they raise the risk of serious complications (Tadimety et al., 2018). Most screening systems cannot detect and identify cancers until they have reached a particular stage of development (Yang et al., 2022). Breast cancer, for example, may have been present for many years before it is discovered through palpation or mammography, and it has the potential to spread to other organs (Akram et al., 2017). There is a pressing need to identify cancer at its earliest stages, particularly before the onset of clinical signs and symptoms. Early breast cancer detection is essential since it provides a more significant number of treatment choices, higher survival, and enhanced quality of life. While there is no foolproof way to avoid breast cancer, early diagnosis gives the most significant opportunity for successful treatment. Early detection and modern treatment are key to avoiding breast cancer fatalities. Early-stage breast cancer is simpler to treat. Regular screenings are the best method to detect breast cancer early (Jatoi and Pinsky, 2021).

* Correspondence to: Delhi Technological University, Main Bawana Road, Delhi 110042, India.
E-mail addresses: asmita1710@gmail.com, asmitadas1710@dce.ac.in (A. Das).

https://doi.org/10.1016/j.compbiolchem.2023.107867
Received 27 July 2022; Received in revised form 22 March 2023; Accepted 1 April 2023
Available online 3 April 2023

Studies of biomarkers from blood, nipple aspirate fluid, perspiration, urine, tears, or breath may diagnose breast cancer early and in a non-invasive manner (Li et al., 2020). A simple blood or breath test may soon be able to identify breast cancer early (Kure et al., 2021). Recent studies imply peripheral blood analyses might provide prognosis and treatment responsiveness (Nixon et al., 2019). Cancer detection using peripheral blood is more straightforward and less invasive (Hou et al., 2020). As a result, generating clinically useful biomarkers requires the study of readily available peripheral blood (Marrugo-Ramírez et al., 2018). The immune system relies on these PMBCs to combat infection and adapts to new threats. Oncogenic cells interact with normal stromal cells and the host immunological defense system to form tumors and prevent apoptosis (Lucchetti et al., 2020). The tumor's ability to evade the immune system also plays a significant role. Immune suppression in the tumor microenvironment by CD4+, CD25+, and FoxP3+ cells, regulatory T cells (Tregs), and other inhibitory peripheral blood mononuclear cells is the primary mechanism of tumor immune evasion (Mattei et al., 2020). Because of this, gene expression profiling of peripheral blood cells has the potential to identify early cancers (Wen et al., 2021). Michael E. Burczynski et al. demonstrated that circulating monocytes of peripheral blood may be utilized as a surrogate monitor for difficult-to-biopsy tissues and/or as an extremely sensitive monitor to check for changes in the physiological condition of the organism (Mattei et al., 2020). Sharma et al. showed that PBMCs might be utilized to build gene expression assays for early diagnosis of breast cancer based on the properties of these cells (Sharma et al., 2005). The process by which malignant development induces distinctive alterations in the blood biochemical environment justifies the use of the PBMC transcriptome gene as a monitor for malignant solid tumors (Čelešnik and Potočnik, 2022). Tumor cells interact with immune cells and change their expression profiling of genes and can escape the immune system of the host easily (Crispen and Kusmartsev, 2020). The transcriptome gene expression of PBMCs may be used as a tumor screening marker since it is conveniently retrieved. Clinical pharmacogenomics might benefit from the use of PBMCs as predictive biomarkers because of the ease with which they can be obtained (Veal et al., 2019).

1.1. Machine learning and XAI

In previous studies, many biomarkers have been identified in the search for genes with more substantial predictive value for PBMCs of Breast tumors via machine learning algorithms (Kothari et al., 2020; Meena and Hasija, 2022). Researchers are presently using AI-based machine learning (ML) approaches to study the genetic diversity of cancers, which may be utilized to enhance diagnostic accuracy, the creation of valuable biomarkers, and the effectiveness of cancer therapeutics (Dlamini et al., 2020). AI is the capacity of a robot to imitate human behavior, which is advantageous when dealing with vast volumes of data. Using AI, robots may learn from their mistakes and successes without being explicitly programmed (Allen, 2022). ML models are learning and improving modeling techniques. These models facilitate the identification of critical components and their interactions (Bender and Cortes-Ciriano, 2021). From a mostly theoretical to an actual application-oriented stage, AI has progressed over the last several years (Bohr and Memarzadeh, 2020). With the rise of AI, there are great hopes for its usage in several fields, especially in cancer research, where ML has already been used to analyze survival and forecast models for pancreatic, advanced nasopharyngeal carcinomas, breast, and other malignancies (Shaheen, 2021). Even though AI algorithms, particularly ML algorithms, appear to be effective in terms of outcomes and predictions, they are afflicted by opacity, which makes it difficult to gain insight into their essential operating processes, which exacerbates the dilemma because putting critical decisions to a system that is incapable of self-explanation carries serious risks (von Eschenbach, 2021). In complex multi-factorial diseases like cancer,even the most powerful learning methods suffer from the fact that, on the one hand, it is difficult to explain the genesis of a result, and on the other hand, they lack robustness. Even the smallest perturbations in the input data can dramatically affect the output, leading to completely different results (Holzinger et al., 2022a). Biological datasets often suffer from high variances as a result of experimental limitations, thus resulting in poor data quality. Biological data acquisition sometimes has interdependent experiments that do not result in independent and identically distributed data sets, particularly in multi-factorial diseases like cancer. Explainability and robustness promote reliability and trust in the results and ensure that humans remain in control (Holzinger, 2021). A paradigm change toward more transparent and intelligible AI is suggested by the eXplainable Artificial Intelligence (XAI) project. Its goal is to develop a set of tactics that will provide better models that can be explained while maintaining high performance (Barredo Arrieta et al., 2020). Legal complications and data protection issues, especially in health care data, suffer from inaccessible black box approaches duly made accessible by eXplainable Artificial Intelligence (XAI) tools (Holzinger et al., 2022b). Shap is a framework for explainable artificial intelligence built from the Shapley values of game theory;one of the benefits is that it allows modeling methods utilizing libraries such as SciKit-Learn, PySpark, TensorFlow, Keras, and PyTorch, among others. The fundamental issue with these commonly used libraries for data modeling is that model outputs are not readily explicable. Using SHAP, we can make the outputs of machine learning models more comprehensible to individuals with fewer machine learning skills. With this capability, we can also utilize SHAP to visualize data (Holzinger et al., 2022b).

Researchers in the XAI field are attempting to enhance their algorithms to make the outputs of AI systems more understandable for humans (Linardatos et al., 2021). There has recently been renewed interest in the notion of XAI in academia and the field of applied artificial intelligence.Models that can justify their output are called explainable or explainability. To add, explainability refers to the model's output is accurately and completely represented (Czerwinska, 2022). Local and global explainability may be categorized according to the context in which the model is used to make a particular choice (Lundberg et al., 2020). Tools that help people understand the behavior of black-box models are becoming more critical because of their capacity to explain their behavior (Rai, 2020). Some XAI frameworks include SHAP, LIME, ELI5, AIX360, and Skaters, with the first two being the most popular and compatible with any deep learning or machine learning model (Linardatos et al., 2021). Currently,Explainable Artificial Intelligence resolves many problems in the diagnosis of cancers (Zhang et al., 2022).

SHAPcanincrease the reliability of the ML model by evaluating each element used for prediction purposes in ML. After the publication of the solution paradigm for the examination in cooperative game theory by Lloyd Shapley in 1951, SHAP was developed (Shapley, 1953). Standard Shapley values for simple model examination are used in an accessible way by SHAP to relate optimum credit allocation with local explanations (Nohara et al.). It's easy to reverse-engineer the outcomes of any prediction algorithm using SHAP, which is a fantastic tool for current ML (Wieland et al., 2021). For more complicated models like gradient boosting, SHAP is often used to better comprehend the model's decisions and verify that they are accurate and faithful. The Shapley values of game theory are the ancestors of SHAP values, which uses the word game to symbolize the prediction model's result and the word players to indicate the model's characteristics. A player's performance may be quantified using Shapley values, which are also known as SHAP values. SHAP values measure each feature's local ability to influence the prediction model (Rodríguez-Pérez and Bajorath, 2020). KernelSHAP and TreeSHAP are two methods for estimating Shapley values. KernelSHAP is used for local surrogate models to explain predictions made by black-box machine learning models, while TreeSHAP is used to explain sophisticated models based on trees (Aas et al., 2021).

In this study, we were trying to identify PBMCs derived biomarkers of Breast cancer based on their gene expression with the help of XAI so that non-invasive and early screening of breast cancer can be achieved.

2. Methodology

2.1. Data retrieval

The datasets for peripheral blood cells from breast tumor patients and normal samples were obtained from NCBI-GEO Database. Two datasets were identified with suitable numbers of samples and matching queries. GSE27562 contains 162 samples. Of them, 31 are from normal women, 57 are from malignant BC patients, 37 are from benign BC patients,and 37 are from patients of other cancers termed ectopic samples. GSE47862 contains 321 samples. Out of them, 52 are from BC patients who had no family history of BC, 43 are from normal women who had no family history either, 106 are from breast cancer patients with a family history, and 120 are from normal women who had a family history of BC.

2.2. Data pre-processing

GSE27562 and GSE47862 GEO datasets were integrated to construct the final dataset. The quality of the dataset must be verified, so for this purpose, batch normalization of the dataset has been done, which was achieved by the gene standardization method, a location-scale method. Gene-wise standardization modifies the values of all genes such that their means equal zero and standard deviations (SDs) equal one. This is performed by removing the mean from each gene's sample data and dividing the resulting value by its standard deviation. Batch normalized expression data was further quantile normalized to remove additional biases from the obtained expression data.Quantile normalization substitutes each attribute (row) in the data with the mean of all attributes across all samples in the same order. The following procedure was employed to normalize a raw high-throughput data collection including multiple samples: Sort the attribute values included inside each sample. (2) Calculate the mean of each attribute's rows. Replace the raw characteristic with its average value. (4) Rearrange all altered values such that they are in the same order as before they were updated.

2.3. Machine learning models implementation

The training and testing sets were made from the dataset randomly in a ratio of 80–20. ML techniques such as SVMs, KNNs, etc., have recently gained more popularity in healthcare fields such as gene expression analysis, drug discovery, omics data analysis, imaging, etc., it was tempting to apply such ML techniques to our dataset and observe the intriguing outcomes.Because of its huge popularity, we have used the XGBoost ML classifier on our training datasets to generate prediction models, and the testing sets were then used to evaluate the performance of the prediction models. All the XGBoost ML models were validated based on their confusion matrix and the accuracy generated using the testing dataset. The XGBoost is a machine learning classifier that is based on decision trees known to boost the performance of the ML model and has been frequently reported to have beaten other ML algorithms, including random forest, decision trees, regression, etc. Despite having compatibility with several computer languages, XGBoost frameworks are most popular for Python and the associated scikit-learn framework.

2.4. Explain the ability of the trained model

The trained XGBoost model was analyzed by the Explainable artificial intelligence (XAI) analysis with the help of the SHAP library. As XAI is concerned with the decision-making process, it helps in the identification of the features significantly impacting the model's prediction.The implementation of XAI analysis will help in identifying the significant genes, and thereafter further identification/classification of the phenotype/condition, such as test or control, will be done by trained models. A local summary plot was formed to exhibit the values indicating the features contributing to the decision confidence with the help of SHAP values. SHAP stands for Shapley Additive exPlanations. The global feature relevance from training data was shown by the SHAP summary plot, and the top 10 genes (top-ranked average SHAP value) features were used to train new XGBoost models again, and the significance of 10 selected genes was validated by comparing new XGBoost models to those previously trained on 16,000 genes.

3. Results

The array data for PBMCs of breast cancer (BC) patients obtained from the GEO database was retrieved in normalized and calibrated form, which can be found in Table 1. Search terms like Breast Cancer and PBMCs were used to obtain the datasets. After retrieval, the datasets were merged based on the attribute common gene symbols, About sixteen thousand such common genes were incorporated along with their values as features.

3.1. Data pre-processing

GSE27562 and GSE47862 GEO datasets were integrated to construct the final dataset and finally, 16,000 common genes were identified in both datasets Their expression profiles were merged and the batch was normalized using the gene standardization method, a location-scale method for batch normalization of data integrated from different datasets. Both datasets are already log-transformed; therefore, quantile normalization was applied to the batch-normalized data to remove further biases from the obtained expression data. Different samples were classified into a binary classification problem: test vs control. The test was the samples of BC patients, and the control was the samples from healthy women.

The normalized expression density plot was created with the help of quantile normalization, shown in Fig. 1.

3.2. XGBoost implementation results

The dataset was randomly divided into a training set (80 %) and a test set (20 %) to apply machine learning. With the help of the scikit-learning library, the XGBoost algorithm was applied.The training dataset trained XGBoost Model for further classification on our test vs control dataset. The performance of the model was then checked using the testing sets. The confusion matrix was implemented to check the model's accuracy using the training sets, and the model's accuracy was obtained using the test set thereafter. There were 28 true positive events, 2 false positive events, 1 false negative event, and 59 true negative events found in the confusion matrix. The accuracy here implies a

Table 1
The table demonstrates the Microarray dataset obtained from the GEO database along with the familial description and the classification of samples that have further been used for ML analysis.

GEO Accession Number	Total Sample	Sample class in the dataset	Sample Size	Classification of samples for ML
GSE27562	162	Malignant	57(test)	Test – 252
		Benign	37(test)	
		Ectopic	37 (eliminated)	
		Normal	31(control)	
GSE47862	321	Breast cancer without a family history	52(test)	Control- 194
		Normal without a family history	43(control)	
		Breast Cancer with a family history	106(test)	
		Normal with family history	120(control)	

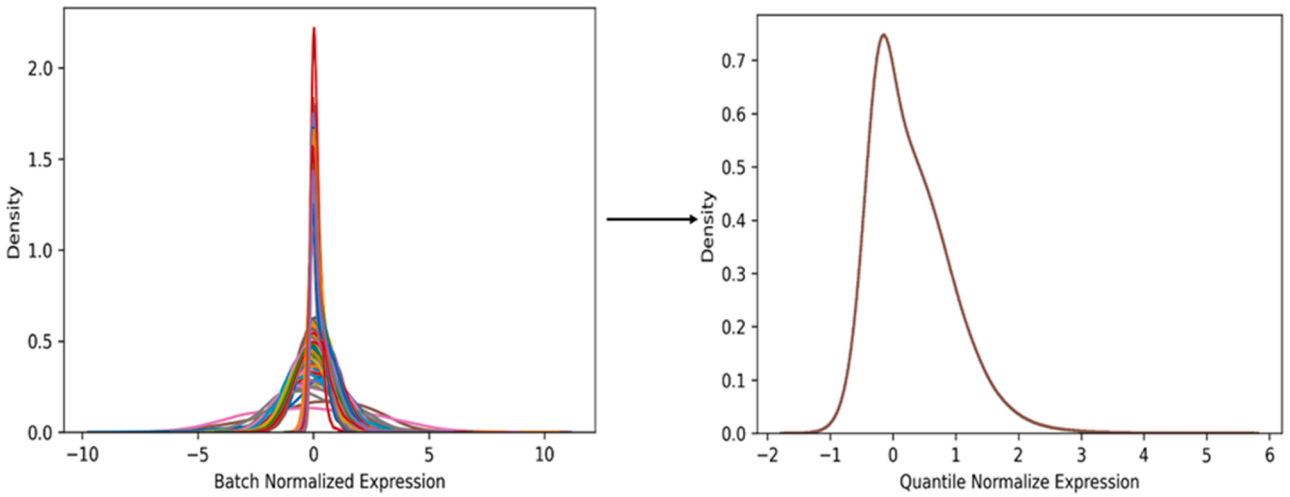


Fig. 1. The figure demonstrates Batch normalized expression data distribution curves followed by quantile normalized expression data curves.

prediction of the model's performance, which stands for the percentage of correct predictions the model has made. For binary classification, the accuracy was calculated in the form of positives and negatives. as described by the following equation:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

A predictive accuracy of 96.67 % for the test vs control dataset was obtained using the XGBoost ML classifier, which implies that the model did well in distinguishing the features of the test and control.

3.3. XGBoost models examination with XAI

With the help of python's SHAPpackage, the XAI analysis was implemented on the XGBoost trained model, which is all about the model's decision-making & identifies the features that influence the model's prediction confidence to a great extent, and this analysis helped in finding out the valuable genes from which trained model can separate the corresponding dataset into test (PBMCs of Cancer Patients) and control (PBMCs of Healthy Women). The corresponding SHAP values representing the respective share of a particular attribute to the accuracy of the model's decision were displayed with the help of a local summary plot.

The global significance for every gene was found as the average absolute value of that particular over all of the given samples, with the help of a global feature importance plot that was obtained by the bar plot function where SHAP values were passed as an array. The inference obtained from this global feature importance plot points out the most significant genes in descending order, suggesting the more contribution of genes on the top towards the model's prediction. The bar plot sorts out the most important genes placed on the top. The gene of utmost significance in our machine learning model was STIV, exhibiting a high predictive value.

With the implementation of SHAP values on the trained models, genes of the highest significance were obtained from the bar plot.The most significant genes in immune cells involved in the progression of Breast Cancer were identified by SHAP listed in Table 2.

The following SHAP summary plot identified that SVIP is the most significant gene in the data set and highly impacted the model's prediction.

3.4. Examination of XAI output

The authenticity of results was checked by applying ML classifier XGBoost on selected genes on the bases of their significance in model prediction. The top ten genes selected by their corresponding significant SHAP values were used to examine the reliability of the results by the ML classifier, namely XGBoost, highlighted in Table 2. The model's accuracy was 94.44 % when trained with the top ten significant genes. Table 3 depicts the accuracy of both the gene sets, i.e., before and after implementing XAI on binary datasets, showing the prediction model's performance in terms of accuracy. The confusion matrix of the model shows that there were 37 true positive events, 48 true negative events,3 false positive, and 2 false negative events in the model's prediction. The confusion matrix of datasets with 16000 genes and the top 10 genes are shown in Fig. 2.

The SHAP plot of the top 10 significant genes, shown in Fig. 3, indicates the contribution of the gene to the model's prediction in descending order, which shows SVIP had the highest impact, followed by BEND3,MDGA2, LEF1-AS1,PRM1,TEX14,MZB1,TMIGD2,KIT,FKBP7 respectively.

Furthermore, to visualize the predictor's positive & negative associations with the respective genes, the SHAP summary plot was also made,as shown in Fig. 4. The inferences obtained from the SHAP summary plots are as follows: -The ranking of genes (vertically) in descending order signifies their attribute importance. The horizontal line depicts the association of the effect of an attribute on the extent of prediction. The color signifies the impact of a particular gene, maximum significance (in red color) or minimum significance (in blue color). The

Table 2
The table shows a list of genes contributing to the model prediction obtained from the merged datasets.

Datasets	Significant Genes
Breast cancer patient's PBMCs vs. Healthy person PBMCs	**SVIP,BEND3,MDGA2,LEF1-AS1,PRM1, TEX14,MZB1,TMIGD2,KIT,FKBP7**,ZNF563, TC2N,LYZ,MAP3K19,GYPE,DSP,ID2,POLR2K, GFPT1,STAM,IRF8,MRPL57,CRYM,SERPIND1, DSG3,APCS,CDH16,HOXD10,TM4SF1,PMEL, COL4A6,MEGF6,HMGB3P1,LRRC20,ZNF668, CLIC3,LRP1B,STK32B,SLC16A10,TSHZ2, PDZRN4,UIMC1,SLC26A6,PIPOX,TMA7, POMGNT2,C19ORF44,CYYR1,DPP10-AS1,

Table 3
The table shows a comparison of accuracy between the prediction model for the 16000 genes set and 10 selected gene sets.

Datasets	Accuracy of 16000 genes set	Accuracy of 10 genes set
Breast cancer patient's PBMCs vs Healthy person PBMCs	96.67 %	94.44 %

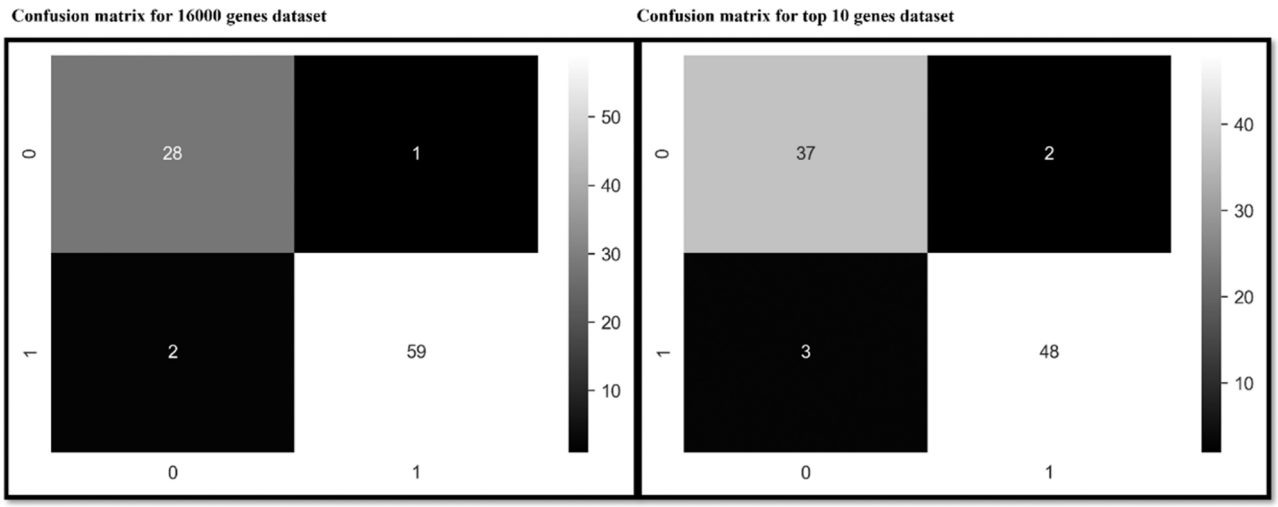


Fig. 2. The figure shows a comparison of the confusion matrix for PBMCs obtained from Breast cancer patients vs the Healthy person dataset for all 16,000 genes and the top 10 genes. True positive, False positive, False negative,and True negative instances are indicated by a grey box, Black box, Black box, and white box respectively.

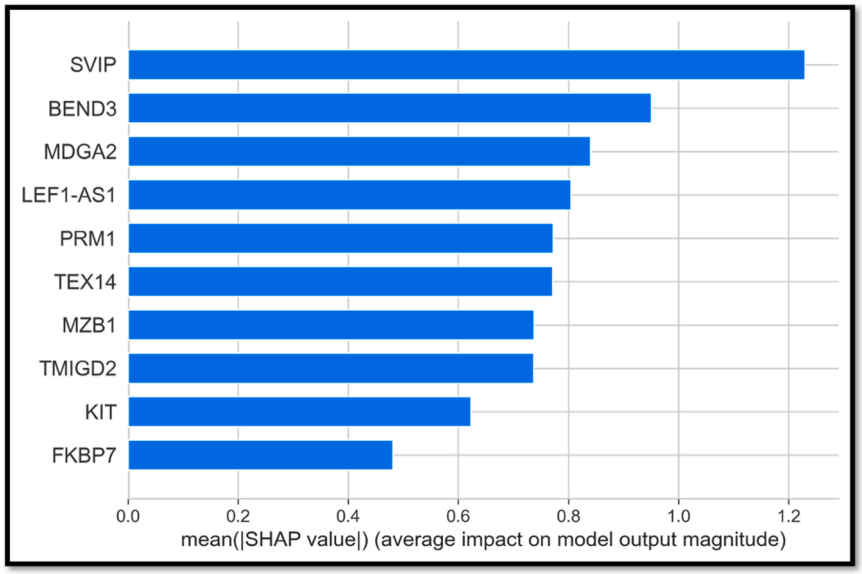


Fig. 3. SHAP Bar plot illustrates the most significant genes and their SHAP values. The x-axis represents the average/mean absolute value for each gene across all the available data, while the y-axis represents the top 10 genes.

strong positive impact of SVIP on the SHAP Summary plot indicates the correlativity of the individual gene, where the X-axis signifies the positive impact and the red color signifies the level (high in this case). Similarly, the inverse connection of BEND3 to the target variable can be ruled out.

3.5. Shortlisted genes statistical significance

The iDEP tool was used to identify key genes differentially expressed in PBMCs during Breast cancer development. P-value ≤ 0.05 was the criteria for identification as statistically significant SVIP, MDGA2, TMIGD2, LEF1-AS1andTEX14 were found to be downregulated while BEND3, FKBP7, MZB1, PRM1, and KIT were found to be upregulated in PBMCs of Breast cancer patients (Table 4).

4. Discussion

Despite the fact that tissue-specific biomarkers, such as aberrant cells, alterations in tumor gene expression, and other malignant abnormalities, may be accurate cancer biomarkers, they have several limitations (Czerwinska, 2022). It is challenging to employ tissue-specific biomarkers to assess therapy response in real-time due to the invasive nature of biopsy collection (Ramos-Medina et al., 2021). TILs may be a valuable prognostic sign for identifying individuals who are most likely to respond to therapy. Biopsies and mammography, which are presently used to identify breast cancer, are painful, costly, and only effective in situations of advanced disease (Mutebi et al., 2020). Mammography may not identify breast cancer immediately since its sensitivity is dependent on tumor size, ranging from 26 % at 5 mm to 91 % at 10 mm (Mann et al., 2019). Breasts with thicker tissue hinder mammography's ability to detect breast cancer (Kressin et al., 2022). A high level of sensitivity and specificity is required for early cancer

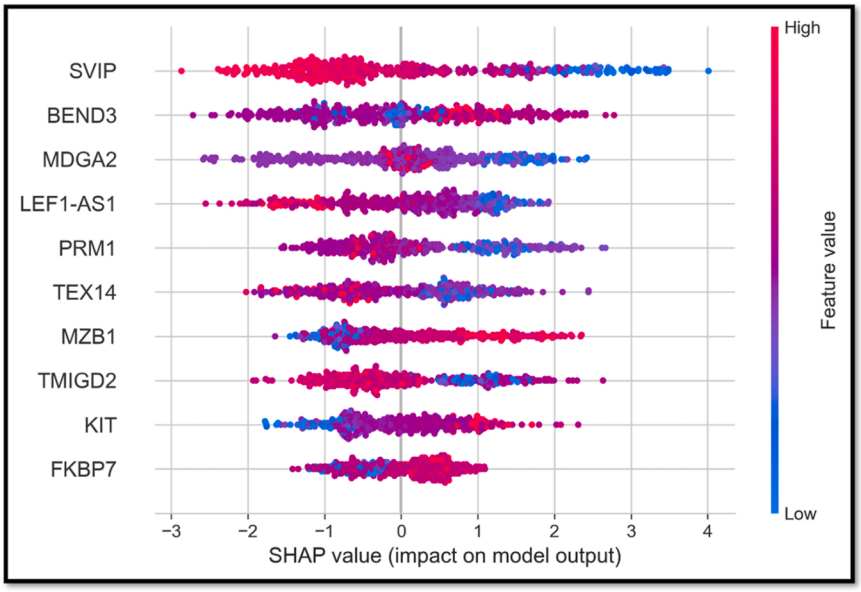


Fig. 4. The figure illustrates the SHAP Summary diagram, which shows the highly significant genes and their influence on the dataset. On the y-axis, selected genes are sorted in descending order, based on the significance of their characteristic. On the other hand, the x-axis shows the influence of genes on the prediction, illustrating the gene's impact on the model output. The color indicates the influence of a particular gene on a prediction, whether it is statistically significant (in red) or low significance values (in blue).

Table 4
The table demonstrates the P-value and log FC value of the top 10 genes.

Genes	P-value	logFC
Dataset: Breast cancer patient's PBMCs vs Healthy person PBMCs		
SVIP	4.82E-14	-1.28E-01
BEND3	1.56E-02	8.31E-01
MDGA2	1.56E-04	-1.29E-01
FKBP7	3.25E-02	7.02E-02
TMIGD2	1.23E-04	-7.31E-02
LEF1-AS1	1.08E-04	-2.14E-01
MZB1	1.27E-03	1.77E-01
TEX14	4.12E-02	-6.66E-01
PRM1	5.63E-03	4.49E-01
KIT	3.67E-04	1.21E-01

detection to increase patient survival rates.

When searching for symptoms of cancer, intrusive tissue collection may be dangerous and may not be the best method for old or delicate individuals (Srivastava et al., 2019). Less invasive and more universally accessible techniques of acquiring biological samples, such as blood collection, may be more acceptable to patients, which might result in a quicker diagnosis (Hu et al., 2021). A high level of sensitivity and specificity is required for early cancer detection to increase patient survival rates.

PBMCs mediate the immunological response of the host to tumor cells; hence, peripheral blood profiling may be used to assess the host's reaction to cancer and offers the possibility of minimally invasive early cancer detection (even before the beginning of clinical symptoms). It can anticipate the prognosis and developmental trajectory of tumors and the clinical outcome. Multiple studies have attempted to identify alterations in PBMC gene expression within breast cancer to categorize subtypes. In individuals with breast cancer, the PBMC transcriptomes correlate poorly with conventional subtypes and are diverse. Using RNA sequencing, Ming et al. determined that ER, PR, and HER2 were not associated with transcriptome-wide PBMC gene expression patterns. The expression of PBMC genes indicates that blood mononuclear cells are immunologically reactive to tumor cells. Therefore, this is not entirely surprising. Similar results were seen for lung cancer patients, who showed high diversity in peripheral blood leucocyte transcriptomes regardless of histological type, with no discernible impact on the peripheral immune system. Therefore, we included PBMC samples from different types of breast cancer patients in our study concerning the stage of cancer,the patient's history of cancer, and different subtypes of breast cancer. 252 breast cancer samples were included in this study. Of them, 37 were associated with benign stage of cancer, 57 were associated with malignant stage of cancer, 106 were from the patient with a family history of breast cancer, and 52 were from patients with no breast cancer history. 194 normal PBMC samples were included in this study for comparison with the tumor PBMC samples. Healthy individuals with a family history and without a family history were also included in the healthy control category.

Machine learning algorithm XGBoost was applied to the binary classified dataset for classification, which is followed by the XAI to identify significant genes based on their contribution to the model's prediction. Ten genes were identified in PBMCs of BC patients, which contribute the highest to the models' prediction. These genes were further analyzed for their biological significance and their involvement in different biological processes and their regulation.

4.1. *Biological significance of the genes*

Each of the top 10 genes was further analyzed for their involvement in biological processes and their regulation to ascertain their impact on cancer progression.

SVIP has tumor suppressor properties, and its restoration is linked to enhanced ER stress and growth inhibition (Llinàs-Arias et al., 2019). According to proteomic and metabolomic studies, mitochondria enzymes and oxidative respiration activity are diminished in tumor cells with SVIP epigenetic deletion (Li, 2021).

BEND3^{+} T cells generated more significant quantities of IL-6 and IL-8 than BEND3^{-} T cells. Multiple inflammatory cells, including neutrophils, basophils, and T lymphocytes, are attracted by IL-8. Activation of BEND3^{+} T cells, which may produce IL-6 and IL-8 in response to TCR/CD3 stimulation, may be essential for the significant and rapid initiation and development of inflammatory responses at the onset of inflammation. BEND3^{+} T cell dysregulation may result in chronic inflammation (Shiheido et al., 2014). BEND3 attaches to the promoters of differentiation-associated factors and important cell cycle regulators, such as CDKN1A, which encodes p21 and represses differentiation-associated gene expression by increasing H3K27me3 expression (Kurniawan et al., 2022).

MDGA2 plays the role of tumor suppressor in many cancers. Hypermethylation of MDGA2 is a prognostic marker in gastric cancer (Wang et al., 2016). MDGA2 knockdown enhances cell viability, boosts

colony formation, and advances the cell cycle but reduces apoptosis. MDGA2-encoded proteins form a new subfamily of the Ig superfamily and have a distinct structural organization consisting of six immunoglobulin chains (Litwack et al., 2004).

Dysregulation of PRM1 was absorbed in different tumor tissues and peripheral blood of cancer patients (Ren et al., 2021a; Meklat et al., 2009; Chen et al., 2018). An abnormal expression of the CTA family gene PRM1 results in a particular humoral immune response (Ren et al., 2021b). It regulates the invasion, migration, and proliferation of cancer cells (Chen et al., 2018).

TEX14 upregulation was associated with the abundance of tumor suppressor protein REST in different cancer so it could be a potential therapeutic target (Karlin et al., 2014). It is essential for kinetochore-microtubule attachment and helps in metaphase to anaphase transition (Mondal et al., 2012).

KIT auto phosphorylates on numerous Y residues that serve as docking sites for downstream effectors once activated. Several downstream mechanisms regulate cell survival and proliferation (Rnnstrand, 2004). SFKs, PI3K p85, phospholipase C-gamma, and adaptors that activate MAP kinase pathways attach to phospho-Y residues on the receptor. KIT mutations are also associated with different types of cancers (Ashman and Griffith, 2013). KIT plays an important part in the activation of different immune cells like Mast cells, dendritic cells, eosinophils, etc (Oriss et al., 2014).

TMIGD2, also known as CD28H (CD28 homolog), expressed in Homo sapiensand monkeys, while not in mice, enhanced angiogenesis when overexpressed in different cancers. It's a naïve T cell expressed stimulatory receptor. TMIGD2 is a member of the Ig superfamily and has an IgV-like domain, transmembrane region, and cytoplasmic tail. TMIGD2 has various functions depending on cell types and signaling pathways. It is a receptor of HHLA2 and hence could be a therapeutic target for various anti-cancer therapies (Janakiram et al., 2015).

LEF1-AS1 (long noncoding RNA) overexpression is associated with the malignant growth of various tumors, and its knockdown inhibits the progression of many cancers. LEF1-AS1 mainly regulates ERK, Akt/mTOR signaling, Wnt/β-Catenin, and Hippo signaling pathways hence playing diverse roles in tumor progression and immune regulation.

FKBP7 could be the therapeutic target for various cancer,especially in case of drug resistance, like the taxane-resistance mTOR pathway can be controlled by targeting FKBP family proteins (Kolos et al., 2018; Garrido et al., 2019).

MZB1 expression is associated with the progression of different cancers and patients' disease-free survival (Watanabe et al., 2020; Kanda et al., 2016). MZB1 is essential for plasma cell differentiation and humoral immune response independent of T-cells by plasma cells (Andreani et al.,) and enhances the secretion of interferon α by dendritic cells (Kapoor et al., 2020).

4.2. Biological processes regulated by genes

Enrichment analysis of the top ten selected genes was achieved by the Funrich tool (Pathan et al., 2015). Biological processes which are statistically significantly regulated by these genes were identified based on their P-value, which should be less than 0.05. It was found that these genes were mainly involved in Apoptosis, Signal transduction, regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism, and Cell communication (Figs. 5 and 6).

4.3. Biological pathways regulated by genes

Biological pathways regulated by these 10 genes were analyzed by the Funrich tool (Pathan et al., 2015) and it was found that KIT signaling, GM-CSF signaling, NOTCH, TGFBR, interleukins signaling, wnt signaling, cytokine signaling in the immune system, CDC42 signaling and EGF receptor signaling were the main pathways regulated by them. The significance of these pathways was analyzed statistically based ontheir P-value which should be less than 0.05.

The significance of these 10 genes was noticed to play a significant role in the regulation of cancer progression and regulation of the immune system that is actively involved in cancer mitigation. They were found to be related to biological processes and pathways that are very much involved in the regulation of cancer metastatic progression.

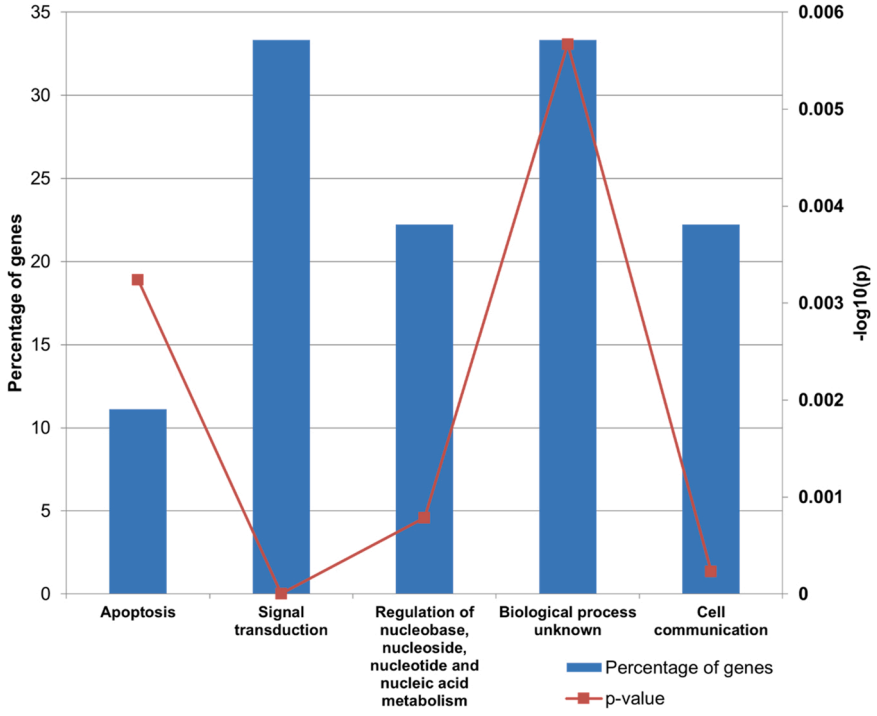


Fig. 5. The figure demonstrates the percentage of the top 10 genes that are involved in different biological processes or their regulation (bar graphs) and the p-value for their statistical significance in each case is represented by line plots.

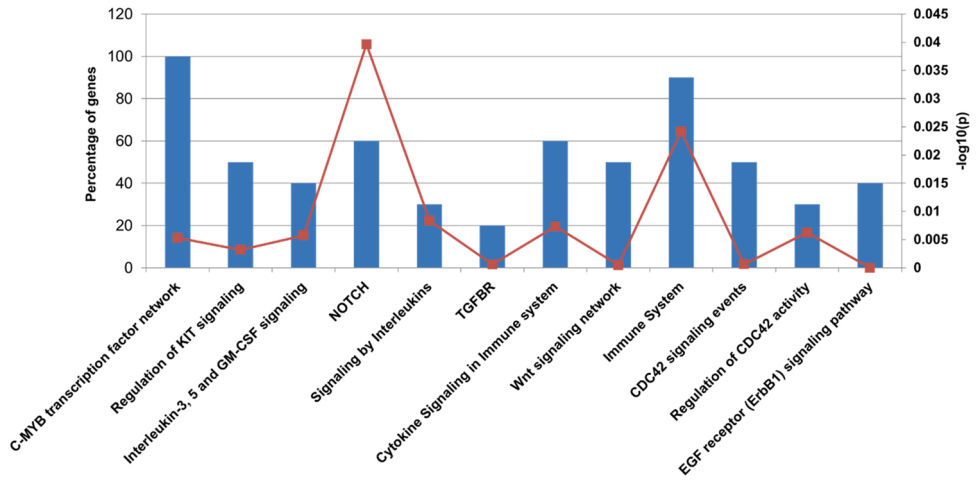


Fig. 6. The figure demonstrates the percentage of the top 10 genes that are involved in different biological pathways or their regulation (bar graphs) and the p-value for their statistical significance in each case is represented by line plots.

Significant evidence was found in the literature proving their immunological role and contribution to cancer progression. Therefore, these genes could be the potential PBMC biomarkers of breast cancer which can help in early detection and could be the non-invasive alternative to breast cancer detection.

5. Conclusion

Early detection of cancer using tumor-derived biomarkers for breast cancer has several lacunae and has been discussed extensively in the present manuscript. Since PBMCs are immune cells in the blood that help the host's immune system respond to tumor cells, peripheral blood profiling can be used for early detection of cancer based on immune marker profiling that alters due to the host immune system's reaction to cancer. Also, it is known that Tumor cells induced modulation of the immune system, which is a potential early detection target as well as a therapeutic target. It also offers the possibility of early cancer detection with minimally invasive methods (even before clinical symptoms appear). It can also be useful for predicting how a tumor will grow and how a patient will fair and the prognosis of clinical progression. However, since these results are based on computational biology, in vivo studies are necessary to validate them. This study promotes the application of XAI on ML models for quantifying & comprehensively examining the predicted findings, particularly in the area of biology, for the development of biomarkers of predictive and prognostic significance.

Conflict of interests

The authors declared that there is no conflict of interest.

CRediT authorship contribution statement

The first author Sunil Kumar contributed to curation of the study and data acquisitions as well as preparation of the manuscript, The corresponding author Dr. Asmita Das contributed by conception of the project, data analysis, Supervision and manuscript editing.

References

Aas, K., Jullum, M., Løland, A., 2021. Explaining individual predictions when features are dependent: more accurate approximations to Shapley values. Artif. Intell. 298, 103502 https://doi.org/10.1016/J.ARTINT.2021.103502.

Akram, M., Iqbal, M., Daniyal, M., Khan, A.U., 2017. Awareness and current knowledge of breast cancer. Biol. Res. 50 (1), 33. https://doi.org/10.1186/s40659-017-0140-9.

G. Allen, Understanding AI Technology A concise, practical, and readable overview of Artificial Intelligence and Machine Learning technology designed for non-technical managers, officers, and executives April 2020. [Online]. Available: ⟨https://www.linkedin.com/company/dod-joint-artificial-intelligence-center/⟩ (Accessed: 3 July 2022).

V. Andreani et al., Cochaperone Mzb1 is a key effector of Blimp1 in plasma cell differentiation and β1-integrin function. doi: ⟨https://doi.org/10.1073/pnas.1809739115⟩.

Ashman, L.K., Griffith, R., 2013. Therapeutic targeting of c-KIT in cancer. Expert Opin. Investig. Drugs 22 (1), 103–115. https://doi.org/10.1517/13543784.2013.740010.

Barredo Arrieta, A., et al., 2020. Explainable Artificial Intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf. Fusion 58, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012.

Bedard, P.L., Hansen, A.R., Ratain, M.J., Siu, L.L., 2013. Tumour heterogeneity in the clinic. Nature 501 (7467), 355–364. https://doi.org/10.1038/nature12627.

Bender, A., Cortes-Ciriano, I., 2021. Artificial intelligence in drug discovery: what is realistic, what are illusions? Part 2: a discussion of chemical and biological data. Drug Discov. Today 26 (4), 1040–1052. https://doi.org/10.1016/J.DRUDIS.2020.11.037.

Benz, C.C., 2008. Impact of aging on the biology of breast cancer. Crit. Rev. Oncol. Hematol. 66 (1), 65–74. https://doi.org/10.1016/j.critrevonc.2007.09.001.

Bohr, A., Memarzadeh, K., 2020. The rise of artificial intelligence in healthcare applications. In: Artificial Intelligence in Healthcare. Elsevier, pp. 25–60.

Čelešnik, H., Potočnik, U., 2022. Peripheral blood transcriptome in breast cancer patients as a source of less invasive immune biomarkers for personalized medicine, and implications for triple negative breast cancer. Cancers 14 (3). https://doi.org/10.3390/CANCERS14030591.

Chen, Z., Shi, C., Gao, S., Song, D., Feng, Y., 2018. Impact of protamine I on colon cancer proliferation, invasion, migration, diagnosis and prognosis. Biol. Chem. 399 (3), 265–275. https://doi.org/10.1515/hsz-2017-0222.

Crispen, P.L., Kusmartsev, S., 2020. Mechanisms of immune evasion in bladder cancer. Cancer Immunol. Immunother. 69 (1), 3–14. https://doi.org/10.1007/s00262-019-02443-4.

Czerwinska, U., 2022. Interpret. Mach. Learn. Models 275–303.

Dlamini, Z., Francies, F.Z., Hull, R., Marima, R., 2020. Artificial intelligence (AI) and big data in cancer and precision oncology. Comput. Struct. Biotechnol. J. 18, 2300–2311. https://doi.org/10.1016/J.CSBJ.2020.08.019.

von Eschenbach, W.J., 2021. Transparency and the black box problem: why we do not trust AI. Philos. Technol. 34 (4), 1607–1622. https://doi.org/10.1007/S13347-021-00477-0.

Feng, Y., et al., 2018. Breast cancer development and progression: risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. Genes Dis. 5 (2), 77–106. https://doi.org/10.1016/j.gendis.2018.05.001.

Garrido, M.F., et al., 2019. Regulation of eIF4F translation initiation complex by the peptidyl prolyl isomerase FKBP7 in taxane-resistant prostate cancer. Clin. Cancer Res. 25 (2), 710–723. https://doi.org/10.1158/1078-0432.CCR-18-0704/87619/AM/REGULATION-OF-EIF4F-TRANSLATION-INITIATION-COMPLEX.

Holzinger, A., 2021. The next frontier: ai we can really trust. Commun. Comput. Inf. Sci. (CCIS) 1524, 427–440. https://doi.org/10.1007/978-3-030-93736-2_33/COVER.

Holzinger, A., et al., 2022a. Information fusion as an integrative cross-cutting enabler to achieve robust, explainable, and trustworthy medical artificial intelligence. Inf. Fusion 79, 263–278. https://doi.org/10.1016/J.INFFUS.2021.10.007.

Holzinger, A., Saranti, A., Molnar, C., Biecek, P., Samek, W., 2022b. Explainable AI Methods - A Brief Overview, Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 13200 LNAI, pp. 13–38. ⟨https://doi.org/10.1007/978-3-031-04083-2_2/FIGURES/3⟩.

H. Hou et al., Peripheral blood transcriptome identifies high-risk benign and malignant breast lesions, 2020. ⟨10.1371/journal.pone.0233713⟩.

Hu, T., Wolfram, J., Srivastava, S., 2021. Extracellular vesicles in cancer detection: hopes and hypes. Trends Cancer 7 (2), 122–133. https://doi.org/10.1016/J.TRECAN.2020.09.003.

Janakiram, M., Chinai, J.M., Zhao, A., Sparano, J.A., Zang, X., 2015. HHLA2 and TMIGD2: new immunotherapeutic targets of the B7 and CD28 families. Oncoimmunology 4 (8), e1026534. https://doi.org/10.1080/2162402X.2015.1026534.

Jatoi, I., Pinsky, P.F., 2021. Breast cancer screening trials: endpoints and overdiagnosis. JNCI J. Natl. Cancer Inst. 113 (9), 1131–1135. https://doi.org/10.1093/jnci/djaa140.

Kanda, M., et al., 2016. Epigenetic suppression of the immunoregulator MZB1 is associated with the malignant phenotype of gastric cancer. Int. J. Cancer 139 (10), 2290–2298. https://doi.org/10.1002/IJC.30286.

Kapoor, T., Corrado, M., Pearce, E.L., Pearce, E.J., Grosschedl, R., 2020. MZB1 enables efficient interferon α secretion in stimulated plasmacytoid dendritic cells. Sci. Rep. 10 (1), 21626. https://doi.org/10.1038/s41598-020-78293-3.

Karlin, K.L., et al., 2014. The oncogenic STP Axis promotes triple-negative breast cancer via degradation of the REST tumor suppressor. Cell Rep. 9 (4), 1318–1332. https://doi.org/10.1016/J.CELREP.2014.10.011.

Kocarnik, J.M., et al., 2022. Cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life years for 29 cancer groups from 2010 to 2019. JAMA Oncol. 8 (3), 420. https://doi.org/10.1001/jamaoncol.2021.6987.

Kolos, J.M., Voll, A.M., Bauder, M., Hausch, F., 2018. FKBP ligands—where we are and where to go. Front. Pharmacol. 9 https://doi.org/10.3389/fphar.2018.01425.

Kothari, C., et al., 2020. Machine learning analysis identifies genes differentiating triple negative breast cancers. Sci. Rep. 10 (1), 10464. https://doi.org/10.1038/s41598-020-67525-1.

Kressin, N.R., et al., 2022. Women's understandings and misunderstandings of breast density and related concepts: a mixed methods study. J. Women's. Health. https://doi.org/10.1089/JWH.2021.0343.

Kure, S., et al., 2021. Breast cancer detection from a urine sample by dog sniffing: a preliminary study for the development of a new screening device, and a literature review. Biology 10 (6), 517. https://doi.org/10.3390/biology10060517.

Kurniawan, F., et al., 2022. BEND3 safeguards pluripotency by repressing differentiation-associated genes. Proc. Natl. Acad. Sci. USA 119 (9). https://doi.org/10.1073/pnas.2107406119.

Li, C., 2021. Unfolded protein response and crohn's diseases: a molecular mechanism of wound healing in the gut. Gastrointest. Disord. 3 (1), 31–43. https://doi.org/10.3390/GIDISORD3010004.

Li, J., et al., 2020. Non-invasive biomarkers for early detection of breast cancer. Cancers 12 (10), 2767. https://doi.org/10.3390/cancers12102767.

Lima, S.M., Kehm, R.D., Terry, M.B., 2021. Global breast cancer incidence and mortality trends by region, age-groups, and fertility patterns. EClinicalMedicine 38, 100985. https://doi.org/10.1016/J.ECLINM.2021.100985.

Linardatos, P., Papastefanopoulos, V., Kotsiantis, S., 2021. Explainable AI: a review of machine learning interpretability methods. Entropy 23. https://doi.org/10.3390/E23010018, 18, 23, no. 1, p. 18, Dec. 2020.

Litwack, E.D., Babey, R., Buser, R., Gesemann, M., O'Leary, D.D., 2004. Identification and characterization of two novel brain-derived immunoglobulin superfamily members with a unique structural organization. Mol. Cell. Neurosci. 25 (2), 263–274. https://doi.org/10.1016/j.mcn.2003.10.016.

Llinàs-Arias, P., et al., 2019. Epigenetic loss of the endoplasmic reticulum–associated degradation inhibitor SVIP induces cancer cell metabolic reprogramming. JCI Insight 4 (8). https://doi.org/10.1172/jci.insight.125888.

Lucchetti, D., Tenore, C.R., Colella, F., Sgambato, A., 2020. Extracellular vesicles and cancer: a focus on metabolism, cytokines, and immunity. Cancers 12 (1). https://doi.org/10.3390/CANCERS12010171.

Lundberg, S.M., et al., 2020. From local explanations to global understanding with explainable AI for trees. Nat. Mach. Intell. 2 (1), 56–67. https://doi.org/10.1038/s42256-019-0138-9.

Mann, R.M., Cho, N., Moy, L., 2019. Breast MRI: state of the art. Radiology 292 (3), 520–536. https://doi.org/10.1148/RADIOL.2019182947.

Marrugo-Ramírez, J., Mir, M., Samitier, J., 2018. Blood-based cancer biomarkers in liquid biopsy: a promising non-invasive alternative to tissue biopsy. Int. J. Mol. Sci. 19 (10), 2877. https://doi.org/10.3390/ijms19102877.

Mattei, F., et al., 2020. The tumor microenvironment: a milieu hindering and obstructing antitumor immune responses. Front. Immunol. 1, 940. https://doi.org/10.3389/fimmu.2020.00940 (|). ⟨www.frontiersin.org⟩.

Meena, J., Hasija, Y., 2022. Application of explainable artificial intelligence in the identification of Squamous Cell Carcinoma biomarkers. Comput. Biol. Med. 146, 105505 https://doi.org/10.1016/j.compbiomed.2022.105505.

Meklat, F., et al., 2009. Identification of protamine 1 as a novel cancer-testis antigen in early chronic lymphocytic leukaemia. Br. J. Haematol. 144 (5), 660–666. https://doi.org/10.1111/J.1365-2141.2008.07502.X.

Mondal, G., Ohashi, A., Yang, L., Rowley, M., Couch, F.J., 2012. Tex14, a Plk1-regulated protein, is required for kinetochore-microtubule attachment and regulation of the spindle assembly checkpoint. Mol. Cell 45 (5), 680–695. https://doi.org/10.1016/j.molcel.2012.01.013.

Mutebi, M., et al., 2020. Breast cancer treatment: a phased approach to implementation. Cancer 126 (Suppl 10), 2365–2378. https://doi.org/10.1002/CNCR.32910.

Nixon, A.B., Schalper, K.A., Jacobs, I., Potluri, S., Wang, I.M., Fleener, C., 2019. Peripheral immune-based biomarkers in cancer immunotherapy: can we realize their predictive potential? J. Immunother. Cancer 7 (1), 1–14. https://doi.org/10.1186/S40425-019-0799-2/TABLES/2.

Y. Nohara, K. Matsumoto, H. Soejima, N. Nakashima, Explanation of Machine Learning Models Using Shapley Additive Explanation and Application for Real Data in Hospital.

Oriss, T.B., Krishnamoorthy, N., Ray, P., Ray, A., 2014. Dendritic cell c-kit signaling and adaptive immunity. Curr. Opin. Allergy Clin. Immunol. 14 (1), 7–12. https://doi.org/10.1097/ACI.0000000000000019.

Pathan, M., et al., 2015. FunRich: an open access standalone functional enrichment and interaction network analysis tool. Proteomics. https://doi.org/10.1002/pmic.201400515.

Rai, A., 2020. Explainable AI: from black box to glass box. J. Acad. Mark. Sci. 48 (1), 137–141. https://doi.org/10.1007/S11747-019-00710-5/TABLES/1.

Ramos-Medina, R., López-Tarruella, S., Del Monte-Millán, M., Massarrah, T., Martín, M., 2021. Technical challenges for CTC implementation in breast cancer. Cancers 13 (18). https://doi.org/10.3390/cancers13184619.

Ren, S., et al., 2021a. The expression, function, and utilization of Protamine1: a literature review. Transl. Cancer Res. 10 (11), 4947–4957. https://doi.org/10.21037/TCR-21-1582.

Ren, S., et al., 2021b. The expression, function, and utilization of Protamine1: a literature review. Transl. Cancer Res. 10 (11), 4947–4957. https://doi.org/10.21037/tcr-21-1582.

Rnnstrand, L., 2004. Signal transduction via the stem cell factor receptor/c-Kit. Cell. Mol. Life Sci. 61 (19–20), 2535–2548. https://doi.org/10.1007/s00018-004-4189-6.

Rodríguez-Pérez, R., Bajorath, J., 2020. Interpretation of machine learning models using shapley values: application to compound potency and multi-target activity predictions. J. Comput. Aided Mol. Des. 34 (10), 1013–1026. https://doi.org/10.1007/s10822-020-00314-0.

Shaheen, M.Y., 2021. Applications of Artificial Intelligence (AI) in healthcare: a review. Sci. Prepr. https://doi.org/10.14293/S2199-1006.1.SOR-.PPVRY8K.V1.

Shapley, L.S., 1953. 17. A Value for n-Person Games, in Contributions to the Theory of Games, Volume II. Princeton University Press, pp. 307–318.

Sharma, P., et al., 2005. Early detection of breast cancer based on gene-expression patterns in peripheral blood cells. Breast Cancer Res. 7 (5) https://doi.org/10.1186/BCR1203.

Shiheido, H., et al., 2014. Human T cells expressing BEND3 on their surface represent a novel subpopulation that preferentially produces IL-6 and IL-8. Immun. Inflamm. Dis. 2 (1), 35–43. https://doi.org/10.1002/iid3.17.

Siegel, R.L., Miller, K.D., Fuchs, H.E., Jemal, A., 2022. Cancer statistics, 2022. CA Cancer J. Clin. 72 (1), 7–33. https://doi.org/10.3322/caac.21708.

Srivastava, S., et al., 2019. Cancer overdiagnosis: a biological challenge and clinical dilemma. Nat. Rev. Cancer 19 (6), 349–358. https://doi.org/10.1038/S41568-019-0142-8.

Sung, H., et al., 2021. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 Cancers in 185 Countries. CA Cancer J. Clin. 71 (3), 209–249. https://doi.org/10.3322/caac.21660.

Tadimety, A., Closson, A., Li, C., Yi, S., Shen, T., Zhang, J.X.J., 2018. Advances in liquid biopsy on-chip for cancer management: Technologies, biomarkers, and clinical analysis. Crit. Rev. Clin. Lab. Sci. 55 (3), 140–162. https://doi.org/10.1080/10408363.2018.1425976.

Veal, G.J., et al., 2019. Pharmacodynamic therapeutic drug monitoring for cancer: challenges, advances, and future opportunities. Ther. Drug Monit. 41 (2), 142–159. https://doi.org/10.1097/FTD.0000000000000606.

Wang, K., et al., 2016. MDGA2 is a novel tumour suppressor cooperating with DMAP1 in gastric cancer and is associated with disease outcome. Gut 65 (10), 1619–1631. https://doi.org/10.1136/gutjnl-2015-309276.

Watanabe, M., et al., 2020. MZB1 expression indicates poor prognosis in estrogen receptor-positive breast cancer. Oncol. Lett. 20 (5), 1. https://doi.org/10.3892/ol.2020.12059.

Wen, G., Zhou, T., Gu, W., 2021. The potential of using blood circular RNA as liquid biopsy biomarker for human diseases. Protein Cell 12 (12), 911–946. https://doi.org/10.1007/S13238-020-00799-3.

Wieland, R., Lakes, T., Nendel, C., 2021. Using Shapley additive explanations to interpret extreme gradient boosting predictions of grassland degradation in Xilingol, China. Geosci. Model Dev. 14 (3), 1493–1510. https://doi.org/10.5194/gmd-14-1493-2021.

Yang, G., et al., 2022. A systematic review of oral biopsies, sample types, and detection techniques applied in relation to oral cancer detection. BioTech 11 (1), 5. https://doi.org/10.3390/biotech11010005.

Zhang, Y., Weng, Y., Lund, J., 2022. Applications of explainable artificial intelligence in diagnosis and surgery. Diagnostics 12 (2). https://doi.org/10.3390/DIAGNOSTICS12020237.

Chinese Journal of Integrative Medicine
Available online at link.springer.com/journal/11655
Journal homepage: www.cjim.cn/zxyjhen/zxyjhen/ch/index.aspx
E-mail: cjim_en@cjim.cn

Original Article

A Cocktail of Natural Compounds Holds Promise for New Immunotherapeutic Potential in Head and Neck Cancer

Sunil Kumar and Asmita Das

ABSTRACT **Objective:** To obtain detailed understanding on the gene regulation of natural compounds in altering prognosis of head and neck squamous cell carcinomas (HNSC). **Methods:** Gene expression data of HNSC samples and peripheral blood mononuclear cells (PBMCs) of HNSC patients were collected from Gene Expression Omnibus (GEO). Differential gene expression analysis of GEO datasets were achieved by the GEO2R tool. Common differentially expressed gerres (DEGs) were screened by comparing DEGs of HNSC with those of PBMCs. The combination was further analyzed for regulating pathways and biological processes that were affected. **Results:** Totally 110 DEGs were retrieved and identified to be involved in biological processes related to tumor regulation. Then 102 natural compounds were screened for a combination such that the expression of all 110 commonly DEGs was altered. A combination of salidroside, ginsenoside Rd, oridonin, britanin, and scutellarein was chosen. A multifaceted, multi-dimensional tumor regression was showed by altering autophagy, apoptosis, inhibiting cell proliferation, angiogenesis, metastasis and inflammatory cytokines production. **Conclusions:** This study has helped develop a unique combination of natural compounds that will markedly reduce the propensity of development of drug resistance in tumors and immune evasion by tumors. The result is crucial to developing a combinatorial natural therapeutic cocktail with accentuated immunotherapeutic potential.

KEYWORDS natural compounds, chemopreventive, gene expression, head and neck cancer, immunomodulator, salidroside, ginsenoside Rd, oridonin, britanin, scutellarein

The genetic variability across different cancer types has impeded the identification of therapeutic targets and the drug design and development against tumors.[1] The most commonly used cancer therapies include surgery, radiation, and chemotherapy, which can be used in isolation or different combinations.[2] However, these techniques have been associated with a high morbidity rate and a significant decline in quality of life.[3] Head and neck squamous cell carcinomas (HNSC) could affect different head and neck areas, including the tongue, pharynx, larynx, nasal cavity, and salivary glands.[4] It is the sixth most common cancer globally, with over 880,000 new cases diagnosed and over 450,000 patients dying yearly.[5] HNSC, with complex pathophysiology and pathogenesis, makes it difficult to determine the optimal treatment strategy.[6]

Additionally, despite monotherapy's specificity and efficiency, cancer cells' molecular flexibility renders ideal lethal effects challenging.[7] HNSC has a poor prognosis due to many patients' high likelihood of recurrence or metastasis following radiation or chemotherapy.[8] This high metastatic rate of HNSC is because of the tumor cell's interactions with the surrounding tissues and immune cells that will form the tumor microenvironment (TME).[9] Host immune cells can recognize and eliminate the tumor cells, but an evasion of immunosurveillance generates an environment that accommodates the progression and survival of tumor cells.[10] Cancer-associated stromal fibroblasts, T cells, B cells, neutrophils, macrophages, myeloid-derived suppressor cells (MDSC), natural killer (NK) cells, and mast cells are all part of the TME.[11] These numerous cell subsets penetrate tumors and interact with cells one another through multiple networks.[12] Tumors progress if they can evade and/or suppress antitumor immune responses.[13] Tumors frequently elude the immune system of their hosts by inhibiting cytotoxic T-cell activity or activating

Department of Biotechnology, Delhi Technological University, Delhi (110042), India
Correspondence to: Dr. Asmita Das, E-mail: asmita1710@gmail.com, asmitadas1710@dce.ac.in
DOI: https://doi.org/10.1007/s11655-023-3694-0

and increasing immunosuppressive cell populations.[14]

Efforts to increase the efficacy of cancer treatment have primarily failed in recent decades, underlining the need for novel techniques such as complementary and alternative medicine.[15] Numerous natural herbal substances have caught the interest of academics and physicians due to their potential to prevent or improve the treatment of chronic diseases, including cancer.[16] Natural chemicals and combinations thereof may be a potential source of synergistic cancer treatments since they can interact with multiple biological targets involved in tumor growth, drug resistance, and metastasis.[17] Through their multitargeting action, natural chemicals may enhance the efficacy of already available cancer treatments or diminish treatment resistance.[18] Cancer treatment tries to eliminate or destroy tumor cells while sparing normal ones. The majority of natural substances are less poisonous, less expensive, have fewer side effects, and have been carefully researched for their carcinogenic potential.[19] Due to the adverse effects and drug resistance associated with conventional therapy, it was evident that natural substances can act as anticancer agents or adjuvants in chemotherapy.[20]

Cancer chemoprevention reverses, suppresses, or prevents cancer initiation, propagation, or advancement using natural or synthetic medications.[21] To be effective in people, a chemopreventive medicine must have an acceptable safety profile and be efficacious at a low enough dose to avoid severe side effects.[22] Natural dietary interventions such as fruits and vegetables show tremendous promise for chemopreventive research due to their potential to prevent and reduce cancer.[23] The chemical diversity of natural chemicals suggests a range of cancer chemoprevention techniques. Chemoprevention appears to be a rational and appealing strategy, as indicated by the success of several recent clinical trials aimed at cancer prevention in high-risk populations.[21]

Combination therapy combines two or more therapeutic drugs and is a crucial component of cancer treatment.[24] In comparison to monotherapy, the combination of anticancer drugs is more effective because it targets important pathways in a synergistic or additive.[25] This method might reduce drug resistance while providing therapeutic anticancer benefits, such as inhibiting mitotically active cells, reducing cancer stem cell populations, and triggering death.[26] Most metastatic tumors still have poor 5-year survival rates, and creating a new anticancer medicine is expensive and time-consuming.[27] As a result, new techniques are being investigated that target survival pathways and give efficient and effective results at a low cost.[28] In TME, the expression of many genes is regulated, affecting cancer prognosis. Thus, designing combinatorial therapy required evidence to reverse those gene regulations and be free of side effects due to concomitant undesirable gene regulation. In this study, the different combinations of natural compounds have been studied for the treatment of HNSC through various computational approaches.

METHODS

Data Collection

Gene expression data of HNSC samples and peripheral blood mononuclear cells (PBMCs) of HNSC patients were collected from Gene Expression Omnibus (GEO)[29] with accession Nos. GSE83519 and GSE39400,[30] respectively. In GSE83519, 22 HNSC tumors and 22 paired normal samples were studied from the same patients. In GSE39400, there are 28 samples of peripheral blood cells of HNSC patients who underwent surgery by means of expression profiling, with a controlled group of 11 patients who underwent surgery in the head and neck region for non-HNSC reasons. RNA was extracted from PBMCs using RNA-bee (Campro Scientific bv., Veenendaal, Netherlands). Microarrays agilent low RNA input fluorescent linear amplification kit and 4 × 44K whole human genome arrays were used for microarray hybridization (Agilent Technologies, Amstelveen, Netherlands).

GSE85871[31] contains gene expression profiles of MCF7 cells cultured in minimum essential medium/ Earle's balanced salt solution (MEM/EBSS, Hyclone), 10% fetal bovine serum, 1 mmol/L sodium pyruvate, and 100 mg/mL streptomycin in an incubator containing 5% CO_2 at 37 ℃ with 102 different molecules in Chinese medicines (CMs), and vehicle control (dimethyl sulfoxide, DMSO). Concentration and duration of compound administration may influence the gene expression patterns. According to the CMap database, the concentration of natural compounds was set to a single dosage of 10 μmol/L for 12 h, an internationally accepted concentration for high-throughput screening.[32] Two biological replicates for each group and the data set included profiles for 212 samples. RNA was isolated

from MCF7 cells using TRIzol after pre-treatment (Life Technologies, Carlsbad, US) and analyzed with Affymetrix Human Genome U133A 2.0 (Santa Clara, US) for gene expression patterns.

Differential Gene Expression Analysis

Differential gene expression analysis of GEO datasets of GSE83519 and GSE39400 were achieved by the GEO2R tool, an R-based tool that provides different high-throughput genetic data analysis packages.[33] Annotated gene tables and graphs were provided to help normalize the data, remove the data error, and visualize differentially expressed genes (DEGs).

Data normalization and differential gene expression analysis of the GSE83519 dataset was achieved with a cut-off of adjusted (adj.) *P*-value ⩽ 0.05, logFC value ⩾1 for differentially upregulated genes and logFC value ⩽−1 for differentially downregulated genes. Differential gene expression analysis of the GSE39400 dataset was also achieved by GEO2R with a cut-off adj. *P*-value ⩽0.05, logFC value >0 for upregulated genes and logFC value <0 for downregulated genes.

Enrichment Analysis

Enrichment analysis of selected upregulated and downregulated genes were achieved by the FunRich tool,[34] a standalone tool used for the functional enrichment analysis of genes. Results can be depicted in various forms like doughnut, Venn, pie, bar etc., and it can handle irrespective of the organism's verity of gene/protein datasets. Users can search either against the default background database or customized database for functional enrichment analysis in biological processes, pathways, etc.

Common Differential Gene Analysis

Common upregulated and downregulated differential genes were selected from the HNSC patient's tumor samples and PBMCs of HNSC patients by comparing their list of DEGs in Microsoft Excel. The set of differential genes from datasets GSE83519 and GSE39400 were compared by using conditional formatting > highlight cells rule > duplicate values.

Differential Gene Expression of Natural Compounds

GEO dataset GSE85871 was analyzed by GEO2R for 102 different natural compounds. Two replicates for each compound were considered a test set and compared with the vehicle control set in which only DMSO was present one by one, and the differential expression profile for all genes was collected. The expression of common upregulated genes was matched with the expression profile of natural compounds in Microsoft Excel and selected compounds that alter the maximum number of gene expressions. Similarly, it was achieved for common downregulated genes.

Screening of Natural Compounds

Natural compounds were screened on the bases of their regulation of common DEGs. The compound which regulates the maximum number of genes from common DEGs was selected first. Then the second compound was selected based on whether it can alter the expression of the maximum number of genes from the remaining ones (which remain unaltered by the first compound). Similarly, the remaining compounds were screened.

Analysis of Selected Combination of Natural Compounds

The selected combination of natural compounds was analyzed with the help of a Venn diagram of genes regulated by selected compounds and Enrichment analysis of the shared genes by the FunRich tool. Different biological processes and pathways regulated by these compounds were also analyzed from the literature (retrieved from PubMed, Google Scholar, SCOPUS, and others).

RESULTS

Differential Gene Expression Analysis

The normalization plot and volcano plot of GSE83519 and GSE39400 are shown in Appendix 1. In dataset of GSE83519, a total of 1,094 genes were found to be differentially upregulated, these genes had high expression in the HNSC tumor sample as compared to the respective normal sample of the same patient, and 889 genes were found to be differentially downregulated; these genes had low expression in the HNSC tumor as compared to the respective normal sample. GSE83519 dataset contains the tumor microenvironment samples, which include not only tumor cells but also other cell mediators like immune cells, fibroblast, blood vessels etc., so when immune cell comes in contact with this tumor microenvironment, they may alter their gene expression profile; therefore

the expression data of PBMCs need to be analyses individually, so we can identify the DEGs in PBMCs, because it is beneficial to target immune cell for the tumor regression along with only targeting tumor cells.

In dataset of GSE39400, a total of 737 genes were found to be upregulated, these genes had high expression in the PBMCs which were retrieved from HNSC patients after surgery as compared to the PBMCs retrieved from patients who got head and neck surgery for a non-HNSC reason, while 1,954 genes were found to be downregulated. These DEGs are mainly present in PBMCs, including dendritic cells, lymphocytes, and monocytes. Therefore they might alter these immune cells' function and help tumor cells escape the immune system. Thus these genes can be further analyzed to screen immunological biomarkers for HNSC.

Screening of Common DEGs

DEGs of HNSC tumors were compared with the DEGs of PBMCs so that common DEGs could be screened. Therefore, 1,094 upregulated DEGs of HNSC were compared with 737 upregulated DEGs of PBMCs and found 46 common genes. These 46 DEGs were upregulated in HNSC tumor samples and PBMCs of HNSC patients. PBMCs may infiltrate the tumor and affect tumor progression. The influence of the tumor microenvironment alters the expression of these genes in the tumor-infiltrating lymphocytes. These genes might be involved in the alteration of gene regulation in the subset of the immune cells in the vicinity of the tumor in HNSC patients due to the complex interplay of cells in the tumor microenvironment. Similarly, 889 downregulated DEGs of HNSC tumor samples were compared with 1,954 downregulated DEGs of PBMCs of HNSC patients, and 64 genes were common in both. The list of common upregulated and downregulated genes with their adj. *P*-value and logFC in HNSC tumor samples and PBMCs of HNSC patients are shown in Appendices 2 and 3. There were 110 common DEGs in HNSC samples and PBMCs, out of which 46 common upregulated genes (Appendix 4).

Enrichment analysis of these 110 DEGs was achieved and found that 46 common upregulated genes were mainly involved in biological processes like signal transduction, cell migration, RNA metabolism, anti-apoptosis, regulation of cell cycle, regulation of gene expression, cell communication, energy pathways, transport, protein metabolism, immune response, cell growth and/or maintenance. These biological processes might help in tumor progression because processes like suppression of apoptosis, cell migration, cell cycle regulation, cell growth and/or maintenance directly support tumor growth. Biological processes like immune response, signal transduction, cell communication, etc., could play an essential role in the tumor microenvironment for tumor progression. Therefore, the overexpression of these genes enhances these biological processes in the tumor microenvironment, which could help in tumor progression. Sixty-four common downregulated genes were mainly involved in biological processes like CGMP-mediated signaling, ribosome biogenesis and assembly, immune response, regulation of signal transduction, RNA metabolism, transcription, DNA repair, signal transduction, cell communication, transport, protein metabolism, energy pathways, metabolism, apoptosis. These biological processes are also linked with tumor progression or regression; therefore, downregulation of these genes could help tumor progression. So those drugs should be screened, which alter the expression of these genes to restore normal expression levels such that normal biological processes are restored. Enrichment analysis of 46 upregulated and 64 downregulated genes are shown in Figure 1.

Differential Expression Analysis of Natural Compounds

The expression profiles of 102 natural compounds already known for their antitumor activity were analyzed with GEO2R, and DEGs were studied. These DEGs were those whose expression were altered after the treatment with a particular natural compound. So, the list of differential genes was made of all the 102 natural compounds with their logFC values. The list of 110 common DEGs (46 upregulated genes and 64 downregulated genes) was compared with the natural compounds' DEGs. The natural compounds were sorted based on their ability to reverse the expression of the maximum number of genes. The list of natural compounds with respective No. of genes whose expression were reversed by that natural compound is shown in Appendix 5.

Combination of Natural Compounds

A combination of natural compounds was selected that alter the expression of the maximum number of genes out of 110 common DEGs. Therefore,

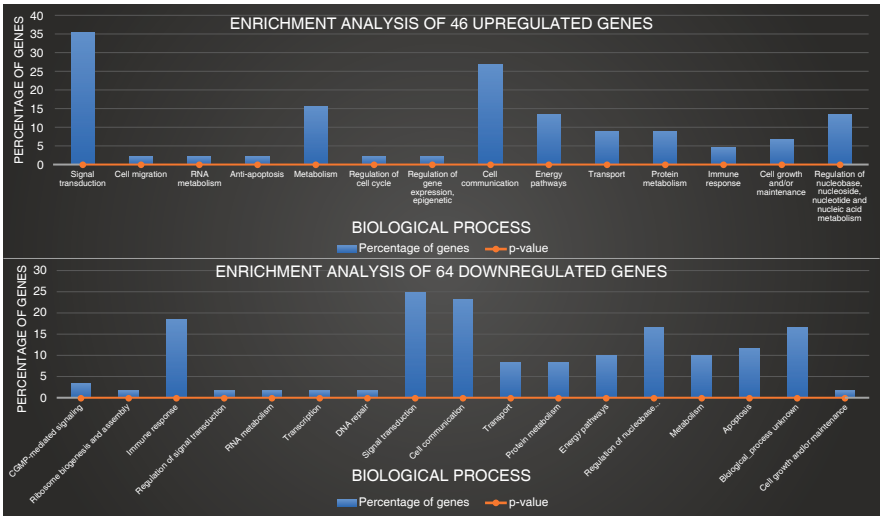


Figure 1. Enrichment Analysis of 46 Common Upregulated Genes and 64 Common Downregulated Genes with Respect to Biological Processes

we first selected that natural compound, which altered the expression of the maximum number of genes, salidroside. Salidroside altered the expression of 66 genes from 110 common DEGs. Then the remaining 44 commonly altered genes were studied for their susceptibility to restorative regulation by other such natural compounds that impacted the expression of the maximum number of the gene. Therefore, ginsenoside Rd was found to regulate the expression of 20 genes. Oridonin was found to regulate 12 genes out of the 24 common DEGs. Britanin was found to regulate 6 genes, and Scutellarein regulated 4 genes. Therefore, these 5 compounds together resulted in the regulation of 108 genes out of 110 common DEGs, restoring the gene expression to that in normal matched tissues. Two genes, GPR15 and CYP2U1, were not suitably regulated by our combination of natural compounds. Appendix 6 shows the screening process of natural compounds for targeting 110 common DEGs.

A combination of 5 compounds was selected and further checked for their combined effect on these common 110 DEGs. Therefore, the expression of these 110 DEGs was compared with the expression of the individual compounds. Salidroside, ginsenoside Rd, oridonin, britanin, and scutellarein individually regulated 66, 60, 58, 52, 62 genes. The gene expression regulation of common upregulated genes by different selected natural compounds separately is listed in Appendix 7. This data analysis showed that many genes were regulated more efficiently in combination rather than isolation; therefore, these compounds might show synergistic effects. The alteration of expression of genes by more than one compound is also beneficial to preventing drug resistance and toxic side effects due to alternative engagement of redundant pathways. The expression of different genes in PBMCs and regulated by the different compounds are shown in the Appendix 8. Venn diagram shows that all the 5 compounds regulated 11 genes, i.e., ZAP70, HFE, TRPM6, RPAIN, PAX9, PCNX2, TRAM2, ARHGEF5, ERN1, SAFB2, and SLC4A10. Twenty-five genes were regulated by any 4 compounds, 24 genes were regulated by any 3 compounds, 23 genes were regulated by any 2 compounds, 25 were regulated only 1 compound out of these 5 compounds (Figure 2).

Enrichment analysis of these 11 compounds was achieved and found that these compounds were involved in biological processes like transport, immune response, signal transduction, and regulation of nucleic acid metabolism. These biological processes are related to cancer progression and the immune system. Therefore, synergistic targeting of these genes would be beneficial for efficient combating HNSC tumors. All the selected compounds were efficient at targeting the expression of these 11 genes that effected the above-mentioned biological processes.

Compound Analysis from Literature

Natural compounds are multitargeting compounds, so targeting genes that positively impact cancer regression may result in undesirable side effects. Hence, to test the synergistic potential for targeting of the selected natural compounds, their mechanisms of

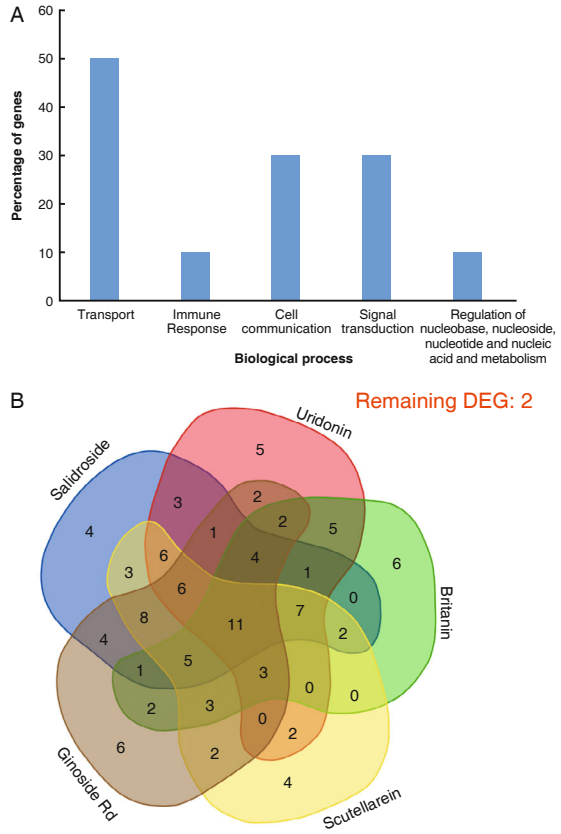


Figure 2. Enrichment Analysis of 11 Genes (A) and Gene Regulation of 110 Common DEGs (B) Regulated by Five Identified Natural Compounds

action were retrieved from the literature and congruence analysis was done. The drug combination of natural compounds must ideally target diverse pathways that converge to result in effective tumor regression.

Salidroside induced autophagy via phosphoinositide 3-kinase/protein kinase B/mammalian target of rapamycin (PI3K/Akt/mTOR) signaling. mTOR is highly upregulated in tumor cells, hence inhibiting autophagy. PI3K/Akt plays an essential role in the regulation of mTOR. Salidroside regulates the PI3K/Akt pathway, decreasing anti-apoptotic factors and increasing pro-apoptotic factors, thus inducing caspase-dependent and mitochondria-mediated apoptotic cell death.[35] Salidroside inhibits proliferation, migration, and invasion of tumor cells by inhibiting reactive oxygen species (ROS), which regulates Src, and downregulates HSP70 via Akt/ERK signaling.[36] Salidroside also reduces the pro-inflammatory cytokine secretion via activating IkB α /NF- κ B pathway and induces apoptosis via p53 and caspase signaling.[37]

Oridonin inhibits angiogenesis via the hypoxic inducible factor 1 α /vascular endothelial growth factor (HIF-1 α /VEGF) pathway and shows anti-migratory, anti-invasive and anti-adhesive properties.[38] Oridonin also inhibits the proliferation and migration of tumor cells via targeting transient receptor potential melastatin 7 (TRPM7) through the inactivation of ERK/AKT signaling.[38] Oridonin induces phagocytosis via activating ERK, which activates NF- κ B.[39]

Ginsenoside Rd reduces metastasis via miR-18a-mediated downregulation of SMAD2.[40] Ginsenoside Rd increases the expression of miR-144-5p, which inhibits the expression of Toll-like receptor 2 (TLR2) hence reducing the proliferation and metastasis of tumor cells.[41] Ginsenoside Rd inhibits VEGF-induced migration, tube formation, and proliferation and suppresses VEGF-induced regulation of Akt/mTOR signaling pathways, inducing apoptosis and inhibiting cell proliferation.[42] Ginsenoside Rd inhibits proliferation, and metastasis mainly reverses epithelial-mesenchymal transition (EMT) via STAT3/JAK2 signaling and STAT3 is the direct target of ginsenoside Rd.[43]

Britanin inhibits NF-kB via downregulation of IKK1/1KK2, controlling tumor cell proliferation and angiogenesis.[44] Britanin shows an anti-inflammatory response via inhibiting NF- κ B signaling.[45] Britanin downregulates cMyc and HIF1 α via upstream effectors like mTOR, reducing the expression of specific proteins, including programmed cell death 1 ligand 1 (PD-L1), leading to the inhibition of angiogenesis and cell proliferation.[46] Britanin induces apoptosis and autophagy via activating AMPK signaling regulated by ROS.[47]

Scutellarin inhibits glutamic-pyruvic transaminase (ALS) and aspartate transaminas (AST) hence regulating the immune system against tumor cells.[48,49] Scutellarin inhibits monocyte chemoattractant protein-1 (MCP1), thus inhibiting cell migration and reducing inflammation.[50] Scutellarin downregulates intercellular cell adhesion molecule 1 (ICAM-1) and inhibits the activation of NF- κ B hence inhibiting adhesion and showing an anti-inflammatory effect.[51] Scutellarin induces vasodilation via eNOS/NO/PKG pathways.[52]

As shown in Figure 3, these compounds regulate many different pathways and can target more than one process. So, our drug combination offers a highly potent multifaceted antitumor and immunomodulatory role and helps in the regression of HNSC cancer. The biological

network of the compounds is shown in Figure 4, which shows the key genes regulated by these compounds.

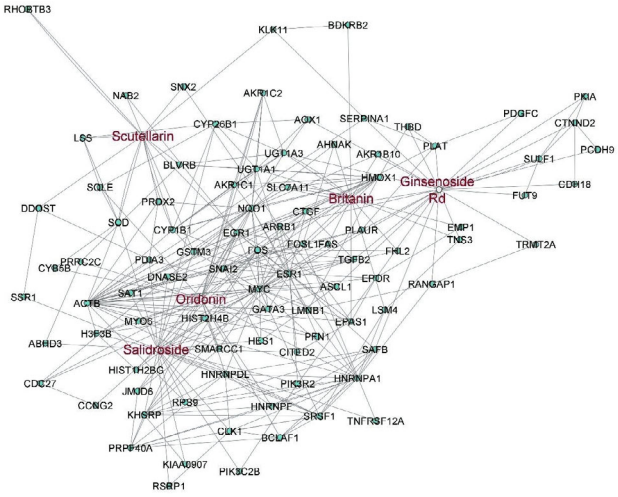


Figure 3. Different Biological Processes and Genes Regulated by Different Natural Compounds Where Different Signs Were Used for Inhibiting, Stimulating, and Activating Genes and Biological Processes

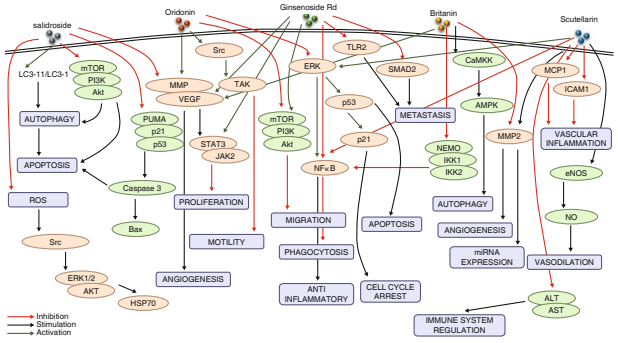


Figure 4. Biological Network of Selected Compounds with Their Regulating Genes

DISCUSSION

Tumor cells change their gene morphology quickly when exposed to single target drugs to find a way to escape the drug effect. So, the problem of drug resistance can be minimized by using a combination of natural compounds because they target multiple pathways, and tumor progression can be effectively reversed. Natural compounds are multitargeting; therefore, they can simultaneously target multiple pathways and many biological processes, helping in tumor regression. Natural compounds are more cost-effective as compared to synthetic compounds. As the cases of HNSC cancer increase in developing countries due to poor lifestyles, it is necessary to find cost-effective drugs for them. So, this study aimed to find a combination of natural compounds that can help in HNSC tumor regression with minimal side effects.

A combination of salidroside, ginsenoside Rd, oridonin, britanin, and scutellarein was chosen such that they can alter the expression of 108 genes out of the selected 110 genes. Salidroside is widely found in *Rhodiola* plants. *Rhodiola sacinehalnsis*, *Rhodiola rosea*, *Rhodibetic tibetica* and large *Rhodiola*. *Ligustrum lucidum*, in the leaves of *Salix triandra* L. and Willow bark, *Vaccinium vitisidaea* L leaves of *Oleaceae*, *Veroniceae* of Veronica minor. Salidroside was found to induce autophagy, inducing caspase-dependent and mitochondria-mediated apoptotic cell death, and inhibiting proliferation, migration, and invasion of tumor cells via PI3K/Akt/mTOR signaling, IkB α /NF- κ B signaling.[35-37] Salidroside is generally deemed safe and effective. In the experimental conditions, salidroside at doses of 0.5, 0.25, and 0.125 g/kg in Sprague-Dawley rats did not cause maternal or embryonic toxicity, nor did it have teratogenic consequences.[53] Genotoxicity testing is critical in drug risk assessment. Salidroside is not genotoxic at a clinical dose of 150 mg/60 kg per day) in humans, according to the Ames test, reverse mutation, chromosomal abnormalities, and mice micronucleus studies.[54] Another study of 60 breast cancer patients found no clinical adverse effects when an effective dose of salidroside (600 mg/kg daily) was given throughout the therapy procedure.[55] The lack of negative effects in pre-clinical and clinical trials suggests salidroside is a safe common clinical medication.

Ginsenoside Rd is mainly found in plants like *P. ginseng*, *Panax notoginseng*, *P. quinquefolius*, *Panax japonicas*, etc. Ginsenoside Rd reduces metastasis, proliferation, migration, inducing apoptosis, reverses EMT via different signaling pathways like Akt/mTOR signaling, STAT3/JAK2 signaling, miR-18a-mediated downregulation of SMAD2.[40-43] Many studies show that ginsenoside Rd has no significant side effects.[56,57]

Oridonin is primarily found in plants like *Rabdosia rubescens*, *Isodon japonicus* Hara, *Isodon trichocarpus*, *Isodon enanderianus*, and *I. lophanthoides*. Oridonin inhibits angiogenesis, migration, invasion and adhesion, proliferation, and phagocytosis properties via HIF-1 α / VEGF, ERK/Akt, ERK/NF κ B signaling.[38,39] Oridonin reduces the side effects of various other cancer drugs when used in combination.[58] Oridonin shows anticancer properties with very low side effects.[59]

Table 1. Plant Sources, Biological Process Regulated, Pathways Regulated and Toxicity of Selected Natural Compounds

Natural compound	Source of origin	Biological process regulated	Biological pathways regulated	Side effects & toxicity
Salidroside	*Rum lucidum* *Salix triandra* L. Willow bark *Vaccinium vitisidaea* L *Oleaceae veroniceae* of Veronica minor	Induces autophagy Induces caspase-dependent and mitochondria-mediated apoptotic cell death Inhibits proliferation Inhibits migration Inhibits invasion of tumor cells	PI3K/Akt/mTOR signaling IkB α /NF- κ b signaling	Salidroside is generally deemed safe and effective[53-55]
Ginsenoside Rd	*Panax ginseng* *Panax notoginseng* *Panax quinquefolius* *Panax japonicas*	Reduces metastasis Inhibits proliferation Inhibits migration Induces apoptosis Reverses emt	Akt/mTOR signaling STAT3/JAK2 signaling miR-18a-mediated downregulation of SMAD2	Ginsenoside Rd have no major side effect[56,57]
Oridonin	*Rabdosia rubescens* *Isodon japonicus* Hara *Isodon trichocarpus* *Isodon enanderianus* *Isodon lophanthoide*	Inhibits angiogenesis Inhibits migration Inhibits invasion Inhibits adhesion Proliferation Phagocytosis properties	HIF-1 α /VEGF signaling ERK/AKT signaling ERK/NF κ B signaling	Very low side effects[59]
Britanin	*Inula linearіifolia* Turcz. (asteraceae) *Inula japonica* *Inula britannica*.	Induces apoptosis Autophagy Inhibits cell proliferation Angiogenesis	IKK1/1KK2 signaling NF- κ B signaling AMPK signaling	Tolerable side effects at low dose administration *in vivo*[60]
Scutellarein	*Scutellaria lateriflora* *Asplenium belangeri* *Mexican oregano* Sweet orange *Scutellaria barbata*	Inhibits cell migration Inhibits adhesion Reduces inflammation Induces vasodilation	eNOS/NO/PKG signaling NF- κ B signaling	No side effects were observed in various studies[61]

Britanin is mainly found in plants like *Inula linearіifolia* Turcz. (Asteraceae), *Inula japonica*, *Inula britannica*. Britanin induces apoptosis and autophagy and inhibits cell proliferation and angiogenesis via regulating different pathways like IKK1/1KK2, NF- κ B, and AMPK signaling.[44,45,47] Britanin shows tolerable side effects at low-dose administration *in vivo*.[60] Scutellarein is found primarily in plants like *Scutellaria lateriflora*, *Asplenium belangeri*, *Mexican oregano*, sweet orange, *Scutellaria barbata*.

Scutellarein inhibits cell migration, adhesion, reducing inflammation, induces vasodilation via regulating different pathways like eNOS/NO/PKG, NF- κ B.[48,49,52] No side effects were absorbed when treated with scutellarein in various studies[61] as shown in Table 1. As this combination was further analyzed, it was found that many biological processes were regulated by more than one compound via different pathways; therefore, it might not be easy for tumor cells to escape this regression mechanism. Further, tumor cells cannot gain drug resistance easily against them.

Immunotherapy is typically associated with side effects that often deter the use of such treatment strategies. Our combination of natural compounds holds a better immunotherapeutic potential without the commonly associated side effects typically seen with chemical immunomodulatory drugs.

In conclusion, our studies showed a multifaceted, multi-dimensional tumor regression by altering autophagy, apoptosis, inhibiting cell proliferation, angiogenesis, metastasis and inflammatory cytokines production. It has helped develop a unique combination of natural compounds that will markedly reduce the propensity of development of drug resistance in tumors and immune evasion by the tumors. Our study has opened a new dimension for developing a combinatorial natural compound cocktail as a potential immunomodulatory drug alternative. Thus, we propose that such a combination could be further analyzed in *in vitro* and *in vivo* studies to develop better treatment for HNSC tumor patients.

Author Contributions

Project conceived and designed by Das A; Data collection, results validation and manuscript written by Kumar S and Das A. Both authors read the manuscript carefully and agreed to submit.

Conflict of Interest

The authors declare that is no conflict of Interest.

Electronic Supplementary Material: Supplementary material is available in the online version of this article at https://doi.org/10.1007/s11655-023-3694-0.

REFERENCES

1. Fisher R, Pusztai L, Swanton C. Cancer heterogeneity: implications for targeted therapeutics. Br J Cancer 2013;108:479.
2. Wadhawan A, Chatterjee M, Singh G. Present scenario of bioconjugates in cancer therapy: a review. Int J Mol Sci 2019;20:5243.
3. Miller KD, Siegel RL, Lin CC, Mariotto AB, Kramer JL, Rowland JH, et al. Cancer treatment and survivorship statistics, 2016. CA Cancer J Clin 2016;66:271-289.
4. Johnson DE, Burtness B, Leemans CR, Lui VWY, Bauman JE, Grandis JR. Head and neck squamous cell carcinoma. Nat Rev Dis Prim 2020;6:1-22.
5. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424.
6. Pisani P, Airoldi M, Allais A, Valletti PA, Battista M, Benazzo M, et al. Metastatic disease in head & neck oncology. Acta Otorhinolaryngol Ital 2020;40(Suppl 1):S1.
7. Zhong L, Li Y, Xiong L, Wang W, Wu M, Yuan T, et al. Small molecules in targeted cancer therapy: advances, challenges, and future perspectives. Signal Transduct Target Ther 2021;6:1-48.
8. Economopoulou P, de Bree R, Kotsantis I, Psyrri A. Diagnostic tumor markers in head and neck squamous cell carcinoma (HNSCC) in the clinical setting. Front Oncol 2019;9:827.
9. Young MRI. Protective mechanisms of head and neck squamous cell carcinomas from immune assault. Head Neck 2006;28:462-470.
10. Beatty GL, Gladney WL. Immune escape mechanisms as a guide for cancer immunotherapy. Clin Cancer Res 2015;21:687.
11. Chen SMY, Krinsky AL, Woolaver RA, Wang X, Chen Z, Wang JH. Tumor immune microenvironment in head and neck cancers. Mol Carcinog 2020;59:766.
12. Baghban R, Roshangar L, Jahanban-Esfahlan R, Seidi K, Ebrahimi-Kalan A, Jaymand M, et al. Tumor microenvironment complexity and therapeutic implications at a glance. Cell Commun Signal 2020;18:1-19.
13. Vinay DS, Ryan EP, Pawelec G, Talib WH, Stagg J, Elkord E, et al. Immune evasion in cancer: mechanistic basis and therapeutic strategies. Semin Cancer Biol 2015;35:S185-S198.
14. Waldman AD, Fritz JM, Lenardo MJ. A guide to cancer immunotherapy: from T cell basic science to clinical practice. Nat Rev Immunol 2020;20:651-668.
15. Buckner CA, Lafrenie RM, Dénommée JA, Caswell JM, Want DA. Complementary and alternative medicine use in patients before and after a cancer diagnosis. Curr Oncol 2018;25:e275.
16. Atanasov AG, Zotchev SB, Dirsch VM, Orhan IE, Banach M, Rollinger JM, et al. Natural products in drug discovery: advances and opportunities. Nat Rev Drug Discov 2021;20:200-216.
17. Sharifi-Rad J, Ozleyen A, Tumer TB, Adetunji CO, Omari N El, Balahbib A, et al. Natural products and synthetic analogs as a source of antitumor drugs. Biomolecules 2019;9:679.
18. Ramsay RR, Popovic-Nikolic MR, Nikolic K, Uliassi E, Bolognesi ML. A perspective on multi-target drug discovery and design for complex diseases. Clin Transl Med 2018;7:1-14.
19. Karimi A, Majlesi M, Rafieian-Kopaei M. Herbal *versus* synthetic drugs; beliefs and facts. J Nephropharmacology 2015;4:27.
20. Huang M, Lu JJ, Ding J. Natural products in cancer therapy: past, present and future. Nat Products Bioprospect 2021;11:5-13.
21. Steward WP, Brown K. Cancer chemoprevention: a rapidly evolving field. Br J Cancer 2013;109:1-7.
22. Mohan Shankar G, Swetha M, Keerthana CK, Rayginia TP, Anto RJ. Cancer chemoprevention: a strategic approach using phytochemicals. Front Pharmacol 2022;12:4044.
23. Kotecha R, Takami A, Espinoza JL. Dietary phytochemicals and cancer chemoprevention: a review of the clinical evidence. Oncotarget 2016;7:52517.
24. Mokhtari RB, Homayouni TS, Baluch N, Morgatskaya E, Kumar S, Das B, et al. Combination therapy in combating cancer. Oncotarget 2017;8:38022-38043.
25. Ayoub NM. Editorial: novel combination therapies for the treatment of solid cancers. Front Oncol 2021;11:2377.
26. Li Y, Wang Z, Ajani JA, Song S. Drug resistance and cancer stem cells. Cell Commun Signal 2021;19:1-11.
27. Eccles SA, Welch DR. Metastasis: recent discoveries and novel treatment strategies. Lancet 2007;369:1742.
28. Yang L, Shi P, Zhao G, Xu J, Peng W, Zhang J, et al. Targeting cancer stem cell pathways for cancer therapy. Signal Transduct Target Ther 2020;5:1-35.
29. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Res 2012;41:D991-D995.
30. Braakhuis B, Graveland A, Dijk F, Ylstra B, van Wieringen W, Leemans C, et al. Expression signature in peripheral blood cells for molecular diagnosis of head and neck squamous cell carcinoma. Oral Dis 2013;19:452-455.
31. Lv C, Wu X, Wang X, Su J, Zeng H, Zhao J, et al. The gene expression profiles in response to 102 traditional Chinese medicine (TCM) components: a general template for research on TCMs. Sci Rep 2017;7:352.
32. Qu XA, Rajpal DK. Applications of Connectivity Map in drug discovery and development. Drug Discov Today 2012;17:1289-1298.
33. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 2004;3:1-25.

34. Pathan M, Keerthikumar S, Ang CS, Gangoda L, Quek CYJ, Williamson NA, et al. FunRich: an open access standalone functional enrichment and interaction network analysis tool. Proteomics 2015;15:2597-601.
35. Rong L, Li Z, Leng X, Li H, Ma Y, Chen Y, et al. Salidroside induces apoptosis and protective autophagy in human gastric cancer AGS cells through the PI3K/Akt/mTOR pathway. Biomed Pharmacother 2020;122:109726.
36. Qi Z, Tang T, Sheng L, Ma Y, Liu Y, Yan L, et al. Salidroside inhibits the proliferation and migration of gastric cancer cells via suppression of Src-associated signaling pathway activation and heat shock protein 70 expression. Mol Med Rep 2018;18:147.
37. Kong YH, Xu SP. Salidroside prevents skin carcinogenesis induced by DMBA/TPA in a mouse model through suppression of inflammation and promotion of apoptosis. Oncol Rep 2018;39:2513-2526.
38. Li C, Wang Q, Shen S, Wei X, Li G. Oridonin inhibits vegf-a-associated angiogenesis and epithelial-mesenchymal transition of breast cancer *in vitro* and *in vivo*. Oncol Lett 2018;16:2289-2298.
39. Liu YQ, You S, Tashiro SI, Onodera S, Ikejima T. Activation of phosphoinositide 3-kinase, protein kinase C, and extracellular signal-regulated kinase is required for oridonin-enhanced phagocytosis of apoptotic bodies in human macrophage-like U937 cells. J Pharmacol Sci 2005;98:361-371.
40. Wang P, Du X, Xiong M, Cui J, Yang Q, Wang W, et al. Ginsenoside Rd attenuates breast cancer metastasis implicating derepressing microRNA-18a-regulated Smad2 expression. Sci Reports 2016 61 2016;6:1-14.
41. Liu GM, Lu TC, Sun ML, Jia WY, Ji X, Luo YG. Ginsenoside Rd inhibits glioblastoma cell proliferation by up-regulating the expression of miR-144-5p. Biol Pharm Bull 2020;43:1534-1541.
42. Zhang E, Shi H, Yang L, Wu X, Wang Z. Ginsenoside Rd regulates the Akt/mTOR/p70S6K signaling cascade and suppresses angiogenesis and breast tumor growth. Oncol Rep 2017;38:359-367.
43. Yang L, Zhang XY, Li K, Li AP, Yang WD, Yang R, et al. Protopanaxadiol inhibits epithelial-mesenchymal transition of hepatocellular carcinoma by targeting STAT3 pathway. Cell Death Dis 2019;10:1-13.
44. Xia L, Tan S, Zhou Y, Lin J, Wang H, Oyang L, et al. Role of the NF κ B-signaling pathway in cancer. Oncotargets Ther 2018;11:2063.
45. Park HH, Kim SG, Park YN, Lee J, Lee YJ, Park NY, et al. Suppressive effects of britanin, a sesquiterpene compound isolated from *Inulae flos*, on mast cell-mediated inflammatory responses. Am J Chin Med 2014;42:935-947.
46. Bailly C. Anticancer targets and signaling pathways activated by britannin and related pseudoguaianolide sesquiterpene lactones. Biomedicines 2021;9:1325.
47. Cui YQ, Liu YJ, Zhang F. The suppressive effects of Britannin (Bri) on human liver cancer through inducing apoptosis and autophagy via AMPK activation regulated by ROS. Biochem Biophys Res Commun 2018;497:916-923.
48. Zhao LY, Yang DD, Ma XK, Liu MM, Wu DH, Zhang XP, et al. The prognostic value of aspartate aminotransferase to lymphocyte ratio and systemic immune-inflammation index for overall survival of hepatocellular carcinoma patients treated with palliative treatments. J Cancer 2019;10:2299.
49. Xiong LL, Du RL, Xue LL, Jiang Y, Huang J, Chen L, et al. Anti-colorectal cancer effects of scutellarin revealed by genomic and proteomic analysis. Chin Med (UK) 2020;15:1-15.
50. Yuan Y, Rangarajan P, Kan EM, Wu Y, Wu C, Ling EA. Scutellarin regulates the Notch pathway and affects the migration and morphological transformation of activated microglia in experimentally induced cerebral ischemia in rats and in activated BV-2 microglia. J Neuroinflammation 2015;12:1-21.
51. Luo P, Tan ZH, Zhang ZF, Zhang H, Liu XF, Mo ZJ. Scutellarin isolated from *Erigeron multiradiatus* inhibits high glucose-mediated vascular inflammation. J Pharm Soc Jap 2008;128:1293-1299.
52. Chen YJ, Chen C, Li MY, Li QQ, Zhang XJ, Huang R, et al. Scutellarin reduces cerebral ischemia reperfusion injury involving in vascular endothelium protection and PKG signal. Nat Products Bioprospect 2021;11:659.
53. Zhong Z, Han J, Zhang J, Xiao Q, Hu J, Chen L. Pharmacological activities, mechanisms of action, and safety of salidroside in the central nervous system. Drug Des Devel Ther 2018;12:1479.
54. Zhu J, Wan X, Zhu Y, Ma X, Zheng Y, Zhang T. Evaluation of salidroside *in vitro* and *in vivo* genotoxicity. Drug Chem Toxicol 2010;33:220-226.
55. Zhang H, Shen WS, Gao CH, Deng LC, Shen D. Protective effects of salidroside on epirubicin-induced early left ventricular regional systolic dysfunction in patients with breast cancer. Drugs R D 2012;12:101-106.
56. Zeng X, Deng Y, Feng Y, Liu Y, Yang L, Huang Y, et al. Pharmacokinetics and safety of ginsenoside Rd following a single or multiple intravenous dose in healthy Chinese volunteers. J Clin Pharmacol 2010;50:285-292.
57. Zhang G, Xia F, Zhang Y, Zhang X, Cao Y, Wang L, et al. Ginsenoside Rd Is Efficacious against acute ischemic stroke by suppressing microglial proteasome-mediated Inflammation. Mol Neurobiol 2016;53:2529-2540.
58. Kazantseva L, Becerra J, Santos-Ruiz L. Oridonin enhances antitumor effects of doxorubicin in human osteosarcoma cells. Pharmacol Reports 2022;74:248-256.
59. Zhou GB, Kang H, Wang L, Gao L, Liu P, Xie J, et al. Oridonin, a diterpenoid extracted from medicinal herbs, targets AML1-ETO fusion protein and shows potent antitumor activity with low adverse effects on t(8;21) leukemia *in vitro* and *in vivo*. Blood 2007;109:3441-3450.
60. Li K, Zhou Y, Chen Y, Zhou L, Liang J. A novel natural product, britanin, inhibits tumor growth of pancreatic cancer by suppressing nuclear factor- κ B activation. Cancer Chemother Pharmacol 2020;85:699-709.
61. Wang L, Ma Q. Clinical benefits and pharmacology of scutellarin: a comprehensive review. Pharmacol Ther 2018;190:105-127.

(Accepted September 5, 2022)
Edited by YUAN Lin

Vegetos
https://doi.org/10.1007/s42535-022-00363-w

RESEARCH ARTICLES

Elucidation of natural compounds Gallic acid and Shikonin for the treatment of HNSC cancer by targeting immune suppressor and tumour progressor genes

Sunil Kumar[1] · **Asmita Das**[1]

Received: 30 December 2021 / Revised: 2 February 2022 / Accepted: 4 February 2022

Abstract
Head and Neck Squamous cell carcinoma is a leading cancer in males, especially in India. Progression of tumour growth in conjunction with immune suppression is the major cause that leads to HNSC cancer. Synthetic drugs targeting tumour cells often trigger tumour cells to acquire resistance against them. Immunotherapy also has its side effects and does not provide an adequate response in all patients, and its inherent variability in patient response often makes them prohibitive. Hence, a concomitant targeting of tumour cells and modulation of immune cell function may be particularly beneficial mechanism for cancer treatment. In the present study, we tried to identify natural compounds that could help in tumour suppression as well as functional immune modulation. The HNSC-associated genes that played role in both tumour growth and immune suppression were identified by enrichment analysis followed by gene expression analysis. 10 such genes were shortlisted, namely Foxp3, CD274, IDO1, IL-10, SOCS1, PRKDC, AXL, CDK6, TGFB1, FADD. CD274 and IDO1 which were found to have the highest degree of interaction based on their network of interactions. Gallic acid and Shikonin were found as the natural compounds that efficiently targeted CD274 and IDO1 respectively. Gallic acid is extracted from leaves of bearberry also found, in pomegranate root bark, gallnuts, witch hazel, both in free-state and as also a part of the tannin molecule, whereas Shikonin is found in the extracts of dried roots of the plant *Lithospermum erythrorhizon*. Studies have demonstrated that both Shikonin and Gallic acid exhibits anti-cancer properties. Expression data analysis of HNSC cancer exhibited 1745 differentially expressed genes. Gallic acid treatment resulted in the downregulation of 120 genes and upregulation of 35 genes while Shikonin treatment resulted in the downregulation of 660 genes and upregulation of 38 genes that are consequential in a positive impact of cancer regression. Thus, combination of these two compounds could be potentially beneficial in effective combinatorial therapy for HNSC.

Keywords Head and neck squamous cell carcinoma (HNSC) · Immunomodulation · Natural compounds · Gallic acid · Shikonin

Introduction

Head and neck squamous cell carcinoma (HNSC) is different from other cancers as for the stand point of its progression. Most of the other cancers lead to death in the metastatic phage of the disease but HNSC stays locoregional for a long time followed by local invasion and lymphatic dissemination. So, the progression of HNSC cancer is highly determined by immunity regulation that plays a critical role in cancer metastasis (Tímár et al. 2005).

Thousands of drugs have been used to treat cancer, but it is still the most abundant cause of fatality in the world. There are different types of therapy used for cancer treatment, such as radiation, surgery, chemotherapy, immunotherapy. However, many chemotherapeutic measures often results in the development of drug resistance in patients. Immune response in every individual is a complex array of immune functionality that are interrelated and regulated in a complex cascade of mechanisms that vary significantly in different individuals, hence patients often have variable tumour immunity and its prognosis. Therefore, the same immunotherapy may have functional variability for every

✉ Asmita Das
asmita1710@gmail.com; asmitadas1710@dce.ac.in

[1] Department of Biotechnology, Delhi Technological University, Main Bawana Road, Delhi 110042, India

Published online: 21 March 2022

patient and can even exhibit variable side effects (Vasan et al. 2019).

Drug resistance to cancer is a very complicated process and depend on different factors, such as mutation at the drug's target site (Bozic and Nowak 2014), alterations in drug metabolism (Zaal and Berkers 2018), resistance due to downregulation of pro-apoptotic signals and upregulation of anti-apoptotic signals (Chen et al. 2018). It may lead to an increase in impaired DNA repair (Salehan and Morse 2013), and may also lead to decrease in drug uptake or increase in drug efflux (Xue and Liang 2012; Alfarouk et al. 2015). Moreover, in the vicinity of the tumour, there are not only uncontrollably proliferating cells, but over the time an immense accumulation of divergent cells that modulate the surrounding environment which is known as the tumour microenvironment emerges. It contains immune cells, extracellular matrix, blood vessels, fibroblasts and signalling molecules (Deng et al. 2018). The immune system plays an important role in the development, establishment, and progression of HNSC. Better treatment for HNSC can be achieved by understanding the dysregulation and evasion of immune system. HNSC cells evade the host immune system through manipulation of their immunogenicity, production of immunosuppressive mediators and promotion of immunomodulatory cell types (Ferris 2015).

Tumour cells and their microenvironment are closely related and are in continuous interaction. Initially, immune cells try to eliminate tumour cells, but as the tumour grows, they over-express certain ligands that bind to immune cells and suppress the immune response. For instance, PD-L1 inhibits T cell response through its binding with PD-1, whereas, Galetin-9 modulate T cell function through its binding with TIM-3 (Keir et al. 2008; Akinleye and Rasool 2019; Sakhnevych 2019). Moreover, interaction of immune cells with each other may inhibit immune response, for instance, CTLA4 binds to CD80 and CD86 present on APCs and suppress T cell function (Chikuma 2017). However, there are numerous antibodies against different immune checkpoints, but they exhibit different response in different individuals (Kearney et al. 2016; Marinelli et al. 2018; Knee et al. 2016). Due to the limitations of chemotherapy and radiation therapy, there is a critical need for early detection and prevention of high-risk premalignant lesions.

Natural compounds have a lot of potential for chemoprevention because of their minimal toxicity, and widespread acceptance as dietary supplements (Amin et al. 2009). However, natural compounds also have several drawbacks like low bioavailability and standardized druggability (Gaston et al. 2020). However, combinatorial therapy-based approaches using natural compounds with highly efficient target specificity and improved bioavailability along with a immunopotentiation role may be an answer to the limitations of chemotherapeutics and immunotherapeutics (George et al. 2021). Combination of EGCG, a natural compound found in green tea with TK1 is found to be a novel chemoprotective substance against HNSCC (Masuda et al. 2011). *Lithospermum erythrorhizon* and *Onosma sericeum* have demonstrated their anti-cancer activities under in vivo and in vitro conditions. For instance, *Lithospermum erythrorhizon* with the help of its secreted product that is acetylshikonin inhibits dihydrofolate reductase and hampers autochthonous mammary carcinogenesis in 16HER2 transgenic mice (Wang et al. 2020). Similarly, *Onosma armeniacum* extract exhibits antioxidants properties as demonstrated in HepG2, A549, and WiDr cell culture model (Demir et al. 2021). Recently, green tea polyphenols have demonstrated the potential anticancerous activities, for instance green tea polyphenols modulate the activity of cancer related signalling molecules like VEGF, cyclin D1, and caspase 3 that are involved in the regulation of angiogenesis, cell cycle, and apoptosis (Miyata et al. 2018). In addition, *Persea americana* seeds exhibits anti-oxidants, anti-inflammatory, and anti-cancerous activities in HePG2 cell line and HCT116 cell culture model (Alkhalaf et al. 2019). Further, to assess the efficacy of these chemo-preventive medicines, new biomarkers with predictive value for clinical disease and risk stratification can be employed for more disease-specific strategy.

From our literature survey, we shortlisted drugs of natural origin based on their dual role of anti-cancerous properties and immune modulation, such as mTOR inhibitor (rapamycin) reduces the expression of PD-L1 in HNSC cancer (Zheng et al. 2019). Drugs such as Statin, Metformin, and Anthracyclines can also enhance the immune system and kill tumour cells more effectively (Matsushita and Kawaguchi 2018; Desai et al. 2018; Pereira et al. 2018). Thalidomide and its derivative drugs, such as lenalidomide, were first used as direct anti-cancer drugs due to their cell-cycle arresting properties but later recognized for their role as immunomodulatory drugs due to their ability to stimulate T cells for secretion of IL-2 and interferon-gamma (Liu et al. 2017).

Here we tried to explore those genes, whose expressions are functionally associated with HNSC disease and are involved in immune suppression. We used an unbiased approach for targeting these genes because of their upregulation associated with both immune suppression and tumour progression. Further, we have explored natural compounds that can inhibit these gene functions because natural compounds are cost effective and have fewer side effects as compared to synthetic compounds. These natural compounds have properties that interfere with initiation, development and progression of tumours through various mechanisms including apoptosis, angiogenesis, metastasis, cell proliferation and cell differentiation (Majolo et al. 2019). So, combinatorial use of natural compounds that interfere with multiple pathways, which results in better therapeutic strategies,

can also address the problem of drug resistance and hence may serve as a better therapeutic strategy (Aung et al. 2017).

Methodology and materials required

Data collection

500 genes associated with HNSC were collected from string disease query database (Szklarczyk et al. 2017). Gene expression data for HNSC was retrieved from NCBI's GEO. Data regarding natural compounds were collected from Npcare database (Choi et al. 2017) and literatures. Gene expression data for the natural compounds were retried from NCBI's GEO.

Functional enrichment analysis

Five hundred genes associated with HNSC cancer were imported in Cytoscape from disease query database and functional enrichment analysis with GO Process. Gene ontology (GO) such as biological process, molecular functions, cellular processes, and protein domain analysis associated with these genes were identified. Biological processes involved in immune system were selected and further filtered for the processes involved in immune suppression. Cytoscape is a web tool containing a collection of applications for visualizing molecular level of biological interactions and biological pathways along with added annotations like gene expression profiles, enrichment analysis and other state of data. Cytoscape core distribution provides a basic set of features for data integration, analysis, and visualization achieved by the core distribution of cytoscape. Adj. p-value ≤ 0.05 was considered as the significantly enriched biological processes.

Gene expression analysis and literature exploration

Gene expression analysis was achieved by the *GEO2R*, which is a tool that allow users to compare two or more groups of samples to identify genes that are differentially expressed across various experimental conditions. Differentially expressed genes are presented as a table ordered by p-value and adjusted p-value significance, along with graphic plots to assess data set quality and visualize differentially expressed genes with their p-value and logFC value. GEOquery and limma R packages from the Bioconductor project are used for comparisons on original submitter-supplied processed data tables. Differential genes were selected based on p-value ≤ 0.05, and |logFC value| ≥ 1.

Gene expression data were checked for these 53 genes which were associated with immune system processes in enrichment analysis so that we could select only up-regulated genes. 21 genes were found up-regulated. Literatures were explored regarding these 21 genes for their evidence as tumour promoter and immune suppressor. Out of these 21 genes, 10 genes had enough evidence for both immune suppressor as well as tumour promotor.

Network analysis of selected genes

Selected genes were subjected as a list in STRING which is a database of known and predicted protein–protein interactions. Both physical and functional associations are include in these interactions, they are curated from interactions aggregated from other (primary) databases, from computational prediction and from knowledge transfered between organisms. Genes with highest degree of interaction were selected from these ten genes.

Natural compounds targeting selected genes

Np care database (Choi et al. 2017) was explored for the selection of natural compounds against these selected genes and their transcription factors. NPCARE, a database for Natural Products-Cancer gene Regulation, provides the level of gene expression and the inhibition of cancer cells in various cancer types by the effect of extract and natural compounds. We further explored the literatures for those genes which were not found in the Np care database.

Gene expression analysis of the selected natural compounds effects

Selected natural compounds were searched in the GEO (Gene Expression Omnibus) and their expression data was curated. Differential genes expression analysis was achieved by GEO2R and iDEP (integrated Differential Expression and Pathway analysis). iDEP is a user-friendly interface for bioinformatic analysis of gene-level data for differential analysis and pathway analysis. User can generate reports to analyse RNA-seq datasets from NCBI's GEO which contains differential expression and enrichment analyses etc.

Functional analysis of natural compounds on differentially expressed genes in HNSC cancer

Differential gene (DGs) of HNSC cancer dataset was compared with the differential gene expression data of both the selected compounds for conforming that these compounds reverse the expression of differentially expressed genes. This

conformation was achieved in the Microsoft-excel by comparing their logFC values in different samples.

Biological pathway analysis of the involved genes

Pathway analysis of the genes whose expression were altered by Gallic acid and Shikonin was achieved by Funrich tool. We divided genes into four categories for pathway analysis i.e. HNSC upregulated genes down regulated by Gallic acid, HNSC upregulated genes down regulated by Shikonin, HNSC downregulated genes upregulated by Gallic acid and HNSC downregulated genes upregulated by Shikonin. Later genes involved in these respective categories were analysed for the pathways in which they are involved in Funrich tool individually. Funrich is a stand-alone tool used for performing enrichment analysis on the basis of biological pathways, biological processes, functional, transcription factors etc.

Results and discussion

Enrichment analysis

An enrichment analysis of 500 genes that are known through experimental validation to be the most explicitly associated with HNSC cancer, was conducted and immune-associated genes were selected. 256 genes related to the immune system were found, out of which 53 genes were found to be associated with the negative regulation of immune system associated processes, such as negative regulation of T cell activation, negative regulation of B cell activation, negative regulation of B cell proliferation etc. as listed in Fig. 1A.

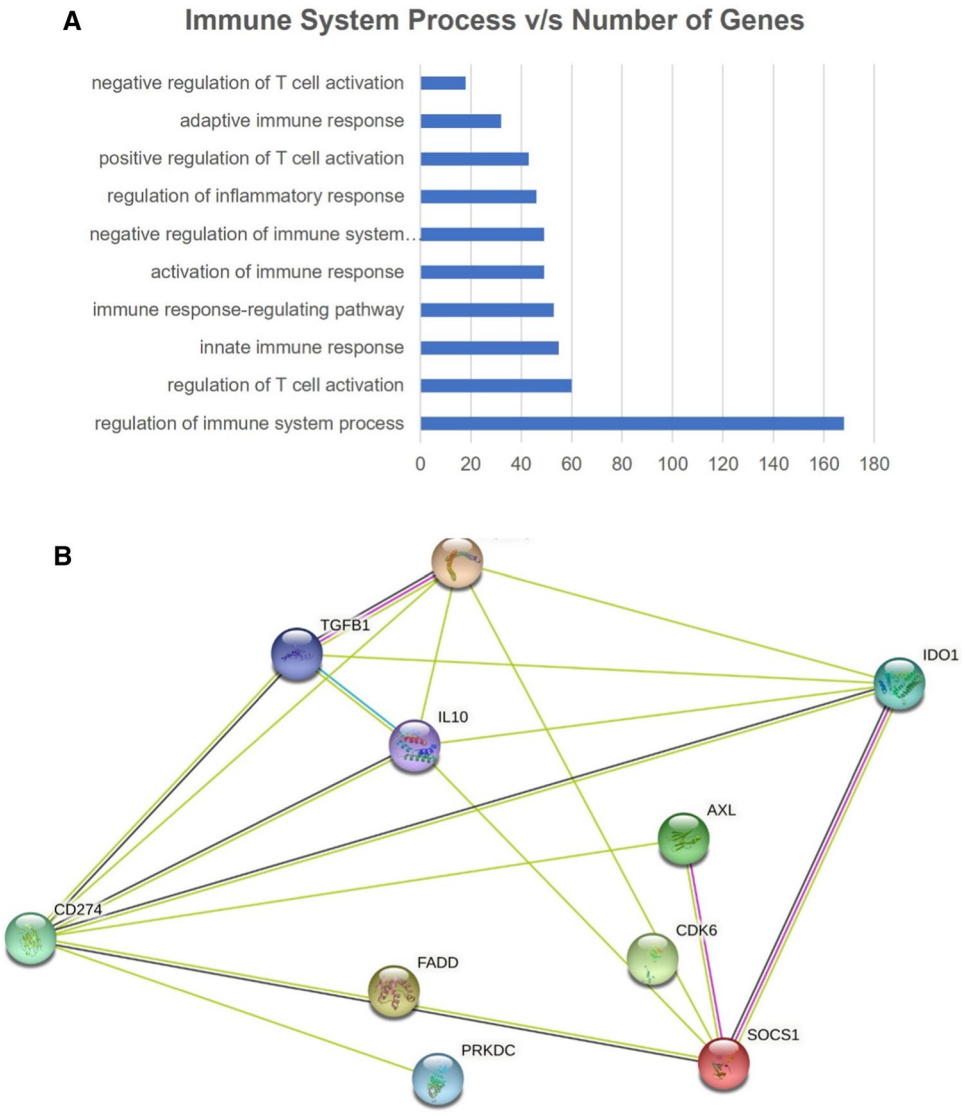


Fig. 1 The following figure demonstrate enrichment analysis data by GO Process involved in immune system related processes which shows number of genes from the HNSC associated involved in different immune system related process no. of gene involved in different immune system related process by enrichment analysis GO Process. **B** interaction network of ten genes which are immune suppressor as well as tumour progressor for checking degree of interaction. As we can see here CD274 have highest degree of interaction 7 and IDO1 have second highest degree of interaction 5

Table 1 List of genes with their tumor progression and immunosuppression roles

S.No	Genes	Tumor progression role	Immunosuppression role	Targeting drugs (FDA approved/in clinical trials)
1	TGFB1	Facilitating EMT (Mahdi et al. 2015; Furler et al. 2018) Metastasis initiation (Calon et al. 2012)	Inhibits anti-tumor responses (T-cell and NK-cell) (Yang et al. 2010)	Lerdelimumab and metelimumab (in clinical trial)
2	IRF1	PD-L1 upregulation in the tumor cell (Shao et al. 2019)		
3	TWSG1	Stimulating tumor-associated angiogenesis and promoting tumor growth and malignant behavior (Xia et al. 2017)		
4	CDK6	Regulates the progression of the cell cycle Plays important role in transcription of tumor angiogenesis pathway (Tadesse et al. 2015)	CDK6 inhibition triggers antitumor immunity (Goel et al. 2017)	Palbociclib (FDA approved)
5	AXL	Tumor proliferation, survival, metastasis, and resistance to cancer therapy (Rankin and Giaccia 2016)	Inhibition of Axl induces tumor immune suppression (Ludwig et al. 2018)	
6	FADD	Mediates cell cycle and proliferation (Papoff et al. 2010)	A negative regulator of T-cell receptor–mediated necroptosis (Osborn et al. 2010)	
7	HAVCR2	Induce EMT by JAK-STAT pathway (Hou et al. n.d.)	Over expression observed in tumor-infiltrating lymphocytes which is further associated with resistance to immunotherapy IL-10 (Huang et al. 2015)	BMS-986258(in clinical trial)
8	PRKDC	Promotes tumor cell growth (Zhang et al. 2019)	A predictive biomarker and target for ICI (Tan et al. 2020)	NU7026 (FDA Approved)
9	IL10	Associated with tumor-related markers such as Bcl-2 (Sheikhpour et al. 2018)	Inhibited T-cell proliferation and function (Turovskaya et al. 2009) TAMs can cause drug resistance via the IL-10 signaling pathway (Sheikhpour et al. 2018; Llanes-Fernández et al. 2006)	GIT 27 (in clinical trials)
10	SOCS1	SOCS1 in downregulated form can inhibit proliferation, via accumulation of G0/G1 phase and reduction of S phase (Zhang et al. 2012)	It inactivated CD8+ T cells mediated anti-tumor response (Chikuma et al. 2017)	
11	MICA		Anti-MICA antibodies can promote the anti-tumor immunity through the induction of direct anti-tumor effects (antibody-dependent cell-mediated cytotoxicity, ADCC) (Torres et al. 2020)	
12	TIGIT		TIGIT suppresses NK cells and CD8+ T cells functionality (Zhou et al. 2018) (Harjunpää and Guillerey 2020)	
13	IDO1	IDO1 expression associated with the progression of tumor (Hornyák et al. 2018) by drug resistance mechanism against response of different drugs (Thaker et al. 2013)	IDO overexpression associated with escape from anti-tumor immunity by suppression of mTOR in T cells (Prendergast et al. 2014)	Indoximod (in clinical trials)
14	CDKN2A	Mutated	Mutated	
15	LAG3		Inactivates the CD4 + T cells Reduces the effector function of CD8 + T cells Promotes the suppressor activity of Tregs (Long et al. 2018)	BI 754,111(in Clinical Trials)
16	CTLA4		Inhibits the activation and proliferation of T cell (Zhao et al. 2018)	Ipilimumab (FDA approved)

Table 1 (continued)

S.No	Genes	Tumor progression role	Immunosuppression role	Targeting drugs (FDA approved/in clinical trials)
17	CD274	Tumor cell growth, migration and invasion using WIP and β-catenin signalling (Yu et al. 2020)	PD-L1 inhibits the T-cell mediated immune response in peripheral tissues (Akinleye and Rasool 2019)	Nivolumab (FDA approved)
18	HLA-A	Highly polymorphic	Highly polymorphic	
19	PDCD1LG2		Overexpression leads to suppression of tumor antigen specific to CD8+ T cells (Tanegashima et al. 2019)	Atezolizumab (FDA approved)
20	FOXP3	FOXP3 overexpression promotes cell proliferation, migration, and invasion (Yang et al. 2017)	FOXP3 is associated with development of Treg cells (Mercer and Unutmaz 2009)	RPG (FDA approved)
21	PDCD1		PDCD1 overexpression suppress the anti-tumor immune regulation (Han et al. 2020)	Avelumab (FDA approved)

Gene expression analysis

Gene expression data was analysed for 53 genes associated with the negative regulation of the immune system, out of which 21 genes were identified for having a LogFc value greater than or equals to 1. These genes could be associated with both immune suppression as well as tumour progression. Complete gene expression data is provided in the supplementary table 1.

These 21 genes were functionally validated by annotation from literatures for their association with tumour progression related process like cell proliferation, metastasis etc. and immune suppression related process like T cell inactivation, development of tumour-associated macrophage etc.

It was found that 10 genes were associated with both of the above-mentioned processes where only 7 genes were associated with immune suppression. 2 genes were associated with tumour progression only and remaining 2 genes were associated with HNSC cancer due to alteration of their function by mutations as shown in Table 1.

Network analysis of selected genes

10 genes were selected that play a crucial role in immune suppression as well as tumour progression. These are Foxp3, CD274, IDO1, IL-10, SOCS1, PRKDC, AXL, CDK6, TGFB1 and FADD. Network analysis in string revealed that CD274 has seven degrees of interaction, IDO1 has five degrees of interactions as shown in the Fig. 1B hence these two genes can serve as the preferred target for modulating immune regulation and will impact multiple immune cells and also determine tumour prognosis. CD274 gene encodes protein PD-L1 which is an immune suppressor ligand. IDO1 expression associates with the progression of tumour (Bilir and Sarisozen 2017). IDO1 expression in tumour cells mediate the catabolism of tryptophan and is a critical factor of immune escape by suppression of anti-tumour immunity (Platten et al. 2012). IDO1 overexpression increases the proliferation of Tregs (Baban et al. 2009).

CD274 regulates tumour growth, proliferation, migration and invasion by targeting WIP along with beta-catenin signalling (Yu et al. 2020). CD274 is a well known immune checkpoint. It helps in tumour cell survival by PD-1/PD-L1 interaction which inhibits T cell activation (Wu et al. 2019) (Fig. 2).

It is expressed in different tissues but mainly expressed in activated T cells and B cells, monocytes, dendritic cells and different tumour cells. The interaction of this ligand with PD1 results in immune escape by the tumour cells, by inhibiting T-cell activation and cytokine production. High expression of this gene is a prognostic marker in many cancers. Indoleamine 2,3-dioxygenase 1 (IDO1) is a heme enzyme that catalyzes the first and rate limiting step in the catabolism

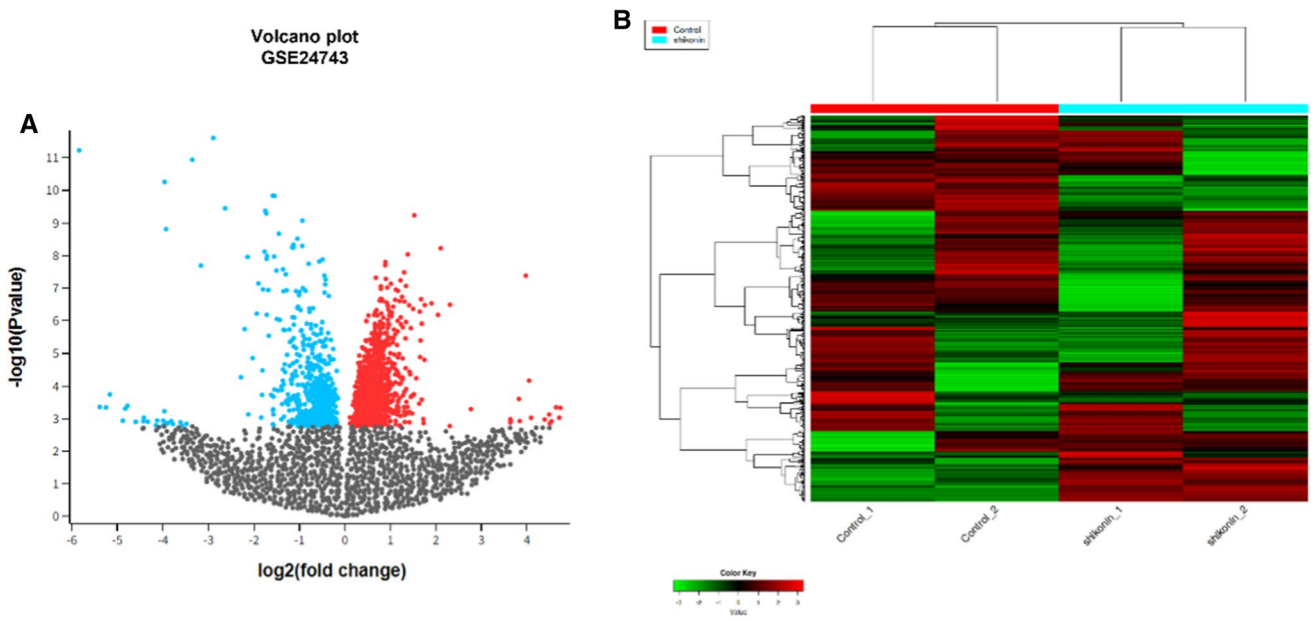


Fig. 2 **A** shows the volcano plot of GEO dataset: GSE24743 Effects of Shikonin on the gene expression, red dots denote the genes which are differentially up-regulated and blue dots denotes the genes which are differentially down-regulated with an adjusted *p*-value less than 0.05. **B** Shows the heatmap of different genes in different control and test samples

Table 2 List of CD274 and IDO1 targeting Natural Compounds and their plant origin

Gene	Natural compounds	Plant origin
CD274	Gallic Acid (3,4,5-trihydroxybenzoic acid)	Banana, walnut, hazelnut, green tea, avocado, guava, mango, mulberry
IDO1	Dihydrotanshinone I	Salvia miltiorrhiza
IDO1	Shikonin	*Lithospermum erythrorhizon*, Alkanna, Arnebia, Onosma, Onosma sericeum Willd
IDO1	9-O-demethyltrigonostemone	Strophioblachia fimbricalyx
CD274	Fisetin	Strawberries, apples, persimmons, onions and cucumbers
CD274	Cosmosiin	Teucrium gnaphalodes
CD274	Kaempferol	Kale, beans, tea, spinach, and broccoli

of tryptophan which changes the behaviour of T-cells. This enzyme plays a role in variety of pathophysiological processes such as antioxidant activity, antimicrobial defence, neuropathology, immunoregulation, antitumour defence. Overexpression of IDO1 is found in different cancers, which is associated with poor prognosis. IDO1 can be inhibited by the cancer-suppression gene bridging integrator 1 (Bin1) and up-regulated by some immune checkpoint molecules and cytokines such as IFN-γ, pathogen-associated molecular patterns (PAMPs, such as Toll-like receptor (TLR) 3, TLR4, TLR7, TLR8, and TLR9), IL-6, prostaglandin E2 (PGE2), damage-associated molecular patterns (DAMPs), immune checkpoint (including PD-1, glucocorticoid-induced TNF receptor-related protein (GITRL), CTLA-4), and TNF-α along with TGF-β to establish an immunosuppressive environment.

Natural compounds selection

We selected natural compound against two selected genes. Gallic Acid (3,4,5-trihydroxybenzoic acid) was inhibitory against CD274 while three compounds, dihydrotanshinone I, shikonin, and 9-O-demethyltrigonostemone were inhibitory against IDO1. Gallic acid is a phenolic acid which is found in sumac, gallnuts, tea leaves, oak bark, witch hazel and other plants. Dihydrotanshinone I(DI) is a natural compound

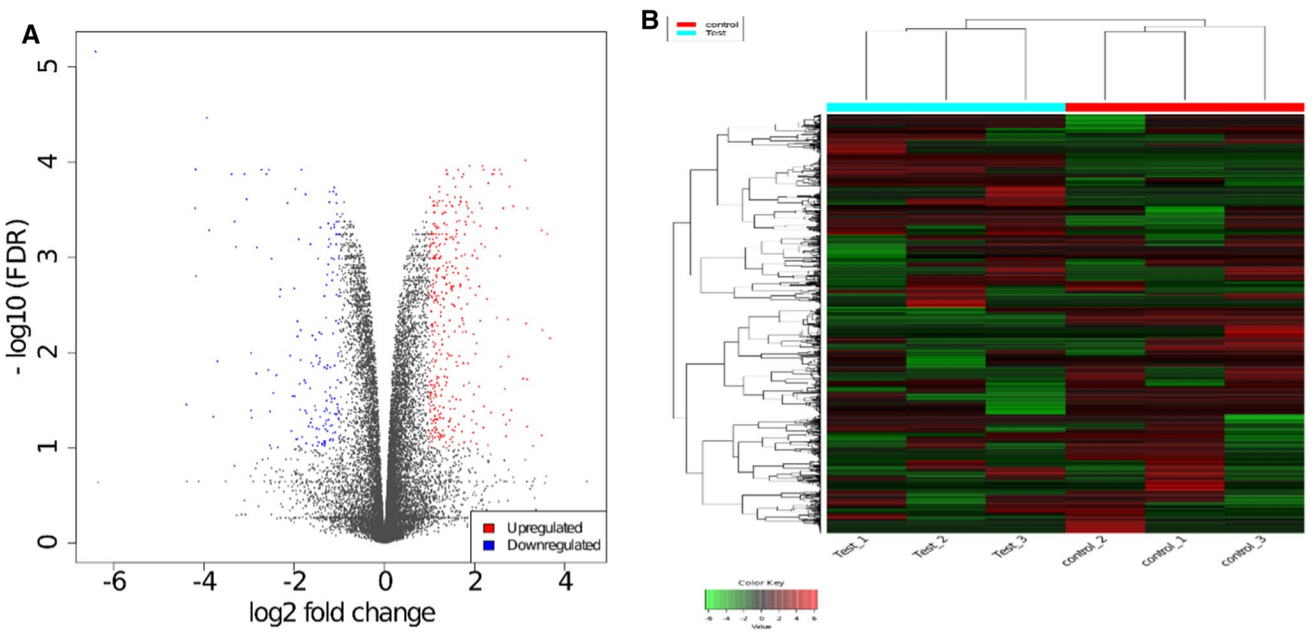
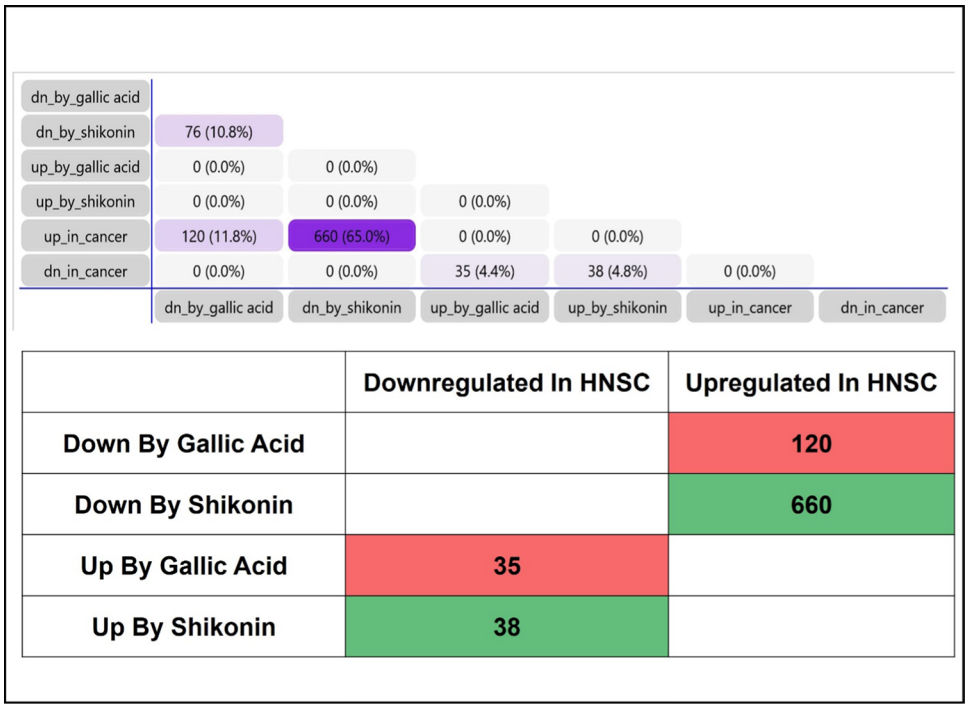


Fig. 3 The **A** shows the volcano plot of GEO dataset: Effects of Gallic acid on the gene expression, blue dots denote the genes which are differentially up-regulated and red dots denotes the genes which are differentially down-regulated with an adjusted *p*-value less than 0.05. **B** Shows the heatmap of different genes in different control and test samples



	Downregulated In HNSC	Upregulated In HNSC
Down By Gallic Acid		120
Down By Shikonin		660
Up By Gallic Acid	35	
Up By Shikonin	38	

Fig. 4 Figure demonstrate that expression of no. of genes altered by Gallic acid and Shikonin from the differentially expressed genes. Gallic acid and Shikonin downregulates 120 genes and 660 genes, respectively that are upregulated in HNSC, whereas Gallic acid and Shikonin upregulates 35 and 38 genes, respectively that are downregulates in HNSC

found in the salvia miltiorrhiza which is a Chinese medicinal plant. It has been reported to have cytotoxicity to variety of tumours. Shikonin is a naphthoquinone compound which is found in the roots of Shikonin plant (*Lithospermum erythrorhizon*) and used as a traditional Chinese medicine. 9-O-demethyltrigonostemone is a natural compound found

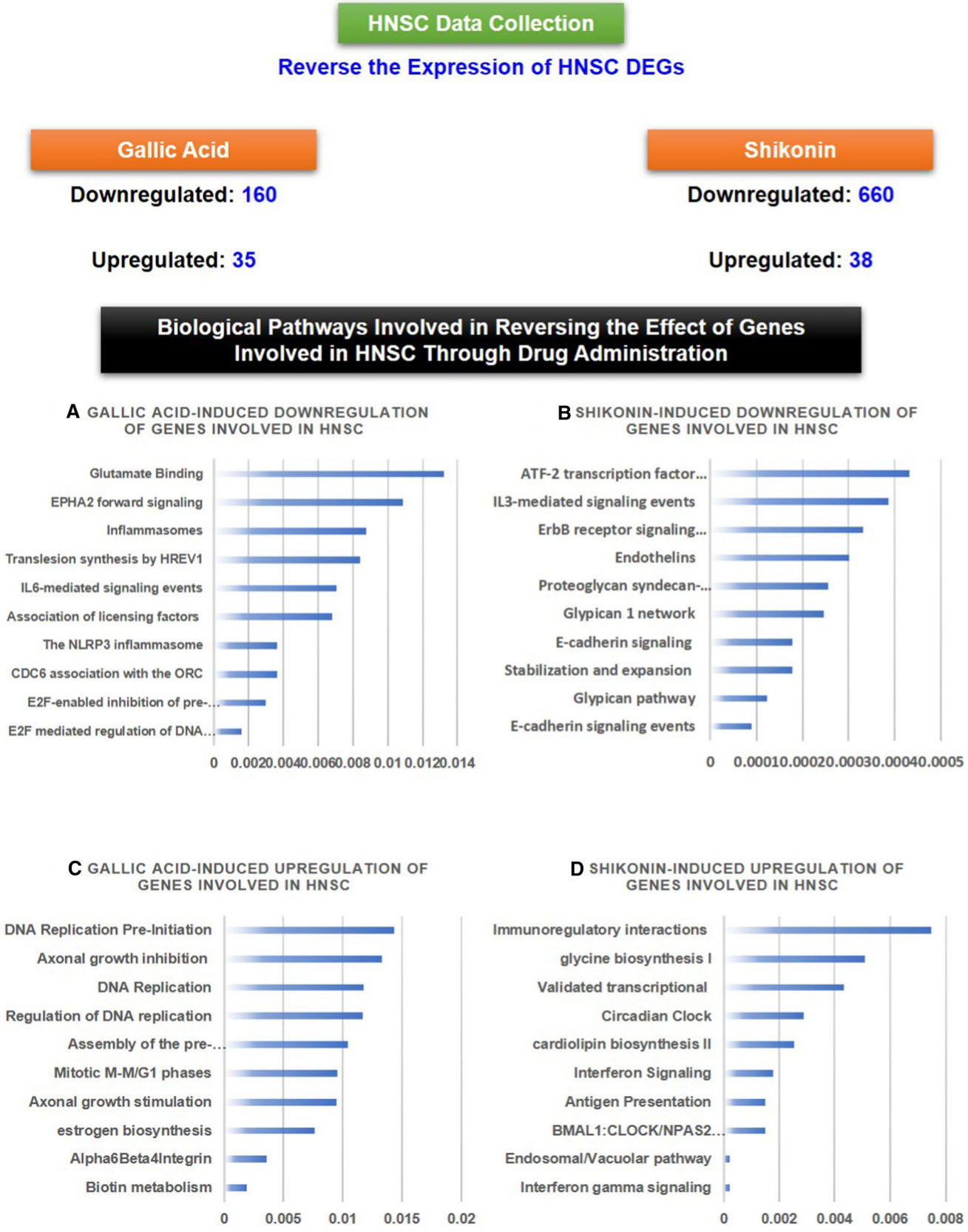
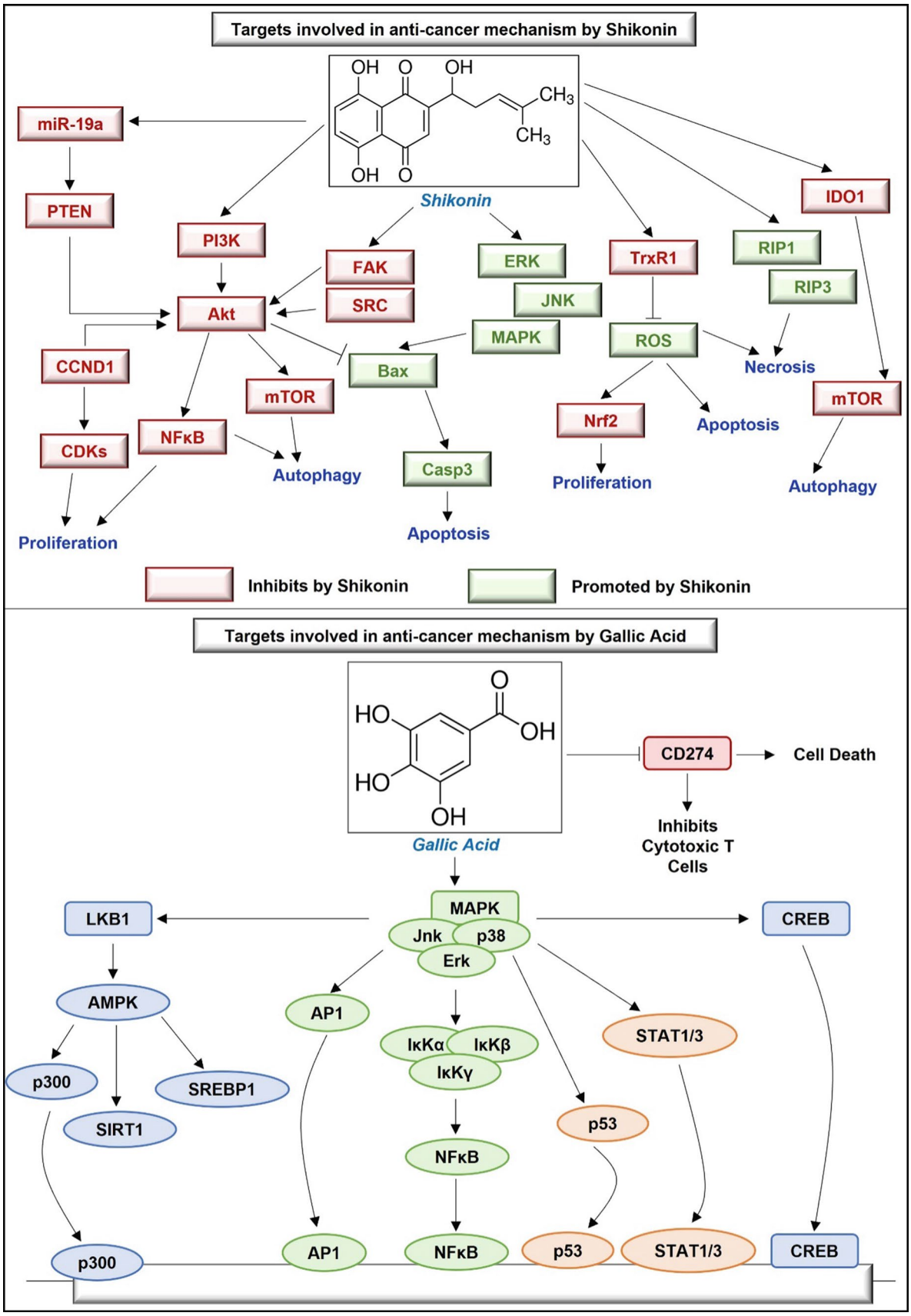


Fig. 5 The figure demonstrates the pathway analysis of differential genes involved in HNSC and give comparative account of the genes whose expressions are reverse by the action of Gallic acid and Shikonin, respectively


Targets involved in anti-cancer mechanism by Shikonin
Shikonin
miR-19a
PTEN
PI3K
Akt
CCND1
CDKs
NFκB
mTOR
Autophagy
Proliferation
FAK
SRC
ERK
JNK
MAPK
Bax
Casp3
Apoptosis
TrxR1
ROS
Nrf2
Proliferation
Apoptosis
RIP1
RIP3
Necrosis
IDO1
mTOR
Autophagy
Inhibits by Shikonin
Promoted by Shikonin
Targets involved in anti-cancer mechanism by Gallic Acid
Gallic Acid
CD274
Cell Death
Inhibits Cytotoxic T Cells
MAPK
Jnk
p38
Erk
LKB1
AMPK
p300
SIRT1
SREBP1
AP1
IκKα
IκKβ
IκKγ
NFκB
p53
STAT1/3
CREB

◂**Fig. 6** Signaling molecules regulated by two phytochemicals, namely Shikonin and Gallic acid. For example, Shikonin Inhibits PI3K, mir-19a, FAK, TrxR1, IDO1 that regulates cell proliferation, autophagy, and apoptosis. In addition, Shikonin upregulates the activity of Erk, Casp3, and RIP1 that are involved in the regulation of autophagy, apoptosis, proliferation, and necrosis. Moreover, Gallic acid targets MAPKs, namely Jnk, p38, and Erk, which regulates the transcriptional status of various signaling molecules, namely CREB, STAT1/3, LKB1, AP1, and p53

in the roots of Strophioblachia fimbricalyx which shows cytotoxic activity against different tumours (Table 2).

We searched NCBI's GEO for four selected compounds and found experimental data corroborating functional inhibitory characteristics of two compounds which are gallic acid and shikonin. Dataset accession number for Gallic Acid is GSE158788. Gene Expression Profile Analysis of Gallic Acid-induced cell death process using Hela cells treated with gallic acid (50 µg/ml) for 0 h (GA0hr), 2 h (GA2hr), 4 h (GA4hr), 6 h (GA6hr), and 9 h (GA9hr) were studied. Dataset accession number for Shikonin is GSE24743 and its effect on the gene expression of human lymphoma U937 cells was studied (Gomes et al. 2003). In this dataset U937 cells were treated with 100 Nm shikonin and followed by incubation for 3 h at 37 °C. The cells treated with dimethyl sulfoxide served as control. The microarray dataset was analysed with Geo2R tool and the results have been shown in the supplementary Table 2. We analysed GSE158788 data with Idep tool for differential gene expression and results are shown in Fig. 3

Expression data of HNSC cancer and both of the natural compounds merged in Microsoft-excel were studied and filtered expression of HNSC cancer with a cut of greater than one. 1016 genes were found differentially over-expressed in HNSC cancer and were compared with the expression data of gallic acid and shikonin. Gallic acid results in downregulation of 120 of these differentially over-expressed genes and shikonin down-regulates 660 genes from these 1016 over-expressed genes. Again, for down regulated gene, expression data of HNSC cancer filtered with a cut off less than or equals to -1 was used. 795 genes were found differentially down expressed in the HNSC cancer and were compared with the expression data of gallic acid and shikonin. Gallic acid results in upregulation of 35 genes and shikonin up-regulates 38 genes from these 795 down-regulated genes. So, this combination of gallic acid and shikonin could be effective for the HNSC cancer treatment (Fig. 4).

Gallic acid shows anti-cancerous activity by its selective cell death effect in various cancer cells but not in normal cells (Gomes et al. 2003). The molecular targets and functions of gallic acid are activation of NF-B inhibition, ATM kinase, UDP-glucose dehydrogenase inhibition, apoptosis induction, Ribonucleotide reductase inhibition, Cyclooxygenase inhibition, GSH depletion and invasion inhibition (Verma et al. 2013). Shikonin deregulates the cellular Ca^{2+} and ROS levels in the mitochondria which leads to breakdown of mitochondrial membrane potential, dysfunction of microtubules, cell-cycle arrest, and ultimately results in induction of apoptosis. Figure 5 shows a comparative account for the effect of both Gallic acid and Shikonin in the multitude of pathways.

Funrich tool was used for the pathway analysis of regulated genes by gallic acid and shikonin from the differentially expressed genes. Results show genes getting downregulated by gallic acid involved in pathways like glutamate binding, inflammasomes, translesion synthesis by HREV1, IL6-mediated signaling events, association of licensing factor, the NLRP3 inflammasome, CDC6 association with the ORC pathways and EPHA2 forward signalling. Genes upregulated by gallic acid are involved in pathways like axonal growth inhibition, DNA replication pre-initiation, axonal growth stimulation, estrogen biosynthesis, biotin metabolism etc. Genes downregulated by shikonin are involved in pathways like ATF-2 transcription factor, IL3-mediated signalling events, ErbB receptor signalling, endothelins, E-cadherin signalling, stabilization and expansion, glypican pathway, E-cadherin signalling events etc. Genes upregulated by shikonin were involved in pathways like immunoregulatory interactions, glycine biosynthesis, validated transcriptional, circadian clock, interferon signalling, antigen presentation etc. (Fig. 6).

Hence, the genes involved in HNSC which contributed to cancer prognosis were shown to be regulated by the natural compounds that can potentially impact cancer progression and immunity related pathways. Therefore, it can be stated that the combination of gallic acid and shikonin could be beneficial for the combinatorial treatment of HNSC cancer.

Many plants, including *Lithospermum erythrorhizon*, Alkanna, Arnebia, Onosma, *Onosma sericeum* Willd, and Echium generate shikonin and research have previously shown that shikonin regulates various functions in these plants, including transgene expression (Yazaki 2017). Shikonin has been used as a red dye for centuries and is reported to possess medicinal properties. It was evaluated as a multi-functional antibacterial and UV protective agent on a silk fabric (Dhandapani and Sarkar 2007), exhibits insulin-like activities by inhibiting phosphatase and tensin homologue deleted on Chromosome 10 (PTEN) (Nigorikawa et al. 2006). Further, the drug has shown various properties, such as anti-viral, anti-tumor, cardiotonic and contraceptive properties (Sharma et al. 2009). Similarly, gallic acid is found in many food sources like banana, walnut, hazelnut, green tea, avocado, guava, mango, mulberry, pomegranate, blackcurrant, cashew, red wine, strawberry, blueberry, apple, grape etc. (Zeb 2021). Gallic acid, is a typical antioxidant tea formulation, and thus considered as

potential natural antioxidant (Xu et al. 2017). Moreover, Gallic acid in addition to its phytochemical activity is also utilised in tanning, ink colours, and paper manufacturing (Valanciene et al. 2020). Gallic acid, commonly known as 3,4,5-trihydroxybenzoic acid, is a phenolic chemical, which can be found both in its free form and as a component of tannins, specifically gallotannin (Al Zahrani et al. 2020). Additionally, gallic acid and its derivatives can be found in almost all parts of the plant, including the bark, wood, leaf, fruit, root, and seed (Daglia et al. 2014).

Conclusion

Our work explores natural compounds that have been shown to interact with key modulators in multiple pathways which may influence tumorigenesis. Our research aims to identify such compounds that have significant immunomodulatory role along with anti-tumour effects. So, we first identified those genes which have potential of tumour progression as well as immunity suppression followed by natural compounds capable of checking their expression data. Gene expression data of natural compounds i.e. gallic acid and shikonin was compared with the differentially expressed genes in HNSC cancer. Expression data of these compounds showed that gallic acid downregulates 120 genes, shikonin downregulates 660 genes that were upregulated in HNSC cancers. Moreover, gallic acid upregulates 48 genes and shikonin upregulates 38 genes that were downregulated in HNSC Cancer. So, this combination of gallic acid and shikonin could be beneficial for the HNSC cancer treatment. We suggest that compounds that can cause immunomodulation along with inhibition of tumour progression would be particularly effective in combinatorial immunotherapy. Combinatorial immunotherapy is a promising therapeutics strategy that would be most effective in cancer therapeutics.

The natural compounds studied possess role of both immune modulators as well as anti-tumour function. There are very limited number of natural compound available for targeting genes which are involved in both immune suppression as well as tumour progression and this necessitates our studies that will pave the way for novel combinatorials. So, these natural compounds could be used in combination with other therapies or with each other for effective treatment of HNSC cancer. Further, we have to study the compounds in clinical studies and in-vitro biopsy samples to ascertain the full ramifications of our results.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s42535-022-00363-w.

Funding This study was funded by Delhi Technological University.

References

Akinleye A, Rasool Z (2019) Immune checkpoint inhibitors of PD-L1 as cancer therapeutics. J Hematol Oncol 12(1):92. https://doi.org/10.1186/s13045-019-0779-5 (**BioMed Central Ltd.**)

Al Zahrani NA, El-Shishtawy RM, Asiri AM (2020) Recent developments of gallic acid derivatives and their hybrids in medicinal chemistry: a review. Eur J Med Chem 204:112609. https://doi.org/10.1016/J.EJMECH.2020.112609

Alfarouk KO, Stock CM, Taylor S, Walsh M, Muddathir AK, Verduzco D, Bashir AHH, Mohammed OY, Elhassan GO, Harguindey S, Reshkin SJ, Ibrahim ME, Rauch C (2015) Resistance to cancer chemotherapy: failure in drug response from ADME to P-gp. Cancer Cell Int 15(1):71. https://doi.org/10.1186/s12935-015-0221-1 (**BioMed Central Ltd.**)

Alkhalaf MI, Alansari WS, Ibrahim EA, ELhalwagy MEA (2019) Antioxidant, anti-inflammatory and anti-cancer activities of avocado (*Persea americana*) fruit and seed extract. J King Saud Univ Sci 31(4):1358–1362. https://doi.org/10.1016/J.JKSUS.2018.10.010

Amin ARMR, Kucuk O, Khuri FR, Shin DM (2009) Perspectives for cancer prevention with natural compounds. J Clin Oncol off J Am Soc Clin Oncol 27(16):2712–2725. https://doi.org/10.1200/JCO.2008.20.6235

Aung TN, Qu Z, Kortschak RD, Adelson DL (2017) Understanding the effectiveness of natural compound mixtures in cancer through their molecular mode of action. Int J Mol Sci. https://doi.org/10.3390/ijms18030656 (**MDPI AG**)

Baban B, Chandler PR, Sharma MD, Pihkala J, Koni PA, Munn DH, Mellor AL (2009) IDO activates regulatory T cells and blocks their conversion into Th17-like T cells. J Immunol 183(4):2475–2483. https://doi.org/10.4049/jimmunol.0900986

Bilir C, Sarisozen C (2017) Indoleamine 2,3-dioxygenase (IDO): only an enzyme or a checkpoint controller? J Oncol Sci 3(2):52–56. https://doi.org/10.1016/j.jons.2017.04.001

Bozic I, Nowak MA (2014) Timing and heterogeneity of mutations associated with drug resistance in metastatic cancers. Proc Natl Acad Sci USA 111(45):15964–15968. https://doi.org/10.1073/pnas.1412075111

Calon A, Espinet E, Palomo-Ponce S, Tauriello DVF, Iglesias M, Céspedes MV, Sevillano M, Nadal C, Jung P, Zhang XHF, Byrom D, Riera A, Rossell D, Mangues R, Massagué J, Sancho E, Batlle E (2012) Dependency of colorectal cancer on a TGF-β-driven program in stromal cells for metastasis initiation. Cancer Cell 22(5):571–584. https://doi.org/10.1016/j.ccr.2012.08.013

Chen L, Zeng Y, Zhou S-F (2018) Role of apoptosis in cancer resistance to chemotherapy. Curr Underst Apoptosis Program Cell Death. https://doi.org/10.5772/intechopen.80056 (**InTech**)

Chikuma S (2017) CTLA-4, an essential immune-checkpoint for T-cell activation. Curr Top Microbiol Immunol 410:99–126. https://doi.org/10.1007/82_2017_61 (**Springer Verlag**)

Chikuma S, Kanamori M, Mise-Omata S, Yoshimura A (2017) Suppressors of cytokine signaling: Potential immune checkpoint molecules for cancer immunotherapy. Cancer Sci 108(4):574–580. https://doi.org/10.1111/cas.13194 (**Blackwell Publishing Ltd.**)

Choi H, Cho SY, Pak HJ, Kim Y, Choi J, Lee YJ, Gong BH, Kang YS, Han T, Choi G, Cho Y, Lee S, Ryoo D, Park H (2017) NPCARE: database of natural products and fractional extracts for cancer regulation. J Cheminformatics 9(1):2. https://doi.org/10.1186/s13321-016-0188-5

Daglia M, Lorenzo A, Nabavi S, Talas Z, Nabavi S (2014) Polyphenols: well beyond the antioxidant capacity: gallic acid and related compounds as neuroprotective agents: you are what you eat! Curr Pharm Biotechnol 15(4):362–372. https://doi.org/10.2174/1389201015014081408251207 37

Demir S, Ayazoglu Demir E, Turan I, Ozgen U (2021) Evaluation of cytotoxic effect of onosma armeniacum extract on various cancer cells. KSU J Agric Nat 24(2):252–259. https://doi.org/10.18016/ksutarimdoga.vi.729814

Deng J, Wang ES, Jenkins RW, Li S, Dries R, Yates K, Chhabra S, Huang W, Liu H, Aref AR, Ivanova E, Paweletz CP, Bowden M, Zhou CW, Herter-Sprie GS, Sorrentino JA, Bisi JE, Lizotte PH, Merlino AA, Wong KK (2018) CDK4/6 inhibition augments anti-tumor immunity by enhancing T-cell activation. Cancer Discov 8(2):216–233. https://doi.org/10.1158/2159-8290.CD-17-0915

Desai P, Wallace R, Anderson ML, Howard BV, Ray R, Wu C, Safford M, Martin LW, Schlecht N, Liu S, Cirillo D, Jay A, Manson JAE, Simon MS (2018) An analysis of the effect of statins on the risk of Non-Hodgkin's Lymphoma in the Women's Health Initiative cohort. Cancer Med 7(5):2121–2130. https://doi.org/10.1002/cam4.1368

Dhandapani R, Sarkar AK (2007) Antibacterial activity and uv property of shikonin on silk substrate. J Text Appar Technol Manag 5(4):1–7

Ferris RL (2015) Immunology and immunotherapy of head and neck cancer. J Clin Oncol 33(29):3293–3304. https://doi.org/10.1200/JCO.2015.61.1509

Furler RL, Nixon DF, Brantner CA, Popratiloff A, Uittenbogaart CH (2018) TGF-β sustains tumor progression through biochemical and mechanical signal transduction. Cancers. https://doi.org/10.3390/cancers10060199 (**MDPI AG**)

Gaston TE, Mendrick DL, Paine MF, Roe AL, Yeung CK (2020) "Natural" is not synonymous with "Safe": toxicity of natural products alone and in combination with pharmaceutical agents. Regul Toxicol Pharmacol RTP. https://doi.org/10.1016/J.YRTPH.2020.104642

George BP, Chandran R, Abrahamse H (2021) Role of phytochemicals in cancer chemoprevention: insights. Antioxidants 10(9):1455. https://doi.org/10.3390/ANTIOX10091455

Goel S, Decristo MJ, Watt AC, Brinjones H, Sceneay J, Li BB, Khan N, Ubellacker JM, Xie S, Metzger-Filho O, Hoog J, Ellis MJ, Ma CX, Ramm S, Krop IE, Winer EP, Roberts TM, Kim HJ, McAllister SS, Zhao JJ (2017) CDK4/6 inhibition triggers anti-tumour immunity. Nature 548(7668):471–475. https://doi.org/10.1038/nature23465

Gomes CA, Girão da Cruz T, Andrade JL, Milhazes N, Borges F, Marques MPM (2003) Anticancer activity of phenolic acids of natural or synthetic origin: a structure−activity study. J Med Chem 46(25):5395–5401. https://doi.org/10.1021/jm030956v

Han Y, Liu D, Li L (2020) PD-1/PD-L1 pathway: current researches in cancer. Am J Cancer Res 10(3): 727–742. http://www.ncbi.nlm.nih.gov/pubmed/32266087

Harjunpää H, Guillerey C (2020) TIGIT as an emerging immune checkpoint. Clin Exp Immunol 200(2):108–119. https://doi.org/10.1111/cei.13407

Hornyák L, Dobos N, Koncz G, Karányi Z, Páll D, Szabó Z, Halmos G, Székvölgyi L (2018) The role of indoleamine-2,3-dioxygenase in cancer development, diagnostics, and therapy. Front Immunol 9:31. https://doi.org/10.3389/fimmu.2018.00151 (**Frontiers Media S.A**)

Huang YH, Zhu C, Kondo Y, Anderson AC, Gandhi A, Russell A, Dougan SK, Petersen BS, Melum E, Pertel T, Clayton KL, Raab M, Chen Q, Beauchemin N, Yazaki PJ, Pyzik M, Ostrowski MA, Glickman JN, Rudd CE, Blumberg RS (2015) CEACAM1 regulates TIM-3-mediated tolerance and exhaustion. Nature 517(7534):386–390. https://doi.org/10.1038/nature13848

Kearney CJ, Ramsbottom KM, Voskoboinik I, Darcy PK, Oliaro J (2016) Loss of DNAM-1 ligand expression by acute myeloid leukemia cells renders them resistant to NK cell killing. OncoImmunology. https://doi.org/10.1080/2162402X.2016.1196308

Keir ME, Butte MJ, Freeman GJ, Sharpe AH (2008) PD-1 and its ligands in tolerance and immunity. Annu Rev Immunol 26(1):677–704. https://doi.org/10.1146/annurev.immunol.26.021607.090331

Knee DA, Hewes B, Brogdon JL (2016) Rationale for anti-GITR cancer immunotherapy. Eur J Cancer 67:1–10. https://doi.org/10.1016/j.ejca.2016.06.028 (**Elsevier Ltd.**)

Liu T, Guo F, Zhu X, He X, Xie L (2017) Thalidomide and its analogues: a review of the potential for immunomodulation of fibrosis diseases and opthalmopathy. Exp Ther Med 14(6):5251–5257. https://doi.org/10.3892/etm.2017.5209 (**Spandidos Publications**)

Llanes-Fernández L, Álvarez-Goyanes RI, del Arango-Prado MC, Alcocer-González JM, Mojarrieta JC, Pérez XE, López MO, Odio SF, Camacho-Rodríguez R, Guerra-Yi ME, Madrid-Marina V, Tamez-Guerra R, Rodríguez-Padilla C (2006) Relationship between IL-10 and tumor markers in breast cancer patients. Breast 15(4):482–489. https://doi.org/10.1016/j.breast.2005.09.012

Long L, Zhang X, Chen F, Pan Q, Phiphatwatchara P, Zeng Y, Chen H (2018) The promising immune checkpoint LAG-3: From tumor microenvironment to cancer immunotherapy. Genes Cancer 9(5–6):176–189. https://doi.org/10.18632/genesandcancer.180

Ludwig KF, Du W, Sorrelle NB, Wnuk-Lipinska K, Topalovski M, Toombs JE, Cruz VH, Yabuuchi S, Rajeshkumar NV, Maitra A, Lorens JB, Brekken RA (2018) Small-molecule inhibition of Axl targets tumor immune suppression and enhances chemotherapy in pancreatic cancer. Can Res 78(1):246–255. https://doi.org/10.1158/0008-5472.CAN-17-1973

Mahdi SHA, Cheng H, Li J, Feng R (2015) The effect of TGF-beta-induced epithelial-mesenchymal transition on the expression of intracellular calcium-handling proteins in T47D and MCF-7 human breast cancer cells. Arch Biochem Biophys 583:18–26. https://doi.org/10.1016/j.abb.2015.07.008

Majolo F, de Oliveira Becker Delwing LK, Marmitt DJ, Bustamante-Filho IC, Goettert MI (2019) Medicinal plants and bioactive natural compounds for cancer treatment: Important advances for drug discovery. Phytochem Lett 31:196–207. https://doi.org/10.1016/j.phytol.2019.04.003

Marinelli O, Nabissi M, Morelli MB, Torquati L, Amantini C, Santoni G (2018) ICOS-L as a potential therapeutic target for cancer immunotherapy. Curr Protein Pept Sci 19(11):1107–1113. https://doi.org/10.2174/1389203719666180608093913

Masuda M, Wakasaki T, Toh S, Shimizu M, Adachi S (2011) Chemoprevention of head and neck cancer by green tea extract: EGCG-the role of EGFR signaling and "lipid raft." J Oncol. https://doi.org/10.1155/2011/540148

Matsushita M, Kawaguchi M (2018) Immunomodulatory effects of drugs for effective cancer immunotherapy. J Oncol. https://doi.org/10.1155/2018/8653489

Mercer F, Unutmaz D (2009) The biology of FoxP3: AKey player in immune suppression during infections, autoimmune diseases and cancer. Adv Exp Med Biol 665:47–59. https://doi.org/10.1007/978-1-4419-1599-3_4

Miyata Y, ScienMatsuoces T, Araki K, Nakamura Y, Sagara Y, Ohba K, Sakai H (2018) Anticancer effects of green tea and the underlying molecular mechanisms in bladder cancer. Medicines 5(3):87. https://doi.org/10.3390/MEDICINES5030087

Nigorikawa K, Yoshikawa K, Sasaki T, Iida E, Tsukamoto M, Murakami H, Maehama T, Hazeki K, Hazeki O (2006) A naphthoquinone derivative, shikonin, has insulin-like actions by inhibiting both phosphatase and tensin homolog deleted on chromosome 10 and tyrosine phosphatases. Mol Pharmacol 70(3):1143–1149. https://doi.org/10.1124/MOL.106.025809

Osborn SL, Diehl G, Han SJ, Xue L, Kurd N, Hsieh K, Cado D, Robey EA, Winoto A (2010) Fas-associated death domain (FADD) is a negative regulator of T-cell receptor-mediated necroptosis. Proc

Natl Acad Sci USA 107(29):13034–13039. https://doi.org/10.1073/pnas.1005997107

Papoff G, Trivieri N, Crielesi R, Ruberti F, Marsilio S, Ruberti G (2010) FADD-calmodulin interaction: a novel player in cell cycle regulation. Biochim Et Biophys Acta Mol Cell Res 1803(8):898–911. https://doi.org/10.1016/j.bbamcr.2010.04.006

Pereira FV, Melo ACL, Low JS, de Castro ÍA, Braga TT, Almeida DC, de Lima AGUB, Hiyane MI, Correa-Costa M, Andrade-Oliveira V, Origassa CST, Pereira RM, Kaech SM, Rodrigues EG, Câmara NOS (2018) Metformin exerts antitumor activity via induction of multiple death pathways in tumor cells and activation of a protective immune response. Oncotarget 9(40):25808–25825. https://doi.org/10.18632/oncotarget.25380

Platten M, Wick W, Van Den Eynde BJ (2012) Tryptophan catabolism in cancer: beyond IDO and tryptophan depletion. Cancer Res 72(21):5435–5440. https://doi.org/10.1158/0008-5472.CAN-12-0569

Prendergast GC, Smith C, Thomas S, Mandik-Nayak L, Laury-Kleintop L, Metz R, Muller AJ (2014) Indoleamine 2,3-dioxygenase pathways of pathogenic inflammation and immune escape in cancer. Cancer Immunol Immunother 63(7):721–735. https://doi.org/10.1007/s00262-014-1549-4 (**Springer Science and Business Media Deutschland GmbH**)

Rankin EB, Giaccia AJ (2016) The receptor tyrosine kinase AXL in cancer progression. Cancers. https://doi.org/10.3390/cancers8110103 (**MDPI AG**)

Sakhnevych S (2019) Tim-3-galectin-9 immunosuppressive pathway in human acute myeloid leukaemia and solid tumour cells and biochemical functions of its crucial components

Salehan MR, Morse HR (2013) DNA damage repair and tolerance: a role in chemotherapeutic drug resistance. Br J Biomed Sci 70(1):31–40. https://doi.org/10.1080/09674845.2013.11669927 (**Step Publishing Ltd.**)

Shao L, Hou W, Scharping NE, Vendetti FP, Srivastava R, Roy CN, Menk AV, Wang Y, Chauvin JM, Karukonda P, Thorne SH, Hornung V, Zarour HM, Bakkenist CJ, Delgoffe GM, Sarkar SN (2019) IRF1 inhibits antitumor immunity through the upregulation of PD-L1 in the tumor cell. Cancer Immunol Res 7(8):1258–1266. https://doi.org/10.1158/2326-6066.CIR-18-0711

Sharma RA, Singh B, Singh D, Chandrawat P (2009) Ethnomedicinal, pharmacological properties and chemistry of some medicinal plants of Boraginaceae in India. J Med Plants Res 3(13): 1153–1175. http://www.academicjournals.org/JMPR

Sheikhpour E, Noorbakhsh P, Foroughi E, Farahnak S, Nasiri R, Neamatzadeh H (2018) A survey on the role of interleukin-10 in breast cancer: a narrative. Rep Biochem Mol Biol 7(1): 30–37. http://www.ncbi.nlm.nih.gov/pubmed/30324115

Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ, Von Mering C (2017) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucl Acids Res 45(D1):D362–D368. https://doi.org/10.1093/nar/gkw937

Tadesse S, Yu M, Kumarasiri M, Le BT, Wang S (2015) Targeting CDK6 in cancer: State of the art and new insights. Cell Cycle 14(20):3220–3230. https://doi.org/10.1080/15384101.2015.1084445 (**Taylor and Francis Inc.**)

Tan KT, Yeh CN, Chang YC, Cheng JH, Fang WL, Yeh YC, Wang YC, Hsu DSS, Wu CE, Lai JI, Chang PMH, Chen MH, Lu ML, Chen SJ, Chao Y, Hsiao M, Chen MH (2020) PRKDC: new biomarker and drug target for checkpoint blockade immunotherapy. J Immunother Cancer. https://doi.org/10.1136/jitc-2019-000485

Tanegashima T, Togashi Y, Azuma K, Kawahara A, Ideguchi K, Sugiyama D, Kinoshita F, Akiba J, Kashiwagi E, Takeuchi A, Irie T, Tatsugami K, Hoshino T, Eto M, Nishikawa H (2019) Immune suppression by PD-L2 against spontaneous and treatment-related antitumor immunity. Clin Cancer Res 25(15):4808–4819. https://doi.org/10.1158/1078-0432.CCR-18-3991

Thaker AI, Rao MS, Bishnupuri KS, Kerr TA, Foster L, Marinshaw JM, Newberry RD, Stenson WF, Ciorba MA (2013) IDO1 metabolites activate β-catenin signaling to promote cancer cell proliferation and colon tumorigenesis in mice. Gastroenterology. https://doi.org/10.1053/j.gastro.2013.05.002

Tímár J, Csuka O, Remenár E, Répássy G, Kásler M (2005) Progression of head and neck squamous cell cancer. Cancer Metastasis Rev 24:107–127

Torres N, Regge MV, Secchiari F, Friedrich AD, Spallanzani RG, Raffo Iraolagoitia XL, Núñez SY, Sierra JM, Ziblat A, Santilli MC, Gilio N, Almada E, Lauche C, Pardo R, Domaica CI, Fuertes MB, Madauss KP, Hance KW, Gloger IS, Zwirner NW (2020) Restoration of antitumor immunity through anti-MICA antibodies elicited with a chimeric protein. J Immunother Cancer 8(1):233. https://doi.org/10.1136/jitc-2019-000233

Turovskaya O, Kim G, Cheroutre H, Kronenberg M, Madan R (2009) Interleukin 10 acts on regulatory t cells to maintain expression of the transcription factor foxp3 and suppressive function in mice with colitis. Nat Immunol 10(11):1178–1184. https://doi.org/10.1038/ni.1791

Valanciene E, Jonuskiene I, Syrpas M, Augustiniene E, Matulis P, Simonavicius A, Malys N (2020) Advances and prospects of phenolic acids production, biorefinery and analysis. Biomolecules 10(6):1–41. https://doi.org/10.3390/BIOM10060874

Vasan N, Baselga J, Hyman DM (2019) A view on drug resistance in cancer. Nature 575(7782):299–309. https://doi.org/10.1038/s41586-019-1730-1 (**Nature Publishing Group**)

Verma S, Singh A, Mishra A (2013) Gallic acid: molecular rival of cancer. Environ Toxicol Pharmacol 35(3):473–485. https://doi.org/10.1016/j.etap.2013.02.011

Wang J, Iannarelli R, Pucciarelli S, Laudadio E, Galeazzi R, Giangrossi M, Falconi M, Cui L, Navia AM, Buccioni M, Marucci G, Tomassoni D, Serini L, Sut S, Maggi F, Dall'Acqua S, Marchini C, Amici A (2020) Acetylshikonin isolated from *Lithospermum erythrorhizon* roots inhibits dihydrofolate reductase and hampers autochthonous mammary carcinogenesis in Δ16HER2 transgenic mice. Pharmacol Res 161:105123. https://doi.org/10.1016/J.PHRS.2020.105123

Wu Y, Chen W, Xu ZP, Gu W (2019) PD-L1 distribution and perspective for cancer immunotherapy—blockade, knockdown, or inhibition. Front Immunol. https://doi.org/10.3389/fimmu.2019.02022 (**Frontiers Media S.A**)

Xia S, Ji R, Xu Y, Ni X, Dong Y, Zhan W (2017) Twisted gastrulation BMP signaling modulator 1 regulates papillary thyroid cancer cell motility and proliferation. J Cancer 8(14):2816–2827. https://doi.org/10.7150/jca.18482

Xu DP, Li Y, Meng X, Zhou T, Zhou Y, Zheng J, Zhang JJ, Li HB (2017) Natural antioxidants in foods and medicinal plants: extraction, assessment and resources. Int J Mol Sci. https://doi.org/10.3390/IJMS18010096

Xue X, Liang XJ (2012) Overcoming drug efflux-based multidrug resistance in cancer with nanotechnology. Chin J Cancer 31(2):100–109. https://doi.org/10.5732/cjc.011.10326 (**BioMed Central**)

Yang L, Pang Y, Moses HL (2010) TGF-β and immune cells: an important regulatory axis in the tumor microenvironment and progression. Trends Immunol 31(6):220–227. https://doi.org/10.1016/j.it.2010.04.002 (**NIH Public Access**)

Yang S, Liu Y, Li MY, Ng CSH, Yang S, Wang S, Zou C, Dong Y, Du J, Long X, Liu LZ, Wan IYP, Mok T, Underwood MJ, Chen GG (2017) FOXP3 promotes tumor growth and metastasis by activating Wnt/β-catenin signaling pathway and EMT in non-small cell lung cancer. Mol Cancer 16(1):1–12. https://doi.org/10.1186/s12943-017-0700-1

Yazaki K (2017) *Lithospermum erythrorhizon* cell cultures: present and future aspects. Plant Biotechnol 34(3):131. https://doi.org/10.5511/PLANTBIOTECHNOLOGY.17.0823A

Yu W, Hua Y, Qiu H, Hao J, Zou K, Li Z, Hu S, Guo P, Chen M, Sui S, Xiong Y, Li F, Lu J, Guo W, Luo G, Deng W (2020) PD-L1 promotes tumor growth and progression by activating WIP and β-catenin signaling pathways and predicts poor prognosis in lung cancer. Cell Death Dis 11(7):1–16. https://doi.org/10.1038/s41419-020-2701-z

Zaal EA, Berkers CR (2018) The influence of metabolism on drug response in cancer. Front Oncol. https://doi.org/10.3389/fonc.2018.00500 (**Frontiers Media S.A**)

Zeb A (2021) Phenolic antioxidants in foods: chemistry, biochemistry and analysis. Phenol Antioxid Foods Chem Biochem Anal. https://doi.org/10.1007/978-3-030-74768-8

Zhang J, Li H, Yu J-P, Wang SE, Ren X-B (2012) Role of SOCS1 in tumor progression and therapeutic application. Int J Cancer 130(9):1971–1980. https://doi.org/10.1002/ijc.27318

Zhang Y, Yang WK, Wen GM, Tang H, Wu CA, Wu YX, Jing ZL, Tang MS, Liu GL, Li DZ, Li YH, Deng YJ (2019) High expression of PRKDC promotes breast cancer cell growth via p38 MAPK signaling and is associated with poor survival. Mol Genet Genomic Med. https://doi.org/10.1002/mgg3.908

Zhao Y, Yang W, Huang Y, Cui R, Li X, Li B (2018) Evolving roles for targeting CTLA-4 in cancer immunotherapy. Cell Physiol Biochem 47(2):721–734. https://doi.org/10.1159/000490025

Zheng A, Li F, Chen F, Zuo J, Wang L, Wang Y, Chen S, Xiao B, Tao Z (2019) PD-L1 promotes head and neck squamous cell carcinoma cell growth through mTOR signaling. Oncol Rep 41(5):2833–2843. https://doi.org/10.3892/or.2019.7053

Zhou XM, Li WQ, Wu YH, Han L, Cao XG, Yang XM, Wang HF, Zhao WS, Zhai WJ, Qi YM, Gao YF (2018) Intrinsic expression of immune checkpoint molecule TIGIT could help tumor growth in vivoby suppressing the function of NK and CD8+T Cells. Front Immunol 9:2821. https://doi.org/10.3389/fimmu.2018.02821

Vegetos
https://doi.org/10.1007/s42535-021-00329-4

RESEARCH ARTICLES

Plant-derived natural compounds aiding SOCS1 mediated JAK1 inhibition, a novel mechanism of combinatorial cancer chemotherapy

Saksham Garg[1] · Sunil Kumar[1] · Ashutosh Anand[1] · Tarunya Menon[1] · Nikita Sharma[1] · Japneet Singh[1] · Siddharth Chawla[1] · Asmita Das[1]

Received: 19 June 2021 / Revised: 30 November 2021 / Accepted: 5 December 2021

Abstract

Numerous drugs have been used in the past to treat HNSC cancer through tumor suppression and immune modulation mechanisms. However, none of them achieved complete tumor remission. Synthetic drugs targeting tumor cells have side effects, and the tumor often acquires resistance against them. A subfamily of tyrosine kinases called Janus Kinases (JAKs) is observed to be over-expressed in various solid tumors, including HNSC. JAKs directly activate a family of transcription factors, Signal Transducers and Activators of Transcription (STATs) and induce a signaling cascade collectively known as JAK/STAT pathways. STATs are responsible for the regulated production of many inflammatory cytokines and growth factors that are beneficial to the tumor cells, favouring them to sustain themselves in a hostile microenvironment. Hence, inhibitors of JAK have been explored previously and SOCS 1 has been shown to be a known direct and most potent inhibitor of JAK1 among the family of SOCSs proteins. The study presented here proposes a mechanism to inhibit the JAK/STAT pathway by inhibiting the JAK1 protein using small molecules of plant origin. The study thereby proposes three inhibitors viz., withaferin A, silymarin, and hypericin, to have significant potential to inhibit JAK1 protein, known to be upregulated in tumors. SOCS1 was also identified to be upregulated in an HNSC tumor samples and is known to inhibit JAK-STAT pathway. Our 3 potent inhibitors, withaferin A, silymarin, and hypericin had the ability to also bind to the SOCS1-JAK1complex thus stabilizing it thus further potentiating the inhibition of JAK-STAT pathway. The three inhibitors explored in the present study can prevent JAK phosphorylation and activation in preventive and therapeutic application. The study proposes a therapy that can be employed in combination with other cancer therapies, thus increasing the overall efficiency of the treatment.

Keywords SOCS1 · JAK1 · Ubiquitination · STAT · HNSC · Structural-based drug discovery · Combinatorial therapy

Abbreviations

ADCC	Antibody-dependent cellular cytotoxicity
BCL2	B cell lymphoma 2
CD274/PDL-1	Programmed death ligand 1
CDK6	Cyclin-dependent kinase
CDKN2A	cyclin-dependent kinase inhibitor 2 A
CTLA4	Cytotoxic T-lymphocyte-associated antigen 4
FADD	Fas-associated protein with death domain
FDA	Food and Drug Administration
FOXP3	Forkhead box P3
HAVCR2	Hepatitis A virus cellular receptor 2
HLA-A	Human leukocyte antigen A
HNSC	Head and neck squamous carcinoma
IL10	Interleukin 10
IRF6	Interferon regulatory factor 6
JAK1	Janus Kinase 1
KIR	Kinase inhibitory region
LAG3	Lymphocyte activation gene 3
MICA	MHC class I polypeptide–related sequence A
PDCD1	Programmed cell death protein 1
PDCD1LG2	Programmed cell death 1 Ligand 2
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
SASA	Solvent accessible surface area
SOCS1	Suppressor of cytokine signalling 1
STAT1/2/3	Signal transducer and activator of transcription 1/2/3
TIGIT	T cell immunoreceptor with Ig and ITIM domains

✉ Asmita Das
asmita1710@gmail.com; asmitadas1710@dce.ac.in

[1] Department of Biotechnology, Delhi Technological University, Main Bawana Road, New Delhi 110042, India

Published online: 12 February 2022

TPM	Transcripts per million.
TWSG1	Twisted gastrulation BMP signalling modulator 1.
TYK2	Tyrosine kinase 2
VEGF	Vascular endothelial growth factor

Introduction

Cancer is a complex and multifactorial disease. In the past various therapies have evolved. Still, due to the complex nature of cancer, a unilateral therapeutic strategy often is insufficient and redundant due to the development of resistance. Recent advances are suggestive towards a combinatorial approach of these therapies to mitigate the cancer cells effectively. The hallmark and the deep-rooted cause for the occurrence of cancer are genetic and proteomic dysregulation at the genetic, transcriptional, or post-transcriptional processing stage (Bradner et al. 2017). While the control and expression of proteins exhibit a high degree of tissue-specificity, it is almost safe to say that cancers have an inherent property to upregulate the proteins that favor their survival while suppressing the genes posing a threat to it (Wuputra et al. 2020; Sur and Taipale 2016). A plethora of proteins is expressed for various biological activities. Among them are few; they are seen as a probable target, i.e., their inhibition can promote proteolytic activity and enhance tumor cell apoptosis. On the other hand, some proteins inherently have tumor-suppressive properties by default. Hence, interest is to promote their activity. Ultimately, both interventions facilitate the elimination of tumor cells (Otto and Sicinski 2017; Liu et al. 2015).

A non-receptor tyrosine kinase, *c-Src*, was the first proto-oncogene identified, and since then, numerous other kinase proteins have been isolated and annotated (Stehelin et al. 1976). Phosphorylation of a target protein is a crucial step in downstream signaling of various pathways, which are essential for cell survival, and the very process is exploited by the cancer cells for their benefit (Schwartz et al. 2018). Kinase receptors have immense involvement in carcinogenesis. Human Genome Project and availability of repositories like the Human cancer genome atlas have enabled the emergence of novel receptors and investigation towards avenues like precision medicine and targeted therapy (Krzyszczyk et al. 2018; Pestell 2003; Sawyers 2002). The sequencing and data analysis efforts have identified kinase receptors either downstream or upstream to essential oncogenes and tumor-suppressive genes, implying their relevance in the molecular pathophysiology of cancer and as an attractive target for drug development (Paul and Mukhopadhyay 2004).

In normal conditions, kinase receptors are generally involved in very crucial developmental and inflammatory responses. One such kinase pathway is the JAK/STAT pathway which regulates cytokine production, interleukin signaling, and growth factor stimulation (Thomas et al. 2015). A cancer cell tends to increase in number and evolve and adapt relative to its proximity. Hence, upregulation of the JAK/STAT pathway provides cancer cells with a favorable mechanism for survival and maturation in a hostile environment. The pathway is extensively exploited to facilitate various molecular stages of cancer. STAT protein is essential for inducing hypoxic conditions and adaptation to the stressful environment (Pawlus et al. 2014). Intense energy requirement in the tumor cell is fulfilled by switching from mitochondrial respiration to glycolysis by releasing pyruvate dehydrogenase kinase 1, which is also mediated by STAT signaling (Demaria et al. 2010). Vasculature development, crucial for tumor survival and metastasis, is induced by vascular endothelial growth factor (VEGF). Transformed cells often use STAT protein as a transcription factor to increase VEGF expression, leading to tumor invasion. Epithelial to mesenchymal transition, which precedes metastasis, is also influenced by JAK signaling through STAT3 (Cho et al. 2013; Huang et al. 2016). Since the ontology of JAK/STAT suggests regulation of cytokine signaling as well as JAK/STAT signaling mediated expression of considerable quantities of interleukins and inflammatory pathways, JAK/STAT pathway is crucial for tumor survival, epithelial to mesenchymal transition, and metastasis of tumor as well as its impact on the immune response to the tumor cells (Sriuranpong et al. 2003; Twyman-Saint Victor et al. 2015; Kortylewski et al. 2005). IRF1 is an interferon regulatory factor-mediated by STAT1 production, and IRF1 facilitates STAT1 binding to DNA, thus forming a positive feedback loop. Knock out of IRF1 showed reduced tumor growth (Shao et al. 2019), while upregulation of IRF1 was followed by a dramatic increase in transcription of decoy receptor 3, promoting migration and poor prognosis (Wei et al. 2019). IFN-β activates JAK1 to produce IL-10, which is infamous for the proliferation of cancer through immunosuppression (Oft 2014; Wang et al. 2011). AXL is another targetable protein that has a role in EMT, survival, anoikis resistance, and invasion of cancer cells, which is also mediated by STAT1 (Wei et al. 2019; Colavito 2020; Lawrence et al. 2015).

As a safety mechanism, the cell employs SOCS1 to keep JAK/STAT pathway in check as prolonged exposure to inflammatory cytokines is detrimental to the cells (Liau et al. 2018). SOCS1 is the most potent SOCS family member and actively regulates the IFNγ production (Larkin et al. 2013). The protein interacts with JAK protein using a short KIR (kinase inhibitory region) motif inhibiting the tyrosine phosphorylation. SOCS1 is classified as a tumor suppressor and therefore is silenced in many human cancers (Vogelstein et al. 2013). SOCS1 is shown to be a direct inhibitor of JAK1, JAK2, and TYK2 (Liau et al. 2018). The SOCS box region of the protein also facilitates binding with

ElonginBC, allowing the E3 ubiquitin ligase mediated ubiquitination. The two scaffolds allow a reasonable inactivation of the JAK protein (Babon et al. 2008; Kamizono et al. 2001]. However, recent literature reveals its tumor-promoting role but fails to elucidate the mechanism and survival (Tobelaim et al. 2015).

The present study has identified a cluster of genes that has an apparent function in both tumor progression and immune regulation in head and neck squamous cancer (HNSC), out of which we identified JAK1 to modulate all the genes (either directly or indirectly). The observation triggered the interest to find a safer plant-based type II JAK1 inhibitor to mitigate JAK-mediated tumor survival, maturation, and immune regulation. The plant-derived compounds selected for the study were already known to have direct anti-tumor functions (Ullrich et al. 2019); the rationale behind this biased selection is that the molecules will have a dual role of direct anti-tumor function and the indirect role of enhancing the SOCS mediated JAK/STAT inhibition response. The analysis also showed SOCS1 protein to be upregulated in HNSC cancer. Hence, we also explore the association between JAK1 and SOCS1 upon ligand binding as the literature suggests that SOCS1 cannot bind with unphosphorylated JAK1 protein due to steric hindrances. However, the conformational change upon ligand binding facilitates the exposure of the activation domain of JAK1 for SOCS1 binding. The proposed ligands are shown to prevent the self-phosphorylation of the tyrosine residues in JAK1, the activation of unphosphorylated JAK1, and promote SOCS1 binding. Thus, by engaging the SOCS1 protein, the proposed compounds effectively reduce the tumor proliferative properties. The study employs a gene expression and computational pipeline to showcase that the JAK1-ligand complex is relatively more stable when in co-association with SOCS1 as compared to un-associated JAK1 keeping binding with ATP as a control measure. This higher affinity of the compounds studied for SOCS1 bound to JAK1 suggests a role in potentiating the inherent suppression of the JAK/STAT pathway by SOCS1 and its anti-tumor function.

Methodology and material

Enrichment and gene expression analysis

Five hundred genes associated with HNSC cancer were imported in Cytoscape from the disease query database and did enrichment analysis with GO Process and found 256 genes related to immune system processes. We further did enrichment analysis and found that 53 genes are associated with immune system suppressor processes.

Using the GEPIA database, multiple genes comparison was carried out on 23 genes isolated based on existing literature for their functional involvement in tumor progression and regulation of immune response for comparing their gene expression in HNSC cells and normal cells (Krasnov et al. 2019). A network analysis was build using the STRING database to identify the cluster of genes (Szklarczyk et al. 2017).

Protein/macromolecule

In our study, the 3-dimensional structure of JAK1-SOCS1 was taken. It was retrieved from RCSB's protein data bank (Berman 2000) in PDB format. The PDB ID assigned to the structure was 6C7Y. Chain A comprises 286 residues and is the Kinase Domain of JAK1 bound to Adenosine-5′-diphosphate at the active site as the native ligand for the said protein, while chain B is the SOCS1 protein of 117 residue length (Fig. 1).

Ligand

The 3-D structures of inhibitors were extracted from the PubChem database in .sdf format (Kim et al. 2016). In total, 56 ligand structures were used and as a control, ATP was employed (Choi et al. 2016). All ligands were converted from .sdf format to .pdb file format using Biovia Discovery Studio Visualizer for docking purposes.

Molecular docking

Molecular docking was then performed in order to obtain each of the protein-ligand binding complexes. For molecular docking, AutoDock 4.2 was used (Morris et al. 2009). Before docking, the structures of the protein and ligands that were downloaded were further prepared. Preparing the optimized protein structure involved removing the water and natural inhibitor molecules, adding polar hydrogen bonds and Kollman charges. Gasteiger charges were also calculated. Then, the energy minimization of the protein was done using the Swiss PDB Viewer. These steps summed up the protein optimization process. A grid box of $54 \times 68 \times 46$ was calibrated as per the active site (ATP-binding site) residues of the JAK1 protein with 0.375 Å spacing with the centers as follows: x center = 9.44, y center = 30.97, and z center = 7.552. Docking Log Files (.dlg) were obtained in the final step to obtain the top 10 binding energies of the docked complexes; here, the Genetic algorithm was kept as the search parameter, and Lamarckian GA was used to run the output. The conformation exhibiting the lowest binding energy was chosen for being converted to a 2D structure to evaluate the binding of the formed protein-ligand interactions.

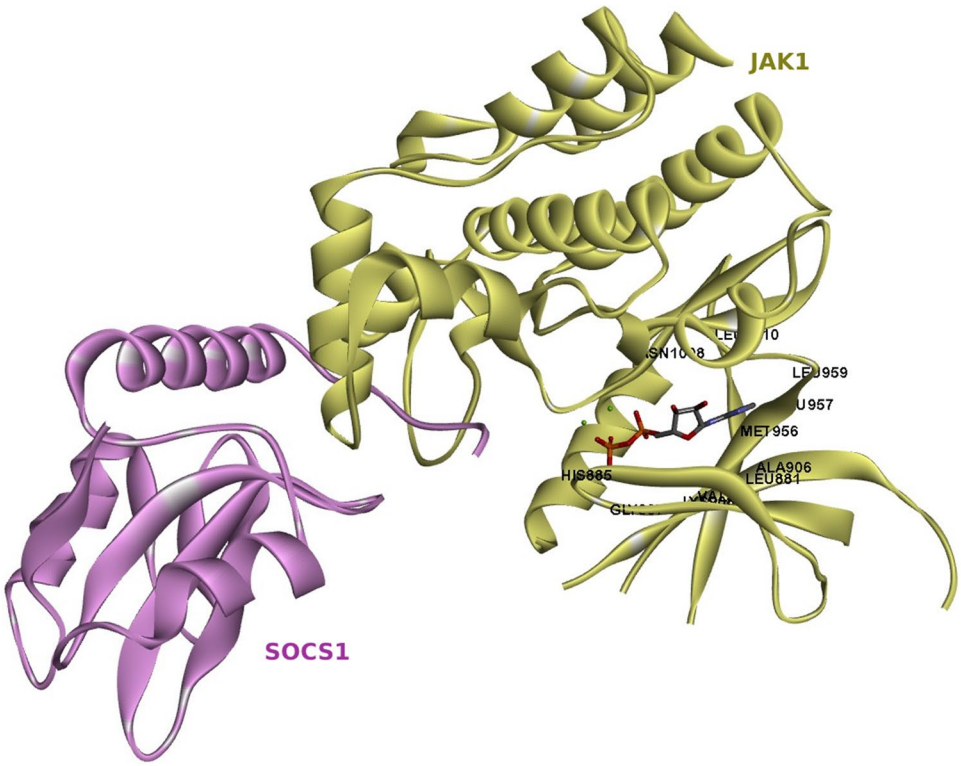


Fig. 1 3D Structure of Kinase Domain of JAK1 protein in complex with SOCS1. ADP is the native ligand bound to the JAK1 structure, and residues forming the active site are labeled with black

Molecular dynamics

A total of 6 dynamic simulations were carried divided into two sets. Set A comprises best-docked conformations of JAK1-SOCS1-ligand complexes, and Set B with the JAK1-ligand complexes, i.e., SOCS1 was omitted using Schrodinger2021 from docked complexes. The segregation allowed a comparative analysis between the two states of JAK1. As a default setting, the Desmond Module of Schrodinger2021 utilizes all the algorithms that support high performance and accurate results. After pre-processing and optimizing, docked complexes were submerged into a solvent environment with water molecules aligning in a TIP3P water model in an orthorhombic box. The water environment was neutralized using sodium ions at 0.15 M concentration. After preparing the system around the complex, all the atoms were subjected to OPLS-AA 2005 force field. A 100ns simulation was done at 300 K and 1 bar pressure. In the pipeline, before simulation onset, the whole system was allowed to relax and minimize. A 1000 frame trajectory was obtained to analyze the interaction and dynamics of protein-ligand complexes (Bowers et al. 2006).

Estimation of biological activity

A prediction of selected compounds' biological activity was carried out. The was accomplished using the PASS web server, which uses multilevel neighbors of atoms descriptors as its core principle to predict bio-activity of the ligand molecules by providing SMILES as the input. The process is solely based on chemical structures (Goel et al. 2011).

Results

Enrichment analysis and gene expression

Enrichment analysis of 500 of the most associated HNSC cancer genes was conducted to select immune-associated genes. 256 genes related to immune regulation were explored, out of which 53 genes were found associated with the suppression of immune system-associated processes, such as regulation of T cell and B cell activation, regulation of B cell proliferation, and many more. Furthermore, the 53 genes were checked in the literature for their association with tumor progression-related processes like cell proliferation, metastasis, etc. It was found that 14 genes were associated with both processes while seven genes were associated with immune suppression only, and two genes were associated with tumor progression only. The remaining two genes were associated with HNSC cancer due to alteration of their function by mutation, as shown in Table 1.

A pan-cancer gene expression analysis of the 14 genes having a role in both tumor-associated functional ontology and immune suppression is shown by Fig. 2, which quantifies the change in their expression value comparing matched

Table 1 List of genes with their tumor progression and immunosuppression roles

S. no.	Genes	Tumor progression role	Immunosuppression role	Targeting drugs (FDA approved/in clinical trials)
1	IRF1	Upregulate PD-L1 in the tumor cell (Shao et al. 2019)	Inhibits CD8+ T cell and NK-cell mediated anti-tumor immune responses Inhibits activation of neutrophils (Yang et al. 2010)	
2	TWSG1	Enhancing tumor growth and malignant cell behavior and stimulating tumor-associated angiogenesis (Xia et al. 2017)		
3	CDK6	Regulates the progression of the cell cycle Transcriptional role in tumor angiogenesis (Tadesse et al. 2015)	CDK6 inhibition triggers anti-tumor immunity (Goel et al. 2017)	Palbociclib (FDA approved)
4	AXL	Tumor proliferation, survival, metastasis, and resistance to cancer therapy (Rankin and Giaccia 2016)	Small-molecule inhibition of AXL targets tumor immune suppression (Ludwig et al. 2018)	
5	FADD	Cell cycle progression and cell proliferation (Papoff et al. 2010)	A negative regulator of T cell receptor-mediated necroptosis (Osborn et al. 2010)	
6	HAVCR2	Induce epithelial-mesenchymal transition by JAK-STAT3 signaling pathway (Hou et al. 2020)	Overexpression of HAVCR2 observed in tumor-infiltrating lymphocytes is associated with adaptive resistance to immunotherapy (Huang et al. 2015)	BMS-986,258(in a clinical trial)
7	PRKDC	Promotes tumor cell growth via p38 MAPK signaling (Zhang et al. 2019)	PRKDC is a predictive biomarker and a drug target for immune checkpoint inhibitors (Tan et al. 2020)	NU7026 (FDA approved)
8	IL10	An association exists between IL-10 expression and tumor-related markers such as Bcl-2 (Sheikhpour et al. 2018)	Inhibited T cell proliferation and function (Hou et al. 2020) It seems that TAMs cause drug resistance via the IL-10/Stat3/Bcl-1/BCL2 signaling pathway (Sheikhpour et al. 2018; Llanes-Fernández et al. 2006)	GIT 27 (in clinical trials)
9	SOCS1	SOCS1 downregulation inhibits cell proliferation via cell cycle progression, resulting in G0/G1 phase accumulation and reduction of the S phase (Zhang et al. 2012)	SOCS1 is involved in the inactivation of CD8+ T cells against tumor cells (Chikuma et al. 2017)	
10	MICA		Anti-MICA antibodies can promote anti-tumor immunity by inducing direct anti-tumor effects (antibody-dependent cell-mediated cytotoxicity, ADCC) (Torres et al. 2020)	
11	TIGIT		It suppresses the Function of NK cells and CD8+ T Cells (Zhou et al. 2018; Harjunpää and Guillerey 2020)	
12	CDKN2A	Mutated	Mutated	
13	LAG3		Inactivates the CD4+ T cells Reduces the effector function of CD8+ T cells It promotes the suppressor activity of T_{regs} (Long et al. 2018)	BI 754,111(in Clinical Trials)
14	CTLA4		Inhibits the activation and proliferation of T cells (Zhao et al. 2018)	Ipilimumab (FDA Approved)
15	CD274	Promotes tumor cell growth, migration, and invasion via WIP and β-catenin signaling (Yu et al. 2020)	CD274 overexpression negatively regulates the T cell mediated immune response in peripheral tissues (Akinleye and Rasool 2019)	Nivolumab (FDA Approved)
16	HLA-A	Highly polymorphic	Highly polymorphic	
17	PDCD1LG2		PDCD1LG2 overexpression suppressed the tumor antigen-specific CD8+ T cells (Tanegashima et al. 2019)	Atezolizumab (FDA approved)
18	FOXP3	FOXP3 overexpression promotes cell proliferation, migration, and invasion (Yang et al. 2017)	FOXP3 plays a vital role in the development of Treg cells (Mercer and Unutmaz 2009)	RPG (FDA approved)

Table 1 (continued)

S. no.	Genes	Tumor progression role	Immunosuppression role	Targeting drugs (FDA approved/in clinical trials)
19	PDCD1		PDCD1 overexpression suppresses the immune response against tumors (Han et al. 2020)	Avelumab (FDA approved)
20	JAK1	JAK1 is involved in inflammatory cytokine signaling like IL-6 and has an essential role in metastatic cancer progression (Wehde et al. 2018)	Small-molecule drugs that inhibit Janus kinases have a significant role in immune suppression in many autoimmune diseases. (Schwartz et al. 2017)	
21	STAT1	STAT1 silencing enhances cell apoptosis in many cancers. Patients with high expression of STAT1 in cancer tissues experience worse clinical outcomes than a low level of STAT1 expression (Zhang and Zhaoyoung 2017)	STAT1 enhances the precision of PD-L1 on tumor cells hence helps in immune suppression (Ahn et al. 2017)	
22	STAT2	STAT2-deficient mice have decreased tumor incidences to carcinogens, suggesting that STAT2 plays a positive role in tumorigenesis (Hu et al. 2020; Gamero et al. 2010)	STAT2 mediates immunosuppression exerted by mesenchymal stromal cells by expression inflammatory cytokines (Yi et al. 2012)	
23	STAT3	Activation of STAT3 plays a crucial role in tumor growth and metastasis. It regulates cellular proliferation, invasion, migration, and angiogenesis that are critical for cancer metastasis (Kamran et al. 2013])	STAT3 promotes the expression of immune suppressive factors while inhibiting Th1 immunostimulatory molecules. By virtue of its ability to promote the expression of many factors that activate STAT3 in diverse cells, STAT3 allows malignant and immune cells to resonate, forming a close partnership for tumor immune evasion, tumor progression, and resistance to therapies (Kortylewski and Yu 2008)	

Fig. 2 Multiple Gene Expression analysis between HNSC sample and a typical sample of 14 genes having dual functions. The gradient is directly proportional to the Log_2 (TPM+1), where TPM is Transcripts per million as the normalized value of gene count

normal tissues with tumor tissues of HNSC. Upregulated genes had a higher Log_2 (TPM+1) in the tumor than typical tissue samples values matched with the TCGA database.

These 14 shortlisted genes were explored for their correlation between them. Input to STRING database provided 2 clusters; 13 genes in first and one gene in another. The cluster with 13 genes became the focus of the study, and JAK1 presented itself as a druggable target interacting with nine genes directly, and others involved downstream or upstream to JAK1 protein in multiple pathways (Fig. 3).

Molecular docking

All the selected ligands were subjected to molecular docking analysis with the JAK1-SOCS1 complex. Table 2 summarizes the results of the docking studies with all the selected phytochemicals for the study. Docked complex with ATP was used a control for the study. Our analysis shows the best three ligands to have significantly lower binding energy than the ATP (− 6.08 kcal/mol).

Among ten different conformations of withaferin A obtained, − 12.34 kcal/mol was the least binding energy obtained. Withaferin A exhibited five different types of bonding with the protein, as shown in Fig. 4 A., namely—van der Waals interaction, H-bond, carbon-hydrogen bond, alkyl bond, and pi-alkyl bond. ASP1003 (A chain), ARG1007 (A chain), and HIS54 (B chain) formed H-bond with withaferin A; PHE886, VAL889, ARG1007, and LEU1010, all residues of chain A, formed alkyl and pi-alkyl interactions; HIS885 and ASN1008, of chain A, formed carbon-hydrogen bond; rest of the residues formed weak or van der Waals interaction with the ligand.

Following Withaferin A, the Hypericin-JAK1-SOCS1 complex exhibited the least binding energy of − 10.15 kcal/mol. As shown in Fig. 4C, hypericin formed six bond types

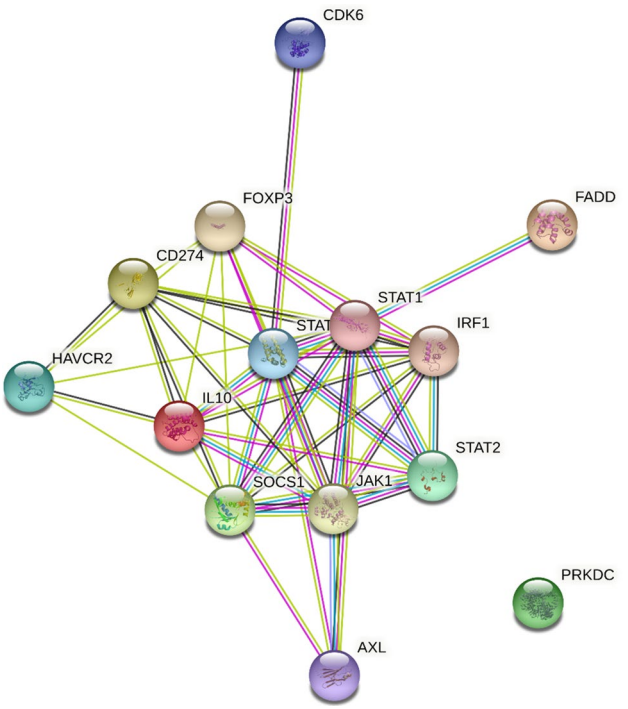


Fig. 3 Network mesh obtained from the STRING database. A clear demarcated cluster comprises 13 genes, while PRKDC itself acts as a second cluster. JAK1 is directly associated with STAT1, STAT2, STAT3, IL10, AXL, SOCS1, FOXP3, CD274, and IRF1, while FADD, CDK6, and HAVCR2 are distantly regulated via mediators in the JAK1 pathway

with the protein, namely—van der Waals, H-bond, carbon-hydrogen bond, pi-sigma bond, alkyl, and pi-alkyl bond. GLU883, MET956, LEU959, GLU966, and ASP1021, all residues of chain A, formed conventional H-bond with the ligand; LEU881, and LEU1010, both of chain A, formed pi-sigma bond; VAL889, ALA906, MET956, and ARG1007,

Table 2 Lists of various docking parameters of all the ligands considered in this study

S. no.	Ligand	Binding energy	Ligand efficiency	Inhibition constant (μM)	Intermolecular energy	Vdw-H bond desolvation energy
1	Carvacrol	− 5.78	− 0.53	57.53	− 6.38	− 6.27
2	Eugenol	− 5.47	− 0.46	98.56	− 6.66	− 6.19
3	Dihydrocarveol	− 5.99	− 0.54	40.94	− 6.58	− 6.47
4	Geraniol	− 5.27	− 0.48	138	− 6.76	− 6.63
5	Nerol	− 4.77	− 0.43	319.93	− 6.26	− 6.11
6	Linalool	− 5.15	− 0.47	167.01	− 6.64	− 6.51
7	1,8-Cineol	− 5.41	− 0.49	107.87	− 5.41	− 5.42
8	3,3'-Diindolylmethane	− 7.5	− 0.39	3.2	− 8.09	− 8.02
9	6-Gingerol	− 5.44	− 0.26	102.79	− 9.02	− 8.87
10	6-Shogaol	− 5.69	− 0.57	67.38	− 5.69	0.01
11	7,4′ Dihydroxyflavonoid	− 7.53	− 0.4	3.01	− 8.43	− 7.84
12	Aconitine	71.41	1.79	N/A	68.43	68.44
13	Albumin tannate	− 7.27	− 0.36	4.68	− 8.46	− 8.26
14	Apigenin	− 5.32	− 0.48	126.45	− 5.62	− 5.6
15	α-Pinene	− 5.93	− 0.54	45.02	− 5.93	− 5.66
16	α-Thujone	− 7.86	− 0.39	1.73	− 9.05	− 8.82
17	Baicalein	− 3.69	− 0.41	1.99	− 5.48	− 5.44
18	β-Carotene	− 4.11	− 0.09	975.39	− 8.28	− 8.64
19	Camphene	− 6.1	− 0.31	33.89	− 9.08	− 8.8
20	Chrysin	− 7.56	− 0.4	2.88	− 8.45	− 8.26
21	Curcumin	− 7.56	− 0.28	2.89	− 10.54	− 10.38
22	Celastrol	− 7.7	− 0.23	2.28	− 8.59	− 7.75
23	Chlorogenic acid	− 7.41	− 0.3	3.69	− 10.69	− 8.49
24	Caffeic acid	− 6.85	− 0.53	9.59	− 8.34	− 5.51
25	Carnosol	− 7.63	− 0.32	2.53	− 8.53	− 8.27
26	Capsaicin	− 6.94	− 0.32	8.14	− 9.93	− 9.77
27	Ellagic acid	− 7.68	− 0.35	2.34	− 8.88	− 8.4
28	Epigallocatechin-3-gallate	− 8.68	− 0.26	0.43556	− 12.26	− 11.9
29	Formononetin	− 6.82	− 0.34	10.04	− 7.71	− 7.3
30	Gallic acid	− 6.27	− 0.52	25.53	− 7.76	− 5.8
31	Genistein	− 7.26	− 0.36	4.79	− 8.45	− 8.2
32	Gossypol	− 7.77	− 0.2	2.02	− 11.05	− 11.1
33	Hypericin	− 10.15	− 0.27	0.03644	− 11.94	− 11.61
34	Hydroxytyrosol	− 5.24	− 0.48	143.38	− 6.73	− 6.23
35	Indole-3-carbinol	− 5.72	− 0.52	64.08	− 6.32	− 5.95
36	Isoliquiritigenin	− 8.01	− 0.42	1.34	− 9.8	− 9
37	Jasmonic acid	− 5.71	− 0.38	64.88	− 7.5	− 7.16
38	Koenimbin	− 7.73	− 0.35	2.17	− 8.03	− 7.96
39	Limonene	− 5.38	− 0.54	114.3	− 5.68	− 5.66
40	Medicarpin	− 7.54	− 0.38	2.96	− 8.14	− 7.99
41	Parthenolide	− 7.56	− 0.42	2.85	− 7.56	− 7.63
42	Piperine	− 7.7	− 0.37	2.27	− 8.6	− 8.67
43	Proanthocyanidins	− 7.17	− 0.17	5.54	− 11.05	− 10.61
44	Plumbagin	− 5.93	− 0.42	45.3	− 6.22	− 6.11
45	Pterostilbene	− 7.25	− 0.38	4.81	− 8.75	− 8.38
46	Resveratrol	− 6.61	− 0.39	14.17	-8.11	-7.53
47	Retinoic acid	− 8.98	− 0.41	0.26356	-10.77	-10.15
48	Sabinene	− 5.36	− 0.54	117.88	− 5.66	− 5.65

Table 2 (continued)

S. no.	Ligand	Binding energy	Ligand efficiency	Inhibition constant (μM)	Intermolecular energy	Vdw-H bond desolvation energy
49	Shikonin	− 6.52	− 0.31	16.56	− 8.31	− 7.98
50	Silymarin	− 10.04	− 0.29	0.04355	− 12.73	− 12.15
51	Sulforaphane	− 4.39	− 0.44	600.64	5.89	− 5.73
52	Triptolide	− 8.48	− 0.33	0.60637	− 9.08	− 9.02
53	Terpineol	− 5.83	− 0.53	53.6	− 6.42	− 6.31
54	Wogonin	− 7.3	− 0.35	4.45	− 8.49	− 7.65
55	Withaferin A	− 12.34	− 0.36	0.000895	− 13.84	− 13.3
56	Podophyllotoxin	− 8.29	− 0.28	0.84385	− 9.78	− 9.26
57	ATP (Control)	− 6.08	− 0.2	34.66	− 10.56	− 8.79

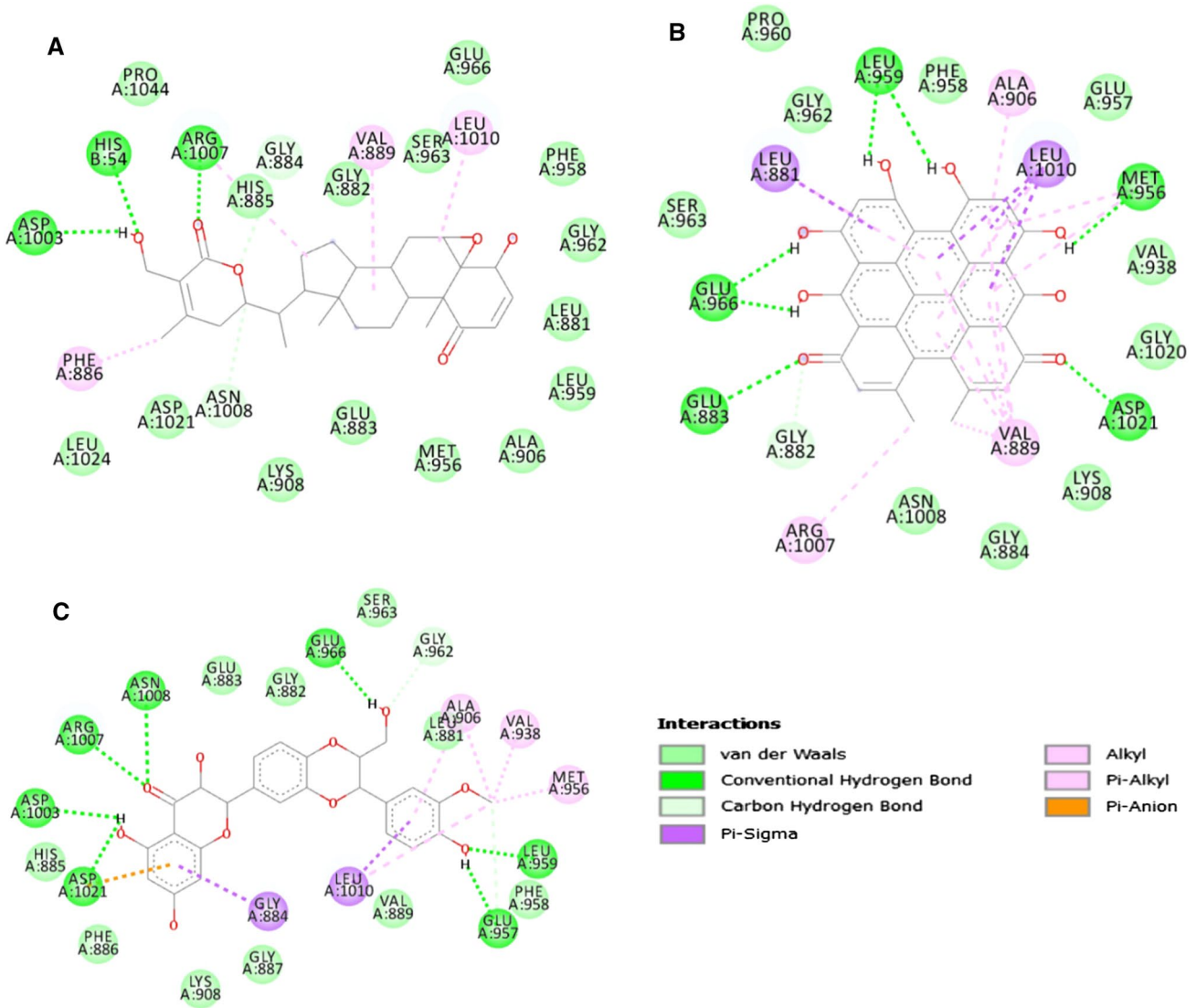


Fig. 4 2D Interactions obtained after molecular docking of **A** Withaferin A, **B** Hypericin, **C** Silymarin

from chain A, formed alkyl and pi-alkyl bond with the ligand; GLY882 of chain A was involved in the formation of carbon-hydrogen bond; rest of the residues were involved in van der Waals interaction with the ligand.

The Silymarin-JAK1-SOCS1 complex had binding energy of − 10.04 kcal/mol. As exhibited by Fig. 4B, protein has formed seven bonds with silymarin; van der Waals, H-bond, carbon-hydrogen bond, pi-anion interaction, pi-sigma bond, alkyl, and pi-alkyl bonds. It can also be seen that all the residues involved in bonding belong to chain A. GLU957, LEU959, GLU966, ASP1003, ARG1007, ASN1008, and ASP1021 formed conventional H-bond; Asp1021 was involved in pi-anion interaction with the ligand; GLY884 and LEU1010 forms pi-sigma bond; ALA906, VAL938, MET956, and LEU1010 were involved in alkyl and pi-alkyl bonding with the ligand; GLU957 and GLY962 formed carbon-hydrogen bond; remaining residues formed van der Waals interactions with the ligand.

Molecular dynamics simulation

In order to understand the conformational stability of the protein-ligand complex under clear water conditions, MD simulations were performed. For further studies, both Set A (Withaferin A-JAK1-SOCS1, Silymarin-JAK1-SOCS1, and Hypericin-JAK1-SOCS1) and Set B (Withaferin A-JAK1, Silymarin-JAK1, and Hypericin-JAK1) complexes were selected based on their docked binding energy and interactions. The complexes were subjected to a 100ns all-atom MD simulation each. Table 3 lists the potential energy and total energies calculated for each complex.

Conformational stability and structural compactness

The structural deviation and compactness become critical to explore the changes in the protein moiety. Hence, upon subjecting the protein-ligand complex to dynamics, four graphical parameters, namely, root means square deviation (RMSD), solvent accessible surface area (SASA), and root means square fluctuations (RMSF), provided us with insights into the stability and compactness of protein-ligand complex.

Comparing the RMSD plots of two sets of complexes with their ATP bound conformation towards the end of the simulation gives us an estimate of the protein and protein-ligand stability. Both sets of complexes with silymarin showcased similar RMSD plots which were lower than the ATP, while with the other two ligands, a considerable amount of difference was observed. Following the removal of SOCS1 protein, the Withaferin A-JAK1 and Hypericin-JAK1 exhibited higher RMSD values, suggestive of instability of the protein-ligand complex (Fig. 5A). The average values in Fig. 5B confirm the difference in RMSD values. Overall, ligand binding to the JAK-SOCS1 bound protein is significantly more stable than binding JAK1, which is not bound to SOCS1.

Residual fluctuations in the peptide chain were calculated and plotted in a root mean square fluctuation (RMSF) graph. As per the convention, for all the complexes, free residues, i.e., residues not involved in secondary structure formation, were seen to fluctuate the most, and residues forming alpha helices or beta sheets were found to show limited fluctuation values. Residual vibrations of the amino acid binding to the ligand molecule obtained the lower set of values suggesting a rigid binding pocket. Figure 5C evaluates the average RMSF value of the complexes pre and post SOCS1 removal from the protein-ligand hybrid.

Solvent accessible surface area (SASA) measures the surface area of a molecule in contact with the surrounding water molecules. The average SASA values calculated at the end of the 100ns simulation for pre and post SOCS1 removed protein complexes with withaferin A, silymarin, and hypericin were calculated (Fig. 5D). The average values suggest complexes with hypericin as the most stable and shielded of the three with the least internal pocket residues interacting with the surrounding solvent molecules.

Secondary structure analysis

Computing and analyzing secondary structures can be used to understand the protein packing and characteristics of folding the protein with different ligands. Table 4 shows the percentage of secondary structures of SOCS1 protein with withaferin A, silymarin, and hypericin ligands.

Table 3 Energy profile of both sets of complexes with three ligands

S. no.	Complex	Total energy (Kcal/mol)	Potential energy(Kcal/mol)
1.	JAK1-SOCS1-Withaferin A	− 127,296.389	− 156,435.100
2.	JAK1-SOCS1-Silymarin	− 127,380.998	− 165,519.864
3.	JAK1-SOCS1-Hypericin	− 127,343.435	− 156,485.226
4.	JAK1-Withaferin A	− 89,222.743	− 109,645.328
5.	JAK1-Silymarin	− 89,326.760	− 109,752.643
6.	JAK1-Hypericin	− 89,241.412	− 109,664.753

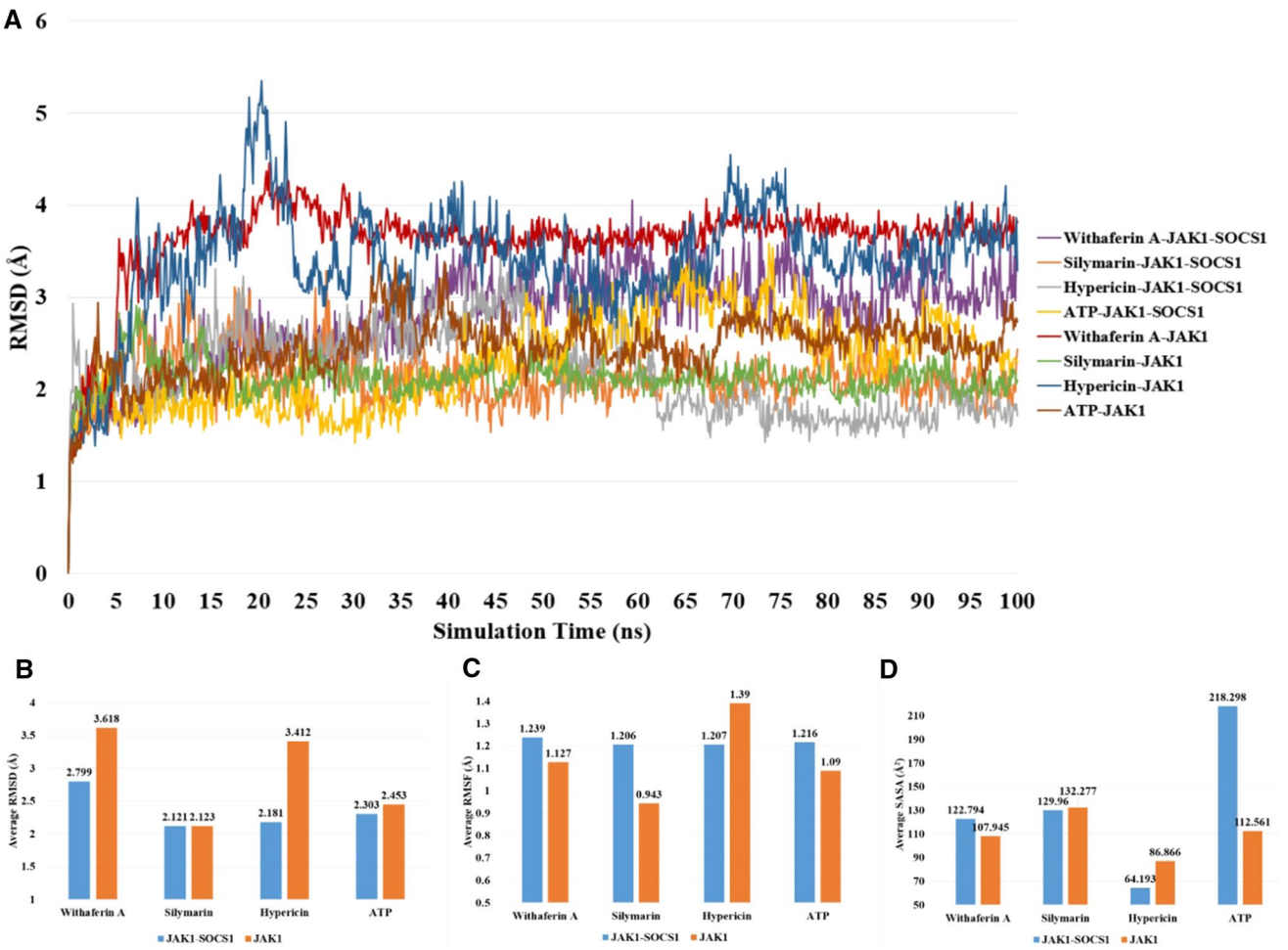


Fig. 5 Post 100ns simulation comparative analysis of conformational stability and compactness in comparison with ATP. **A** RMSD. **B** Average RMSD. **C** Average RMSF. **D** Average SASA

Table 4 Secondary Structure analysis of all six protein-ligand complexes

Complex	% Helix	% Strand	% Total SSE
JAK1-SOCS1-Withaferin A	28.80	11.95	40.75
JAK1-SOCS1-Silymarin	28.85	11.75	40.60
JAK1-SOCS1-Hypericin	29.01	11.39	40.40
JAK1-Withaferin A	29.95	8.62	38.57
JAK1-Silymarin	30.29	8.94	39.23
JAK1-Hypericin	29.61	8.33	37.95

Analysis of interaction dynamics

After the 100ns simulation of the three complexes, the histogram of all the interactions with their ligands was analyzed. Withaferin A-JAK1-SOCS1 complex showed the lowest number of H-bond interactions. However, ASP1003 interacted with the ligand using the H-bond for most of the simulation time. Other interactions complementing this binding were water bridges involving LYS908 and GLY1020 and hydrophobic bond formations by PHE886 and LEU1010 (Fig. 6A).

The histogram plot of Silymarin-JAK1-SOCS1 presented the highest number of residues interacting with ligand (Fig. 6B). LEU959, ASP1003, ASP1021, and GLU957 complemented each other for stable binding of the ligand in the binding pocket by forming a solid H-bond for 100% of the period studied. LYS965, HIS885 formed all three kinds of bonds, while others like GLU883 and ASN1008 formed conventional hydrogen bonds and water bridges during the simulation. Other residues are supported by forming hydrophobic and water bridge interactions with the ligand in a scattered nature.

In a Hypericin-JAK1-SOCS1 complex, seven amino acids formed H-bonds, out of which LEU959 and GLU956 held the ligand in place for the most duration of the simulation. Other residues contributed to the ligand-protein interaction

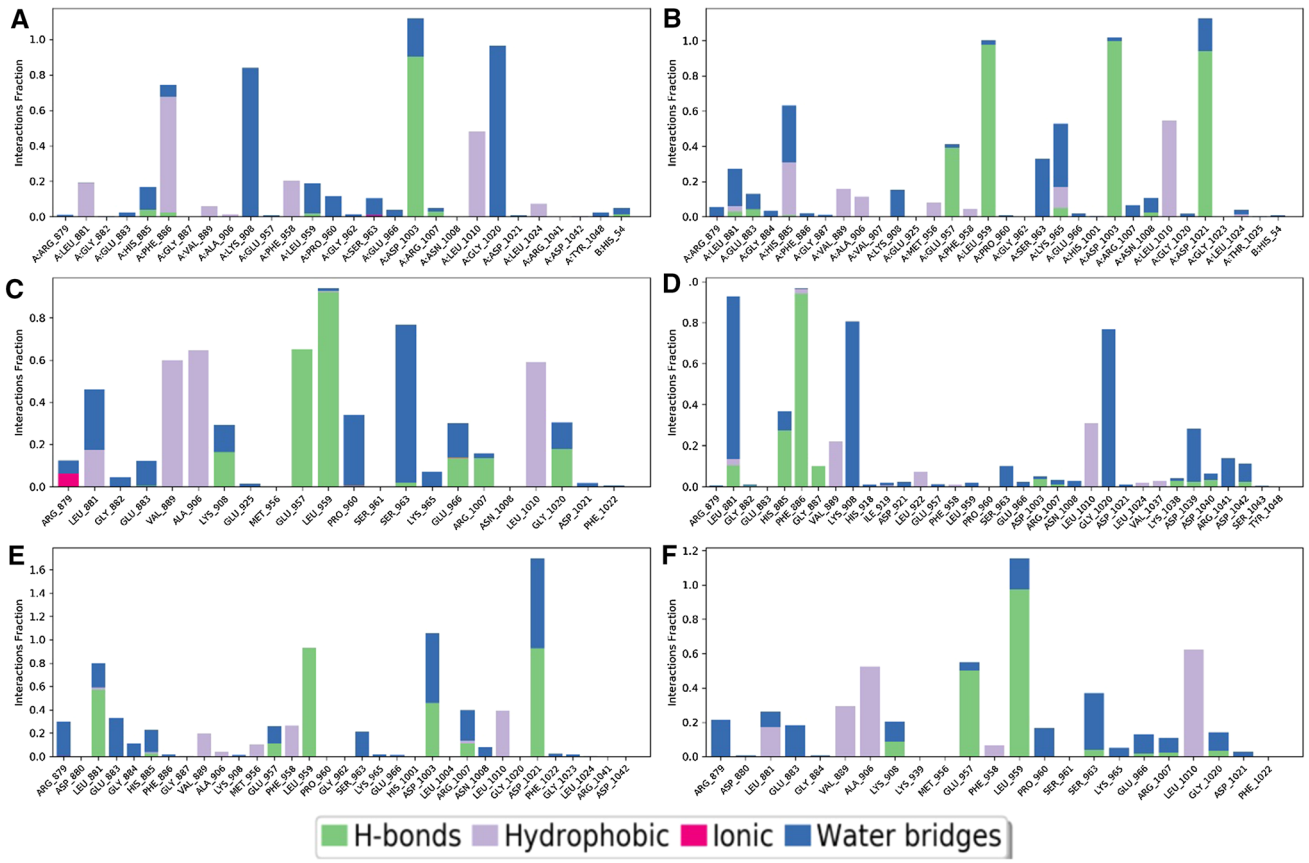


Fig. 6 Post dynamic analysis of protein-ligand interaction **A** Withaferin A-JAK1-SOCS1. **B** Silymarin-JAK1-SOCS1. **C** Hypericin-JAK1-SOCS1. **D** Withaferin A-JAK1. **E** Silymarin-JAK1. **F** Hypericin-JAK1

by hydrophobic bonds (VAL889, ALA906), water bridges (SER963, PRO960), and ionic bonds like ARG879. Some residues like LEU881, LYS908, GLY996, and GLY1020 exhibited dual nature of bond formation with the complex (Fig. 6C).

Withaferin A-JAK1 complex interacted with a total of 34 amino acids, out of which PHE886 and HIS885 formed string H-bond more than 20% of the simulation time. While LEU881, GLY1020, LYS908, ASP1039, and other residues complemented the H bonds quite well anchoring, cumulatively anchoring the ligand at the binding site (Fig. 6D).

Bar plot of Silymarin-JAK1 complex showed an appreciable mixed type of interactions throughout the simulation. ASP1021, LEU959, HIS1003, and ASP880 acted as the hook to the ligand while other residues such as ARG879, GLU883, ARG1007, LEU1010, and other 25 residues interacted with either a combination of bonds or with a single type, ultimately contributing towards the ligand binding at different instances of time (Fig. 6E).

A total of 23 residues of the JAK1 protein interacted with hypericin. LEU959 and GLU957 stabilized the binding using the string H-bond at least 60% of the simulation time. ALA1010, ALA906, VAL889, and LEU881 showed appreciable hydrophobic interactions while remaining other formed water bridges either in combination with other interactions or solely. Hypericin ligand was well embedded and shielded in the active site by the protein residues (Fig. 6F).

Biological activity

Providing the SMILES as the input to the server-generated prediction for the three ligands. All three ligands have been found to have significant anticancer activities mediated by different pathways. The probability ranges from 0.797 to 0.936 when $P_i < P_a$. Obtaining a similar biological activity suggests their strong candidature for the therapy (Table 5).

Discussion and conclusion

A compelling amount of evidence suggests that JAK/STAT pathway is generally activated in solid tumors, and the upregulated pathway contributes to the malignancy of cancer cells. Given the extensive role of the JAK/STAT pathway in molecular processes and its exploitation by cancer cells for their survival, proliferation, protection, adaptation, and

Table 5 Prediction of Biological Activity

Phytochemical	Bio-activity	P_a	P_i
Withaferin A	Antineoplastic	0.916	0.005
Silymarin	TP53 expression enhancer	0.936	0.002
	Chemopreventive	0.797	0.004
Hypericin	TP54 expression enhancer	0.846	0.008
	Antineoplastic	0.836	0.008

P_a activity probability, P_i inactivity probability

mobility, the pathway becomes an exciting and promising target to develop a novel therapy to kill tumor cells, prevent their metastasis and potentiate the anti-tumor response of the body.

The JAK1 protein and related STAT proteins are shown to be upregulated in HNSC cancers in multiple studies. JAK1 becomes a promising target as it is involved in tumor progression, and immune regulation. It was interesting to see that JAK1 was overexpressed in cancer cells, and SOCS1, a JAK1 inhibitor, was also highly expressed in the HNSC samples. In our study, plant-derived ligands were shown to bind to JAK1, and it could preclude the phosphorylation of JAK1 and thus prevent JAK/STAT pathway activation. Along with testing the stability of JAK1 binding with the plant derived ligands, binding of SOCS1-JAK1 complex with our ligands was extensively studied, and their stabilities were compared. JAK1 mediates most of the pathways via STAT isoforms such as STAT1, STAT2, STAT3, STAT4, and cytokine production.

JAK-STAT pathway resulted in a dramatic increase in tumor migration and IFN-β activated IL-10 production, which is crucial for immunosuppression and enabling tumor proliferation and metastasis. The present study identifies three potent plant-based inhibitors, withaferin A, silymarin, and hypericin, from an extensive dataset of ligands with anti-tumor functions. The RMSD obtained, post-simulation, emphasized the stability of all three protein ligand complexes. The same plot also compares the stability of JAK1-ligand complexes in SOCS1 bound and unbound forms. Our studies have shown that the complex is relatively more stable when conjugating with SOCS1 bound form, as is evident by average values. Although SOCS1 has a high affinity to bind to JAK1 via the KIR domain, it is observed that SOCS1 cannot interact with the activation loop of JAK1 in its inactive form. Presence of any ligand molecule induces a conformational change in the protein, and the ligands proposed in the present study have been shown to do the same and provide the interface to SOCS1 protein to bind to JAK1 protein, which is the probable cause of higher stability of JAK1-SOCS1-ligand complexes. Since, ATP acts as a substrate during phosphorylation, the three drug molecules were also compared with the ATP bound conformations. Silymarin clearly showcased its potentiality in competing for the binding site in both SOCS1 bound and non-bound states thus preventing the initial phosphorylation and the subsequent ones as well. The three drug molecules are predicted to have a high probability of acting as either anti-neoplastic or chemopreventive agents, which along with interaction with JAK1 has the unique property of stabilizing the inherent JAK inhibitor SOC1 that is present in the tumor microenvironment. Our studies show that the three ligands can act as potent inhibitors to both activated and inactivated JAK1 receptors, thus exhibiting both treatment and preventive nature of these compounds. The SOCS1 scaffold, on the other hand, has two binding surfaces, the KIR domain, and the SOCS box. KIR domain can proficiently bind to JAK1 even with its dephosphorylated form, preventing the phosphorylation of the second kinase molecule while enabling the degradation of bound JAK by inducing ubiquitination via SOCS box domain. Use of our inhibitors should act as a fail-proof mechanism to inhibit the JAK1 associated pathways and regulation by aiding the existing SOCS1 protein binding to JAK1 and also by directly competing with ATP at the ATP-binding site of JAK1 in HNSC cancers. Thus the ligands will prevent JAK1 phosphorylation thereby activation of the JAK/STAT pathway, also, stabilize and aid the SOCS1 mediated inhibition of JAK1 by bringing about a conformational change to the activation domain of JAK1 protein. This method also has potential to be combined with other therapies such as PDL-1 based immunotherapy, AXL inhibitory chemotherapy, and radiotherapy. Therefore, our study encourages exploring withaferin A, silymarin, and hypericin to further evaluate their efficacy and safety in cell and animal models and their use in combinatorial therapies in various solid tumors.

Funding All studies have been conducted with the financial and infrastructural supportof Delhi Technological University.

Data availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors declare no conflict of interest.

Ethical approval The present study does not contain any human participants or animals and has been carried out in the Department of Biotechnology, Delhi Technological University, following all ethical principles of the university.

References

Ahn R, Sabourin V, Bolt AM, Hébert S, Totten S, De Jay N, Festa MC, Young YK, Im YK, Pawson T, Koromilas AE, Muller WJ,

Mann KK, Kleinman CL (2017) Ursini-Siegel, the Shc1 adaptor simultaneously balances Stat1 and Stat3 activity to promote breast cancer immune suppression. Nat Commun 81:1–14. https://doi.org/10.1038/ncomms14638

Akinleye A, Rasool Z (2019) Immune checkpoint inhibitors of PD-L1 as cancer therapeutics. J Hematol Oncol 12:92. https://doi.org/10.1186/s13045-019-0779-5

Babon JJ, Sabo JK, Soetopo A, Yao S, Bailey MF, Zhang J-G, Nicola NA, Norton RS (2008) The SOCS box domain of SOCS3: structure and interaction with the ElonginBC-Cullin5 ubiquitin ligase. J Mol Biol 381:928–940. https://doi.org/10.1016/j.jmb.2008.06.038

Berman HM (2000) The Protein Data Bank. Nucleic Acids Res 28:235–242. https://doi.org/10.1093/nar/28.1.235

Bowers KJ, Chow DE, Xu H, Dror RO, Eastwood MP, Gregersen BA, Klepeis JL, Kolossvary I, Moraes MA, Sacerdoti FD, Salmon JK, Shan Y, Shaw DE (2006) Scalable algorithms for molecular dynamics simulations on commodity clusters. In: ACM/IEEE SC IEEE, pp 43–43. https://doi.org/10.1109/SC.2006.54

Bradner JE, Hnisz D, Young RA (2017) Transcriptional addiction in cancer. Cell 168:629–643. https://doi.org/10.1016/j.cell.2016.12.013

Chikuma S, Kanamori M, Mise-Omata S, Yoshimura A (2017) Suppressors of cytokine signaling: Potential immune checkpoint molecules for cancer immunotherapy. Cancer Sci 108:574–580. https://doi.org/10.1111/cas.13194

Cho KH, Jeong KJ, Shin SC, Kang J, Park CG, Lee HY (2013) STAT3 mediates TGF-β1-induced TWIST1 expression and prostate cancer invasion. Cancer Lett 336:167–173. https://doi.org/10.1016/j.canlet.2013.04.024

Choi H, Cho SY, Pak HJ, Kim Y, Choi J, Lee YJ, Gong BH, Kang YS, Han T, Choi G, Cho Y, Lee S, Ryoo D, Park H (2017) NPCARE: database of natural products and fractional extracts for cancer regulation. J Chem Inform 9:2. https://doi.org/10.1186/s13321-016-0188-5

Colavito SA (2020) AXL as a target in breast cancer therapy. J Oncol. https://doi.org/10.1155/2020/5291952

Demaria M, Giorgi C, Lebiedzinska M, Esposito G, D'Angeli L, Bartoli A, Gough DJ, Turkson J, Levy DE, Watson CJ, Wieckowski MR, Provero P, Pinton P, Poli V (2010) A STAT3-mediated metabolic switch is involved in tumour transformation and STAT3 addiction. Aging (Albany NY) 2:823–842. https://doi.org/10.18632/aging.100232

Gamero AM, Young MR, Mentor-Marcel R, Bobe G, Scarzello AJ, Wise J, Colburn NH (2010) STAT2 contributes to promotion of colorectal and skin carcinogenesis. Cancer Prev Res (Phila) 3:495. https://doi.org/10.1158/1940-6207.CAPR-09-0105

Goel RK, Singh D, Lagunin A, Poroikov V (2011) PASS-assisted exploration of new therapeutic potential of natural products. Med Chem Res 20:1509–1514. https://doi.org/10.1007/s00044-010-9398-y

Goel S, DeCristo MJ, Watt AC, BrinJones H, Sceneay J, Li BB, Khan N, Ubellacker JM, Xie S, Metzger-Filho O, Hoog J, Ellis MJ, Ma CX, Ramm S, Krop IE, Winer EP, Roberts TM, Kim H-J, McAllister SS, Zhao JJ (2017) CDK4/6 inhibition triggers anti-tumour immunity. Nature 548:471–475. https://doi.org/10.1038/nature23465

Han Y, Liu D, Li L (2020) PD-1/PD-L1 pathway: current researches in cancer. Am J Cancer Res 10:727–742. http://www.ncbi.nlm.nih.gov/pubmed/32266087

Harjunpää H, Guillerey C (2020) TIGIT as an emerging immune checkpoint. Clin Exp Immunol 200:108–119. https://doi.org/10.1111/cei.13407

Hou X, Zhou G, Fan Y, Zhang Q, Xiang C, Cao F, Yao S (2020) The high expression of CD276/HAVCR2 and CD163 is an adverse immune subtype of glioblastoma and is closely related to epithelial-mesenchymal transition. https://doi.org/10.21203/rs.3.rs-31174/v1

Hu Y, Sun H, Hu J, Zhang X (2020) LncRNA DLX6-AS1 promotes the progression of neuroblastoma by activating STAT2 via targeting miR-506-3p. Cancer Manag Res 12:7451. https://doi.org/10.2147/CMAR.S252521

Huang Y-H, Zhu C, Kondo Y, Anderson AC, Gandhi A, Russell A, Dougan SK, Petersen B-S, Melum E, Pertel T, Clayton KL, Raab M, Chen Q, Beauchemin N, Yazaki PJ, Pyzik M, Ostrowski MA, Glickman JN, Rudd CE, Ploegh HL, Franke A, Petsko GA, Kuchroo VK (2015) Blumberg, CEACAM1 regulates TIM-3-mediated tolerance and exhaustion. Nature 517:386–390. https://doi.org/10.1038/nature13848

Huang Q, Duan L, Qian X, Fan J, Lv Z, Zhang X, Han J, Wu F, Guo M, Hu G, Du J, Chen C, Jin Y (2016) IL-17 promotes angiogenic factors IL-6, IL-8, and Vegf production via Stat1 in lung adenocarcinoma. Sci Rep 6:36551. https://doi.org/10.1038/srep36551

Kamizono S, Hanada T, Yasukawa H, Minoguchi S, Kato R, Minoguchi M, Hattori K, Hatakeyama S, Yada M, Morita S, Kitamura T, Kato H, Nakayama K, Yoshimura A (2001) The SOCS Box of SOCS-1 accelerates ubiquitin-dependent proteolysis of TEL-JAK2. J Biol Chem 276:12530–12538. https://doi.org/10.1074/jbc.M010074200

Kamran MZ, Patil P, GudeRP, (2013) Role of STAT3 in cancer metastasis and translational advances. Biomed Res Int. https://doi.org/10.1155/2013/421821

Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, Wang J, Yu B, Zhang J, Bryant SH (2016) PubChem substance and compound databases. Nucleic Acids Res 44:D1202–D1213. https://doi.org/10.1093/nar/gkv951

Kortylewski M, Yu H (2008) Role of Stat3 in suppressing anti-tumor immunity. Curr Opin Immunol 20:228. https://doi.org/10.1016/J.COI.2008.03.010

Kortylewski M, Kujawski M, Wang T, Wei S, Zhang S, Pilon-Thomas S, Niu G, Kay H, Mulé J, Kerr WG, Jove R, Pardoll D, Yu H (2005) Inhibiting Stat3 signaling in the hematopoietic system elicits multicomponent antitumor immunity. Nat Med 11:1314–1321. https://doi.org/10.1038/nm1325

Krasnov GS, Kudryavtseva AV, Snezhkina AV, Lakunina VA, Beniaminov AD, Melnikova NV, Dmitriev AA (2019) Pan-cancer analysis of TCGA data revealed promising reference genes for qPCR normalization. Front Genet. https://doi.org/10.3389/fgene.2019.00097

Krzyszczyk P, Acevedo A, Davidoff EJ, Timmins LM, Marrero-Berrios I, Patel M, White C, Lowe C, Sherba JJ, Hartmanshenn C, O'Neill KM, Balter ML, Fritz ZR, Androulakis IP, Schloss RS, Yarmush ML (2018) The growing role of precision and personalized medicine for cancer treatment. Technology 6:79–100. https://doi.org/10.1142/S2339547818300020

Larkin J, Ahmed CM, Wilson TD, Johnson HM (2013) Cells T. Front Immunol. https://doi.org/10.3389/fimmu.2013.00469

Liau NPD, Laktyushin A, Lucet IS, Murphy JM, Yao S, Whitlock E, Callaghan K, Nicola NA, Kershaw NJ, Babon JJ (2018) The molecular basis of JAK/STAT inhibition by SOCS1. Nat Commun 9:1558. https://doi.org/10.1038/s41467-018-04013-1

Liu Y, Hu X, Han C, Wang L, Zhang X, He X, Lu X (2015) Targeting tumor suppressor genes for cancer therapy. BioEssays 37:1277–1286. https://doi.org/10.1002/bies.201500093

Llanes-Fernández L, Álvarez-Goyanes RI, Arango-Prado MdelC, Alcocer-González JM, Mojarrieta JC, Pérez XE, López MO, Odio SF, Camacho-Rodríguez R, Guerra-Yi ME, Madrid-Marina V, Tamez-Guerra R (2006) Rodríguez-Padilla, Relationship between IL-10 and tumor markers in breast cancer patients. The Breast 15:482–489. https://doi.org/10.1016/j.breast.2005.09.012

Long L, Zhang X, Chen F, Pan Q, Phiphatwatchara P, Zeng Y, Chen H (2018) The promising immune checkpoint LAG-3: from tumor

microenvironment to cancer immunotherapy. Genes Cancer 9:176–189. https://doi.org/10.18632/genesandcancer.180

Lawrence MS, Sougnez C, Lichtenstein L, Cibulskis K, Lander E, Gabriel SB, Getz G, Ally A, Balasundaram M, Birol I, Bowlby R, Brooks D, Butterfield YSN, Carlsen R, Cheng D, Chu A, Dhalla N, Guin R, Holt RA, Jones SJM, Lee D, Li HI, Marra MA, Mayo M, Moore RA, Mungall AJ, Robertson AG, Schein JE, Sipahimalani P, Tam A, Thiessen N, Wong T, Protopopov A, Santoso N, Lee S, Parfenov M, Zhang J, Mahadeshwar HS, Tang J, Ren X, Seth S, Haseley P, Zeng D, Yang L, Xu AW, Song X, Pantazi A, Bristow CA, Hadjipanayis A, Seidman J, Chin L, Park PJ, Kucherlapati R, Akbani R, Casasent T, Liu W, Lu Y, Mills G, Motter T, Weinstein J, Diao L, Wang J, Hong Fan Y, Liu J, Wang K, Auman JT, Balu S, Bodenheimer T, Buda E, Hayes DN, Hoadley KA, Hoyle AP, Jefferys SR, Jones CD, Kimes PK, Liu Y, Marron JS, Meng S, Mieczkowski PA, Mose LE, Parker JS, Perou CM, Prins JF, Roach J, Shi Y, Simons JV, Singh D, Soloway MG, Tan D, Veluvolu U, Walter V, Waring S, Wilkerson MD, Wu J, Zhao N, Cherniack AD, Hammerman PS, Tward AD, Pedamallu CS, Saksena G, Jung J, Ojesina AI, Carter SL, Zack TI, Schumacher SE, Beroukhim R, Freeman SS, Meyerson M, Cho J, Noble MS, DiCara D, Zhang H, Heiman DI, Gehlenborg N, Voet D, Lin P, Frazer S, Stojanov P, Liu Y, Zou L, Kim J, Muzny D, Doddapaneni HV, Kovar C, Reid J, Morton D, Han Y, Hale W, Chao H, Chang K, Drummond JA, Gibbs RA, Kakkar N, Wheeler D, Xi L, Ciriello G, Ladanyi M, Lee W, Ramirez R, Sander C, Shen R, Sinha R, Weinhold N, Taylor BS, Aksoy BA, Dresdner G, Gao J, Gross B, Jacobsen A, Reva B, Schultz N, Sumer SO, Sun Y, Chan TA, Morris LG, Stuart J, Benz S, Ng S, Benz C, Yau C, Baylin SB, Cope L, Danilova L, Herman JG, Bootwalla M, Maglinte DT, Laird PW, Triche T, Weisenberger DJ, Van Den Berg DJ, Agrawal N, Bishop J, Boutros PC, Bruce JP, Byers LA, Califano J, Carey TE, Chen Z, Cheng H, Chiosea SI, Cohen E, Diergaarde B, Egloff AM, El-Naggar AK, Ferris RL, Frederick MJ, Grandis JR, Guo Y, Haddad RI, Harris T, Hui ABY, Lee JJ, Lippman SM, Liu FF, McHugh JB, Myers J, Ng PKS, Perez-Ordonez B, Pickering CR, Prystowsky M, Romkes M, Saleh AD, Sartor MA, Seethala R, Seiwert TY, Si H, Van Waes C, Waggott DM, Wiznerowicz M, Yarbrough WG, Zhang J, Zuo Z, Burnett K, Crain D, Gardner J, Lau K, Mallery D, Morris S, Paulauskis J, Penny R, Shelton C, Shelton T, Sherman M, Yena P, Black AD, Bowen J, Frick J, Gastier-Foster JM, Harper HA, Leraas K, Lichtenberg TM, Ramirez NC, Wise L, Zmuda E, Baboud J, Jensen MA, Kahn AB, Pihl TD, Pot DA, Srinivasan D, Walton JS, Wan Y, Burton RA, Davidsen T, Demchok JA, Eley G, Ferguson ML, Mills Shaw KR, Ozenberger BA, Sheth M, Sofia HJ, Tarnuzzer R, Wang Z, Yang L, Zenklusen JC, Saller C, Tarvin K, Chen C, Bollag R, Weinberger P, Golusiński W, Golusiński P, Ibbs M, Korski K, Mackiewicz A, Suchorska W, Szybiak B, Curley E, Beard C, Mitchell C, Sandusky G, Ahn J, Khan Z, Irish J, Waldron J, William WN, Egea S, Gomez-Fernandez C, Herbert L, Bradford CR, Chepeha DB, Haddad AS, Jones TR, Komarck CM, Malakh M, Moyer JS, Nguyen A, Peterson LA, Prince ME, Rozek LS, Taylor EG, Walline HM, Wolf GT, Boice L, Chera BS, Funkhouser WK, Gulley ML, Hackman TG, Hayward MC, Huang M, Rathmell WK, Salazar AH, Shockley WW, Shores CG, Thorne L, Weissler MC, Wrenn S, Zanation AM, Brown BT, Pham M (2015) Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature 517:576–582. https://doi.org/10.1038/nature14129

Ludwig KF, Du W, Sorrelle NB, Wnuk-Lipinska K, Topalovski M, Toombs JE, Cruz VH, Yabuuchi S, Rajeshkumar NV, Maitra A, Lorens JB, Brekken RA (2018) Small-molecule inhibition of Axl targets tumor immune suppression and enhances chemotherapy in pancreatic cancer. Cancer Res 78:246–255. https://doi.org/10.1158/0008-5472.CAN-17-1973

Mercer F, Unutmaz D (2009) The biology of FoxP3: A key player in immune suppression during infections. In: Autoimmune diseases and cancer, pp 47–59. https://doi.org/10.1007/978-1-4419-1599-3_4

Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem 30:2785–2791. https://doi.org/10.1002/jcc.21256

Oft M (2014) IL-10 master switch from tumor-promoting inflammation to antitumor immunity. Cancer Immunol Res 2:194–199. https://doi.org/10.1158/2326-6066.CIR-13-0214

Osborn SL, Diehl G, Han S-J, Xue L, Kurd N, Hsieh K, Cado D, Robey EA, Winoto A (2010) Fas-associated death domain (FADD) is a negative regulator of T cell receptor-mediated necroptosis. Proc Natl Acad Sci 107:13034–13039. https://doi.org/10.1073/pnas.1005997107

Otto T, Sicinski P (2017) Cell cycle proteins as promising targets in cancer therapy. Nat Rev Cancer 17:93–115. https://doi.org/10.1038/nrc.2016.138

Papoff G, Trivieri N, Crielesi R, Ruberti F, Marsilio S, Ruberti G (2010) FADD–calmodulin interaction: a novel player in cell cycle regulation. Biochim Biophys Acta Mol Cell Res 1803:898–911. https://doi.org/10.1016/j.bbamcr.2010.04.006

Paul MK, Mukhopadhyay AK (2004) Tyrosine kinase: role and significance in cancer. Int J Med Sci. https://doi.org/10.7150/ijms.1.101

Pawlus MR, Wang L, Hu C-J (2014) STAT3 and HIF1α cooperatively activate HIF1 target genes in MDA-MB-231 and RCC4 cells. Oncogene 33:1670–1679. https://doi.org/10.1038/onc.2013.115

Pestell KE (2003) Paul Workman on the challenges of cancer drug development. Drug Discov Today 8:775–777. https://doi.org/10.1016/S1359-6446(03)02838-1

Rankin E, Giaccia A (2016) The receptor tyrosine kinase AXL in cancer progression. Cancers (Basel) 8:103. https://doi.org/10.3390/cancers8110103

Sawyers C (2002) Rational therapeutic intervention in cancer: kinases as drug targets. Curr Opin Genet Dev 12:111–115. https://doi.org/10.1016/S0959-437X(01)00273-8

Schwartz DM, Kanno Y, Villarino A, Ward M, Gadina M, O'Shea JJ (2017) JAK inhibition as a therapeutic strategy for immune and inflammatory diseases. Nat Rev Drug Discov 17:78. https://doi.org/10.1038/NRD.2017.267

Schwartz DM, Kanno Y, Villarino A, Ward M, Gadina M, O'Shea JJ (2018) Erratum: JAK inhibition as a therapeutic strategy for immune and inflammatory diseases. Nat Rev Drug Discov 17:78–78. https://doi.org/10.1038/nrd.2017.267

Shao L, Hou W, Scharping NE, Vendetti FP, Srivastava R, Roy CN, Menk AV, Wang Y, Chauvin J-M, Karukonda P, Thorne SH, Hornung V, Zarour HM, Bakkenist CJ, Delgoffe GM (2019) Sarkar, IRF1 inhibits antitumor immunity through the upregulation of PD-L1 in the tumor cell. Cancer Immunol Res 7:1258–1266. https://doi.org/10.1158/2326-6066.CIR-18-0711

Sheikhpour E, Noorbakhsh P, Foroughi E, Farahnak S, Nasiri R, Neamatzadeh H (2018) A survey on the role of interleukin-10 in breast cancer: a narrative. Rep Biochem Mol Biol 7:30–37

Sriuranpong V, Park JI, Amornphimoltham P, Patel V, Nelkin BD, Gutkind JS (2003) Epidermal growth factor receptor-independent constitutive activation of STAT3 in head and neck squamous cell carcinoma is mediated by the autocrine/paracrine stimulation of the interleukin 6/gp130 cytokine system. Cancer Res 63:2948–2956. http://www.ncbi.nlm.nih.gov/pubmed/12782602

Stehelin D, Varmus HE, Bishop JM, Vogt PK (1976) DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. Nature 260:170–173. https://doi.org/10.1038/260170a0

Sur I, Taipale J (2016) The role of enhancers in cancer. Nat Rev Cancer 16:483–493. https://doi.org/10.1038/nrc.2016.62

Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ, von Mering C (2017) The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. Nucleic Acids Res 45:D362–D368. https://doi.org/10.1093/nar/gkw937

Tadesse S, Yu M, Kumarasiri M, Le BT, Wang S (2015) Targeting CDK6 in cancer: state of the art and new insights. Cell Cycle 14:3220–3230. https://doi.org/10.1080/15384101.2015.1084445

Tan KT, Yeh C-N, Chang Y-C, Cheng J-H, Fang W-L, Yeh Y-C, Wang Y-C, Hsu DS-S, Wu C-E, Lai J-I, Chang PM-H, Chen M-H, Lu M-L, Chen S-J, Chao Y, Hsiao M, Chen M-H (2020) PRKDC: new biomarker and drug target for checkpoint blockade immunotherapy. J Immunother Cancer 8:e000485. https://doi.org/10.1136/jitc-2019-000485

Tanegashima T, Togashi Y, Azuma K, Kawahara A, Ideguchi K, Sugiyama D, Kinoshita F, Akiba J, Kashiwagi E, Takeuchi A, Irie T, Tatsugami K, Hoshino T, Eto M, Nishikawa H (2019) Immune suppression by PD-L2 against spontaneous and treatment-related antitumor immunity. Clin Cancer Res 25:4808–4819. https://doi.org/10.1158/1078-0432.CCR-18-3991

Thomas SJ, Snowden JA, Zeidler MP, Danson SJ (2015) The role of JAK/STAT signalling in the pathogenesis, prognosis and treatment of solid tumours. Br J Cancer 113:365–371. https://doi.org/10.1038/bjc.2015.233

Tobelaim WS, Beaurivage C, Champagne A, Pomerleau V, Simoneau A, Chababi W, Yeganeh M, Thibault P, Klinck R, Carrier JC, Ferbeyre G, Ilangumaran S, Saucier C (2015) Tumour-promoting role of SOCS1 in colorectal cancer cells. Sci Rep 5:14301. https://doi.org/10.1038/srep14301

Torres N, Regge MV, Secchiari F, Friedrich AD, Spallanzani RG, Raffo Iraolagoitia XL, Núñez SY, Sierra JM, Ziblat A, Santilli MC, Gilio N, Almada E, Lauche C, Pardo R, Domaica CI, Fuertes MB, Madauss KP, Hance KW, Gloger IS, Zylberman V, Goldbaum FA, Zwirner NW (2020) Restoration of antitumor immunity through anti-MICA antibodies elicited with a chimeric protein. J Immunother Cancer 8:e000233. https://doi.org/10.1136/jitc-2019-000233

Twyman-Saint Victor C, Rech AJ, Maity A, Rengan R, Pauken KE, Stelekati E, Benci JL, Xu B, Dada H, Odorizzi PM, Herati RS, Mansfield KD, Patsch D, Amaravadi RK, Schuchter LM, Ishwaran H, Mick R, Pryma DA, Xu X, Feldman MD, Gangadhar TC, Hahn SM, Wherry EJ, Vonderheide RH, Minn AJ (2015) Radiation and dual checkpoint blockade activate non-redundant immune mechanisms in cancer. Nature 520:373–377. https://doi.org/10.1038/nature14292

Ullrich CI, Aloni R, Saeed MEM, Ullrich W, Efferth T (2019) Comparison between tumors in plants and human beings: mechanisms of tumor development and therapy with secondary plant metabolites. Phytomedicine 64:153081. https://doi.org/10.1016/j.phymed.2019.153081

Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW (2013) Cancer genome landscapes. Science 339:1546–1558. https://doi.org/10.1126/science.1235122

Wang H, Brown J, Garcia CA, Tang Y, Benakanakere MR, Greenway T, Alard P, Kinane DF, Martin M (2011) The role of glycogen synthase kinase 3 in regulating IFN-β–mediated IL-10 production. J Immunol 186:675–684. https://doi.org/10.4049/jimmunol.1001473

Wehde BL, Rädler PD, Shrestha H, Johnson SJ, Triplett AA, Wagner K-U (2018) Janus kinase 1 plays a critical role in mammary cancer progression. Cell Rep 25:2192. https://doi.org/10.1016/J.CELREP.2018.10.063

Wei Y, Chen X, Yang J, Yao J, Yin N, Zhang Z, Li D, Zhu D, Zhou J (2019) DcR3 promotes proliferation and invasion of pancreatic cancer via a DcR3/STAT1/IRF1 feedback loop. Am J Cancer Res 9:2618–2633. http://www.ncbi.nlm.nih.gov/pubmed/31911850

Wuputra K, Ku C-C, Wu D-C, Lin Y-C, Saito S, Yokoyama KK (2020) Prevention of tumor risk associated with the reprogramming of human pluripotent stem cells. J Exp Clin Cancer Res 39:100. https://doi.org/10.1186/s13046-020-01584-0

Xia S, Ji R, Xu Y, Ni X, Dong Y, Zhan W (2017) Twisted gastrulation BMP signaling modulator 1 regulates papillary thyroid cancer cell motility and proliferation. J Cancer 8:2816–2827. https://doi.org/10.7150/jca.18482

Zhang Y, Zhaoyoung L (2017) STAT1 in cancer: friend or foe? Discov Med 24:19–29. https://www.discoverymedicine.com/Ying-Zhang/2017/08/stat1-in-cancer-friend-or-foe/

Yang L, Pang Y, Moses HL (2010) TGF-β and immune cells: an important regulatory axis in the tumor microenvironment and progression. Trends Immunol 31:220–227. https://doi.org/10.1016/j.it.2010.04.002

Yang S, Liu Y, Li M-Y, Ng CSH, Yang S, Wang S, Zou C, Dong Y, Du J, Long X, Liu L-Z, Wan IYP, Mok T, Underwood MJ, Chen GG (2017) FOXP3 promotes tumor growth and metastasis by activating Wnt/β-catenin signaling pathway and EMT in non-small cell lung cancer. Mol Cancer 16:124. https://doi.org/10.1186/s12943-017-0700-1

Yi T, Lee D-S, Jeon M-S, Kwon SW, Song SU (2012) Gene expression profile reveals that STAT2 is involved in the immunosuppressive function of human bone marrow-derived mesenchymal stem cells. Gene 497:131–139. https://doi.org/10.1016/j.gene.2012.01.073

Yu W, Hua Y, Qiu H, Hao J, Zou K, Li Z, Hu S, Guo P, Chen M, Sui S, Xiong Y, Li F, Lu J, Guo W, Luo G, Deng W (2020) PD-L1 promotes tumor growth and progression by activating WIP and β-catenin signaling pathways and predicts poor prognosis in lung cancer. Cell Death Dis 11:506. https://doi.org/10.1038/s41419-020-2701-z

Zhang J, Li H, Yu J-P, Wang SE, Ren X-B (2012) Role of SOCS1 in tumor progression and therapeutic application. Int J Cancer 130:1971–1980. https://doi.org/10.1002/ijc.27318

Zhang Y, Yang W, Wen G, Tang H, Wu C, Wu Y, Jing Z, Tang M, Liu G, Li D, Li Y, Deng Y (2019) High expression of PRKDC promotes breast cancer cell growth via p38 MAPK signaling and is associated with poor survival. Mol Genet Genomic Med. https://doi.org/10.1002/mgg3.908

Zhao Y, Yang W, Huang Y, Cui R, Li X, Li B (2018) Evolving roles for targeting CTLA-4 in cancer immunotherapy. Cell Physiol Biochem 47:721–734. https://doi.org/10.1159/000490025

Zhou X-M, Li W-Q, Wu Y-H, Han L, Cao X-G, Yang X-M, Wang H-F, Zhao W-S, Zhai W-J, Qi Y-M, Gao Y-F (2018) Intrinsic expression of immune checkpoint molecule TIGIT could help tumor growth in vivo by suppressing the function of NK and CD8+ T cells. Front Immunol. https://doi.org/10.3389/fimmu.2018.02821

SUNIL KUMAR, PhD

Phone: +91-9996674089

E-Mail: princeboora@gmail.com

Corresponding Address:

Immunology and immunotherapeutic Laboratory, Department of Biotechnology, Delhi Technological University, Shahbad Daulatpur, Bawana Road, Delhi: 110042

EDUCATIONAL BACKGROUND

2019-2023	Delhi Technological University	Ph.D.	Immunoinformatic, Biotechnology
2016-2018	Delhi Technological University	M.Tech.	Bioinformatics
2012-2016	Deen Bandhu Sir Chhotu Ram University of Science and Technology, Muthal, Haryana	B.Tech.	Biotechnology

PERSONAL STATEMENT

My scientific research interests involve the Design of combinatorial therapy for tumor diagnostic and prognosis by different bioinformatic techniques and Biostatistics pipelines. I understand Cancer diseases through the involvement of multiple signaling pathways and involvement of immune system at different phases of cancer. My academic training, research experience, teaching assistance, and scientific training experience have provided me with excellent background on multiple discipline, such as computational biology, immunoinformatic, chemoinformatic, drug designing, drug discovery, proteomic studies, genetics, and molecular biology. As a master's student, I was feeling lucky to work under the supervision of Dr. Asmita Das on the prediction of novel epitopes for design of vaccine against cancer. I gained the expertise in the machine learning models, network biology, and transcriptional regulation of disease progression, Docking and MD Simulation. As a doctoral student under the supervision of Dr. Asmita Das, I was able to implement my experience of computational biology, machine learning, bioinformatics tools, network and structural biology in Designing of combinatorial therapy for tumor treatment. In my doctoral training, I published several papers in major peer reviewed journals. In my recent publication, I identified immunological biomarkers and their importance in Head and Neck Cancer and Design a cocktail of natural compounds to reverse the expression of immunological biomarkers. Different tumor samples and Peripheral Blood mononuclear cells samples were analyzed for the identification of Differentially expressed genes. Different natural compounds were screened for the reversal of expression of these genes. A combination of natural compounds was formed based on their regulating signaling pathways. I also aim to identify Diagnostic Biomarker for the early detection of breast cancer from blood samples using machine learning and explainable artificial intelligence (XAI) by analyzing more than 500 patients blood samples. Later on, we aim to identify potential indirect target SOCS1 and screening natural compound by molecular docking and MD Simulation approach that will be act as potential therapeutic agent in Cancer. In my doctoral training, I assisted in different laboratory and teaching class like Animal Biotechnology, immunotherapeutic, immunology for M.Tech., M.Sc., B.Tech. students in Delhi Technological University. For my future career training, I

will continue to build on my previous training on regulation of disease progression through involvement of translational modification and multiple signaling pathways.

SKILLS AND EXPERTISE

COMPUTATIONAL BIOLOGY AND BIOINFORMATICS:
Sequence Homology, Network Biology, Molecular Docking, Virtual Screening, Bioinformatic Tools, Blast, Data Processing and Visualization, Sequence Alignment, Multi-Omics Data Integration and Analysis, Microbiome Data Analysis, Microarray and RNA-Seq Data Analysis, Machine Learning Classification and Regression, Epigenetics Data Analysis, Artificial Intelligence Algorithms and explainable artificial intelligence.

MOLECULAR BIOLOGY:
Mammalian Cell Culture, Nucleic Acid and Protein Extraction, Nucleic Acid and Protein Quantification, MTT Assay, Western Blotting, Drug toxicity assay, Cytotoxicity Assay.

POSITIONS, TRAINING, AND CERTIFICATIONS

Teaching Assistant:

- **ANIMAL BIOTECHNOLOGY LABORATORY (BT304)**
  The subject focusses on the techniques in animal cell culture, cell revival, suspension and adherent cell culture and passaging, cell freezing, MTT Assay, media preparation, cell counting, establishing primary and secondary cell culture.
- **IMMUNOTHERAPUTIC LABORATORY (BIO503)**
  The subject was taught to master students that focusses on the techniques of In-Silico Vaccine Design like epitope prediction for B-cell and T-cell, antigenicity prediction, population coverage, Protein-protein docking and simulation, protein structure modelling.

Certifications:

1. Introduction To the Biology of Cancer
2. Understanding Cancer Metastasis
3. Programming For Everybody (Getting Started with Python)
4. Whole Genome Sequencing of Bacterial Genomes- Tools and Application
5. Cancer Immunology

CONTRIBUTION TO SCIENCE

1. ***Early Career*:** My early career focusses on identification of novel therapeutic epitope prediction for novel cancer vaccines in which we predict B cell epitopes, T cytotoxic cell epitope, T helper cell epitopes, their immunogenicity, toxicity, population coverage, protein interaction check by docking. Identification of oncogenes which can be used as a candidate of vaccine.
2. ***Doctoral Career***: In doctoral career, I forwarded my initially studies to a next level and working on the project entitled "Combinatorial therapy for tumor treatment" and "Machine learning and artificial intelligence in Biomarkers prediction and drug discovery.

I am interested in identifying differential regulated genes in the tumor patient and trying to reverse their expression with the help of a cocktail of natural compounds. We try to identify blood-based biomarkers which can be used for easy and early diagnostic of cancer with the help of machine learning and explainable artificial intelligence.

PUBLICATIONS AND PRESENTATIONS

Publications

1. **Sunil Kumar**, Asmita Das A cocktail of natural compounds holds promise for new immunotherapeutic potential in head and neck cancer. Chinese Journal of Integrative Medicine September 2022. DOI: **10.1007/s11655-023-3694-0**
2. **Sunil Kumar**, Asmita Das Peripheral Blood Mononuclear Cell derived Biomarker detection using eXplainable Artificial Intelligence (XAI) provides better diagnosis of Breast Cancer. Computational Biology and Chemistry. DOI: **10.1016/j.compbiolchem.2023.107867**
3. Saksham Garg, **Sunil Kumar**, Ashutosh Anand, Tarunya Menon, Nikita Sharma, Japneet Singh, Siddharth Chawla, Asmita Das Plant-derived natural compounds aiding SOCS1 mediated JAK1 inhibition, a novel mechanism of combinatorial cancer chemotherapy. Vegetos journal February 2022 DOI: **https://doi.org/10.1007/s42535-021-00329-4**.
4. **Sunil Kumar**, Asmita Das Elucidation of natural compounds Gallic acid and Shikonin for the treatment of HNSC cancer by targeting immune suppressor and tumour progressor genes. Vegetos journal March 2022 DOI: **https://doi.org/10.1007/s42535-022-00363-w**.

Presentations in National/International Conferences

- A combinatorial Therapy of Natural Compounds Targeting Tumor Progression and Immune Suppression for Breast Cancer accepted for publication in international conference on innovations in Biotechnology and Life Science (ICIBLS 2020).
- A Natural Compounds Cocktail for Breast Cancer presented on 35th conference on Preventive Oncology and Diagnostic cancer (April 2021)

DECLARATION

I hereby declare that the given information is accurate to the best of my knowledge and belief and can be supported with reliable documents when needed.

Sunil Kumar
Date: 30.07.2023
Place: New Delhi