University of Southampton

Faculty of Environmental and Life Sciences

School of Biological Sciences

Effects of active ingredients from *Limonium Sinense* (Girard) Kuntze extracts: a potential link with hypoxia-inducible factor (HIF) activation

by

Hualong Zhao

ORCID ID: 0000-0002-5009-1444

Thesis for the degree of Doctoral of Philosophy

February 2024

University of Southampton

Abstract

Faculty of Environmental and Life Sciences

School of Biological Sciences

Doctor of Philosophy

Effects of active ingredients from *Limonium Sinense* (Girard) Kuntze extracts:
a potential link with hypoxia-inducible factor (HIF) activation

by

Hualong Zhao

Limonium Sinense (Girard) Kuntze is a traditional Chinese medicinal herb with a wide range of therapeutic uses, including fever, hepatitis, haemostasis, anaemia, menorrhagia, irregular menstruation, and other disorders. Recent studies revealed the antioxidant, anti-tumour, anti-hepatitis, anti-viral, and immunomodulatory activities of *Limonium Sinense* amongst others. However, the precise mechanisms underlying its pharmacological activities remain largely unknown.

This study aimed to investigate the effects of bioactive ingredients from *Limonium Sinense* and elucidate their underlying mechanisms. The therapeutic effects of *Limonium Sinense* were predicted using a network pharmacology approach. Integrated analysis including bioinformatic analysis and *in vitro* experiments was employed to explore the effects of bioactive ingredients from *Limonium Sinense*. A structure-based virtual screening approach was conducted to identify potential inhibitors of prolyl hydroxylase domain protein 2 (PHD2) from Hydroxybenzoic acids and their derives. The network pharmacology analysis revealed the involvement of the breast cancer pathway and several cancer-related pathways in the biological effects of *Limonium Sinense*. Water extracts from *Limonium Sinense* (LSW) showed a strong growth inhibitory effect on multiple cells in both 2D and 3D cultures. Global transcriptomic profiling and connectivity map (CMap) analysis identified several similarly acting therapeutic candidates, including Tubulin inhibitors and hypoxia-inducible factor (HIF) modulators. The effect of LSW on the cell cycle was verified with flow cytometry showing a G2/M phase arrest. Integrated analysis suggested a role for gallic acid in mediating HIF activation. Additionally, three potent PHD2 inhibitors were identified, which can potentially bind to the 2OG binding pocket of PHD2, leading to the inactivation of the enzyme.

Taken together, this study provides novel insights into the bioactive ingredients in *Limonium Sinense*, highlighting the rich natural resource and therapeutic values of herbal plants.

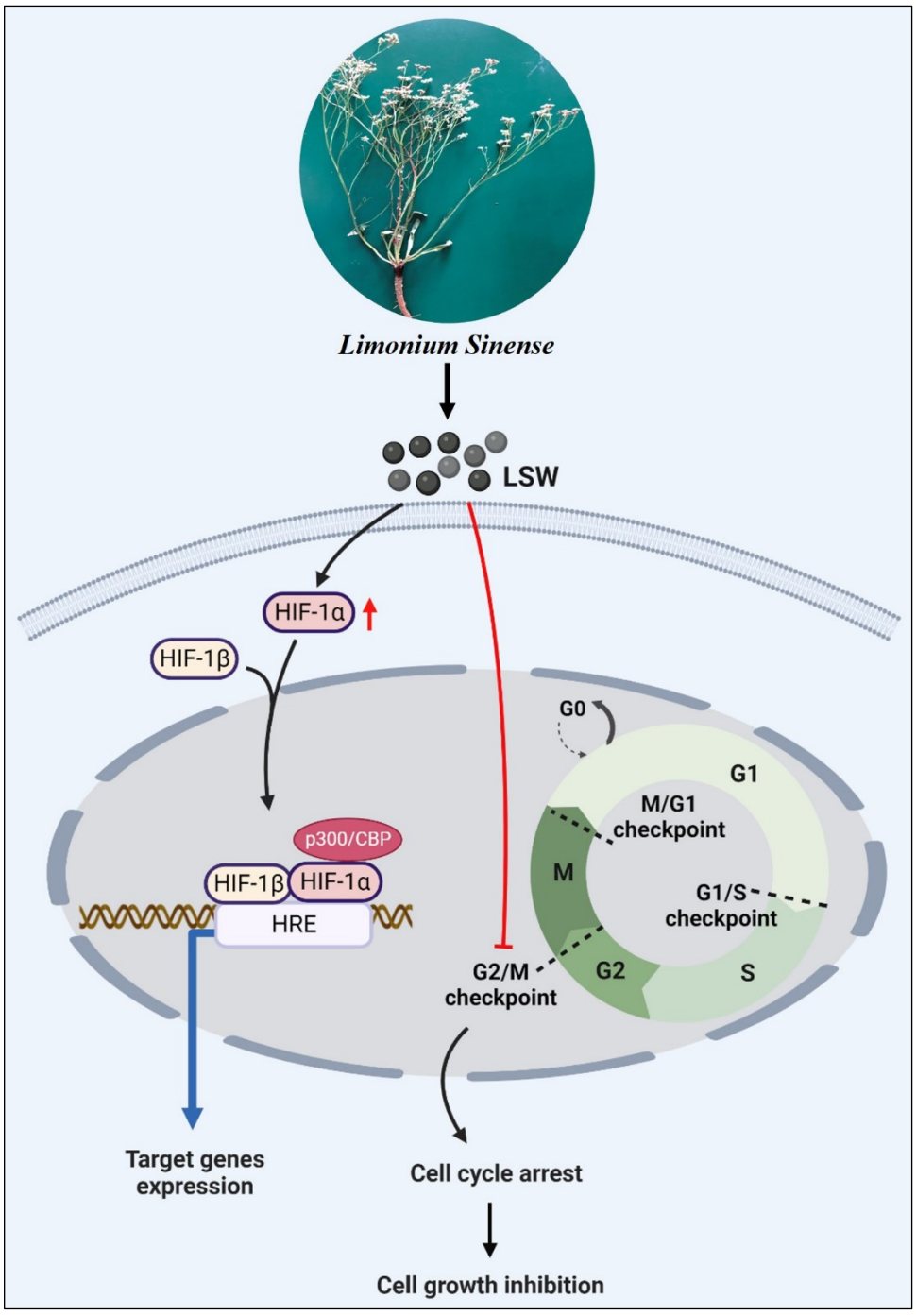


Figure 1. Graphic Abstract: *Limonium Sinense* (Girard) Kuntze and its biological effects in activating HIF and inducing cell cycle arrest. The whole plants of *Limonium Sinense* were collected from the coastal region in Jiangsu, eastern China (33°09'33.0" N, 120°46'40.4" E). The water extracts from *Limonium Sinense* (LSW) exhibit a significant capability to induce cell cycle arrest at G2/M checkpoint, leading to the inhibition of cell growth. On the other hand, LSW prominently stimulates the upregulation of HIF-1α at the protein level, leading to its accumulation. The HIF-1α then translocates into the nucleus, where it binds with co-factors HIF-1β, p300/CBP and HRE, ultimately activating the transcription of its target genes. CBP, CREB-binding protein; HRE, hypoxia response element; p300, E1A binding protein p300; HIF-1α, Hypoxia-inducible factor 1 alpha; HIF-1β, Hypoxia-inducible factor 1 beta; LSW: *Limonium Sinense* water extract.

Table of Contents

Table of Tables

Table of Figures

Research Thesis: Declaration of Authorship

Print name: Hualong Zhao

Title of thesis: Effects of active ingredients from *Limonium Sinense* (Girard) Kuntze extracts: a potential link with hypoxia-inducible factor (HIF) activation

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:

Zhao, H., et al., Integrated analysis reveals effects of bioactive ingredients from *Limonium Sinense* (Girard) Kuntze on hypoxia-inducible factor (HIF) activation. Front Plant Sci, 2022. 13: p. 994036. doi: 10.3389/fpls.2022.994036.

Zhao H, Wang S, et al. Integrated network pharmacology and cellular assay reveal the biological mechanisms of *Limonium sinense* (Girard) Kuntze against breast cancer. BMC Complement Med Ther 2023; https://doi.org/10.21203/rs.3.rs-3068701/v1 (Under view)

Signature: Date:

Acknowledgements

I would like to express my heartfelt gratitude to my supervisor Dr. Yihua Wang for his dedicated guidance and invaluable support throughout my PhD life. His insightful suggestions helped me to think and work as a scientist. I am truly grateful for the opportunity to work with him and for his guidance in achieving my research goals. I would also like to thank my co-supervisors Prof. Rob. Ewing and Prof. P.T. Williamson for their guidance throughout this project. Their expertise and assistance were invaluable whenever I encountered challenges.

I appreciate the help of the Biological Sciences PhD Studentship, and the Faculty of Environmental and Life Sciences Gradate School. I deeply thank Lorraine Prout, Lindy Holden-Dye and Herman Wijnen for helping me overcome tremendous financial pressures and provided me with the opportunity to continue my studies. Thank Medical Research Council [grant no: MR/S025480/1] and an Academy of Medical Sciences/the Wellcome Trust Springboard Award [grant no: SBF002/1038] to support my research.

I would like to thank Yancheng Teachers' University for granting me the chance to study in Southampton and for their support in this project.

I would like to thank Dr. Matthias Baud from the Chemistry department for providing valuable advice on the isolation of natural ingredients. Thanks to Michael Mccoy for his assistance in extracting compounds from the *Limonium Sinense* plant. I appreciate Prof. Paul Skipp for his advice on connectivity map analysis. Thanks to Luis Coy for generously providing breast cells for my experiments. I would like to thank Ayse Ertay for her helpful suggestions at the beginning of my PhD career, thank Steve John for his kindly help in both experiments and daily life, thank Noor Shamki for the help of flow cytometry, thank Lloyd Steele-Nicholson for the help of microscopy assay. I thank Yilu Zhou for his assistance in bioinformatic analysis. I would like to thank all the staff in the 85 building for creating an excellent working environment and for their valuable training in experimental techniques.

I would like to thank all the members of Dr. Yihua Wang's group, Juanjuan, Charlotte, Liudi, Ayse, Yilu, Zijian, Siyuan, Beatriz, Zhe Chen, Tao Guo and Xi Li, for their help and the positive working atmosphere in the lab. I really appreciate working all of you.

Finally, I want to sincerely thank my family, my parents, my wife, my friends, and especially my daughter, for their endless support and encouragement throughout my studies. Although experienced a tough time, I am grateful that you all still standing by my side and supporting me all the time, and I wouldn't be where I am today without you.

Definitions and Abbreviations

2D..... two-dimensional

2OG ..... 2 oxoglutarate

3D..... Three-dimensional

ABTS ..... 2,2'-Azino-bis-(3-ethylbenzothiazoline-6-sulfonic acid) diammonium salt

AI..... aromatase inhibitor

AlogP ..... octanol/water partition coefficient

ALT ..... alanine transaminase

ARNT ..... aryl hydrocarbon receptor nuclear translocator

ASR..... Age-Standardized Rate

AST ..... aspartate aminotransferase

BA..... benzoic acid

BC..... Betweenness Centrality

Bcl-2 ..... B-cell lymphoma 2

BH ..... Benjamini-Hochberg

bHLH..... basic helix-loop-helix

BL ..... basal-like

BSA..... Bovine serum albumin

C1..... 2-[[4-Hydroxy-2-oxo-1-[(2,4,5-trifluorophenyl)methyl]-5,7-dihydrofuro[3,4-b]pyridine-3-carbonyl]amino]acetic acid

C2..... 2-(4-hydroxy-2-oxo-1-(m-tolyl)-1,2,5,7-tetrahydrofuro[3,4-b]pyridine-3-carboxamido)acetic acid

C3..... 2-(2-(2-Fluorophenyl)-5-hydroxy-8-methyl-7-oxo-7,8-dihydropyrido[2,3-d]pyrimidine-6-carboxamido)acetic acid

CAF..... Cancer-associated fibroblast

CBP..... CREP-binding protein

CC..... Closeness Centrality

CCl4	carbon tetrachloride
CDK	cyclin-dependent kinase
CKD	chronic kidney disease
CMap	Connectivity map
CT network	Compound-Target network
C-TAD	COOH-terminal transactivation domain
CTP network	Compound-Target-Pathway network
DAG	directed acyclic graph
DAPI	4'6-Diamidino-2-Pheylindole
DCIS	ductal carcinoma in situ
DEG	differentially expressed gene
D-GalN	D-galactosamine
DHBA	dihydroxybenzoic acid
DMOG	Dimethyloxalylglycine
DMSO	dimethylsulfoxide
DPPH	1,1-diphenyl-2-picryl-hydrazyl
DSB	double-strand break
EDHB	Ethyl-3,4-dihydroxybenzoate
EGFR	epidermal growth factor receptor
EMT	epithelial-mesenchymal transition
EPO	erythropoietin
ER	estrogen receptor
ES	Enrichment Score
ESA	erythropoiesis-stimulating agent
ETCM	Encyclopedia of Traditional Chinese Medicine
FDA	Food and Drug Administration
FDR	false discovery rate
FGF-2	fibroblast growth factor 2

FIH	Factor inhibiting HIF
FUH	Functional Uterine Haemorrhage
GA	Genetic algorithm
GCO	Global Cancer Observation
GEO	Gene Expression Omnibus
GFR	glomerular filtration rate
GI	gastrointestinal
GLUT-1	glucose transporter 1
GM-CSF	Granulocyte-macrophage colony-stimulating factor
GO	Gene Ontology
GOLD	Genetic Optimisation for Ligand Docking
GSEA	Gene Set Enrichment Analysis
GSVA	Gene Set Variation Analysis
HBA	Hydroxybenzoic acid
HCV	hepatitis C virus
HER2	human epithelial growth factor receptor 2
HGF	Hepatocyte growth factor
HIF	hypoxia-inducible factor
HPLC	High-performance liquid chromatography
HRE	hypoxia response element
HSP	Heat Shock Protein
HSV-1	herpes simplex virus type 1
IARC	International Agency for Research on cancer
ICH	intracerebral haemorrhage
IDC	invasive ductal carcinoma
IFN-γ	interferon-γ
IL-2	interleukin 2
IL-4	interleukin-4

ILC	invasive lobular carcinoma
IM	immunomodulatory
IPF	idiopathic pulmonary fibrosis
KEGG	Kyoto Encyclopedia of Genes and Genomes
LAR	luminal androgen receptor
LDHA	lactate dehydrogenase A
LSE	*Limonium Sinense* ethanol extracts
LSP	*Limonium Sinense* polysaccharides
LSW	*Limonium Sinense* water extracts
LZIP	leucine zipper
MA	matching algorithm
MAPK	Mitogen activated protein kinase
MDSC	Myeloid-derived suppressor cell
MEK	Mitogen-Activated Protein kinase
MES	mesenchymal
MFE	mammosphere forming efficiency
MHBA	monohydroxybenzoic acid
MOA	mechanism of action
MR	molar refractivity
MSigDB	molecular signature database
mTOR	mammalian target of rapamycin
MW	molecular weight
NCBI	National Centre for Biotechnology Information
NES	normalized enrichment score
NF-κB	nuclear factor kappa-light-chain-enhancer of activated B cells
nHA	number of hydrogen bond acceptor
nHD	number of hydrogen bond donor
NSC	neural stem cell

N-TAD	NH2-terminal transactivation domain
ODD	oxygen-dependent degradation domain
OGD/R	oxygen-glucose deprivation/reperfusion
OH·	hydroxyl radicals
PAI-1	plasminogen activator-1
PAS	Per-ARNT-Sim
PBS	phosphate buffered saline
PCNA	proliferating cell nuclear antigen
PDGF	platelet-derived growth factor
PentaHBA	pentahydroxybenzoic acid
PFA	paraformaldehyde
PGK	phosphoglycerate kinase
PHD	prolyl hydroxylase domain
PI	Propidium Iodide
PI3K	phosphatidyl inositol-4,5-bisphosphate-3-kinase
PKC	Protein kinase C
PPI	Protein-protein interaction
PT network	Pathway-Target network
PTEN	Phosphatase and tensin homolog gene
RB	rotatable bond
RNA-seq	RNA sequencing
RNS	reactive nitrogen species
ROS	reactive oxygen species
Rox	Roxadustat
RTK	receptor tyrosine kinase
Sam B	Samarangenin B
SD	standard deviation
SDS-Page	sodium dodecyl sulphate polyacrylamide gel electrophoresis

SERD..... selective estrogen receptor down-regulator

SERM..... selective estrogen receptor modulator

TCMSP..... Traditional Chinese Medicine Systems Pharmacology Database and Analysis Platform

tetraHBA..... tetrahydroxybenzoic acid

TGF-β1..... transforming growth factor-beta 1

Th1..... Type 1 T helper

THBA ..... trihydroxybenzoic acid

TKI..... tyrosine kinase inhibitor

tMCAO..... transient middle cerebral artery occlusion

TNBC ..... triple-negative breast cancer

TPSA..... topological polar surface area

t-SNE ..... t-distributed stochastic neighbour embedding

Ub ..... Ubiquitin

VDAC..... voltage-dependent silver ion channel

vdW..... van der Waal

VEGF..... Vascular endothelial growth factor

VHL..... von Hippel-Lindau protein

Chapter 1 Introduction

1.1 Overview of *Limonium Sinense* (Girard) Kuntze

Limonium species belonging to the *Plumbaginaceae* family, are primarily perennial herbs specially adapted to extreme coastal environments. There are more than 600 *Limonium* species widely distributed throughout the world, with most of them thriving in coastal and saline grasslands, maritime cliffs, salt marshes, lagoons, meadows, steppes, and deserts of the continental interior [1].

Many *Limonium* species exhibit various pharmacological activities and have been extensively used in both traditional and modern medicine. As examples, *Limonium tetragonum* water extract has shown the ability to enhance endurance exercise capacity by improving mitochondrial biosynthesis and increasing the formation of oxidative slow-twitch fibres in C57BL/6 male mice [2]. The n-BuOH extract of *Limonium Duriusculum* has exhibited a growth inhibitory effect on HCT116 cancer cells, with its molecular mechanism associated with the suppression of Mitogen-activated protein kinase (MAPK) pathway and activation of the p53 pathway [3]. Moreover, the polar extract from the roots of *Limonium Brasiliense* (Boiss.) Kuntze has displayed excellent bacteriostatic and antioxidant effects [4, 5]. *Limonium Wrightii* (Hance) Kuntze has also been reported to possess cardioprotective effects against myocardial ischemia-reperfusion injury in isolated rat hearts [6].

There are approximately 17 known species of the *Limonium* genus found in China, mainly distributed in the Northeast, North, Northwest of China and coastal mud flats (Table 1.1, Flora of China, www.eFloras.org). Most of these *Limonium* species have medicinal properties and can be used for the treatment of various human disorders, including bleeding, blood circulation issues, cancer, and other diseases.

Table 1.1. Types and distribution of *Limonium* species in China

Species	Other Names	Distribution	Uses
Limonium aureum (Linn.) Hill	*Limonium erythrorrhizum* Ikonn.-Gal. ex Lincz., *Statice aurea* L.	Northeast, North, Northwest provinces of China	Relief pain, reduce inflammation, blood tonic, used for women with irregular menstruation, nosebleeds and abnormal leucorrhea.
Limonium bicolor (Bag.) Kuntze	*Statice bicolor* Bunge, *Statice bungeana* Boiss., *Statice florida* Kitag., *Statice sinensium* Gand., *Statice varia* Hance	Northeast China, Northern Jiangsu Province, Yellow River Basin Provinces, seashore	Scattered stasis hemostasis, epigastric pain, dyspepsia, Women with irregular menstruation and leakage, nosebleeds, haemorrhoids.
Limonium chrysocomum (Kar.et Kir.) Kuntze	*Statice sedoides Regel*, *Statice chrysocephala Regel*	Northern and Western Xinjiang Province	Unknown
Limonium coralloides (Tausch) Lincz	*Statice aphylla* Poir., *Statice coralloides* Tausch, *Statice decipiens* Ledeb.	Northern Xinjiang Province	Unknown
Limonium callianthum (Peng) Kamelin	*Statice latissima*, *Statice myriantha*	Southern Xinjiang Province	Unknown
Limonium flexuosum (Linn.) Kuntze	*Statice flexuosa* L.	Inner Mongolia Hulunbeier Plateau	Unknown
Limonium franchetii (Debx.) Kuntze	*Statice franchetii* Debeaux, *Statice tchefouensis* Gand.	Liaoning Province, Shandong Peninsula, Northeastern Jiangsu Province	Unknown
Limonium gmelinii (Willd.) Kuntze	*Statice gmelinii*	Inner Mongolia, Northern Xinjiang Province, Nrotheast China,	Disperses blood stasis and reduces swelling, Dysfunctional uterine bleeding, Cervical cancer, and other bleeding.
Limonium kaschgaricum (Rupr.) lk.-Gal	*Statice holtzeri* Regel, *Statice kaschgarica* Rupr.	Xinjiang Province	Unknown

Species	Other Names	Distribution	Uses
Limonium leptolobum (Regel) Kuntze	*Statice leptoloba* Regel, *Statice leptoloba* Regel var. subaphylla Regel	Xinjiang Province	Unknown
Limonium myrianthum (Schrenk) Kuntze	*Statice latissima* Kar. et Kir., *Statice myriantha* Schrenk	North Xinjiang Province	Unknown
Limonium otolepis (Schrenk) Kuntze	*Statice otolepis* Schrenk	North Xinjiang Province, West Hexi of Gansu Province	Unknown
Limonium lacostei (Danguy) Kamelin	*Statice otolepis* Schrenk	Southwestern Xinjiang Province	Unknown
Limonium sinense (Girard) Kuntze	*Statice fortunei* Lindl., *Statice sinensis* Girard, *Limoniumspp*	Jiangsu Province, Fujian Province, Taiwan and other regions of China coastal cities	Hemostasis, used for the treatment of clearing damp-heat, relieving stool bleeding, preventing rectal prolapse, reducing the symptoms of menorrhagia and leucorrhea, curing carbuncle, Functional Uterine Hemorrhage (FUH) and cervical cancer.
Limonium suffruticosum (L.) Kuntze	*Statice suffruticosa* L.	Northern Xinjiang Province	Unknown
Limonium tenellum (Turcz.) Kuntze	*Statice tenella* Turcz.	Gansu Province, Ningxia Province, Inner Mongolia	Unknown
Limonium wrightii (Hance) Kuntze	*Statice arbuscula* Maxim., *Statice wrightii* Hance	Fujian Province, Taiwan Province, Diaoyu islands	Relief the pain of rheumatic, control high blood pressure.

Among these species, *Limonium Sinense* (Girard) Kuntze (*Limonium Sinense*), also known *as Latouchea Fokiensis* or *Limoniumspp*, is an endemic and perennial herb in Chinese medicine [7]. *Limonium Sinense* predominantly grows in the seashores and salt marshes regions of mainland China, western Taiwan, and Ryukyus islands in Japan [8, 9] (Figure 1.1). Historically, *Limonium Sinense* has been widely used as a remedy for bleeding, piles, fever, hepatitis, diarrhoea, bronchitis and cancers in Chinese folk medicine [8, 10]. The "Chinese Materia Medica Dictionary" records that *Limonium Sinense* can clear heat and detoxify the body, enhance the blood circulation and remove dampness phlegm [11]. According to the “Chinese Herbal Medicine Handbook in Xinjiang”,

Limonium Sinense is known for its properties to stop bleeding, promote blood circulation, resolve blood stasis, and it can also be used for the treatment of Functional Uterine Haemorrhage (FUH) and cervical cancer diseases [12].

In recent decades, due to the rapid growth of the population in China's coastal areas and the reclamation and utilization of large areas of coastal wetlands, the number of coastal saltmarsh plants has sharply decreased, and many plants are on the verge of extinction, including *Limonium Sinense*, which was once distributed widely in China. According to the order of the Ministry of Agriculture of the People's Republic of China, *Limonium Sinense* was listed as the New Agricultural Plant Variety Protection of PRC in 2001 [7]. It's time to take measures to protect this rare plant, and there is also vast potential for the extensive medicinal value of *Limonium Sinense* in further developing its clinical applications.

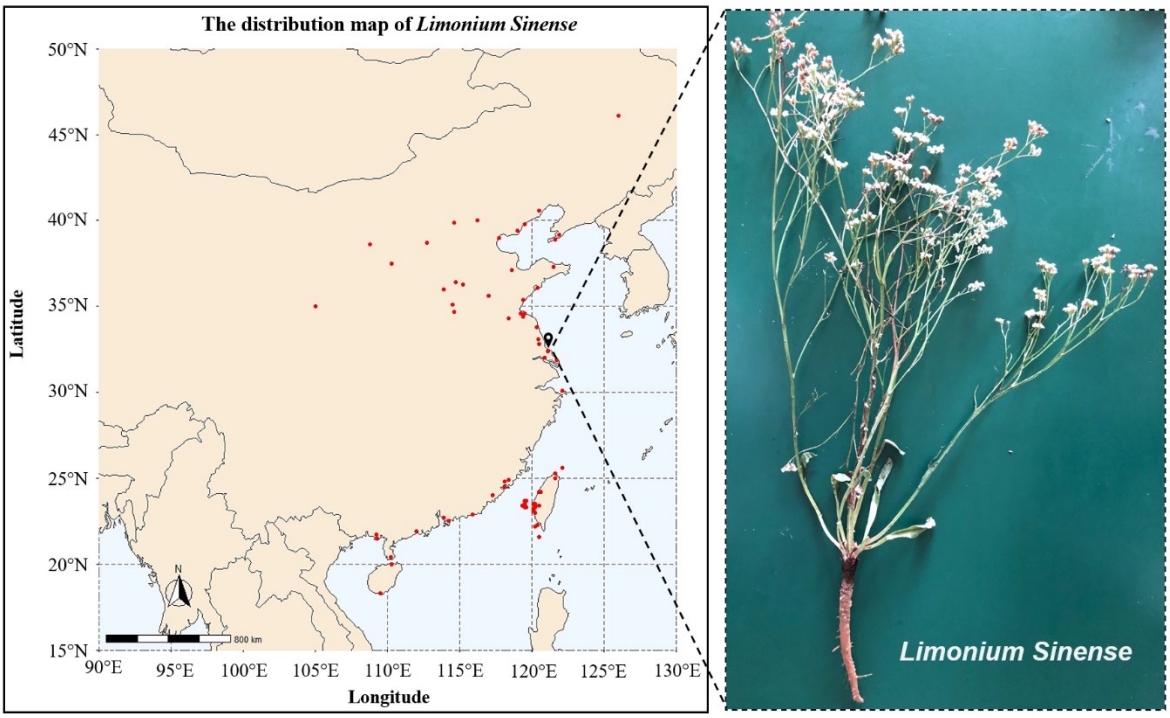


Figure 1.1. Distribution of *Limonium Sinense*. Map showing that the *Limonium Sinense* is mainly distributed along seashores and salts marshes in southern China, and western Taiwan. Distribution data from GBIF: https://www.gbif.org. The picture in the frame showed the shape of *Limonium Sinense*, which we collected from Yancheng City, Jiangsu Province, China (33°09'33.0" N, 120°46'40.4" E).

1.2 Phylogenetic analysis of *Limonium Sinense*

Limonium Sinense is an insect-pollinated and outcrossing species, its population and genetic variability are highly susceptible to change in the surrounding environment. Most species of *Limonium* are self-incompatible which is an important factor for maintaining the high genetic variability. Compared with other *Limonium* species, *Limonium Sinense* showed a relatively high genetic diversity at species level but lower genetic diversity at the population level [13]. This is probably because *Limonium Sinense* is a perennial herb and this long-lived habit may contribute to the high level of genetic diversity of this species [14].

The geographical distribution is also a very important factor that may affect the genetic diversity level of a species. In general, widespread species often considered to have higher level of genetic variability than narrowly distributed ones. From the Phylogenetic tree (Figure 1.2), *Limonium Sinense* may share an immediate common ancestor with *Limonium aureum*, which is another *Limonium* species mainly distributed in Northwest areas of China with the rare rainfall environment [8]. The outgroup species *Reaumuria trigyna* and *Trigastrotheca stricta* also show a closer relationship with *Limonium Sinense*. The inner joint neighbours and the outgroup species of *Limonium Sinense* were found to be annual or perennial plants mostly distributed in coastal saline-alkali land and relatively arid inland areas.

Limonium Sinense is a perennial halobiotic plant, and this physiological characteristic also restrict it to growing in salty, arid areas such as coastal areas, tidal flats, wetlands, and deserts. Hence, the seed germination rate of *Limonium* is very low due to lack of contact with soil because of the seeds being packed by the calyx, contributing to the continued decline of this species [13]. Although *Limonium Sinense* has extensive ornamental and medicinal value, its research development and clinical utilization are still relatively weak. For the past few years, most studies on *Limonium Sinense* are concentrated on the isolation and identification of its active components and its biological activities, however, the specific underlying molecular mechanism is still unknown and urgently need to be explored.

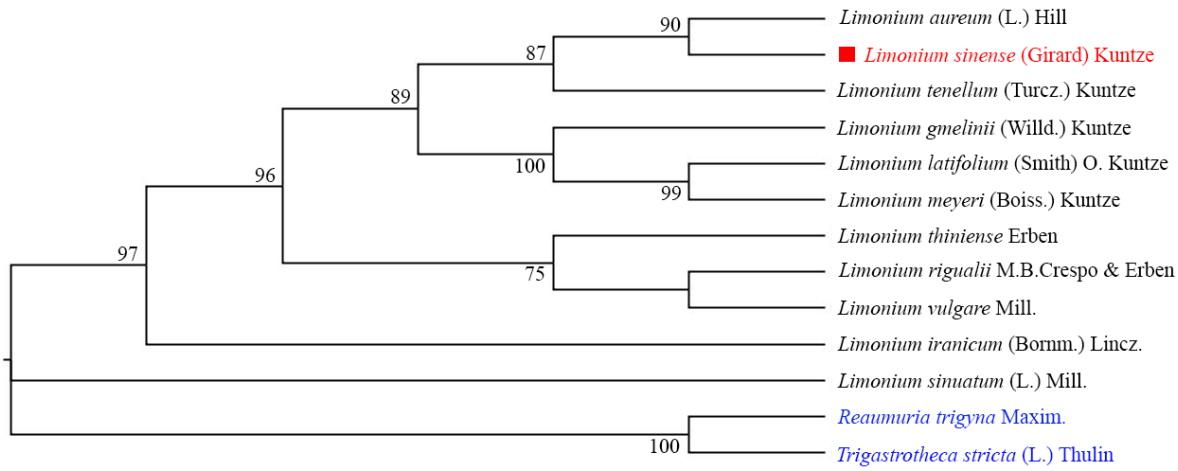


Figure 1.2. Phylogenetic tree of *Limonium Sinense*. The evolutionary history was inferred using the Neighbour-Joining method based on protein sequences and 1,000 bootstrap replicates. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test are shown next to the branches. No values are given for groups with bootstrap values less than 70%. The evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site. This analysis involved 13 amino acids sequences with a total of 3,866 positions in the final dataset. The *Reaumuria trigyna* and *Trigastrotheca stricta* were used as outgroup and coloured blue.

1.3 Pharmacological functions of *Limonium Sinense*

Medicinal plants have been used to treat human diseases for thousands of years. Natural plant compounds isolated from natural sources (plants) indicated are a source of therapeutic substances and a structural basis for drug elaboration [15]. *Limonium Sinense* has a long history of therapeutic use in China due to its broad biological and pharmacological properties with its richness compounds. Modern pharmacological studies have revealed that *Limonium Sinense* can affect the body through many aspects.

1.3.1 Blood enrichment

Limonium Sinense is known for its remarkable blood-enriching properties and its traditional use in treating anaemia [7]. It has been reported that *Limonium Sinense* is rich in various inorganic elements such as potassium (K), sodium (Na), calcium (Ca), and magnesium (Mg), as well as several trace elements including iron (Fe), zinc (Zn), cobalt (Co), nickel (Ni), and chromium (Cr). The plant is also abundant in vitamins such as vitamin B2, vitamin B12, vitamin C, vitamin D and carotenoids. These elements collectively contribute to creating a favourable haematopoietic environment within

the body. Of particular importance is the element of cobalt (Co) in *Limonium Sinense*, which acts as a major active ingredient of vitamin B12, has important physiological effects on iron metabolism, haemoglobin synthesis, red blood cell development, and the maturation of red blood cells. Moreover, it exhibits a special therapeutic effect on haemorrhagic anaemia, further highlighting its beneficial properties in treating anaemia [16].

1.3.2 Antioxidant

Oxidative stress is caused by an imbalance between reactive oxygen species (ROS) and is believed to play a significant role in pathogenesis of various diseases, including cardiovascular diseases, neurodegeneration, cancers, immune disorders, diabetes, aging, and other related conditions [17, 18]. In the body, the antioxidant network works harmoniously through multiple mechanisms, such as ROS scavenging, termination of lipid peroxidation, and metal chelation. Numerous active ingredients extracted from plants, such as polysaccharides, flavonoids, polyphenols, and vitamins, have demonstrated robust antioxidant capacities [19, 20].

It was reported that the polysaccharides extracted from the root of *Limonium Sinense* exhibit obvious scavenging abilities against 1,1-diphenyl-2-picryl-hydrazyl (DPPH), 2,2'-Azino-bis-(3-ethylbenzothiazoline-6-sulfonic acid) diammonium salt (ABTS) radicals, and hydroxyl radicals (OH·), suggesting the potential role of *Limonium Sinense* in combating ROS [21]. The flavonoids present in *Limonium Sinense*, such as isorhamnetin, quercetin, isoquercetin, luteolin and apigenin, also demonstrated strong scavenging activities against DPPH free radicals [22]. In addition, the polyphenols obtained from the root, rhizome, leaf and flower of *Limonium Sinense* showed potent scavenging abilities against DPPH, OH· and Superoxide Anion Radical (O^{2-}) free radicals, with the clearance rate showing a certain dose-dependent trend [23].

1.3.3 Anti-virus

Many herbs have been widely utilized to prevent and treat viral infectious diseases. Samarangenin B (Sam B), isolated from *Limonium Sinense*, has been observed to exhibit inhibitory effects on herpes simplex virus type 1 (HSV-1) multiplication in Vero cells. Its mode of action potentially involves interference with structural proteins and DNA synthesis, DNA polymerase mRNA transcription, and gene expression of HSV-1, and all without reducing cell viability and growth [24]. This antiviral activity of Sam B may also be helpful in the treatment of microorganism infections.

Moreover, the water extract of *Limonium Sinense* root and its purified phytochemical constituent gallic acid have demonstrated antiviral potential by targeting the early viral entry of the hepatitis C virus (HCV). Reports indicate that the water extract of *Limonium Sinense* root and gallic acid can

specifically target the early stages of HCV infection without affecting viral replication, translation, and cell-to-cell transmission. They effectively block viral attachment and post-attachment entry/fusion steps. Additionally, both the water extract of *Limonium Sinense* root and gallic acid exhibit the ability to suppress HCV infection in primary human hepatocytes [25].

1.3.4 Anti-hepatitis

Reports have highlighted the significant role of *Limonium Sinense* in the prevention of hepatitis and liver cancer. Modern pharmacological studies have revealed that both the aqueous extract and the chloroform fraction of the ethanol extract derived from the leaves of *Limonium Sinense* possess remarkable effectiveness in inhibiting the elevation of serum alanine transaminase (ALT) and aspartate aminotransferase (AST) activity in rat livers damaged by carbon tetrachloride (CCl_4) or D-galactosamine (D-GalN). Furthermore, the *Limonium Sinense* extract also demonstrates significant inhibition of the decrease in serum triglyceride levels [9]. These findings suggest that the anti-liver injury effect of *Limonium Sinense* extract exhibits a protective effect on mitochondria, thereby contributing to its anti-liver injury effect. In studies conducted on mice with liver damage caused by CCl_4, it was observed that the extract of *Limonium Sinense* improved mitochondrial swelling, restored membrane structure and tendon deformation, and reduced hepatocyte rupture and disorder. Simultaneously, the *Limonium Sinense* extract enhanced the voltage-dependent silver ion channel (VDAC) in the damaged liver, indicating its potent hepatoprotective activity [26].

1.3.5 Anti-tumour

Various natural active ingredients such as polysaccharides, flavonoids and alkaloids have been confirmed to possess certain anti-tumour activity. These components derived from higher plants are generally non-toxic or low-toxic, exhibiting minimal side effects, making them promising candidates for therapeutics with anti-tumour application.

Extracts of *Limonium Sinense*, especially the water root extract, have demonstrated significant inhibitory effect on the proliferation of human hepatoma cell HepG2 and have shown remarkably prevention of the tumour growth, along with increased spleen and thymus weights in B16 melanoma-bearing mice [27]. LSP21, a polysaccharide isolated from the root of *Limonium Sinense*, showed a strong anti-tumour activity. LSP21 prominently inhibits the growth of transplanted mouse tumours and the proliferation of HepG2 cells in a dose-dependent manner. Further research has revealed that LSP21 induces cell body shrinkage, chromatin condensation and tumour cells reduction, suggesting that its cytotoxicity on tumour cells may be related to both the inhibition on cell proliferation and induction of cell death [28].

Flavonoids isolated from *Limonium Sinense* such as isorhamnetin, quercetin, isoquercetin, luteolin and apigenin have also displayed significant inhibition of HepG2 and Hela tumour cell proliferation in *vitro* [22]. Other studies have reported that polyphenols from *Limonium Sinense* inhibit the proliferation of HL-60 human leukaemia cells and increase the apoptosis rate in a dose-dependent manner. The molecular mechanism underlying this effect may associate with the regulation of B-cell lymphoma 2 (Bcl-2) and Bax expression [29].

1.3.6 Immunomodulatory

Reports have demonstrated that *Limonium Sinense* shows the ability to stimulate the proliferation of spleen and thymus cells and improve the function of immune organs, indicating the immunomodulatory activity of *Limonium Sinense*. Moreover, *Limonium Sinense* polysaccharides (LSP), derived from the water root extract of *Limonium Sinense*, has been found to significantly increase the production of interferon-γ (IFN-γ) and interleukin 2 (IL-2), while decrease the production of interleukin-4 (IL-4), suggesting that LSP can directly activate T lymphocytes through the Type 1 T helper (Th1) response, thereby enhancing the phagocytic activity of macrophages [30]. The immunomodulatory activity of *Limonium Sinense* also makes it a potential candidate for development of adjuvants for chemotherapy drugs.

1.4 Natural compounds in *Limonium Sinense*

In recent years, there has been increasing attention on natural plant active ingredients due to their low toxicity and excellent antioxidant, antibacterial, anti-tumour and anti-virus properties, and these active ingredients have a broad application prospect in the fields of drug development, food nutrition and health care products. *Limonium Sinense* contains a diverse array of compounds, including flavonoids, polysaccharides, alkaloids, phenolic acids, tannins, sterols and pentacyclic triterpenoids. It is speculated that the physiological functions of *Limonium Sinense* are attributed to the presence of these chemical compounds.

1.4.1 Flavonoids

Flavonoids are the predominant and main active ingredients of *Limonium Sinense*. A variety kind of flavonoids have been isolated from various parts of the plant, especially the roots and leaves. The flavonoids found in *Limonium Sinense* were identified as myricetin, myricetin 3-O-α-rhamnopyranoside, myricetin 3-O-β-galactopyranoside, myricetin 3-O-β-arabinopyranoside, myricetin 3-O-(2''-O-p-hydroxybenzoyl)-α-rhamnopyranoside, myricetin 3-O-(2''-O-galloyl)-α-rhamnopyranoside, myricetin 3-O-(3''-Ogalloyl)-α-hamnopyranoside, myricetin 3-O-(4''-O-galloyl)-

α-rhamnopyranoside, myricetin 3-O-(6''-O-galloyl)-β-galactopyranoside, myricetin 3-O-β-D-glucoside, Myricetin-3-O-α-L-rhamnoside, quercetin, quercetin-3-O-β-D-glucoside (isoquercetin), quercetin 3-O-α-L-rhamnoside, quercetin3-O-α-rhamnopyranoside, quercetin 3-O-(2''-O-galloyl)-α-rhamnopyranoside, isorhamnetin, isorhamnetin-3-rutinoside, eriodictyol, homoeriodictyol, morin, apigenin, luteolin, naringenin, epigallocatechin 3-O-gallate, epigallocatechin 3-O-(3'-O-methyl)-gallate, epigallocatechin 3-O-(3', 5'-di-O-methyl)-gallate, isodihydrosyringetin, kaempferol, kaempferol-3-O-α-L-rhamnopyranoside and catechin [31-35].

Flavonoids and their subclasses, such as anthocyanins, flavonols, flavan-3-ols, flavanones and isoflavones, have been extensively studied and have demonstrated a wide range of pharmacological activities and beneficial effects on health [36].

1.4.2 Saccharides

Saccharides play a key role in the vital activities of organisms and serve as the main source of energy required for sustaining life. Generally, saccharides are classified into three main types: monosaccharide, disaccharide, and polysaccharide. Polysaccharides are one of the most abundant compounds in the roots of *Limonium Sinense*. Numerous polysaccharides have been reported to possess various biological activities, including anti-inflammatory, anti-tumour, and anti-fibrotic properties [37, 38]. Polysaccharides extracted from *Limonium Sinense* have demonstrated significant antioxidant, anti-tumour, and immunomodulatory effects [28, 30].

Monosaccharides, on the other hand, are the simplest form of saccharide and serve as the basic units of carbohydrates. Monosaccharides can convert into disaccharides and polysaccharides by condensation in the cell for quickly using and play a pivotal role in metabolism. Mannitol, a type of monosaccharide, has been isolated from *Limonium Sinense* [32]. Mannitol is widely used in the food, pharmaceutical, medical, and chemical industries [39], and showed obvious anti-inflammatory effect due to its ability to inhibit the lipid peroxidation and nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) complex formation [40]. Moreover, it has shown high effectiveness in suppressing the inflammatory response in perihematomal and contralateral tissues in rat model of intracerebral haemorrhage (ICH) [41].

1.4.3 Alkaloids

Alkaloids have a long history of medicinal use and show an enormous potential for the development of cancer therapies and management. Alkaloids are important secondary metabolites generally extracted from plants, bacteria, and a few animal sources. Many kinds of alkaloids have been

traditionally used in cancer therapy due to their wide range of pharmacological activities, including anticancer, antibacterial, anti-inflammatory, and antiangiogenic properties [42].

Currently, there are two alkaloid compounds isolated from *Limonium Sinense*, namely N-trans-caffeoyltyramine and N-trans-feruloyltyramine [31]. These two alkaloids are both exhibit significant antioxidant, antimicrobial, and anticancer activities [43, 44]. However, further research is needed to fully understand their specific functions within *Limonium Sinense*.

1.4.4 Phenolic acids

Phenolic acids are a major class of phenolics and many of them exert significant antioxidant activity by scavenging ROS and acting as chain-breaking antioxidants and reducing agents [45]. Numerous phenolic acids shown strong inhibitory effects on cancer invasion and metastasis, making them potential candidates for therapeutic drugs against cancer.

Two kinds of phenolic acids, ethyl gallate and gallic acid, were identified from *Limonium Sinense* [31, 32]. It was reported that gallic acid isolated from *Limonium Sinense* exhibits noteworthy antiviral and hepatoprotective effects [25]. It is a major phenolic acid found in various plants and fruits, and evidences show that gallic acid possesses multiple biological activities such as antioxidant, anti-microbial, anti-inflammatory and anti-cancer properties [46]. Ethyl gallate is another phenolic compound found in *Limonium Sinense*, also exhibits antioxidant and anti-cancer activities. Studies have shown that ethyl gallate can suppress the growth of esophageal cancer by inhibiting ERK signalling pathway both in *vitro* and in *vivo* [47]. Through blocking the expression of MMP-2/9 and modulating the PI3K/AKT/NF-κB signalling pathway, ethyl gallate inhibits proliferation and invasion in human breast cancer cells [48].

1.4.5 Tannins

Tannins are heterogeneous group of water-soluble polyphenols and exert a broad range of pharmacological activities including antioxidant, antimicrobial and anti-cancer [49, 50]. A tannin compound known as Sam B has been identified from *Limonium Sinense*, which exerts anti-viral properties in Vero cells without reducing cell viability and growth activity [24].

1.4.6 Sterols

Sterols are a type of lipid found in the membranes of plants, fungi and animal organisms, and can be used for the treatment of common conditions such as hypercholesterolemia, metabolic syndrome and diabetes [51]. Sterols in plants are often consisted with 3 major phytosterols, namely

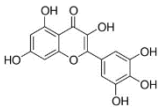
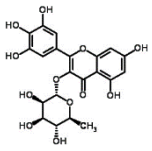
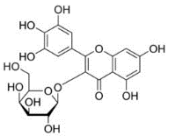
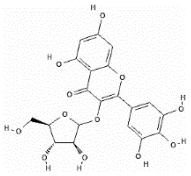
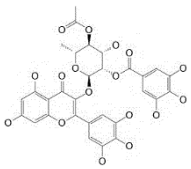
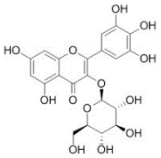
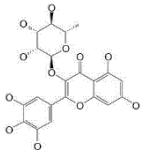
β-sitosterol, campesterol and stigmasterol. The sterol isolated from *Limonium Sinense* was identified as β-sitosterol, which is the most abundant type of phytosterol present in the plant [32]. Studies have demonstrated the ability of β-sitosterol to inhibit cell proliferation and induce apoptosis in a variety type of cancer cell lines including gastric, colon, prostate, lung, and breast cancer cell [52], though the specific biological functions of β-sitosterol in *Limonium Sinense* remain to be explored.

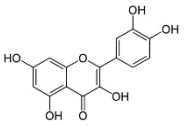
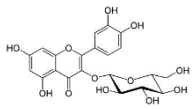
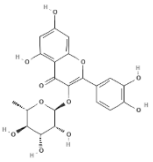
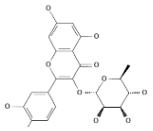
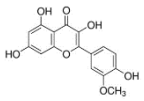
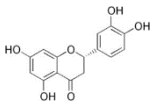
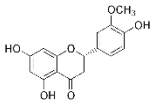
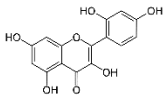
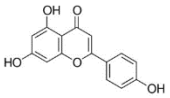
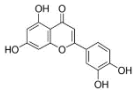
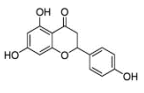
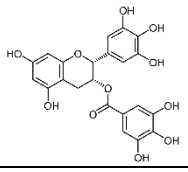
1.4.7 Pentacyclic triterpenoids

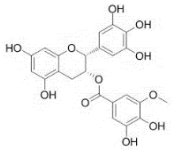
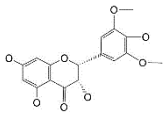
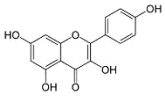
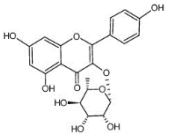
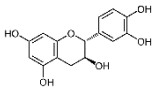
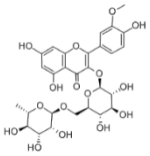
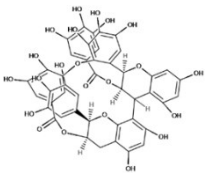
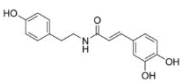
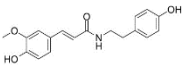
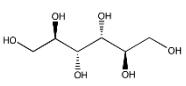
Pentacyclic triterpenoid is a major class of chemical compounds found in large amounts of plants and resinous natural materials [53]. Increasing numbers of reports have demonstrated the diverse biological activities of pentacyclic triterpenoids, including anti-inflammatory, anti-cancer and hepatoprotective effects [54, 55].

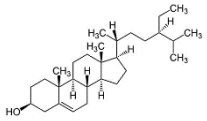
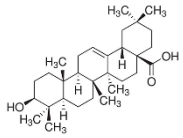
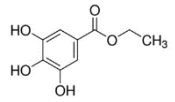
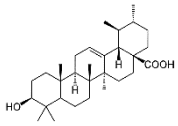
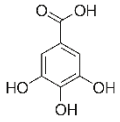
There are two types of pentacyclic triterpenoids in *Limonium Sinense* were identified, namely oleanolic acid and ursolic acid [32]. Oleanolic acid is an important member of the pentacyclic triterpenoid compounds and has demonstrated several promising pharmacological activities, such as hepatoprotective effects, anti-inflammatory, antioxidant and anti-tumour [56]. Similarly, ursolic acid, another pentacyclic triterpenoid found in *Limonium Sinense*, exhibits similar biological activities to oleanolic acid, including anti-inflammatory, anti-tumour, antidiabetic, antioxidant, and antibacterial effects [57]. However, their specific functions within *Limonium Sinense* are still needed to be investigated. In summary, the natural components isolated and identified from *Limonium Sinense* are displayed in Table 1. 2.

Table 1.2. Natural compounds identified from *Limonium Sinense*

Compound Name	Chemical Structure	Compound Type	Extract Part	Reference
myricetin		Flavone	Leaves, Roots	[31] [33]
myricetin-3-O-α-rhamnopyranoside		Flavonol glycoside	Leaves, Roots	[31] [33]
Myricetin 3-O-β-galactopyranoside		Flavonol glycoside	Leaves	[31]
Myricetin 3-O-β-arabinopyranoside		Flavonol glycoside	Leaves	[31]
Myricetin 3-O-(2″-O-p-hydroxybenzoyl)-α-rhamnopyranoside	—	Flavonol glycoside	Leaves	[31]
Myricetin 3-O-(2″-O-galloyl)-α-rhamnopyranoside		Flavonol glycoside gallate	Leaves	[31]
Myricetin 3-O-(3″-O-galloyl)-α-rhamnopyranoside	—	Flavonol glycoside gallate	Leaves	[31]
Myricetin 3-O-(4″-O-galloyl)-α-rhamnopyranoside	—	Flavonol glycoside gallate	Leaves	[31]
Myricetin 3-O-(6″-O-galloyl)-β-galactopyranoside	—	Flavonol glycoside gallate	Leaves	[31]
Myricetin 3-O-β-D-glucoside		Flavone	Whole plant	[32]
Myricetin-3-O-α-L-rhamnoside		Flavone	Leaves	[35]

Compound Name	Chemical Structure	Compound Type	Extract Part	Reference
Quercetin		Flavonol	Whole plant	[31] [32] [35]
Quercetin-3-O-β-D-glucoside (Isoquercetin)		Flavonoid	Whole plant	[32] [34]
Quercetin 3-O-α-L-rhamnoside		Flavonoid	Leaves	[35]
Quercetin 3-O-α-rhamnopyranoside		Flavonol glycoside	Leaves, Roots	[31] [33]
Quercetin 3-O-(2″-O-galloyl)-α-rhamnopyranoside	—	Flavonol glycoside gallate	Leaves	[31]
Isorhamnetin		Flavonol	Whole plant	[32]
Eriodictyol		Flavanone	Leaves	[31]
Homoeriodictyol		Flavanone	Leaves	[31]
Morin		Flavonol	Whole plant	[34]
Apigenin		Flavone	Leaves	[31]
Luteolin		Flavone	Leaves	[31]
Naringenin		Flavanone	Leaves	[31]
(-)-Epigallocatechin 3-gallate		Flavan-3-ol	Leaves, Roots	[31] [33]

Compound Name	Chemical Structure	Compound Type	Extract Part	Reference
Epigallocatechin 3-O-(3'-O-methyl)-gallate		Flavan-3-ol	Leaves	[31]
Epigallocatechin 3-O-(3',5'-di-O-methyl)-gallate	—	Flavan-3-ol	Leaves	[31]
Isodihydrosyringetin		Flavanone	Roots	[33]
Kaempferol		Flavonol	Whole plant	[32]
Kaempferol-3-O-α-L-rhamnopyranoside		Flavonoid	Whole plant	[32]
(+)-Catechin		Flavan-3-ol	Whole plant	[32]
Isorhamnetin-3-rutinoside		Flavonoid-3-o-glycoside	Whole plant	[32]
Samarangenin B		Tannin	Roots	[33]
N-trans-caffeoyltyramine		Alkaloid	Roots	[33]
N-trans-feruloyltyramine		Alkaloid	Roots	[33]
Mannitol		Monosaccharide	Whole plant	[32]
LSP21	—	Polysaccharide	Roots	[28]

Compound Name	Chemical Structure	Compound Type	Extract Part	Reference
β-sitosterol		Sterol	Whole plant	[32]
Oleanolic acid		Pentacyclic triterpenoid	Whole plant	[32]
Ethyl gallate		Phenolic acid	Whole plant	[32]
Ursolic acid		Pentacyclic triterpenoid	Whole plant	[31]
Gallic acid		Phenolic acid	Leaves, Roots	[31] [33]

1.5 Potential therapeutic applications of *Limonium Sinense*

As previously mentioned, *Limonium Sinense* has a long history of traditional use in treating conditions such as anaemia, endometritis, cervical erosion, gastric cancer, and cervical cancer [25], modern pharmacological studies have demonstrated its diverse range of activities, further supporting it potential as a therapeutic agent.

The anti-inflammatory effect of *Limonium Sinense* plays a vital role in the treatment of multiple diseases. Inflammation is closely associated with the occurrence of cancer and organ fibrosis. Environmental stimuli, infectious agents and endogenous irritants can trigger an inflammatory response, which is a crucial mechanism for repairing damaged tissues and restoring the body's microenvironment balance. However, failure to resolve acute inflammation can lead to chronic inflammation. Immune cells including macrophages and neutrophils are one of the first lines of defence against invading pathogens in the inflammatory phase. Macrophages release cytokines that stimulate the inflammatory response, regulate inflammation during infection, and help maintain homeostasis once the infection is cleared [58]. Chronic inflammation disrupts the homeostatic interactions among epithelial cells, stromal cells and immune cells, resulting in the transformation of tissue-resident fibroblasts, stellate cells or bone marrow-derived fibrocytes into myofibroblast-like cells, which subsequently leads to organ fibrosis [59].

Limonium Sinense has been shown to enhance the functions of immune organs via stimulating the proliferation of spleen and thymus cells. It increases the production of IFN-γ and IL-2 while decreases the levels of IL-4, thereby improving the phagocytic activity of macrophages [27]. Studies have shown that IFN-γ inhibits fibroblastic activity and eliminates Th2 immune response. Patients with idiopathic pulmonary fibrosis (IPF) had lower IFN-γ levels in the lungs [60, 61]. The pathogenesis of pulmonary fibrosis is closely associated with T-cell subsets, including Th1, Th2, and Th17 T-cells. The Th1 subset produces profibrotic factors such as IL-1, TNF-γ, platelet-derived growth factor (PDGF) and transforming growth factor-beta 1 (TGF-β1). The Th2 subset, particularly the cytokine IL-4, induces the production of IL-5, IL-13 and TGF-β1, leading to the recruitment of macrophages, mast cells, eosinophils, mesenchymal cells, as well as the activation of fibroblasts. The Th17 subset indirectly promotes fibrosis by increasing TGF-β1 levels [60]. Although lacking clear evidence, these analyses suggest a potential role for *Limonium Sinense* in treating of immune-related diseases such as fibrosis.

ROS plays a crucial role in the body's response to the external stimuli. The over-production of ROS, known as oxidative stress, can lead to significant damage to cell structure and functions. Imbalances in oxidative stress have been implicated in the initiation and progression of various diseases and disorders, such as cardiovascular diseases, inflammation, cancers, viral pathogenesis, drug-induced tissue injury, hypertension and the formation of drug-resistant mutants [62]. In cancer cells, ROS levels are generally higher than in the normal cells, highlighting the critical role of ROS in carcinogenesis and cancer progression. Interestingly, high levels of ROS can also be toxic to cancer cells and induce cell death, making ROS an important target for the development of potential anti-cancer therapies that aim to modulate cellular redox levels [63]. Furthermore, ROS can regulate chronic inflammation [64] and contribute to the process of epithelial-mesenchymal transition (EMT). Inhibition of EMT induced by oxidative stress has been shown to attenuate the development and progression of various cancers and fibrotic diseases [65, 66].

Many ingredients in *Limonium Sinense* exhibit strong antioxidant activity, especially the flavonoid compounds such as quercetin, luteolin and isorhamnetin. As examples, quercetin inhibits EMT in various ways, thereby attenuating tumour proliferation and invasion. Quercetin has also demonstrated inhibitory effects on neurological disorders, cardiovascular disease, idiopathic pulmonary disease, and organ fibrosis [67-70]. Luteolin has shown neuroprotective, reno-protective, and anti-fibrosis, anti-tumour, and anti-inflammatory properties [71-73]. These findings suggest that *Limonium Sinense* may have potential not only for the treatment of tumours but also for the treatment of neurological, cardiovascular, and fibrosis-related diseases.

It is well known that multiple chemical components with diverse pharmacological activities are involved in the process of disease treatment by drug. Natural compounds and their derivatives have long history been used as a source of therapeutic substances and as a structural basis for drug development, especially as anticancer agents [74]. Flavonoids, which are naturally occurring polyphenolic metabolites widely distributed in foods and medicinal herbs [75], have been used as anti-inflammatory, anti-cancer and anti-neurodegenerative agents to treat various chronic diseases, cancers and brain disorders [76-78]. Polysaccharides derived from herbal plants are ideal candidates for therapeutics due to their relatively low toxicity, immunomodulatory and anti-tumour properties. Polysaccharides such as lentinan, schizophyllan and krestin have been used as immunoceuticals in clinical trials [79]. Alkaloids display diverse pharmacological activities and have been used to treat various diseases and tumours for a long history. Some alkaloids have also shown neuroprotective effects. Galantamine and rivastigmine, two alkaloid compounds derived from plants, have been approved by US Food and Drug Administration (FDA) as acetylcholinesterase inhibitors for Alzheimer's disease [80].

Obviously, the abundant active ingredients of *Limonium Sinense* contribute to its possession of multiple pharmacological functions. However, the treatment of diseases by plant active products in the body is extremely complex that involved in the cooperation of a variety of active ingredients, microenvironment changes, immune responses, cell proliferation and migration, signalling pathways regulation and other factors activation. The underlying mechanisms still needed further exploration.

In summary, as a traditional Chinese folk medicinal plant, *Limonium Sinense* has very high medicinal value and promising development prospects. The outstanding biological activities of *Limonium Sinense* make it valuable in application for the treatment of cancers and other diseases. With its richness active compounds, *Limonium Sinense* exerts a wide range of physiological and pharmacological effects, which make it a potential adjuvant for chemotherapy. However, the specific genetic and cellular mechanisms of *Limonium Sinense* are not yet fully understood, more continued work such as conduct standard clinical trials and possible side effects observation should be undertaken.

1.6 Hypothesis and Aims

The hypothesis of this project is mainly based on these key questions below:

1) Does *Limonium Sinense* show effects on both normal and cancer cells?
2) What are the transcriptomic changes that occur in cells following treatment with extracts from *Limonium Sinense*?
3) What are the underlying mechanisms by which *Limonium Sinense* extracts affect cellular behaviour and function?

By addressing these questions, the project aims to enhance our understanding of *Limonium Sinense*'s therapeutic potential and shed light on its cellular mechanisms of action.

The aim of this project is to investigate the biological effects and cellular mechanisms underlying the therapeutic applications of *Limonium Sinense*. The interdisciplinary research project encompasses a wide range of areas including molecular and cell biology, structural and system biology, proteomics and bioinformatics, biochemistry, and pharmacology. One aspect of the project involves exploring the effects of *Limonium Sinense* extracts on both normal and cancer cell lines through 2D and 3D culturing. The goal is to understand how extracts influence cellular behaviour and viability. RNA sequencing analysis will be employed to examine the changes in gene expression patterns, and protein-protein interactions, relative signalling pathways, apoptosis, and cell cycle regulation in response to treatment with *Limonium Sinense* extracts will be fully analysed. This analysis will provide insights into the cellular programs affected by the extracts. Additionally, High-Performance Liquid Chromatography (HPLC) will be utilized to identify the bioactive compounds present in *Limonium Sinense*. The project will also employ structural methods such as molecular docking, to confirm the underlying mechanisms of action of the bioactive compound(s).

Chapter 2 Materials and Methods

2.1 Preparation of *Limonium Sinense* extracts

Healthy whole plants of *Limonium Sinense* (Girard) Kuntze were collected from the coastal region in Jiangsu, eastern China (33°09'33.0" N, 120°46'40.4" E). The whole plants were washed, and oven-dried at 60 °C until the weight was constant.

For the water extracts, the dried plants were crushed and extracted with distilled water at 95 °C for 2 hours at 1:200 (mass/volume, m/v) ratio. The extract process was repeated 3 times under the same conditions. The extracts were then put to suction filtration, rotary evaporated, and freeze-dried to powder. The powder was aliquoted and stored at -20 °C until future use when extracts were diluted with sterile water and filtered through a 0.45 µm filter (Millipore filter membranes, Merck, UK).

For the ethanol extracts, the dried plants were crushed and extracted with 95% ethanol using an ultrasonic extractor (Fisher Scientific, UK) at a ratio of 1:400 (m/v) with a processing time of 50 seconds and an elapsed time of 20 seconds, for a total extraction time of 30 minute. The extract process was repeated 3 times under the same conditions. The extracts were then put to suction filtration, rotary evaporated, and freeze-dried to powder. The powder was aliquoted and stored at -20 °C until future use when extracts were diluted with dimethylsulfoxide (DMSO) and filtered through a 0.45 µm filter (Millipore filter membranes, Merck, UK).

2.2 Cell culture

2.2.1 Cell lines

All cell lines used were listed in Table 2.1. The vial of cells was removed from the -80 °C freezer and thawed in a 37 °C water bath. Once the contents had thawed, the outside of the vial were disinfected and moved to a tissue culture hood. The cells were re-suspended to a new falcon tube with pre-warmed media and centrifuged to remove freezing media. Old media was removed, and fresh media was re-suspended with the cell pellet. The appropriate volume of cell suspension was then added to a cell culture flask in the incubator at 37 °C, 5% CO_2.

Table 2.1. Summary of cell lines

Cell line	Originated from	Media	Supplemented
MDA-MB-157	Human mammary gland/breast	Dulbecco's Modified Eagle Media (DMEM) (Life technologies, UK)	1% Penicillin /streptomycin, 10% FBS (Life technologies, UK)
MDA-MB-231	Human mammary gland/breast	Dulbecco's Modified Eagle Media (DMEM) (Life technologies, UK)	1% Penicillin /streptomycin, 10% FBS (Life technologies, UK)
MDA-MB-468	Human mammary gland/breast	Dulbecco's Modified Eagle Media (DMEM) (Life technologies, UK)	1% Penicillin /streptomycin, 10% FBS (Life technologies, UK)
BT-20	Human mammary gland/breast	Dulbecco's Modified Eagle Media (DMEM) (Life technologies, UK)	1% Penicillin /streptomycin, 10% FBS (Life technologies, UK)
HCC1395	Human mammary gland/breast	Roswell Park Memorial Institute (RPMI) 1640 (Life technologies, UK)	1% Penicillin /streptomycin, 10% FBS (Life technologies, UK)
HCC1806	Human mammary gland/breast	Roswell Park Memorial Institute (RPMI) 1640 (Life technologies, UK)	1% Penicillin /streptomycin, 10% FBS (Life technologies, UK)
HCC1937	Human mammary gland/breast	Roswell Park Memorial Institute (RPMI) 1640 (Life technologies, UK)	1% Penicillin /streptomycin, 10% FBS (Life technologies, UK)
MCF10A	Human mammary gland/breast	Dulbecco's Modified Eagle Media (DMEM) (Life technologies, UK)	1% Penicillin /streptomycin, 10% FBS (Life technologies, UK)
HEK293T	Human embryonic kidney	Dulbecco's Modified Eagle Media (DMEM) (Life technologies, UK)	1% Penicillin /streptomycin, 10% FBS (Life technologies, UK)

2.2.2 Passaging cells

Cells were incubated at 37 °C and supplied with 5% CO_2. Adherent cells were passaged 2-3 times per week. To passage cells, media was removed, cells were washed with 1 × phosphate buffered saline (PBS, Thermo Fisher Scientific, UK) twice. Cells were detached using 0.05% trypsin-EDTA (Thermo Fisher Scientific, UK) and placed in the incubator at 37 °C, 5% CO_2 for 5-10 mins. Cells were then checked under the microscope to confirm they were detached. An appropriate volume of media was added (1:1 ratio media:trypsin) and cell suspension was re-suspended. Required cell suspension was then added to falcon tubes and centrifuged at 500 g for 5 mins. Old media was removed, and appropriate volume of fresh media was added. Resuspended cells were plated in

appropriate plate and put in the incubator for further use. Routine mycoplasma testing was conducted on all cell lines and no mycoplasma was detected.

2.2.3 Freezing cells

To freeze cells, old media was removed, and the cells were washed with 1 × PBS. Cells were detached using 0.05% trypsin-EDTA and placed in the incubator at 37 °C, 5% CO_2 for 5-10 mins. Cells were then checked under the microscope to confirm they were detached. An appropriate volume of media was added (1:1 ratio media:trypsin) and cell suspension was re-suspended. The required cell suspension was then added to falcon tubes and centrifuged for 5 mins. Old media was then carefully removed from falcon tube. 1 ml of freezing media (10% DMSO, 90% FBS) was then added to the tube, gently re-suspending to prevent bubbles forming for around 1 minute. Then the suspension was added to a labelled cryogenic vial (Merck) and placed into Nalgene Freezing container (Merck) and moved to the -80 °C freezer. After 2 days the cryogenic vial was moved into a storage box in the -80 °C for storage.

2.2.4 3D culture

Mammosphere serum free media was prepared which containing 1% L-glutamine, 1% Penicillin /streptomycin, 30% F12, 2% B27, 20 ng/ml EGF and 20 ng/ml FGFb. Old media was removed from the adherent cells. The cells were washed twice with 1 × PBS and detached using 0.05% trypsin-EDTA and placed in the incubator at 37 °C, 5% CO_2 for 5-10 mins. Cells were then checked under the microscope to confirm they were detached. An appropriate volume of media was added (1:1 ratio media:trypsin) and cell suspension was re-suspended. Then cells were centrifuged at 200 g at room temperature for 5 minutes. The supernatant was removed, and cells were re-suspended in 1-5 ml of mammosphere media. Pellets were pipetted up and down until to obtain single cell suspension. Cells were checked by using haemocytometer under the microscope to ensure a single cell suspension. The number of viable cells per ml were then calculated using trypan blue. Cells were re-suspended in appropriate volume of mammosphere media in ultra low-attachment 6/12/24 well plate (Seeding density 500-4000 cells/cm^2) and incubated at 37 °C and 5% CO_2 for 5-10 days. After the culture period, the mammospheres (diameter greater than 40 µm) were counted under a microscope at 40 × magnification (Digitally image 5 random fields, the mammosphere size was determined using acquisition software). The mammosphere forming efficiency (MFE %) was calculated as follows:

MFE (%) = (number of mammospheres per well) / (number of cells seeded per well) × 100

2.2.5 Cell treatment

Limonium Sinense water extracts (LSW) or *Limonium Sinense* ethanol extracts (LSE) were diluted in a concentration gradient as 125, 250 and 500 μg/ml, respectively. In the dose course experiment, the cells were treated with different concentrations of extracts for 24 hours and 48 hours, and in the time course experiment, the cells were treated with specific concentrations of extract for 0, 1, 3, 6, 12 and 24 hours.

2.3 Cell viability assay

Cell viability was detected by the CellTiter-Glo Luminescent Cell Viability Assay (Promega). Control wells containing media without cells were prepared to obtain a value for background luminescence. The test compound was added to the experimental wells and incubated according to culture protocol. The plate and its contents were equilibrated at room temperature for approximately 30 minutes. An equal volume of CellTiter-Glo reagent was added to each well (e.g., 100 μl of reagent was added to 100 μl of medium containing cells for a 96-well plate), then contents were mixed for 2 minutes on an orbital shaker to induce cell lysis. The plate was incubated at room temperature for 10 minutes to stabilize luminescent signal. Luminescence was recorded (Note: Instrument settings depend on the manufacturer. An integration time of 0.25~1 second per well should serve as a guideline).

2.4 Cell cycle assay

Cells were washed with 1 × PBS and lysed with 0.05% trypsin. The lysates were centrifuged at 4,000 rpm for 2 minutes at 4 °C, supernatants were removed, and cell pellets were washed with 1 × PBS by pipette up and down to make the cells are single. Then cells were centrifuged at 4,000 rpm for 2 minutes at 4 °C and the supernatants were removed by pouring on tissue. Cell pellets were resuspended with 1 × PBS (300 μl) and cold 100% ethanol (700 μl). Then samples were spin down at 4,000 rpm for 2 minutes at 4 °C, the supernatants were removed, and cell pellets were resuspended with 500 μl 0.25% TritonX-100 (diluted with 1 × PBS) and incubated on ice for 15 minutes. Then samples were centrifuged at 4,000 rpm for 2 minutes at 4 °C, the supernatants were removed, and cell pellets were resuspended in 500 μl 1 × PBS which containing 200 μg/ml RNase A and 50 μg/ml Propidium Iodide (PI) solution (Sigma, UK). Samples were incubated for 30 minutes at room temperature avoiding directly light. After incubation, cell cycle was detected by using Guava software (*version 3.0*, Merck Millipore).

2.5 RNA extraction

Cells total RNA was isolated by RNeasy Mini Kit (cat. nos. 74104 and 74106, QIAGEN) according to the manufacturer's instructions. Cells were harvested at confluency in 6-well plates and lysed in Buffer RLT. Up to 700 µl of lysates were added to the RNeasy Mini spin columns provided after re-suspending in 70% ethanol (Thermo Fisher Scientific, UK). The columns were spun at 30 seconds at full speed. Discard the flow-through. Add 700 µl Buffer RW1 to the RNeasy spin column. Three wash steps were conducted using initially 700 µl of Buffer RW1, followed by two washes using 500 µl of the Buffer RPE with centrifugation of full speed for 30 seconds. The columns were then added to the new 2 ml collection tubes and centrifuged for 1 min at 14,000 g to dry the membrane. The columns were placed in the new 1.5 ml collection tubes and added up to 50 µl RNase-free water and centrifuged for 1 min at 14,000 g to elute the RNA. RNA concentration was quantified using the NanoDrop Spectrophotometers 2000c (Thermo Fisher Scientific, UK). RNA was then store at -20 °C or -80 °C for later use.

2.6 Quantitative real-time PCR (qRT-PCR) analysis

RNA was diluted to 20 ng/µl finally concentration using RNase free water. Real-time PCR was carried out using gene-specific primers. The qRT-PCR reaction system and set up conditions were listed in Table 2.2 and Table 2.3. Triplicates was performed for each sample in a 96-well qPCR plate on an ABI PRISM 7,000 Sequence Detection Systems (Thermo Fisher Scientific, UK). QuantiNova™ SYBR Green RT-PCR kits (Qiagen) were used with *CA9* (QT00011697), *VEGFA* (QT01010184) and *ACTB* (β-actin, QT00095431) gene specific primers (QuantiTect Primer Assays, Qiagen). Relative mRNA levels of target genes were normalized to *ACTB* (β-actin).

Table 2.2. Real-time PCR reaction system

Name of component	Volume (µl)
RNA (100 ng/µl)	0.5
Primer	0.5
2 × QN SG RT-PCR MM (Qiagen, UK)	5
QN SG RT Mix (Qiagen, UK)	0.1
QN ROX dye (Qiagen, UK)	0.5
RNase free H_2O (Qiagen, UK)	3.4

Table 2.3. Real -time PCR reaction conditions

Condition	Temperature	Time	Cycles
Initial Denaturation	95 °C	15 min	1
Denaturation	94 °C	15 s	35
annealing	55 °C	15 s	
extension	72 °C	30 s	

2.7 Plasmid DNA extraction

For small scale plasmid DNA extraction (< 20 µg total), the QIAprep Spin Miniprep Kit (Qiagen, UK) was used according to manufacturer's instructions. DNA was eluted in 50 µl of H_2O, EB (elution buffer) or TE (10 mM Tris-Cl, 1 mM EDTA, pH 8.0) and obtained by centrifuging at 13,000 rpm for 50 seconds. The amount of DNA was measured by Nanodrop Spectophotometer 2000c (Thermo Fisher Scientific, UK). If the final concentration of the DNA was lower than 50 ng/µl, up to 15 ml of overnight bacterial culture could be used to increase the yield.

2.8 Lipofectamine 3000 transfection

Cells were transfected according to the Lipofectamine 3000 Reagent protocol (Life technologies, UK). Cells were 70-90% confluent at the time of transfection. Lipofectamine 3000 reagent was diluted in Opti-MEM Medium and mixed well. DNA-lipid complexes were prepared before transfection, the transfection amounts were made following by Table 2.4 according to the protocol. Master mix of DNA was prepared by diluting DNA in Opti-MEM Medium, then the P3000 Reagent was added and mixed well. Diluted DNA was added to each tube of diluted Lipofectamine 3000 Reagent at 1:1 ratio and incubated for 10-15 minutes at room temperature. Then the complexes were added directly to cells in culture medium, in the presence or absence of serum/antibiotic. During transfection, it is not necessary to remove complexes or change/add medium. The cells were incubated for 2-4 days at 37 °C, then appropriate concentration of LSW was added to the cells. After treatment for 24 hours, cells were tested and analysed.

Table 2.4. Transfection amounts of Lipofectamine 3000 Reagent

Component	96-well	24-well	6-well
DNA-lipid complex	10 µl	50 µl	250 µl
DNA per well	100 ng	500 ng	2500 ng
P3000™ Reagent per well	0.2 µl	1 µl	5 µl
Lipofectamine 3000 Reagent per well	0.15 and 0.3 µl	0.75 and 1.5 µl	3.75 and 7.5 µl

2.9 Luciferase reporter assay

Cells were transfected using Lipofectamine 3000 (Life technology) in a 96 well plate with 100 ng of Renilla along with 100 ng of HRE motif per well. Cells were washed with 1 × PBS and lysed by trypsin (0.05% trypsin, Gibco), then centrifuged at 500 g for 5 minutes. The cell pellet was then re-suspended in certain count of complete media before plating on a 96 well plates (Usually 100 µl of medium for each well) at 70%-80% confluency. For each well transfected, 0.1 µl of lipofectamine 3000 reagent (Life technology) was diluted in 5 µl of Opti-MEM medium. Mixed well reagent was made and quickly vortexed. Diluted HRE DNA and lipofectamine 3000 were mixed by pipetting up and down and the lipid-DNA mixture was incubated at room temperature for 15 minutes. Cells were transfected at 37 °C for 48 hours before analysis. Then cells were treated with *Limonium Sinense* water extracts for 24 hours. The transcriptional assay was carried out using the Dual-Luciferase reporter assay system (Promega, UK) following the manufacturer's protocol. Transfected cells from *Limonium Sinense* water extracts treatment were washed with 1 × PBS prior to lysis in a 96 well plate. Cells were lysed in 100 µl of passive lysis buffer and put on a room temperature shaker for 15 minutes. Freezing lysates at -20 °C facilitated the lysis. 5 µl of lysate was analysed for each well in a 96 well white plate. Triplets were used for each transfection and 25 µl LAR II was first added and mixed by pipetting to measure the firefly luciferase activity. Another 25 µl of stop and go reagent was then added to help identify the Renilla activity. The final Dual-Luciferase Reporter activity was normalized based on both two measurements.

2.10 DharmaFECT RNAi knockdown

Cells were transfected using the protocol provided by Dharmacon (Life Sciences, UK). The siRNA oligos were prepared to a final concentration of 20 nM using a 1:50 dilution in HBSS (Thermo Fisher Scientific, UK) and added to each well. DharmaFECT reagent was diluted in HBSS (2:100) and incubated for 10 minutes at room temperature. Prior to transfection cell suspensions centrifuged for 5 minutes. 200 µl of diluted DharmaFECT reagent was added to each well, mixed and left for 20 minutes at room temperature. Following this, a total volume of 2,000 µl was required per well. Cells were seeded to give 40% confluence in fresh media, then total volume per well was made up to 2,000 µl per well. Cells were left for 2-3 days in 37 °C, 5% CO_2 incubator for transfection to complete.

2.11 Protein extraction

Cells were washed twice with 1 × PBS, a certain volume (100-200 µl) of 8 M urea lysis buffer was added to lysis the cells. Lysed cells were incubated on ice for 30 minutes and then centrifuged at 14,000 g, 4 °C for 20 minutes. The lysate supernatant was collected and stored at -20 °C. Protein lysates were quantified using Bradford assay (Thermo Fisher Scientific, UK) as per manufacturer's instructions. 1 µl of samples was loaded in triplicates to 96-well plate. Bovine serum albumin (BSA) dilutions (0, 0.5, 1, 2, 4 µl) in 200 µl Bio-Rad Protein Assay Dye Reagent Concentrate (Bio-Rad, UK) reagent dye was used to produce a protein concentration standard curve. 1 µl of sample lysate was added to 200 µl of diluted reagent dye into each well at a ratio of 1:5 dH_2O. The absorbance of each well was read at 584 nm by using a FLUOstar Optima microplate reader (BMG Labtech, UK) and normalized by linear regression to form the standard curve using MARS software (BMG Labtech, UK) to determine protein levels of the samples.

2.12 Western Blot

2.12.1 Gel preparation

Appropriate percentage acrylamide gels for sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-Page) were prepared according to the protein size (typically 8%, 10% or 12%, 15 ml total) (Table 2.5). Samples were prepared for loading using 4 × NuPAGE loading buffer (Thermo Fisher Scientific, UK) with 5% β-Mercaptoethanol (Thermo Fisher Scientific, UK) boiled at 100 °C for 5 min and centrifuged 14,000 rpm for 2 minutes. Equal quantity of protein was loaded into each well, along with a molecular weight marker, HyperPAGE pre-stained protein marker (Bioline, UK). Gels were run at 80 V for 30 minutes and then 150 V for 90 minutes in a mini-PROTEAN tank (Bio-Rad, UK).

Table 2.5. Polyacrylamide Gels for SDS-Page

Component	8% resolving gel (ml)	10% resolving gel (ml)	12% resolving gel (ml)	Stacking gel (ml)
dH_2O	6.9	5.9	4.9	10.2
30% Acrylamide mix (Seven Bioscience, UK)	4.0	5.0	6.0	2.55
1.5 M Tris-HCl	3.8	3.8	3.8	-
1.0 M Tris-HCl	-	-	-	1.9
10% SDS	0.15	0.15	0.15	0.15
10% Ammonia Persulphate (APS) (Thermo Fisher Scientific, UK)	0.15	0.15	0.15	0.15
TEMED (Thermo Fisher Scientific, UK)	0.009	0.006	0.006	0.015

2.12.2 Transferring

After SDS-Page separation, proteins were transferred onto a nitrocellulose membrane (GE Healthcare Life Science, UK) in a transfer cassette (Bio-Rad, UK), and then placed in a transfer tank (Amersham Biosciences, UK) filled with transfer buffer, transfer at 80 V, 4 °C, 20 mA for 3 hours or 1 mA for overnight.

2.12.3 Antibody incubation

Membrane was briefly blotted in ddH_2O and then stained in Ponceau S solution (Sigma Aldrich, UK) to check loading and transfer quality. Membrane was then blocked in 5% milk in TBST for 1 hour on a room temperature shaker. Membrane was then incubated with primary antibody (listed in Table 2.6) at dilution indicated by company at 4 °C for overnight. Membranes was then washed for 10 minutes in TBST for 3 times on a room temperature shaker. Membrane was then incubated in appropriate secondary antibody at 1:5000 dilution for 1 hour at room temperature in the dark. The IRDye secondary antibody were either Anti-Mouse IgG (M680, LICOR bioscience, UK) or Anti-Mouse IgG (M800, LICOR bioscience, UK) or Anti-Rabbit IgG (R800, LICOR bioscience, UK). The membrane was finally washed 3 × 10 minutes in TBST.

Table 2.6. List of antibodies

Antibody	Company	Cat-No	Species	Size (kDa)	Working dilutions
Primary Antibody					
Anti-β-tubulin	Cell Signalling Technology	86298	Mouse	55	1:1000
Anti-β-tubulin	Abcam	Ab6046	Rabbit	52	1:500
Anti-β-actin	Santa Cruz	sc-47778	Mouse	43	1:500
Anti-HIF-1α	Cell Signalling Technology	79233	Mouse	120	1:1000
Anti-HIF-2α	Cell Signalling Technology	7096s	Rabbit	120	1:1000
Anti-Hydroxy-HIF-1α (Pro564)	Cell signalling Technology	3434	Rabbit	120	1:1000
Secondary Antibody					
800CW	LI-COR	926-32211	Rabbit		1:5000
680CW	LI-COR	926-68020	Mouse		1:5000

2.12.4 Imaging and analysis

LI-COR Odyssey® CLx was used to image the membranes and images were analysed by Image Studio Lite Ver 5.2 software (Odyssey, UK). Then, ImageJ was used to quantify images with the normalisation to β-actin or β-tubulin loading control by comparing the density of each band.

2.13 Immunofluorescence microscopy

The media was removed when the cells reached 80-90% confluency and cells were gently washed with 1 × PBS twice. 1 ml 4% paraformaldehyde (PFA) (Thermo Fisher Scientific, UK) in 1 × PBS was added to fix the cells for 15 minutes. PFA was removed and cells were washed with 1 × PBS. For permeabilization of cells, 500 µl of 0.1% TironX-100 (Thermo Fisher Scientific, UK) in 1 × PBS was added to the each well of 12 well plate and the slide was transferred from 6 well plate into 12 well

plate and incubated in 0.1% TironX-100 for 5 minutes on ice. This was followed by washing the slides with 1 × PBS twice. Then, cells on the slides were blocked in 0.2% Fish Skin Gelatine (Sigma Aldrich, UK) in 1 × PBS for 60 minutes at room temperature. Meanwhile, anti-HIF-1α primary antibody was prepared in blocking buffer with 1:50 dilution and parafilm was put on the foil wrapped container. To moisture the container, wet tissues were put into the side of box. 45 µl of primary antibody (BD Transduction Laboratories™, HIF-1α, Cat No: 610959; 1:100) was put on the parafilm and the excess buffer was got rid of from the slides and slides were put on the antibody upside down for 60 minutes at room temperature. 60 minutes after the primary antibody incubation, slides were flipped and put into 12 well plate and washed with 1 × PBS 3 times, each time for 15 minutes on the rocker. Then, the secondary antibody with 4'6-Diamidino-2-Pheylindole (DAPI) (Invitrogen, UK) was prepared in 1 × PBS with the dilution of 1:400 and 1:1000, respectively. New parafilm was put into the box, 95 µl of secondary antibody was put onto parafilm, and slides were put onto parafilm upside down and incubated at room temperature for 60 minutes. Slides were washed with 1 × PBS as previously by avoiding light. 8 µl of mounting solution was added to the cover slip and slide was put on the cover slip upside down and left to air dry overnight by avoiding light. Protein expression was detected using Alexa Fluor (1:400, Molecular Probes) for 20 minutes. Immunostained cells were analysed and photographed using an Olympus IX83 inverted fluorescence microscope.

2.14 High-performance liquid chromatography (HPLC) analysis

Chromatography analysis for the identification of gallic acid in LSW was conducted on a Shimadzu® HPLC system (LC-20 AT, SHIMADZU, Japan) equipped with C18 column (Shim-pack GIS: 5 µm particle size; 4.6 × 250 mm, P/N: 227-30106-08). LSW samples were diluted with 50% methanol at a concentration of 20 mg/ml. The chromatographic separation was carried out using a mobile phase with phosphoric acid: water 0.1% (pH 3.0) as solvent A and methanol as solvent B at a flow rate of 1 ml/min. An isocratic system was used as follows: 40% A/60% B, 20 minutes. Peaks were detected at 271 nm using an UV-Vis detector (SPD-20A), and the peak for gallic acid was identified by comparing the retention time with its standard.

2.15 RNA sequencing (RNA-seq) analysis

A total amount of 3 µg RNA per sample was used as input material for library construction. Sequencing libraries were generated using NEBNext® UltraTM RNA Library Prep Kit for Illumina® (NEB, Ipswich, Massachusetts, USA) following manufacturer's instruction. Libraries were pooled in equimolar and sequenced using the paired-end strategy (2 × 150) on the Illumina NovaSeq 6000

platform following the standard protocols (Novogene, UK). Raw read counts were imported into RStudio (*version 4.2.0*) and analysed by using R packages. Transcripts with low abundance (under 10 counts across all samples) were removed. Genes with |Log2FoldChange| above 1 and *P* values adjusted by using Benjamini-Hochberg (BH) method less than 0.05 were considered as differentially expressed genes (DEGs).

2.16 t-SNE analysis

The t-distributed stochastic neighbour embedding (t-SNE) plot was generated from Omics Playground which is an online self-service analytics platform (https://bigomics.ch/omics-playground/).

2.17 Enrichment analysis

Gene Set Enrichment Analysis (GSEA) Hallmark analysis was performed via GSEA software (*version 4.1.0*) (with registration). GSEA is a computational method designed to ascertain whether a predefined set of genes displays statistically significant, consistent variations between two distinct biological states (For example phenotypes) [81, 82]. GSEA operates by evaluating experiments containing genome-wide expression profiles derived from samples belonging to two classes. Genes are ranked based on the correlation between their expression levels and the class distinction, utilizing a suitable metric. The inputs required for GSEA include: **1**) Expression data set *D* with *N* genes and *k* samples. **2**) A ranking procedure to generate the Gene List *L*, involving a correlation (or other ranking metric) and a phenotype or profile of interest *C*. **3**) An exponent *p* that regulates the weighting of the step. **4**) A separately derived Gene Set *S* of N_H genes, such as a pathway, a cytogenetic band, or a Gene Ontology (GO) category. It's noteworthy that in the aforementioned analyses, only gene sets with a minimum of 15 members were utilized to emphasize robust signals (accounting for 78% of MSigDB).

Enrichment Score (ES) reflects the degree to which a gene set *S* is overrepresented at the extremes (top or bottom) of the entire ranked list *L*. The process to calculate ES involves: **1**) Rank order the *N* genes in *D* to form $L = \{g_1, \ldots, g_N\}$c according to the correlation, $r(g_j) = r_j$, of their expression profiles with *C*. **2**) Assessing the ratio of genes in *S* (''hits'') weighted by their correlation, alongside the ratio of genes not in *S* (''misses'') present up to a given position *i* in *L*.

$$P_{\text{hit}}(S, i) = \sum_{\substack{g_j \in S \\ j \le i}} \frac{|r_j|^p}{N_R}, \quad \text{where } N_R = \sum_{g_j \in S} |r_j|^p$$

$$P_{\text{miss}}(S, i) = \sum_{\substack{g_j \notin S \\ j \le i}} \frac{1}{(N - N_H)}.$$

The ES represents the maximum deviation from zero of P_{hit} - P_{miss}. When dealing with a randomly distributed *S*, *ES* tends to be relatively small. However, if the gene set exhibits concentration either at the top or bottom of the list, or if it demonstrates a non-random distribution, the *ES* will correspondingly be higher. For parameter values, when $p = 0$, *ES* simplifies to the standard Kolmogorov–Smirnov statistic. Conversely, when $p = 1$, the genes in S were weighted by their correlation with *C*, normalized by the sum of correlations across all genes in *S*. The significance of an observed *ES* was assessed by comparing it with the set of scores ES_{NULL} computed with randomly assigned phenotypes: **1**) Randomly assign the original phenotype labels to samples, reorder genes, and recalculate the *ES*. **2**) Repeat step 1 for 1,000 permutations, creating a histogram of the corresponding enrichment scores ES_{NULL}. **3**) Estimate the nominal *P* value for *S* from ES_{NULL} by using the positive or negative portion of the distribution, matching the sign of the observed *ES*.

For multiple hypothesis testing, the methodology includes: **1**) Computing *ES* for each gene set in the collection or database. **2**) Reordering the genes in *L* for each *S* and 1,000 fixed permutations π of the phenotype labels to determine *ES* (*S*, π). **3**) Adjusting for gene set size variance by normalizing *ES* (*S*, π) and the observed *ES* (*S*), separately rescaling the positive and negative scores by dividing by the mean of the *ES* (*S*, π) to yield the normalized scores *NES* (*S*, π) and *NES* (*S*). **4**) Calculating the False Discovery Rate (FDR) to control the ratio of false positives among all gene sets attaining a fixed level of significance for positive (negative) *NES* (*S*) and *NES* (*S*, π). Create a histogram of all *NES* (*S*, π) over all S and π. Use this null distribution to compute an FDR *q* value for a given *NES* (*S*) = *NES** ≥ 0. The FDR is the ratio of the percentage of all (*S*, π) with *NES* (*S*, π) ≥ 0, where *NES* (*S*, π) $\geq$ *NES**, divided by the percentage of observed *S* with *NES* (*S*) ≥ 0, where *NES* (*S*) $\geq$ *NES**, and similarly if *NES* (*S*) = *NES** ≤ 0.

2.18 Connectivity map (CMap) analysis

A public online tool CMap (http://clue.io/) (with registration) was performed to determine which target drugs might have similar or opposing expression signatures with *Limonium Sinense* extracts, the mechanism of actions (MOA) and drug-target were investigated as well. Top 150 up regulated genes and top 150 down regulated genes were submitted in the CMap Query tool in December 2021, and compounds with connectivity score ≥ 90 were considered as similar compounds, whereas compounds with connectivity score ≤ -90 were considered as opposite compounds.

2.19 Gene Set Variation Analysis (GSVA) analysis

GSVA computes gene set enrichment scores on a per-sample basis by considering genes both inside and outside the gene set, similar to a competitive gene set test. Moreover, it evaluates the

variability of gene set enrichment across samples independently of any predefined class labels [83]. The GSVA process begins by assessing whether a gene *i* is highly or lowly expressed in sample *j* in the context of the sample population distribution. Probe effects can lead to variations in hybridization intensities in microarray data, causing significant differences in expression values between two non-expressed genes. To standardize disparate expression profiles onto a common scale, an expression-level statistic is calculated as follows. For each gene expression profile, a non-parametric kernel estimation of its cumulative density function is performed. In the case of microarray data, a Gaussian kernel is used:

$$\widehat{F}_{h_i}\left(x_{ij}\right) = \frac{1}{n}\sum_{k=1}^{n}\int_{-\infty}^{\frac{x_{ij}-x_{ik}}{h_i}} \frac{1}{\sqrt{2\Pi}} e^{-\frac{t^2}{2}}\, dt$$

where h_i is the gene-specific bandwidth parameter that controls the resolution of the kernel estimation, which is set to $h_i = s_i/4$, where s_i is the sample standard deviation of the *i*-th gene.

To reduce the impact of potential outliers, the enrichment score was assessed using the Kolmogorov-Smirnov like random walk statistic:

$$\widehat{F}_r\left(x_{ij}\right) = \frac{1}{n}\sum_{k=1}^{n}\sum_{y=0}^{x_{ij}} \frac{e^{-(x_{ik}+r)}\left(x_{ik}+r\right)^y}{y!}$$

where τ represents a parameter determining the weight of the tail in the random walk (default τ = 1), γ_k signifies the *k*-th gene set, $I(g_{(i)} \in \gamma_k)$ is the indicator function denoting whether the *i*-th gene (corresponding to the *i*-th ranked expression-level statistic) belongs to the gene set γ_k, $|\gamma_k|$ is the number of genes in the *k*-th gene set, and *p* indicates the total number of genes in the data set. In essence, Equation 3 generates a distribution across genes, aiming to assess whether the genes within the gene set are more inclined towards either tail of the rank distribution. The enrichment statistic (ES), also called GSVA score, quantifies the maximum deviation from zero of the random walk concerning the *j*-th sample in relation to the *k*-th gene set:

$$ES_{jk}^{\max} = \nu_{jk}\left[\arg\max_{\ell=1,\ldots,p}\left|\nu_{jk}(\ell)\right|\right]$$

For each gene set *k*, this approach produces a distribution of enrichment scores that is bimodal.

In GSEA, it's noted that the empirical null distribution, acquired through permutations of sample labels, often displays a bimodal pattern. Consequently, significance is assessed separately using the

positive and negative sides of the null distribution. In this case, an alternative score that produces an *ES* distribution approximating this requirement:

$$ES_{jk}^{\text{diff}} = \left|ES_{jk}^{+}\right| - \left|ES_{jk}^{-}\right| = \max_{\ell=1,\ldots,p}\left(0, \nu_{jk}\left(\ell\right)\right) - \min_{\ell=1,\ldots,p}\left(0, \nu_{jk}\left(\ell\right)\right)$$

where ES_{jk}^{+} and ES_{jk}^{-} are the largest positive and negative random walk deviations from zero, respectively, for sample *j* and gene set *k*.

The HIF score was calculated by using GSVA method, briefly, a 15-gene expression signature (*ACOT7*, *ADM*, *ALDOA*, *CDKN3*, *ENO1*, *LDHA*, *MIF*, *MRPS17*, *NDRG1*, *P4HA1*, *PGAM1*, *SLC2A1*, *TPI1*, *TUBB6* and *VEGFA*), which enables classification of hypoxia-inducible factor (HIF) activity [84, 85].

2.20 Statistical analysis

Experiments were validated in triplicates. Data were plotted as mean ± SD (standard deviation). Statistical analysis of single comparisons of two groups utilised Student's *t*-test or Mann-Whitney U-test for parametric and non-parametric data, respectively. A Dunnett's multiple comparisons test was used for comparing multiple samples. Results were regarded statistically significant if $P < 0.05$. * referred to $P < 0.05$, ** referred to $P < 0.01$, *** referred to $P < 0.001$, **** referred to $P < 0.0001$. Unless cited, all data and graphs were analysed by GraphPad Prism (*version 9.5.1*) (GraphPad Software Inc). Both test methods and the number of biological replicates were detailed in each figure legend.

Chapter 3 Integrated network pharmacology reveals the biological mechanisms of *Limonium Sinense* (Girard) Kuntze against breast cancer

Part of the background of provided in this chapter is from my original draft of the following publication with support from my supervisors:

***Hualong Zhao**, Siyuan Wang, Philip T F Williamson, Rob M Ewing, Xinhui Tang, Jialian Wang, Yihua Wang. Integrated network pharmacology and cellular assay reveal the biological mechanisms of Limonium sinense (Girard) Kuntze against breast cancer. BMC Complement Med Ther, 2023, 23(1):408. doi: 10.1186/s12906-023-04233-z.*

Authors' contributions

*Yihua Wang and Jialian Wang designed the study. **Hualong Zhao** performed the experiments. Yihua Wang, Jialian Wang, **Hualong Zhao** and Siyuan Wang analysed the data and finalised the figures, and all other authors contributed to data interpretation.*

3.1 Abstract

Limonium Sinense (Girard) Kuntze (*Limonium Sinense*) has been widely used for the treatment of anaemia, bleeding, cancer, and other disorders in Chinese folk medicine. The aim of this study is to predict the therapeutic effects of *Limonium Sinense* and investigate the potential mechanisms using integrated network pharmacology methods. The active ingredients of *Limonium Sinense* were collected from published literature, and the potential targets related to *Limonium Sinense* were obtained from public databases. Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) and DisGeNET enrichment analyses were performed to explore the underlying mechanisms. Molecular docking and Gene Expression Omnibus (GEO) datasets were employed to further evaluate the findings. A total of 15 active ingredients of *Limonium Sinense* and their corresponding 389 targets were obtained. KEGG enrichment analysis revealed that the biological effects of *Limonium Sinense* were primarily associated with Pathways in cancer. DisGeNET enrichment analysis highlighted the potential role of *Limonium Sinense* in the treatment of breast cancer. Apigenin and Luteolin within *Limonium Sinense* showed promising potential against cancer. Additionally, analysis of GEO datasets validated the involvement of breast cancer and several cancer-related pathways upon treatment with Apigenin or Luteolin in human cancer cells. This study predicts the biological activities of *Limonium Sinense* and reveals its potential for breast cancer treatment, highlighting the promising role of *Limonium Sinense* in the field of cancer therapy.

3.2 Introduction

3.2.1 Network pharmacology

The concept of network pharmacology was developed to analyse the comprehensive interactions between natural compounds and the human body [86]. It challenges the assumption of “one drug for one target for one disease”, which has influenced various aspects of drug discovery, such as disease classification, target identification, screening, drug design and clinical trial design. However, systems biology advancement have indicated that complex diseases may not be effectively treated by targeting single nodes alone [87].

It has been recognised in recent years that many effective drugs in different therapeutic areas, including oncology, psychiatry and anti-infectives, exert their effects on multiple targets rather than a single one [88]. Traditionally, medicinal chemists have been cautious about designing ligands with multiple activities, fearing complex structure-activity relationships or high molecular weights of conjugated ligands [89, 90]. However, the polypharmacology of approved drugs demonstrated the potential benefits of an opportunistic approach that exploits multi-target activity. By combining

chemogenomics with network biology, a new network-pharmacology approach to drug discovery can be facilitated [90-92].

In network pharmacology, the initial step involves collecting the active compounds and their potential targets. Protein-protein interaction (PPI) analysis is performed to identify the hub proteins, which are further subjected to biological enrichment analysis using Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) analysis. These analyses help to predict the related pathways, molecular functions, biological processes, and cellular components specific to the species under study [93, 94]. Molecular docking is employed to assess the binding affinity between proteins and ligands, which aids in identifying probable key ligands and hub proteins for validation, especially when dealing with a large number of differential proteins and potential ligands. In recent years, molecular dynamics simulations have been increasingly utilized to validate the binding mechanisms between proteins and ligands based on the docking pose. The results obtained from network pharmacology are further validated through *in vitro* and *in vivo* experiments, RNA-sequencing (RNA-seq) approaches, gene chip analysis, or other techniques. Additionally, the integration of multi-omics approaches, such as transcriptomics, proteomics, and metabolomics, with network pharmacology enables a comprehensive and holistic exploration of the mechanisms underlying traditional herbal medicine or natural compounds. In summary, network pharmacology greatly contributes to the prediction of the biological activities of the active compounds.

3.2.2 ADMET

Drug discovery and development is an extremely complex and expensive process that involves several stages, including disease selection, target identification and validation, lead discovery and optimization, preclinical and clinical trials [95]. In recent years, *in silico* ADMET prediction, which refers to the prediction of chemical absorption, distribution, metabolism, excretion, and toxicity, has become a preferred method in early drug discovery and plays a vital role in drug development [96, 97].

3.2.2.1 Absorption

The absorption of orally administered compounds depends on their solubility, dissolution within the gastrointestinal (GI) tract, and ability to permeate the intestinal membrane. The compound's physicochemical properties like solubility, hydrophobicity, ionization, and molecular weight (MW), along with the characteristics of the GI tract, collectively determine the rate and extent of compound absorption. Consequently, these factors strongly influence the bioavailability of the compound after oral administration [98]. The process of oral drug absorption is complex and affected by the drug's properties, dosage form, and the patient's GI tract condition [99].

Simultaneously, the GI mucosal membrane's nature and surface area, which changes from the stomach to the rectum, as well as the physicochemical attributes of the contents in the digestive tract, also impact absorption. As the drug traverses the GI tract, either in liquid or solid form, at various speeds with the digestive content, it encounters different environments [100, 101]. Permeability assessment is vital for identifying potential drug candidates that might face challenges in preclinical and clinical development. Generally, compounds with over 90% oral absorption are considered highly permeable, while those with under 50% absorption are classified as poorly permeable, although this does not rule out their potential as drugs. Compounds with very poor permeability are more likely to have limited absorption in the body. Various *in vitro* assays are available to assess intestinal absorption and permeability across different membrane barriers. These assays including solubility, log*P* and log*D* measurements, and Caco-2 permeability assays, greatly aid in designing and interpreting *in vivo* experiments [98].

3.2.2.2 Distribution

Once in the bloodstream, a drug spreads across diverse tissues or organs. This pattern is influenced in part by factors like how well the drug dissolves in fat or water, its ionization, and physiological processes including protein binding and tissue absorption. As a result, distribution depends on protein binding, pH conditions, blood flow, barrier permeability like the blood-brain barrier and placental barrier, and body composition [102]. After moving from the digestive system to the bloodstream, the drug enters various tissue and organ spaces, including interstitial and intracellular areas, which is crucial for targeting. Multiple factors affect distribution dynamics, including cardiac output, regional blood flow, changes in capillary permeability, interactions with plasma proteins, and tissue binding affinity. Furthermore, active transporters and metabolic enzymes also play roles in determining the distribution and disposition of the drug within tissues and organs [99].

3.2.2.3 Metabolism

The liver is recognized as the primary site for drug metabolism. However, the first-pass effect of orally administered drugs is also impacted by intestinal drug metabolism, and organs like kidneys, lungs, blood cells, placenta, and even the brain can also contribute to drug metabolism. Drug metabolism typically follows phase I or phase II reactions. Phase I reaction includes processes like oxidation, reduction, and hydroxylation, while phase II reaction involves joining molecules to enhance water solubility [102]. Metabolism converts the parent compound into metabolites, which are subsequently excreted primarily through kidneys or bile [98]. Notably, certain enzymes involved in metabolism have genetic variation, such as CYP2C9, CYP2D6, CYP2C19, UDP-glucuronosyl transferase, and N-acetyltransferase [103]. A prominent group of drug-metabolizing enzymes is the microsomal cytochrome P450 (CYP450) family. Among them, CYP3A4 is the most abundant P450

enzyme in the liver and metabolizes around 40% to 50% of clinically utilized drugs, and CYP2D6, present in the liver and brain, metabolizes the second most drugs [104]. CYP2C9 is present in multiple body tissues, including the GI tract, and shows a predilection for both neutral and acidic substrates. Although most substrates are weakly acidic compounds, CYP2C9 also participates in the N-demethylation of different basic drugs [105].

3.2.2.4 Excretion

Renal excretion is vital for eliminating compounds and their metabolites. Compounds that undergo renal excretion are usually water-soluble, have low MW, or metabolize slowly in the liver. This elimination involves steps including glomerular filtration, activate tubular secretion, tubular reabsorption, and renal metabolism. Initially, plasma drugs filter through the glomerular membrane into the renal tubule. Transport systems in the tubular membrane aid drug excretion into the tubule. Furthermore, lipophilic drugs can passively diffuse back into the bloodstream from the tubule's distal part. Each of these processes matures independently, especially from a developmental standpoint [102]. Glomerular filtration is a size-selective process, eliminating small, unbound compounds (MW < 500). It's crucial for water-soluble drugs and their metabolites. The glomerular filtration rate (GFR) is a key metric for assessing renal function. Tubular secretion happens in proximal tubules with epithelial cells and is an active carrier-mediated transport process [106]. It's capacity-limited and can saturated. Factors such as renal blood flow, unbound plasma fraction, and intrinsic renal tubular secretion influence active tubular secretion clearance. Tubular reabsorption mainly occurs in distal tubules via passive diffusion. The extent of reabsorption depends on a compound's lipophilicity and ionization. Additionally, proximal tubules also have uptake and efflux transporters that could impact reabsorption process [107].

3.2.2.5 Toxicology

Toxicology ensures the safety of molecules, confirming their harmlessness. Toxicology's progression is constrained by regulatory standards. Innovation in toxicology is limited by the need to meet regulations, often leading to a rigid "box-ticking" approach with mandated tests, from molecular to animal studies. Regulatory authorities value ADME data, especially regarding liver enzyme activity, which could potentially influence drug interactions [108].

3.2.3 Drug-likeness assessment rules

One of the earliest applications of candidate property evaluation in drug discovery is drug-likeness assessment. The "Rule of five", initially propose by *Lipinski* and co-workers in 1997, is the most well-known rule-based filter for assessing drug-likeness. It has been used to select orally bioavailable

compounds based on four simple rules related to molecular properties [109]. The rules are as follows: molecular weight (MW) ≤ 500, octanol/water partition coefficient (Alog*P*) ≤ 5, number of hydrogen bond donors (nHDs) ≤ 5 and number of hydrogen bond acceptors (nHAs) ≤ 10. According to the Rule of Five, a molecule would likely not be orally active if it violates two or more of these rules. However, the rules may not be suitable for complex natural products because they were derived from relatively simple small molecules.

Several drug-likeness rules or filters similar to the "Rule of Five" have been proposed recently [110-115]. For instance, *Ghose* and co-workers proposed that more than 80% of compounds meet the following limits: -0.4 ≤ Alog*P* ≤ 5.6, 160 ≤ MW ≤ 480, 40 ≤ MR (molar refractivity) ≤ 130, and the total number of atoms between 20 and 70. These drug-likeness filters based on physicochemical properties help expedite the drug development process [110]. *Veber* and co-workers proposed another set of criteria. They suggested that compounds meeting only two criteria of (1) having 10 or fewer rotatable bonds (RBs) and (2) a topological polar surface area (TPSA) equal to or less than 140 $Å^2$ (or having 12 or fewer HDs and HAs), would have a high probability of good oral bioavailability [115].

3.2.4 Molecular docking

Molecular docking has emerged as a crucial tool in drug discovery, particularly in the realm of structure-based drug design, and has been extensively utilized since the early 1980s [116]. The molecular docking approach enables the modelling of interactions between small molecules and proteins at the atomic level, which allow the users to characterize the behaviour of small molecules within the binding sites of target proteins and elucidate essential biochemical processes [117]. The docking process involves two basic steps: predicting the conformation, position, and orientation of the ligand within these sites (often referred to as pose) and evaluating the binding affinity.

Efficiency in the docking process significantly improves when the location of the binding site is known beforehand. In many cases, the binding site is already known before the introduction of ligands for docking. Additionally, information about these sites can be obtained by comparing the target protein with a family of proteins sharing a similar function or with proteins co-crystallized with other ligands. Essentially, the aim of molecular docking is to predict the structure of the ligand-receptor complex through computational methods. Ideally, sampling algorithms should be replicate the experimental binding mode, and the scoring function should rank it as the highest among all generated conformations.

The foundational concept of docking stems from the "lock and key" principle in rational drug design. However, the specific algorithms employed to fit the "key" (the ligand) into the "lock" (the receptor

protein) vary among different software programs [118]. Several software tools have been developed for molecular docking, including AutoDock [119], Autodock Vina [120], GOLD [121, 122], FlexX [123] and Glide [124]. Consequently, diverse sampling algorithms have been created and widely utilized in molecular docking software. For instance, matching algorithms (MA) [125-127] rely on molecular shape to map a ligand into a protein's active site based on shape features and chemical information. Both the protein and ligand are represented as pharmacophores, and the distances between the pharmacophore elements within the protein and ligand are computed for a suitable match. During this process, considerations can include chemical properties such as hydrogen-bond donors and acceptors. Matching algorithms have the advantage of speed and are available in DOCK programs [116]. In addition to MA, the Genetic algorithms (GA) [119, 121, 128] explore conformational space by randomly modifying a ligand's conformation or a population of ligands. The ligand's degrees of freedom are encoded as binary strings known as genes, which collectively form the "chromosome" representing the ligand's pose. Genetic operators, such as mutation and crossover, are employed in GA. Mutation introduces random changes to the genes, while crossover involves exchanging genes between two chromosomes. As a result of these genetic operations on the genes, a new ligand structure is generated. These newly created structures undergo evaluation via a scoring function, and those that meet certain criteria can progress to the next generation. Genetic algorithms have been used in AutoDock [119] and GOLD [122].

More than 40 chemical compounds have been identified from *Limonium Sinense*, most of which are flavonoids [31, 33]. However, there are few studies investigating the biological activities of these active ingredients and their associated mechanisms in *Limonium Sinense*. By applying *Lipinski*'s Rule of five and *Veber*'s rule, this study intended to screen the active compounds in *Limonium Sinense* and elucidate its potential underlying biological mechanisms through network pharmacology methods.

3.3 Aims

This study aims to predict the therapeutic effects of *Limonium Sinense* (Girard) Kuntze and explore the potential mechanisms through integrated network pharmacology methods. By leveraging diverse databases and computational tools, a comprehensive network will be constructed to understand the potential therapeutic effects of *Limonium Sinense*. Furthermore, the utilization of Gene Expression Omnibus (GEO) datasets, which encompass gene expression profiles from different experimental conditions, will be utilized to validate the network pharmacology results. This integrated approach will contribute to uncover the medicinal potential of *Limonium Sinense*'s and provide valuable insights for further research and development of this plant. The detailed technical strategy of this study is shown in Figure 3.1.

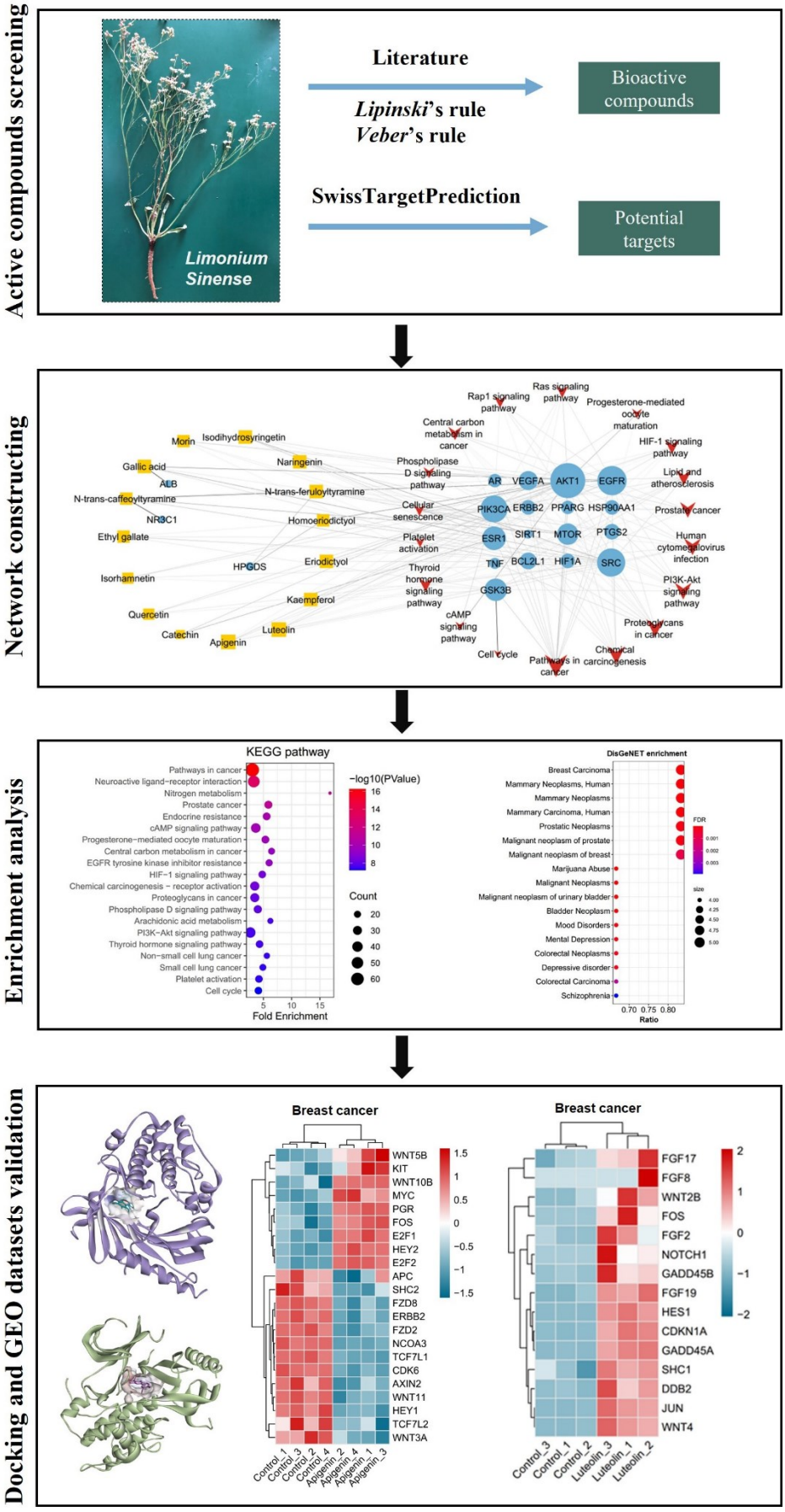


Figure 3.1. Flow chart of the network pharmacology study for *Limonium Sinense*.

3.4 Methods

3.4.1 Screening of active compounds and the targets of *Limonium Sinense*

Limonium Sinense is not included in the Traditional Chinese Medicine Systems Pharmacology Database and Analysis Platform (TCMSP) or the Encyclopedia of Traditional Chinese Medicine (ETCM) platform, which allow users to explore the relationships or build networks among drugs, targets and diseases [129, 130]. The compounds of *Limonium Sinense* were therefore obtained from the published literature. Briefly, we initiated our search on relevant reports on *Limonium Sinense* by querying "*Limonium Sinense*" in NCBI PubMed Database (https://www.ncbi.nlm.nih.gov) and its Chinese name "Zhong Hua Bu Xue Cao" in CNKI (www.cnki.net). This search yielded a total of 28 results in NCBI and 115 results in CNKI. These reports encompassed most of the available research on *Limonium Sinense* to date. After carefully screening the reported compounds in *Limonium Sinense*, 41 natural compounds were identified (Shown in Table 1.2). Active compounds were screened using the SwissADME database (http://swissadme.ch), based on the *Lipinski*'s rule of five [110] and *Veber*'s rule [115]. Then, the corresponding targets of the active compounds of *Limonium Sinense* were obtained from SwissTargetPrediction. The workflow for collecting *Limonium Sinense* compounds is illustrated in the Figure 3.2.

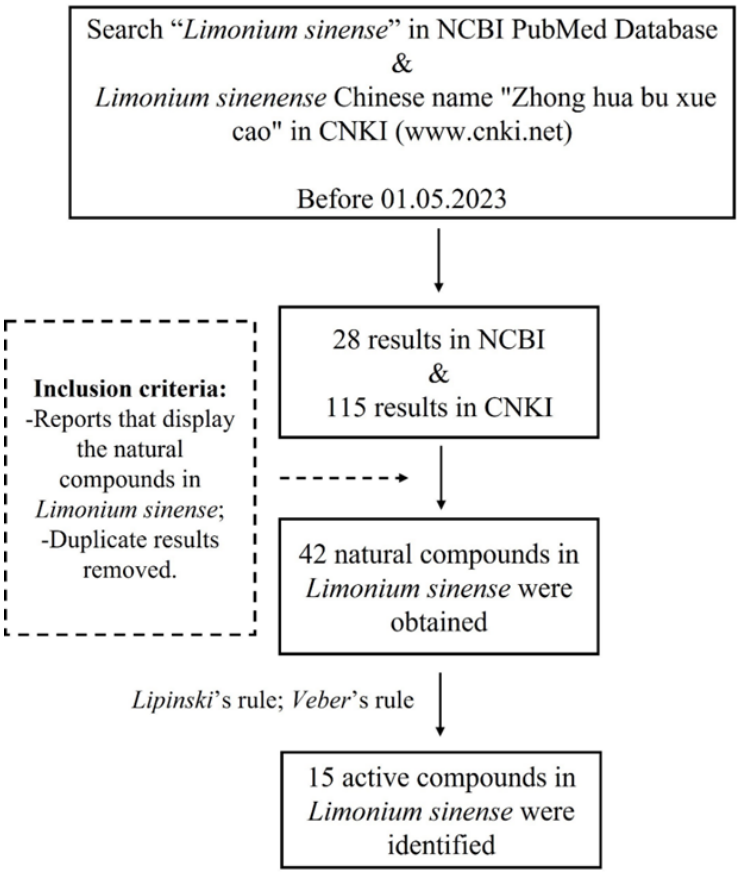


Figure 3.2. Workflow for compound screening of *Limonium Sinense*.

3.4.2 Construction of Protein-Protein interaction (PPI) network

The STRING (https://cn.string-db.org) database was utilized to analyse protein-protein interaction data. The species were limited to "Homo sapiens", and medium confidence value > 0.4 was selected as the minimum required interaction score to construct the PPI network. Then, the obtained PPI network from the STRING database was visualized by using Cytoscape *v3.9.1* software. The core nodes were selected based on the median values of three parameters in the interaction network by using NetworkAnlayzer plug-in in Cytoscape:

(a) "Degree" reflects how often one node interacts with other nodes [131]. The node degree of a node *n* is the number of edges linked to *n*. A self-loop of a node is counted like two edges for the node degree. The node degree distribution gives the number of nodes with degree *k* for *k* = *0*, *1*, …. This node degree distribution is to distinguish between random and scale-free network topologies [132].

(b) "Betweenness Centrality" which quantifies the extent to which a node lies on paths between other nodes [133]. The algorithm for the computation of node Betweenness Centrality has a complexity of *O* (*NM*), *N* being the number of nodes and *M* being the number of edges in the network [134]. The Betweenness Centrality $C_b(n)$ of a node *n* is computed as follows:

$$C_b(n) = \sum_{s \neq n \neq t} (\delta_{st}(n) / \delta_{st})$$

where *s* and *t* are nodes in the network different from *n*, δ_{st} denotes the number of shortest paths from *s* to *t*, and $\delta_{st}(n)$ is the number of shortest paths from *s* to *t* that *n* lies on. The betweenness value for each node *n* is normalized by dividing by the number of node pairs excluding *n*: (*N*-1)(*N*-2)/2, where *N* is the total number of nodes in the connected component that *n* belongs to. Thus, the Betweenness Centrality of each node is a number between 0 and 1.

(c) "Closeness Centrality" is a measure of how fast information spreads from a given node to other reachable nodes in the network [135, 136]. The Closeness Centrality $C_c(n)$ of a node *n* is defined as the reciprocal of the average shortest path length and is computed as follows:

$$C_c(n) = 1 / \text{avg}(L(n,m))$$

where L(*n*,*m*) is the length of the shortest path between two nodes *n* and *m*. The Closeness Centrality of each node is a number between 0 and 1. The Closeness Centrality of isolated nodes is equal to 0.

The level of the three parameters represents the topological importance of the nodes in the interaction network, with more important nodes outputting higher values in the network [137].

3.4.3 Enrichment analysis

GO term and KEGG enrichment analysis were performed through DAVID website tools. The GO project is a collaborative effort aimed at establishing a comprehensive controlled vocabulary to characterize gene product attributes across various organisms [138]. GO consists of three interconnected vocabularies (ontologies) that describe gene products based on their associated biological processes, cellular components, and molecular functions. One of the most common applications of the GO vocabulary is enrichment analysis, which involves pinpointing GO terms that exhibit significant overrepresentation within a specific gene set [139]. This enrichment aids in uncovering potential functional characteristics of the analysed set. For instance, observing enriched GO terms within a gene set significantly overexpressed in a particular condition might indicate potential regulatory mechanisms or activated functional pathways specific to that condition. KEGG is a knowledge base for systematic analysis of gene functions, linking genomic information with higher order functional information. KEGG consists of three databases: PATHWAY for representation of higher order functions in terms of the network of interacting molecules; GENES for the collection of gene catalogues for all the completely sequenced genomes and some partial genomes, and LIGAND for the collection of chemical compounds in the cell, enzyme molecules and enzymatic reactions.

The DAVID agglomeration method was developed to cluster genes or terms into functional groups (biological modules) based on similarity distance measures, while gene-gene functional relationships were identified through *kappa* statistics [140]. Initially, a gene-term annotation matrix was constructed in binary mode, using numerous annotation terms across 14 categories within the DAVID knowledgebase, including GO classifications such as Biological Process, Molecular Function, and Cellular Component, along with KEGG Pathways, BioCarta Pathways, Swiss-Prot Keywords, BBID Pathways, SMART Domains, NIH Genetic Association DB, UniProt Sequence Features, COG/KOG Ontology, NCBI OMIM, InterPro Domains, and PIR Super Family Names. *Kappa* statistics, a measure correcting for chance agreement between two sets of categorized data, were employed to assess the statistical co-occurrence of annotations for any given gene pairs. Given the annotation profiles of genes *m* and *n*, O_{mn} represents the observed cooccurrence, A_{mn} represents chance cooccurrence, and K_{mn} is the kappa value representing the degree of annotation cooccurrence between genes *m* and *n*.

$$K_{mn} = \frac{O_{mn} - A_{mn}}{1 - A_{mn}}$$

Where K_{mn} is 1 for perfect cooccurrence and 0 for cooccurrence no better than random chance.

P values adjusted by BH method were used to estimate the statistical significance, $P \leq 0.05$ was defined as significant for KEGG and GO enrichment.

Diseases associated with the hub targets were enriched by using the disgenet2r package (*version 0.99.3*) in R [141], the specific R codes for performing the enrichment analysis of GO, KEGG and disgenet2r are available in the Appendix A section.

3.4.4 Network construction

In this process, network construction was performed as follows: First, the Compound-Target network (CT network) was built based on the active compounds of *Limonium Sinense* and their potential targets. Next, the Pathway-Target network (PT network) was constructed by selecting the top 20 enriched KEGG pathways and their associated targets. Finally, a Compound-Target-Pathway network (CTP network) was established by integrating the CT network, PT network and the core targets identified from the PPI network. All visualized network graphs were created using Cytoscape *v3.9.1* software. The hub network was screened by using Cytohubba [142], a degree algorithm based Cytoscape plug-in.

3.4.5 Molecular docking

The crystal structure of *AKT1* (PDB ID: 3O96), *EGFR* (PDB ID: 1XKK), *SRC* (PDB ID: 3EL8), *ESR1* (PDB ID: 5ACC) and *GSK3B* (PDB ID: 3I4B) were downloaded from Protein Data Bank (https://www.rcsb.org/). Compound structure of Apigenin and Luteolin were downloaded from PubChem (https://pubchem.ncbi.nlm.nih.gov/). The docking studies were performed by PyRx Autodock VINA tool (*v0.8*). The scoring function formula of AutoDock is as follows:

$$V = W_{vdw}\sum_{i,j}\left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right) + W_{hbond}\sum_{i,j}E(t)\left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}}\right) + W_{elec}\sum_{i,j}\frac{q_i q_j}{\epsilon(r_{ij})r_{ij}} + W_{sol}\sum_{i,j}\left(S_i V_j + S_j V_i\right)e^{\left(-r_{ij}^2/2\sigma^2\right)}$$

For two atoms *i*, *j*, the pair-wise atomic energy is evaluated by the sum of van der Waals, hydrogen bond, coulomb energy and desolvation. *W* are weighted factors for calibrate the empirical free energy. A grid box was set to cover the active site of crystal structure with a default exhaustiveness of 8. The best compound with highest binding affinity (kcal/mol) was selected and visualized by using Discovery Studio (*version 2021 Client*) and PyMOL.

3.4.6 GEO datasets analysis

GEO datasets on human cancer cells treated with Apigenin or Luteolin were screened, with publication dates prior to 01/05/2023 in the National Centre for Biotechnology Information (NCBI)

GEO platform. The obtained datasets were further screened with the following criteria: **1)** mRNA expression data; **2)** Homo sapiens samples; **3)** Human cancer cells; **4)** minimum 3 biological replicates. Duplicate datasets were removed. Datasets with fewer than 10,000 genes were excluded to balance the number of analysed genes and sample size. Microarray probe IDs were translated to gene symbols according to the GPL annotation files provided in the GEO database. Probes mapped to multiple gene symbols were removed and genes mapped to multiple probe IDs were summarized by calculating the mean. (R codes for obtaining the raw data of GEO datasets are shown in Appendix A). Genes with *P* value less than 0.05 and |Log2FoldChange| above 1 were considered as DEGs. DEGs of each GEO datasets were obtained by GEO2R in the GEO platform.

3.5 Results

3.5.1 Identification of active compounds and targets of *Limonium Sinense*

As described in Chapter 1, based on the available published literature, a total of 41 natural compounds have been identified from *Limonium Sinense* (Table 1.2). Among these compounds, 15 have been found to meet the criteria of *Lipinski*'s rule of five and *Veber*'s rule, making them the active compounds of *Limonium Sinense*. These active compounds include Gallic acid, Ethyl gallate, Apigenin, Naringenin, Luteolin, Kaempferol, Eriodictyol, (+)-Catechin, N-trans-caffeoyltyramine, Quercetin, Morin, Homoeriodictyol, N-trans-feruloyltyramine, Isorhamnetin, Isodihydrosyringetin (Table 3.1). It is worth noting that the majority of these active compounds belong to the flavonoid group. Additionally, 389 targets associated with these 15 active compounds were obtained after removing duplicate values (Appendix B, Supplementary Table 3.1). The Compound-Target (CT) network revealed that all 15 bioactive compounds exhibited relatively high values of Degree, Betweenness Centrality (BC) and Closeness Centrality (CC) within the network (Figure 3.3; Appendix B, Supplementary Table 3.2), showing significant connectivity and potential importance of these compounds in the overall network.

Table 3.1. Detailed information on active compounds in *Limonium Sinense*

Compound Name	Molecular weight	nRB	nHA	nHD	TPSA	LogP
Gallic acid	170.12	1	5	4	97.99	-0.16
Ethyl gallate	198.17	3	5	3	86.99	0.49
Apigenin	270.24	1	5	3	90	0.52
Naringenin	272.25	1	5	3	86.99	0.71
Luteolin	286.24	1	6	4	111.13	-0.03
Kaempferol	286.24	1	6	4	111.13	-0.03
Eriodictyol	288.25	1	6	4	107.22	0.16
(+)-Catechin	290.27	1	6	5	110.38	0.24
N-trans-caffeoyltyramine	299.32	6	4	4	89.79	1.65
Quercetin	302.24	1	7	5	131.36	-0.56
Morin	302.24	1	7	5	131.36	-0.56
Homoeriodictyol	302.28	2	6	3	96.22	0.41
N-trans-feruloyltyramine	313.35	7	4	3	78.79	1.89
Isorhamnetin	316.26	2	7	4	120.36	-0.31
Isodihydrosyringetin	348.3	3	8	4	125.68	-0.66

nRB: Number of Rotatable bonds, optimal: 0-10.

nHA: Number of Hydrogen bond acceptors, optimal: 0-10.

nHD: Number of Hydrogen bond donors, optimal: 0-5.

TPSA: Topological Polar Surface Area, optimal: 0-140.

LogP: Log of the octanol/water partition coefficient, optimal: ≤ 5.

***Lipinski* Rule**: nHD ≤ 5, nHA ≤ 10, MW ≤ 500, LogP ≤ 5.

***Veber* Rule**: nRB ≤ 10, TPSA ≤ 140 $Å^2$.

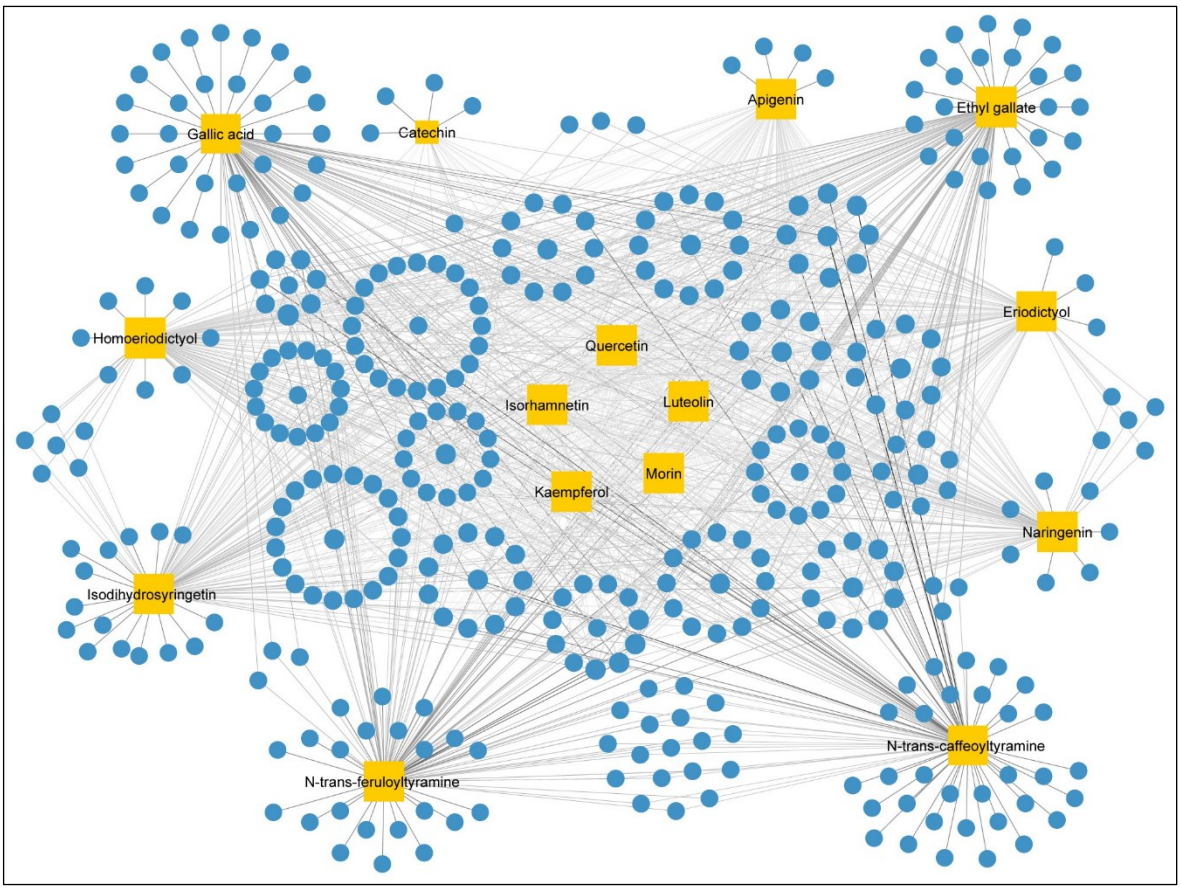


Figure 3.3. Compound-Target (CT) network of *Limonium Sinense*. Graph consists of 15 compounds and 389 compound-related targets. The orange rectangles represent the small molecular components in *Limonium Sinense*. The blue circles represent the relevant targets. The edges represent the relationship between compounds and target nodes. The node size is proportional to the node degree in the network, and the width and colour of the edges is proportional to the edge betweenness centrality.

3.5.2 Construction of PPI network and core targets of *Limonium Sinense*

A PPI network consisting of 389 nodes and 4,947 edges was established (Figure 3.4A). Based on the threshold values of Degree, BC, and CC, with the first screening using Degree ≥ 26, BC ≥ 0.00363 and CC ≥ 0.4228, the PPI network was pruned to 73 nodes and 1,049 edges (Figure 3.4B). Subsequently, nodes with Degree ≥ 29, BC ≥ 0.0085 and CC ≥ 0.6324 were screened as the second screening threshold values and resulted in 20 core nodes and 178 edges (Figure 3.4C). These 20 targets, namely *AKT1*, *SRC*, *ALB*, *EGFR*, *HSP90AA*, *ESR1*, *VEGFA*, *TNF*, *MTOR*, *HIF1A*, *ERBB2*, *AR*, *SIRT1*, *BCL2L1*, *PPARG*, *GSK3B*, *PTGS2*, *NR3C1*, *PIK3CA* and *HPGDS*, were defined as the core targets of *Limonium Sinense* (Table 3.2), which likely to contribute to its pharmacological activities.

Table 3.2. Details of the core targets identified from PPI network

Target name	BetweennessCentrality	ClosenessCentrality	Degree
VEGFA	0.00452684	1	19
BCL2L1	0.00452684	1	19
SRC	0.00452684	1	19
HIF1A	0.00452684	1	19
PTGS2	0.00452684	1	19
HSP90AA1	0.00452684	1	19
AKT1	0.00452684	1	19
ERBB2	0.003765132	0.95	18
ALB	0.003378135	0.95	18
AR	0.003378135	0.95	18
EGFR	0.003765132	0.95	18
ESR1	0.003378135	0.95	18
TNF	0.003378135	0.95	18
NR3C1	0.002567694	0.904761905	17
SIRT1	0.002567694	0.904761905	17
MTOR	0.002567694	0.904761905	17
GSK3B	0.002567694	0.904761905	17
PPARG	0.002647141	0.863636364	16
PIK3CA	0.002647141	0.863636364	16
HPGDS	0.001879699	0.826086957	15

BetweennessCentrality: Average length of a shortest path between n and any other node. **ClosenessCentrality**: The mean distance from a node to other nodes. **Degree**: The number of links to one node. The data list is organized here from maximal to minimal Degree.

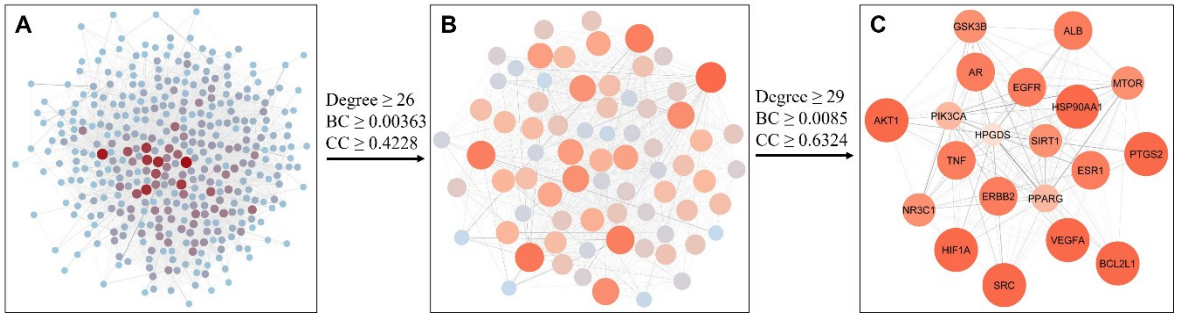


Figure 3.4. The process of topological screening for the PPI network. Each node represents a protein and each edge refers an interaction. The size and colour of the nodes represent the node degree in the network, and the width and colour of the edges represent the edge betweenness centrality. The core targets were screened based on the Degree, betweenness centrality (BC) and closeness centrality (CC) values in the network.

3.5.3 Enrichment analysis of compound targets

The enrichment analysis was performed with DAVID online tool (*v2023q1*) on the 389 compound targets, with a screening threshold of a *P* value less than 0.05. The results retrieved 152 KEGG pathways and 1,018 GO terms. The GO terms were further grouped into Biological Process (GO_BP, 691 items), Cellular Component (GO_CC, 109 items) and Molecular Function (GO_MF, 218 items) (Top 20 enriched KEGG pathways and top 5 enriched GO terms of each category were displayed in Table 3.3).

Among the top 20 ranked enrichment results, several cancer-related pathways were significantly identified in the KEGG pathways, including Pathways in cancer, Prostate cancer, Central carbon metabolism in cancer, Proteoglycans in cancer, Non-small cell lung cancer and Small cell lung cancer (Figure 3.5A). In addition, signalling pathways including cAMP signalling pathway, HIF-1 signalling pathway, Phospholipase D signalling pathway, PI3K-AKT signalling pathway and Thyroid hormone signalling pathway were also significantly identified.

The GO term analysis predicted that the targets of *Limonium Sinense* are mainly related with peptidyl-tyrosine phosphorylation in Biological Process, plasma membrane in Cellular Component, and protein serine/threonine/tyrosine kinase activity in Molecular Function (Figure 3.5B). These results indicating the involvement of various cancer-related pathways in the biological effects of *Limonium Sinense*.

Table 3.3. Details of the KEGG pathways (Top 20) and GO terms (Top 5 of each category)

Category	Pathway Name	Count	*P*value	Fold Enrichment	FDR
KEGG	Pathways in cancer	68	5.95E-17	3.04560714	9.34E-15
KEGG	Neuroactive ligand-receptor interaction	51	5.72E-14	3.304940173	4.49E-12
KEGG	Nitrogen metabolism	12	5.94E-12	16.78772379	2.60E-10
KEGG	Prostate cancer	24	6.63E-12	5.884356791	2.60E-10
KEGG	Endocrine resistance	23	6.15E-11	5.581632653	1.93E-09
KEGG	cAMP signalling pathway	34	1.32E-10	3.658862876	3.18E-09
KEGG	Progesterone-mediated oocyte maturation	23	1.42E-10	5.362745098	3.18E-09
KEGG	Central carbon metabolism in cancer	19	3.26E-10	6.455279503	6.34E-09
KEGG	EGFR tyrosine kinase inhibitor resistance	20	3.64E-10	6.020913594	6.34E-09
KEGG	HIF-1 signalling pathway	22	3.36E-09	4.800159553	5.05E-08
KEGG	Chemical carcinogenesis - receptor activation	31	3.54E-09	3.477645611	5.05E-08
KEGG	Proteoglycans in cancer	30	6.59E-09	3.48038176	8.62E-08
KEGG	Phospholipase D signalling pathway	25	9.28E-09	4.01733255	1.12E-07
KEGG	Arachidonic acid metabolism	16	1.95E-08	6.238061297	2.19E-07
KEGG	PI3K-AKT signalling pathway	40	2.27E-08	2.687300418	2.31E-07
KEGG	Thyroid hormone signalling pathway	22	2.36E-08	4.324110672	2.31E-07
KEGG	Non-small cell lung cancer	17	3.11E-08	5.615338164	2.87E-07
KEGG	Small cell lung cancer	19	3.54E-08	4.911625709	3.04E-07
KEGG	Platelet activation	22	3.68E-08	4.219495091	3.04E-07
KEGG	Cell cycle	22	5.67E-08	4.119821979	4.45E-07
GO_BP	peptidyl-tyrosine phosphorylation	45	1.28E-40	16.02724981	4.30E-37
GO_BP	protein phosphorylation	71	2.12E-39	7.188044917	3.54E-36
GO_BP	protein autophosphorylation	48	4.20E-39	13.27545757	4.69E-36
GO_BP	transmembrane receptor protein tyrosine kinase signalling pathway	36	2.35E-29	13.40022691	1.97E-26
GO_BP	peptidyl-serine phosphorylation	38	1.61E-26	10.28001621	1.05E-23
GO_CC	plasma membrane	193	1.96E-23	1.900252424	4.97E-21
GO_CC	integral component of plasma membrane	89	2.44E-23	3.250409795	4.97E-21
GO_CC	receptor complex	32	2.17E-18	7.67384214	2.95E-16
GO_CC	cytosol	183	1.65E-17	1.740765777	1.68E-15
GO_CC	cytoplasm	185	2.53E-17	1.724494777	2.07E-15
GO_MF	protein serine/threonine/tyrosine kinase activity	89	1.07E-61	9.759143089	8.88E-59
GO_MF	protein tyrosine kinase activity	42	1.86E-40	18.09279337	7.74E-38
GO_MF	protein kinase activity	61	1.01E-35	7.785963989	2.80E-33
GO_MF	ATP binding	114	2.23E-34	3.57388511	4.63E-32
GO_MF	transmembrane receptor protein tyrosine kinase activity	24	3.25E-25	21.84790142	5.40E-23

"**Category**" is the term in the annotation cluster; "**Gene Count**" is the number of DEGs belonging to an annotation term; "***P*value**" is the modified Fisher Exact *P*-value, referring to one-tail Fisher Exact probability value used for gene-enrichment analysis. The smaller, the more enriched; "**Fold Enrichment**" is the ratio of the input genes involved in specific annotation term proportion on background associated genes in the same term proportion; "**FDR**" in DAVID requests adaptive linear step-up adjusted *P*-values for approximate control of the false discovery rate, as discussed in Benjamini and Hochberg (2000). Use the lowest slope method to estimate the number of true NULL hypotheses.

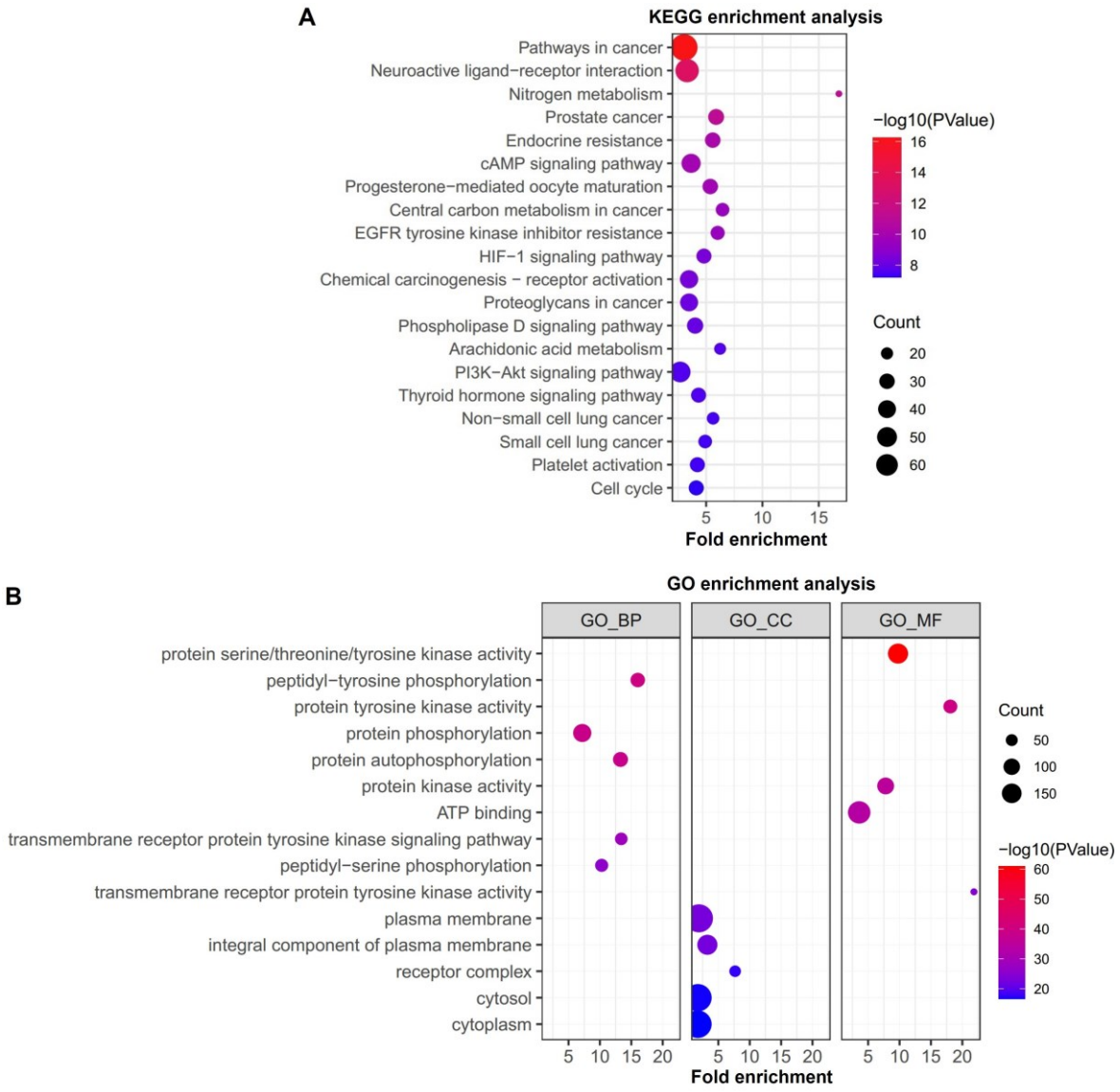


Figure 3.5. Enrichment analysis from *Limonium Sinense* targets. (**A** and **B**) Scatter plots showing the Kyoko Encyclopedia of Genes and Genomes (KEGG) pathways (Top 20) and Gene Ontology (GO) terms (Top 5 of each category) enriched by the *Limonium Sinense* potential targets. The GO term was further categorised into Biological Process (GO_BP), Cellular Component (GO_CC) and Molecular Function (GO_MF). The sizes of circles represent the percentage of genes in the gene set, and the colours of circles represent the -log10(*P*value). The Fold Enrichment is defined as the ratio of the two proportions.

3.5.4 CTP network construction and hub connections identification

The top 10 KEGG pathways and their associated targets were selected to construct the TP network (Figure 3.6). It was shown that the TP network contains 160 nodes and 296 edges. Analysis of the network revealed that all the selected 10 pathways exhibited higher degree, CC, and BC values (Appendix B, Supplementary Table 3.3). Notably, the top-ranked pathway was "Pathways in cancer", suggesting a potential role of *Limonium Sinense* in regulating cancer. The "Pathways in cancer" map in KEGG encompasses a collection of interconnected molecular pathways involved in the development and progression of various types of cancer. These pathways encompass crucial cellular processes such as cell growth, cell proliferation, cell survival, cell apoptosis, angiogenesis, and metastasis. The map integrates information on diverse signalling pathways, including the MAPK signalling pathway, PI3K-AKT signalling pathway, Wnt signalling pathway, Notch signalling pathway, TGF-β signalling pathway, HIF-1 signalling and cell cycle, as well as many other signalling pathways. This finding highlights the significance of *Limonium Sinense* in influencing cancer-related processes within the pathways associated with tumorigenesis and progression.

Subsequently, a CTP network was built by incorporating the 15 active compounds, 20 core targets and the top 10 KEGG pathways. The resulting CTP network consisted of 43 nodes and 130 edges (Figure 3.7A). From the CTP network, a hub network was identified, comprising 15 nodes and 38 edges (Figure 3.7B, C). Within the hub network, five hub targets, namely *AKT1*, *EGFR*, *SRC*, *ESR1* and *GSK3B*, were found to be directly regulated by Apigenin and Luteolin, two of the active compounds of *Limonium Sinense*. Furthermore, four signalling pathways were found to be involved in the hub network, including Pathways in cancer, EGFR tyrosine kinase inhibitor resistance, Endocrine resistance and HIF-1 signalling pathway. These pathways play crucial roles in cancer development, drug resistance, and cellular adaptation to hypoxia, suggesting that *Limonium Sinense* may through its active compounds to exert the therapeutic effects by modulating these specific pathways and targeting key regulatory proteins.

To investigate the diseases associated with the potential targets of *Limonium Sinense*, an analysis was conducted using the DisGeNET database (https://www.disgenet.org/ accessed on 15 May 2023) (*version 7.0*), employing the disgenet2r package (*version 0.99.3*) in R. The results uncovered a significant correlation with breast carcinoma (Figure 3.8A). Interestingly, the disease enrichment analysis of the hub targets also revealed a strong association with breast carcinoma, as well as mammary neoplasms, and mammary carcinoma (Figure 3.8B). These findings suggest a potential therapeutic role of *Limonium Sinense* in the treatment of breast cancer.

To further investigate the potential relevance of these hub targets in cancer, we examined their mRNA expression in the TCGA cohort using UCSCXenaShiny (*v1.1.9*) (A comprehensive description

of UCSCXenaShiny methodology can be found in reference [143]). The results revealed significant up- or down-regulation of these hub targets in most of the 33 TCGA cancer types at the mRNA level (Figure 3.9A-E). In addition, these target genes were found to predict overall survival in multiple cancer types (Figure 3.9F). These findings suggest that Apigenin and Luteolin in *Limonium Sinense* may play a crucial role in regulating cancer process and disease-related signalling pathways.

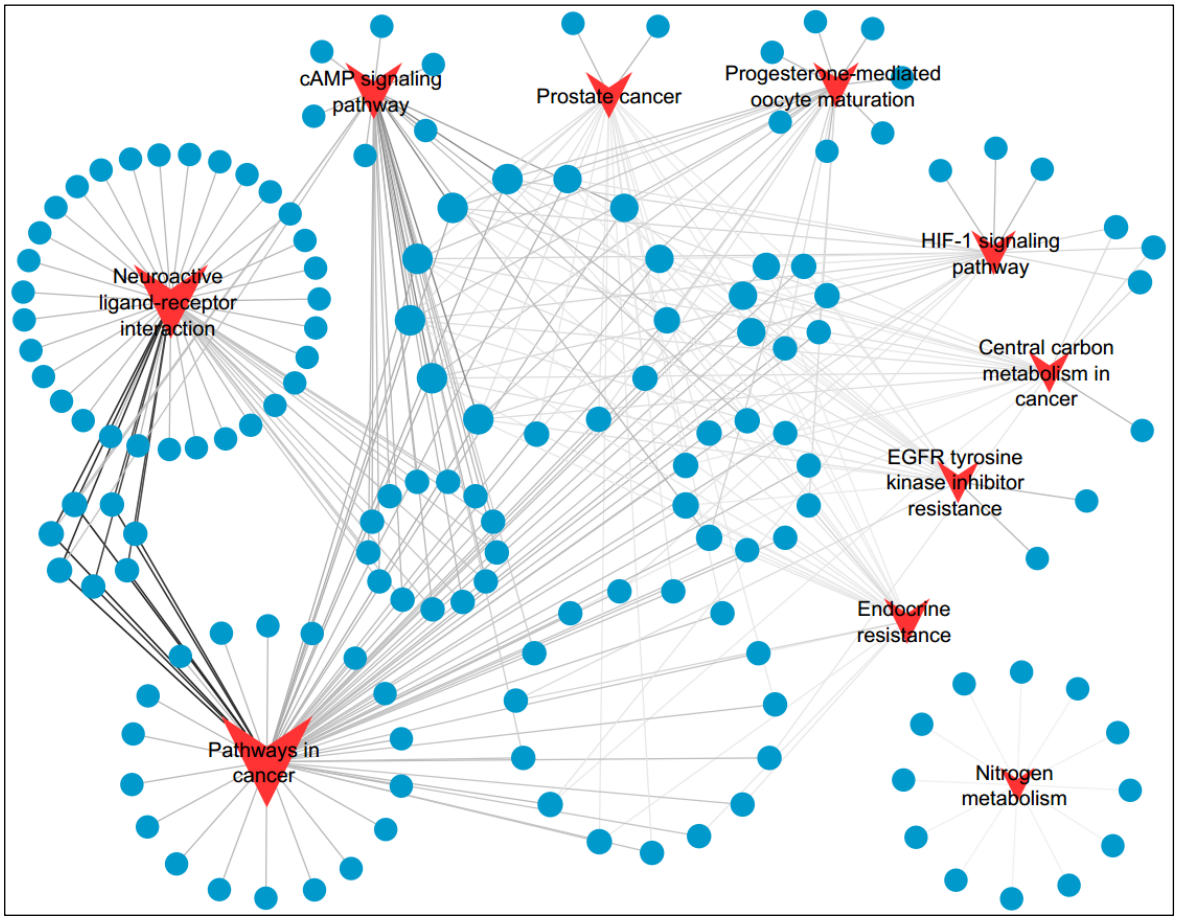
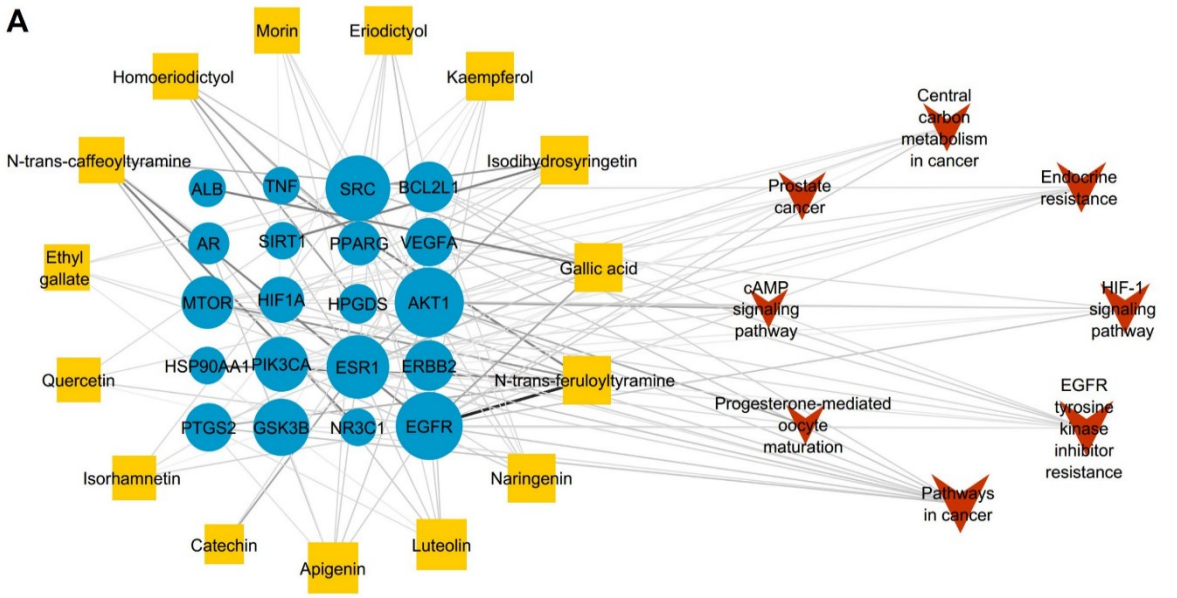
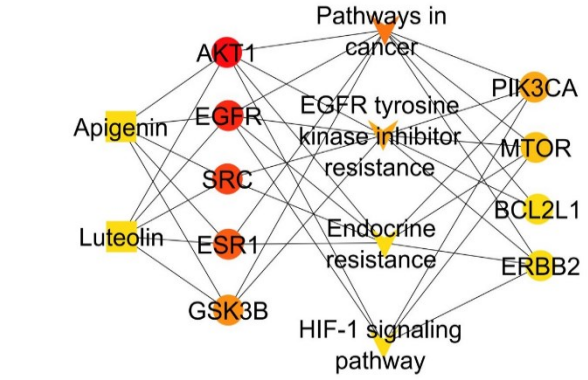


Figure 3.6. Target-Pathway (TP) network of *Limonium Sinense*. Graph consists of 10 signalling pathways and 150 pathway-related targets. The red V nodes represent the top 10 enriched KEGG pathways by the potential targets of *Limonium Sinense*. The blue circles represent the relevant targets. The edges represent the relationship between signalling pathways and target nodes. The node size is proportional to the node degree in the network, and the width and colour of the edges is proportional to the edge betweenness centrality.





C

Rank	Name	Score
1	AKT1	16
2	EGFR	15
3	SRC	14
4	ESR1	13
5	Pathways in cancer	11
6	GSK3B	10
7	EGFR tyrosine kinase inhibitor resistance	9
7	PIK3CA	9
9	MTOR	8
10	BCL2L1	7
10	Endocrine resistance	7
10	ERBB2	7
10	HIF-1 signaling pathway	7
10	Apigenin	7
10	Luteolin	7

Figure 3.7. Compound-Target-Pathway (CTP) and Hub network analysis. (**A**) Compound-Target-Pathway (CTP) network of *Limonium Sinense*. The orange rectangles represent the small molecular components in *Limonium Sinense*. The blue circles represent the relevant targets. The red V nodes represent the relevant signalling pathways. The edges represent the relationship among compounds, targets, and pathways. The node size is proportional to the node degree in the network, and the width and colour of the edges is proportional to the edge betweenness centrality. (**B**) Hub network identified from the CTP network using the CytoHubba. Colour node indicates the degree value of each node in the network. (**C**) The 15 nodes and their scores in the hub network. The colour intensity was decreased based on the score of each node.

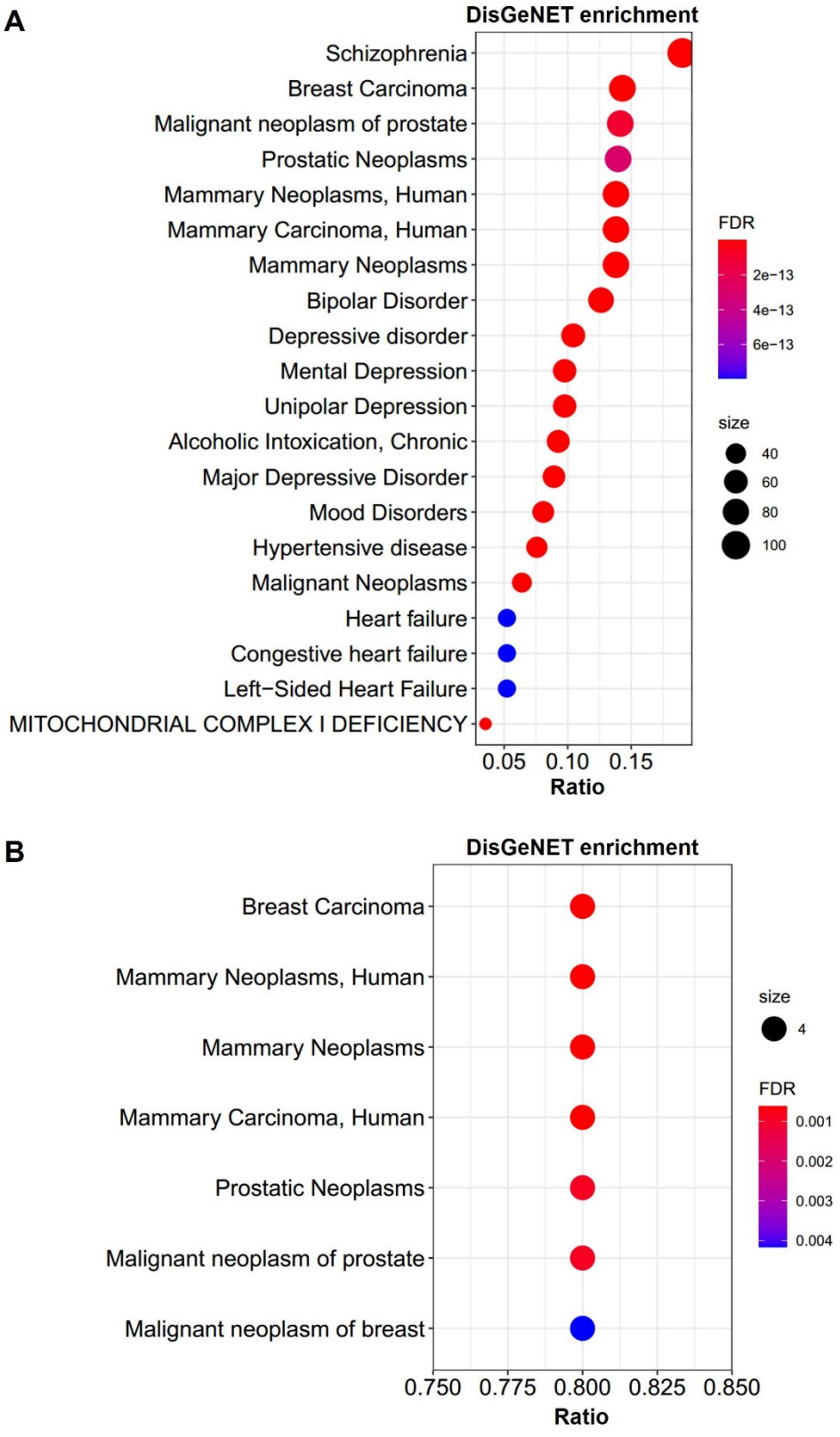
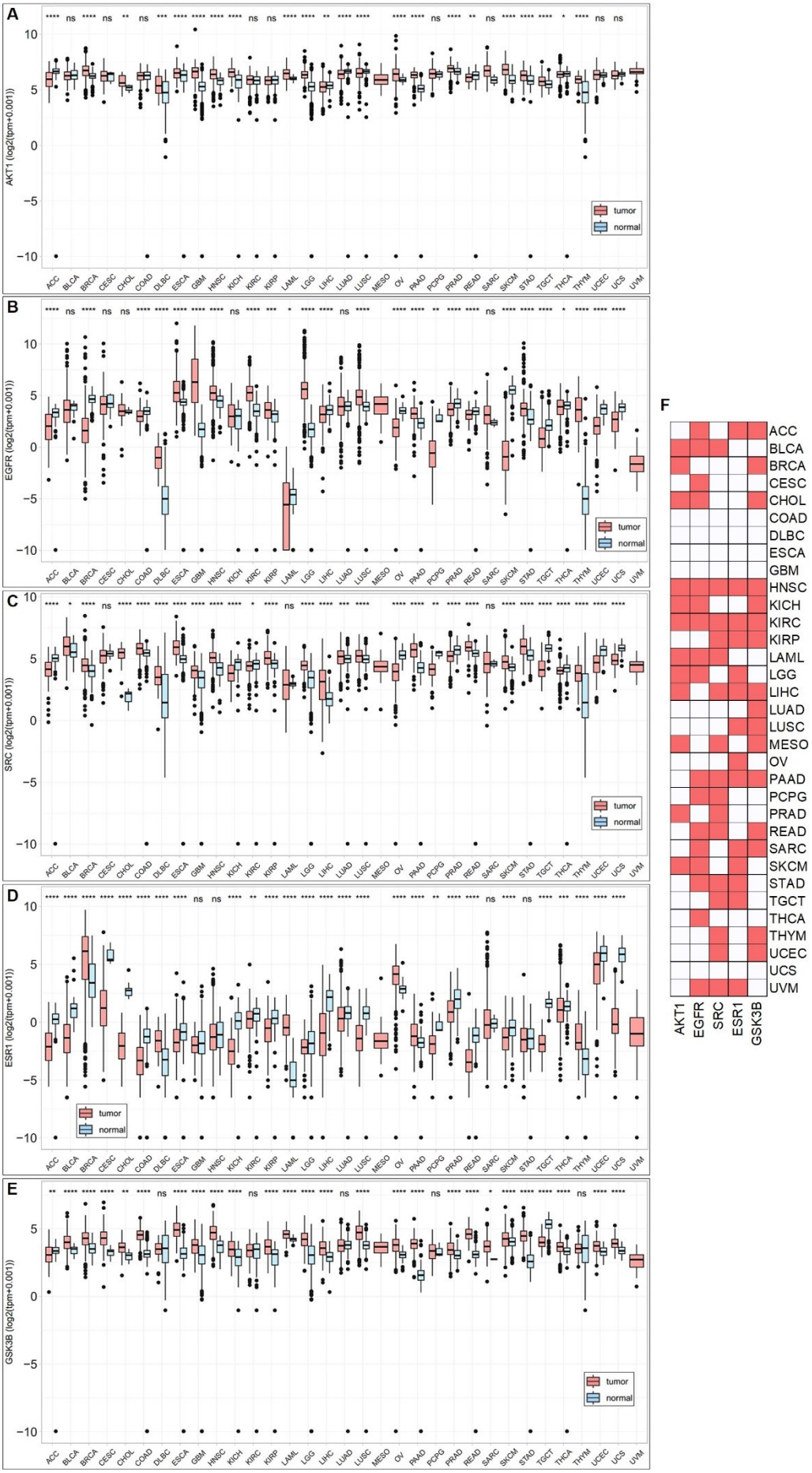


Figure 3.8. DisGeNET enrichment analysis. Graphs showing the disease enrichment results of the potential targets (**A**) and hub targets (**B**) of *Limonium Sinense*. The size of the dots refers to the number of genes enriched for that disease (the greater the number of genes, the larger the dot). The FDR is defined by a colour scale, the closer it is to red, the greater the FDR value and the greater the association.



(Figure legend is provided on the following page)

Figure 3.9. Relative mRNA expression and Overall Survival of the 5 hub target genes in TCGA cancers. (**A**-**E**) Graphs showing the relative mRNA expression of each target gene obtained from the hub network in 33 TCGA cancers. ns, not significant, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. (**F**) Heatmap showing the Overall Survival of each target gene obtained from the hub network in 33 TCGA cancers. Red colour represents significant correlated with overall survival, while white colour represents not significant correlated with overall survival.

3.5.5 Molecular docking between Apigenin and hub targets

Notably, the hub network and DisGeNET analysis provided insights into the top enriched diseases associated with the hub targets, suggesting the potential role of Apigenin and Luteolin in cancer treatment. To further validate the interactions between Apigenin/Luteolin and the five hub target proteins, molecular docking was employed.

The results of molecular docking demonstrated favourable binding affinities of Apigenin and Luteolin with the hub targets. The docking affinities of Apigenin and Luteolin with the five hub targets were all less than -5 kcal/mol (Table 3.4), indicating a strong interaction between the ligands and target proteins. Among the interactions of Apigenin with the hub targets, AKT1 exhibited the highest binding activity with Apigenin (affinity = -9.6), followed by ESR1(affinity = -8.9), PTGS2 (affinity = -8.8), SRC (affinity = -8.6), EGFR (affinity = -8.5), and GSK3B (affinity = -8.4). Similarly, AKT1 showed the strongest binding activity with Luteolin (affinity = -9.8), followed by ESR1(affinity = -9), EGFR (affinity = -8.9), GSK3B (affinity = -8.5) and SRC (affinity = -8.5). These results indicate the potential of Apigenin and Luteolin to interact with and modulate the activity of the hub targets in the network. Interaction residues and interaction bonds between Apigenin/Luteolin and the hub targets were also displayed in Table 3.4.

Table 3.4. Binding affinities and interactions of Apigenin/Luteolin with hub targets

Compound	Protein	Binding affinity (kcal/mol)	Interaction residues	Interaction bonds
Apigenin	AKT1	-9.6	Ser205, Asp292, Trp80, Leu210, Leu264, Val270, Lys268	Conventional Hydrogen Bond, Pi-Anion, Pi-Pi Stacked, Pi-Alkyl
	EGFR	-8.5	Lys745, Met766, Thr854, Leu777, Val726, Ala743, Cys775, Leu844	Conventional Hydrogen Bond, Unfavourable Donor-Donor, Pi-Lone Pair, Pi-Alkyl
	SRC	-8.6	Asp404, Gly406, Met314, Lys295, Val323, Ala403	Conventional Hydrogen Bond, Carbon Hydrogen Bond, Unfavourable Acceptor-Acceptor, Pi-Anion, Pi-Lone Pair, Pi-Alkyl
	ESR1	-8.9	Glu353, Arg394, His524, Phe404, Leu346, Leu349, Ala350, Leu387, Leu391, Met421, Ile424	Conventional Hydrogen Bond, Carbon Hydrogen Bond, Pi-Pi Stacked, Pi-Alkyl
	GSK3B	-8.4	Cys199, Tyr134, Ile62, Val70, Ala83, Leu188	Pi-Sulfur, Pi-Pi Stacked, Pi-Alkyl
Luteolin	AKT1	-9.8	Asn54, Gln79, Val271, Thr211, Val270, LEU264, Ser205, Trp80, Leu210, Asp292	Conventional Hydrogen Bond, Unfavourable Donor-Donor, Unfavourable Acceptor-Acceptor, Pi-Anion, Pi-Sigma, Pi-Pi Stacked, Pi-Alkyl
	EGFR	-8.9	Met793, Asp855, Lys745, Val726, Leu844, Ala743, Leu718	Conventional Hydrogen Bond, Pi-Anion, Pi-Sigma, Pi-Alkyl
	SRC	-8.5	Met341, Asp404, Lys295, Thr338, Val281, Ala293, Leu393, Ala403	Conventional Hydrogen Bond, Pi-Anion, Pi-Donor Hydrogen Bond, Pi-Alkyl
	ESR1	-9	Leu387, Arg394, Thr347, Glu353, Phe404, Leu346, Leu391, Leu525, Ala350	Conventional Hydrogen Bond, Unfavourable Acceptor-Acceptor, Pi-Sigma, Pi-Pi T-shaped, Pi-Alkyl
	GSK3B	-8.6	Tyr134, Val135, Arg141, Ile62, Leu188, Val70, Cys199, Ala83	Conventional Hydrogen Bond, Unfavourable Donor-Donor, Pi-Sigma, Pi-Sulfur, Pi-Alkyl

The binding structures showed that Apigenin and Luteolin are deeply entered the binding site of each hub target. In AKT1, Apigenin bound to the Ser205 site via a conventional hydrogen bond, while it formed pi-anion, pi-pi stacked and pi-alkyl bonds with the Asp292, Trp80, Leu210, Leu264, Lys268 and Val270 sites (Figure 3.10A). Apigenin bound to the Lys745 and Met766 sites in EGFR via conventional hydrogen bonds, and it also formed a pi-lone pair bond with the Thr854 site and an unfavourable donor-donor bond with the Leu777 site. Furthermore, it established pi-alkyl bonds with the Val726, Ala743, Cys775, Leu844 sites (Figure 3.10B). In SRC, Apigenin bound to the Asp404 via conventional hydrogen bonds, as well as an unfavourable acceptor-acceptor and pi-anion bonds. It formed a carbon hydrogen bond with the Gly406 site and a via pi-lone pair bond with the Met314 site. Additionally, it established pi-alkyl bonds with the Lys295, Val323, Ala403 sites (Figure 3.10C). In ESR1, Apigenin bound to the Arg394 and Glu353 sites via conventional hydrogen bonds, and it formed a carbon hydrogen bond with the His524 site and a pi-pi stacked bond with the Phe404 site. Moreover, it established pi-alkyl bonds with the Leu346, Leu349, Ala350, Leu387, Leu391, Met421 and Ile424 sites (Figure 3.10D). In GSK3B, Apigenin bound to the Cys199 site through a conventional hydrogen bond. It also formed a pi-pi stacked bond with the Tyr134 site and established pi-alkyl bonds with the Ile62, Ala83, Val70, Leu188 and Met421 sites (Figure 3.10E).

On the other hand, Luteolin bound to the Gln79, Asn54, Val271 and Thr211 sites in AKT1 via conventional hydrogen bonds. It also formed pi-alkyl bonds with the Leu210 and Val270 sites, and established pi-pi stacked, pi-anion, pi-sigma and unfavourable donor-donor/unfavourable acceptor-acceptor bonds with the Trp80, Leu264, Asp292 and Ser205 sites, respectively (Figure 3.11A). In EGFR, Luteolin bound to the Asp855 and Met793 sites via conventional hydrogen bonds. It formed a pi-cation bond with the Lys745 site and pi-alkyl bonds with the Ala743 and Leu718 sites. Additionally, it established pi-alkyl and pi-sigma bonds with the Val726 and Leu844 sites (Figure 3.11B). Luteolin bound to the Asp404 and Met341 sites in SRC via conventional hydrogen bonds, and it formed a pi-cation bond with Lys295 site, a pi-donor hydrogen bond with Thr338 site. Furthermore, it established pi-alkyl bonds with the Val281, Leu393, Ala293 and Ala403 sites in SRC (Figure 3.11C). In ESR1, Luteolin bound to the Arg394, Leu387 and Thr347 sites via conventional hydrogen bonds, and formed an unfavourable acceptor-acceptor bond with the Glu353 site, and a pi-pi T-shaped bond with the Phe404 site. Moreover, it established pi-alkyl bonds with the Leu346, Leu391, Ala350 and Leu525 sites (Figure 3.11D). In GSK3B, Luteolin bound to the Tyr134 and Val135 sites via conventional hydrogen bonds. It formed a pi-sulfur bond with the Cys199 site and established pi-sigma bonds with the Ile62, Val70 and Leu188 sites. In addition, it bound to Ala83 site through pi-alkyl bond (Figure 3.11E).

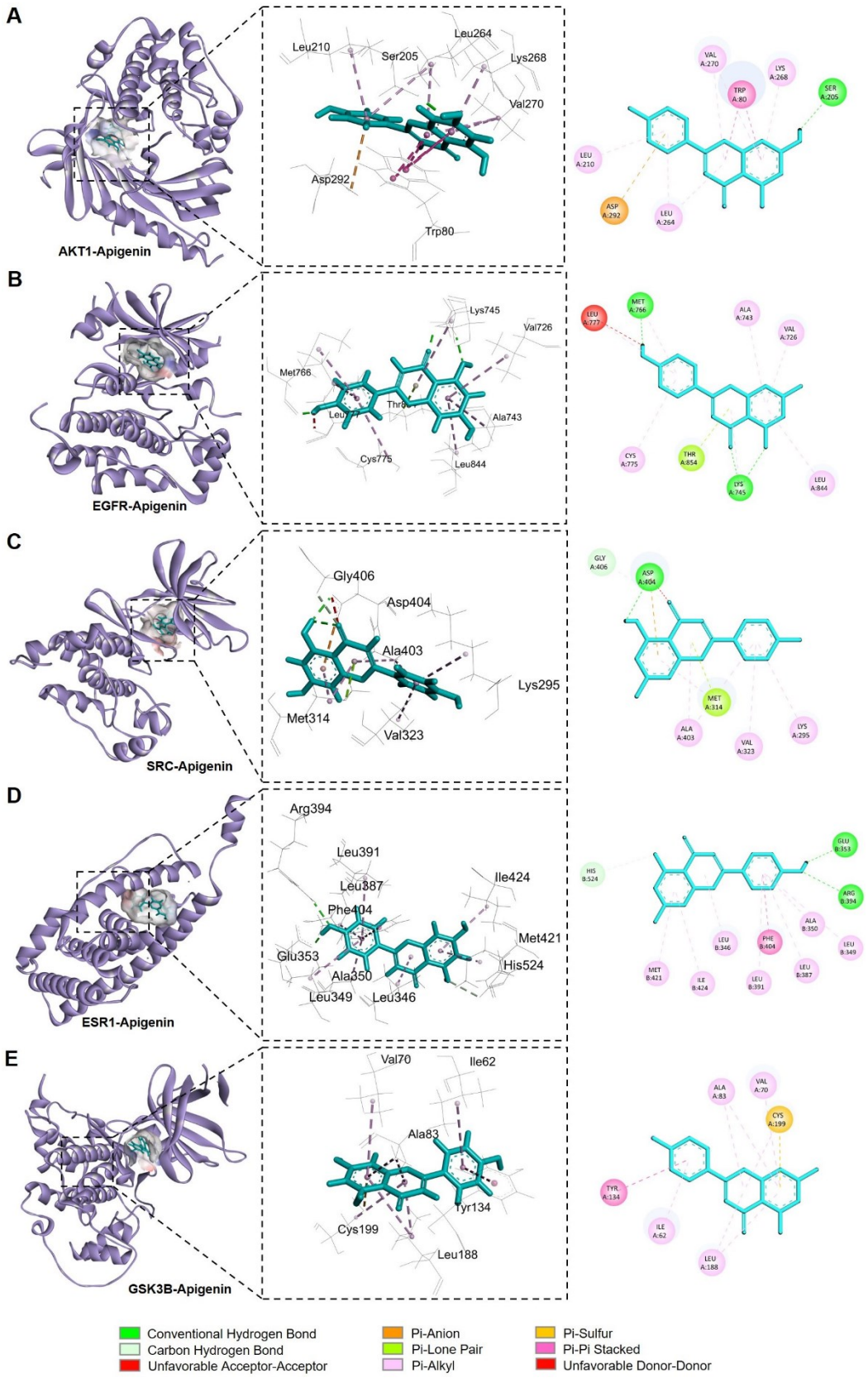


Figure 3.10. Molecular docking analysis of Apigenin with hub target proteins. (**A-E**) Graphs showing the binding mode and molecular interactions of Apigenin with AKT1, EGFR, SRC, ESR1 and GSK3B, respectively. The 3D graphs showing the binding model of Apigenin and each target. The 2D graphs showing the specific interactions between Apigenin and the hub targets.

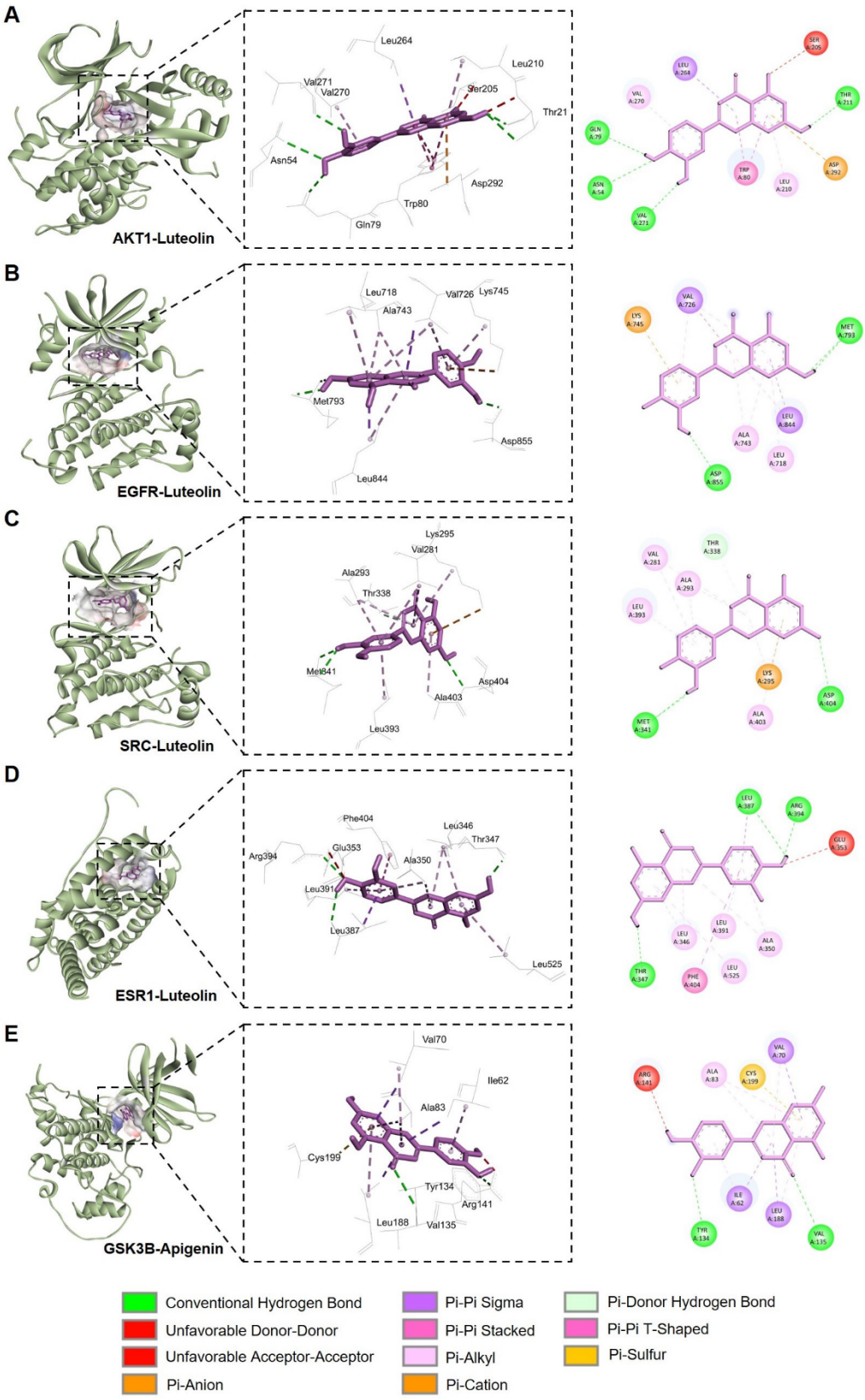


Figure 3.11. Molecular docking analysis of Luteolin with hub target proteins. (**A**-**E**) Graphs showing the binding mode and molecular interactions of Luteolin with AKT1, EGFR, SRC, ESR1 and GSK3B, respectively. The 3D graphs showing the binding model of Luteolin and each target. The 2D graphs showing the specific interactions between Luteolin and the hub targets.

3.5.6 GEO dataset analysis in Apigenin and Luteolin

The network pharmacology results revealed that Apigenin and Luteolin play a crucial role in regulating cancer-related signalling pathways, suggesting their potential as therapeutic agents in treating cancer. To gain further insights into the mechanisms of Apigenin and Luteolin against cancer, we utilized publicly available GEO datasets on human cancer cells with the treatment of Apigenin or Luteolin. By applying predetermined screening criteria, three GEO datasets were selected, comprising a total of 26 samples derived from two distinct human breast cancer cell lines (MCF7 cells and MDA-MB-231 cells) and two different types of human ovarian cancer cells (A2780 cells and ES2 cells) (Table 3.5).

Table 3.5. Details of the GEO datasets collected for Apigenin or Luteolin treatment

Datasets	Title	Cells	Platform	Induced sample information
GSE119552 [144]	Effect of estradiol, zearalenone and apigenin on ER-positive breast cancer cells MCF-7	MCF7 cells	Agilent-039494 SurePrint G3 Human GE v2 8x60K Microarray 039381 (Feature Number version)	MCF-7_Control: 4 MCF-7_Apigenin: 4
GSE120550 [145]	Transcriptomic profiling of MDA-MB-231 cells treated with TNFα [40 ng/ml] ± Apigenin [40 µM]	MDA-MB-231 cells	[HuGene-2_1-st] Affymetrix Human Gene 2.1 ST Array [transcript (gene) version]	MDA-MB-231_Control: 3 MDA-MB-231_Apigenin: 3
GSE212598	Transcriptome analysis for luteolin or siVRK1 treated ovarian cancer cells	A2780 cells ES2 cells	DNBSEQ-G400 (Homo sapiens)	A2780_Control: 3 A2780_Luteolin: 3 ES2_Control: 3 ES2_Luteolin: 3

DEGs were screened from the obtained GEO datasets, genes with *P* value less than 0.05 and $|Log_2FoldChange|$ above 1 were considered as DEGs. In MCF7 cells upon treatment with Apigenin, it was identified 856 up-regulated DEGs and 593 down-regulated DEGs (Figure 3.12A). Similarly, in MDA-MB-231 cells treated with Apigenin, there are 280 up-regulated DEGs and 220 down-regulated DEGs were identified (Figure 3.12B). For Luteolin treatment, A2780 cells showed 747 up-regulated DEGs and 2,143 down-regulated DEGs (Figure 3.12C), while ES2 cells exhibited 389 up-regulated DEGs and 997 down-regulated DEGs (Figure 3.12D). These differentially expressed genes in human cancer cells upon treatment with Apigenin or Luteolin were further utilized for enrichment analysis to explore the underlying mechanisms.

KEGG enrichment analysis revealed a significant involvement of Breast cancer in both Apigenin- and Luteolin-treated cancer cells, highlighting the crucial role of Apigenin and Luteolin in the treatment of breast cancer. Moreover, Pathways in cancer was also commonly enriched in both treatment groups (Figure 3.13A, B; Figure 3.14A, B). The treatment of Apigenin or Luteolin resulted in significant alterations in the expression levels of genes associated with Breast cancer and Pathways in cancer (Figure 3.13C, D and E; Figure 3.14C, D and E). These findings suggest the potential anti-cancer effects of Apigenin and Luteolin and provide further support for their role in breast cancer treatment.

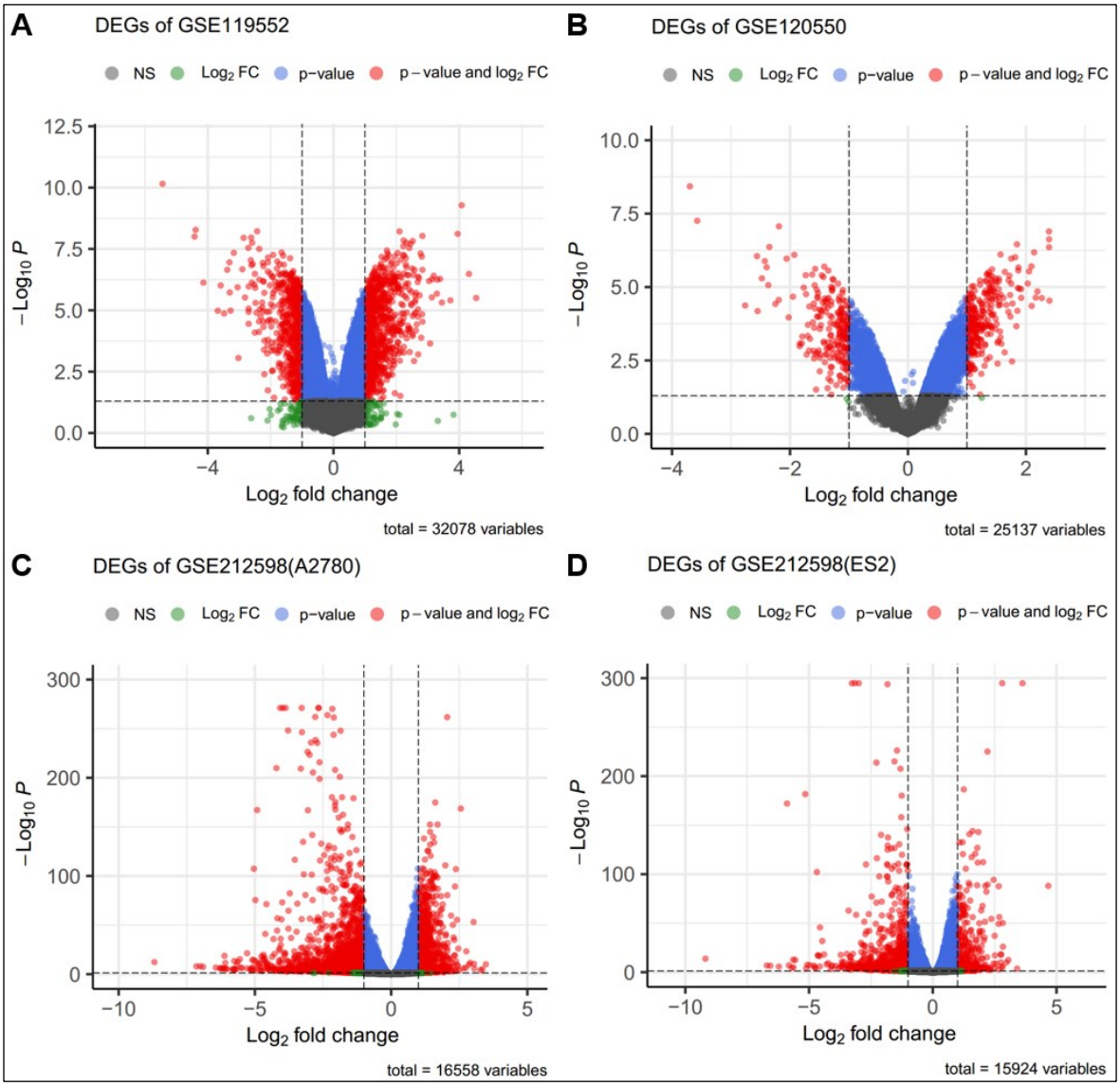


Figure 3.12. Global transcriptomic changes in human cancer cells exposed to Apigenin or Luteolin. (**A** and **B**) Volcano plots showing the differentially expressed genes (DEGs) in Apigenin-treated MCF7 cells and MDA-MB-231 cells within indicated GEO datasets. (**C** and **D**) Volcano plots showing the DEGs in Luteolin-treated A2780 cells and ES2 cells within indicated GEO datasets. DEGs were defined as *P* value less than 0.05 and |Log$_2$FoldChange| above 1. Log$_2$FoldChange in the *x*-axis and -Log$_{10}$*P* in the *y*-axis. Genes with different colours represent indicated changes in the figure legends.

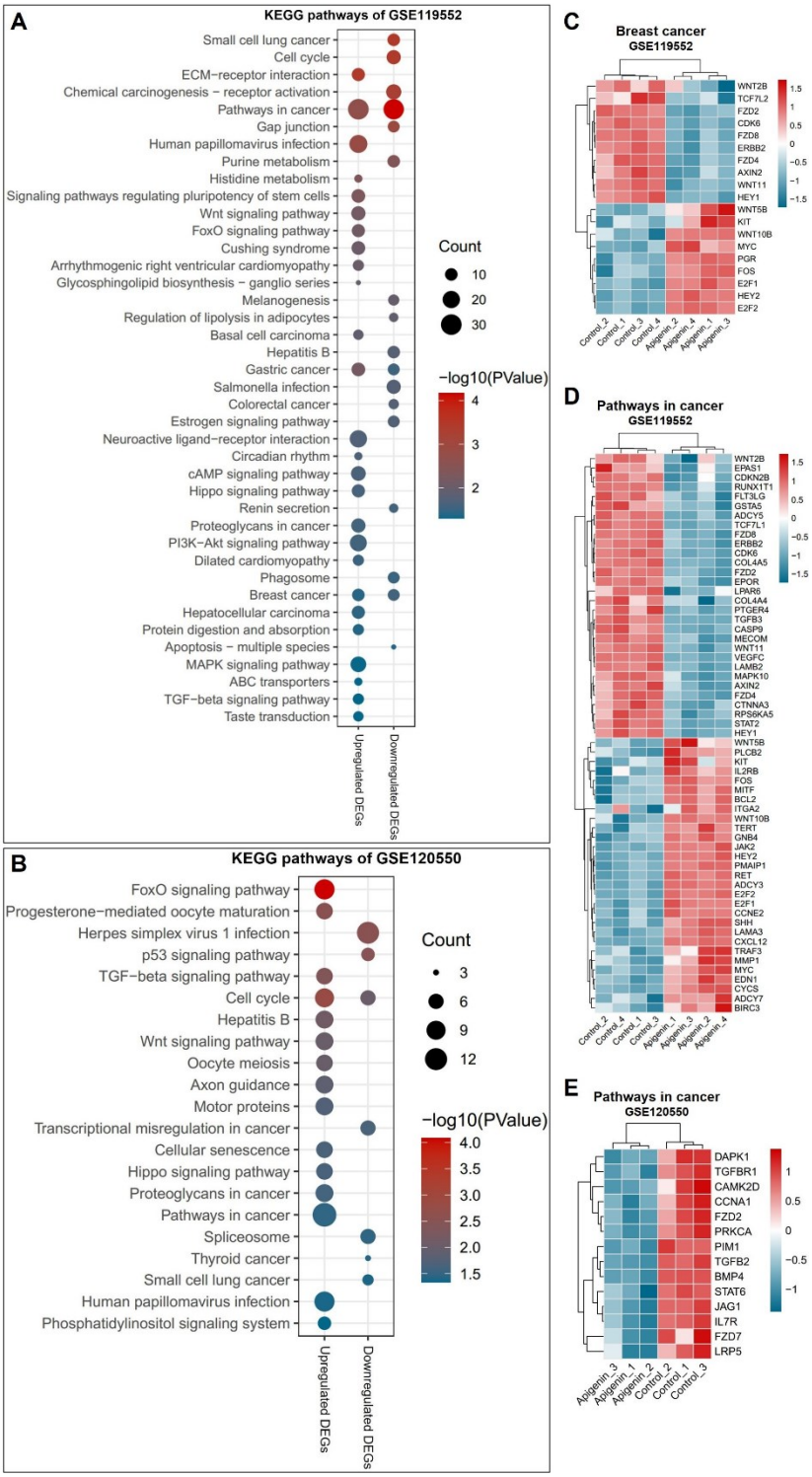


Figure 3.13. KEGG enrichment analysis of MCF7 and MDA-MB-231 cells exposed to Apigenin. (**A** and **B**) Scatter plots illustrating the significant KEGG pathways enriched by the up- and down-regulated differentially expressed genes (DEGs) within GSE119552 and GSE120550 datasets. The *y*-axis represents the name of pathway, and the *x*-axis represents the up- and down-regulated DEGs. Dot size represents the gene count, and the colour indicates the -$\log_{10}$(*P*value). (**C**, **D** and **E**) Heatmaps showing the expression patterns of the associated genes within the indicated signalling pathways.

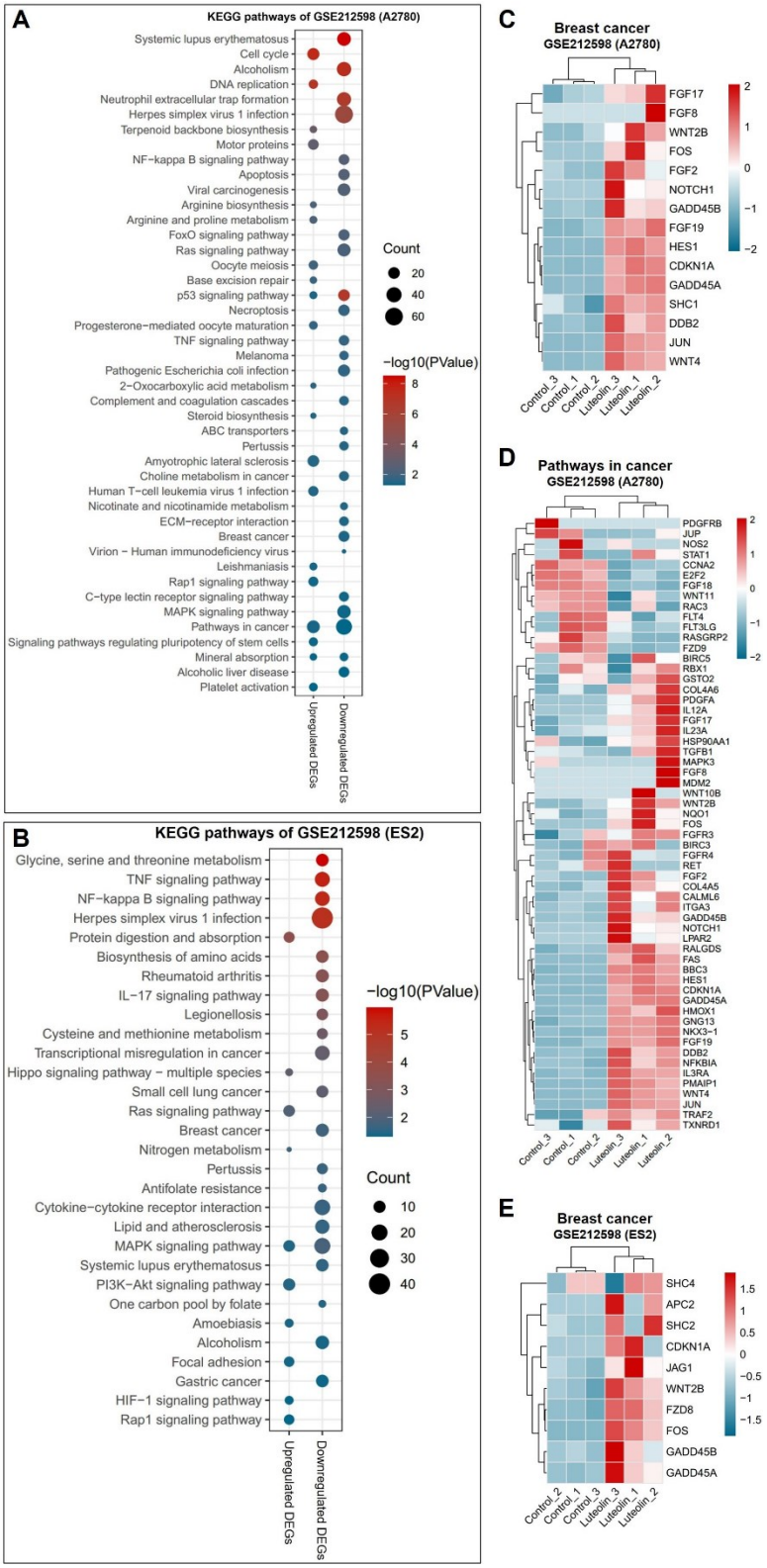


Figure 3.14. KEGG enrichment analysis of A2780 and ES2 cells exposed to Luteolin. (**A** and **B**) Scatter plots illustrating the significant KEGG pathways enriched by the up- and down-regulated differentially expressed genes (DEGs) within GSE212598 dataset. The *y*-axis represents the name of pathway, and the *x*-axis represents the up- and down-regulated DEGs. Dot size represents the gene count, and the colour indicates the -$\log_{10}$(*P*value). (**C**, **D** and **E**) Heatmaps showing the expression patterns of the associated genes within the indicated signalling pathways.

3.6 Discussion

The anti-tumour activity of *Limonium Sinense* has been previously discussed to be associated with its immunomodulatory activity [27]. In line with these findings, our current analysis further emphasizes the strong association between the biological activity of *Limonium Sinense* and cancer-related pathways, suggesting a potential role of *Limonium Sinense* in cancer treatment. Through network pharmacology analysis, we identified several enriched pathways related to cancer regulation, with "Pathways in cancer" being the top ranked KEGG pathway. Additional cancer-related signalling pathways, including Prostate cancer, Central carbon metabolism in cancer, Proteoglycans in cancer, Non-small cell lung cancer and Small cell lung cancer, were also significantly enriched by the targets of *Limonium Sinense*. These findings underscore the importance of *Limonium Sinense* in cancer regulation.

Interestingly, our analysis also revealed significant enrichment of the cAMP signalling pathway, HIF-1 signalling pathway and cell cycle by the *Limonium Sinense* targets, further supporting the potential efficacy of *Limonium Sinense* in cancer treatment. Additionally, the molecular events and interactions within the pathways in cancer involve various biomolecules, such as signalling molecules, receptors, transcription factors, kinases, and other key players. These intricate mechanisms regulate cell behavior and contribute to the characteristic features of cancer. Our findings suggest that the regulation of *Limonium Sinense* in cancer possess is multifaceted, which may involve the cooperation of multiple levels including signalling pathways, genes, proteins, and other factors.

Many active compounds isolated from *Limonium Sinense* have demonstrated significant anti-cancer activity, such as Apigenin [146], Naringenin [147], Luteolin [77], Kaempferol [148] and Quercetin [149], which are belong to the flavonoid group commonly found in herbal plants and fruits. In our analysis, we identified a hub network consisting of Apigenin and Luteolin interacting with five target genes and four signalling pathways. Apigenin is a naturally occurring plant flavone, has gained recognition as a cancer chemopreventive agent. It exhibits remarkable antioxidant, anti-mutagenic, anti-inflammatory, antibacterial and antiviral effects [150-153]. Recent studies have extensively investigated Apigenin for its anti-cancer activities and low toxicity. It has shown broad anti-cancer effects across various types of cancer, including colorectal cancer, breast cancer, liver cancer, lung cancer, melanoma, prostate cancer, and osteosarcoma [154-159]. The mechanisms underlying Apigenin's suppression of different human cancers involve multiple biological effects, such as triggering cell apoptosis and autophagy, inducing cell cycle arrest, suppressing cell migration and invasion, and stimulating an immune response [150, 160, 161]. During the regulation of these biological processes, various signalling pathways were involved, including PI3K/AKT signalling

pathway, MAPK/ERK signalling pathway, JAK/STAT signalling pathway, NF-κB signalling pathway and Wnt/β-catenin pathway [154, 162-166]. These findings indicate the potential of Apigenin as a dietary supplement or adjuvant chemotherapeutic agent for cancer therapy. Luteolin, another flavone compound, is found in various plants, including fruits, vegetables, and medicinal herbs. It exerts diverse biological effects such as antioxidant and anti-inflammatory properties. Luteolin has also been shown to act as an anticancer agent against multiple types of human cancers, including lung cancer, breast cancer, glioblastoma cancer, prostate cancer, colon cancer, and pancreatic cancer [167]. In the context of carcinogenesis, Luteolin hampers cancer progression by inhibiting cell transformation, metastasis, invasion, and angiogenesis through multiple mechanisms, including kinase suppression, cell cycle regulation, apoptotic cell death induction, and transcription factors reduction [168]. For instance, Luteolin induces G1 phase cell cycle arrest in various human cancer cell lines, such as gastric, prostate and melanoma cells. This effect is associated with the suppression of CDK2 activity in colorectal cancer and melanoma cells [168, 169]. Additionally, in breast cancer cells, Luteolin exhibits anti-cancer properties via suppressing EMT-mediated RAS/RAF/MEK/ERK signalling [170].

Four signalling pathways were identified from the hub network, namely Pathways in cancer, EGFR tyrosine kinase inhibitor resistance, Endocrine resistance and HIF-1 signalling pathway, suggesting a strong association of *Limonium Sinense* biological effects with these signalling pathways. As mentioned earlier, Pathways in cancer represents a comprehensive collection of interconnected molecular pathways involved in the development and progression of various types of cancer. The pathway provides valuable insights into the molecular mechanisms underlying carcinogenesis and the hallmarks of cancer.

The epidermal growth factor receptor (EGFR) belongs to the receptor tyrosine kinases family [171-174], and the EGFR family consists of EGFR1, EGFR2, EGFR3 and EGFR4 proteins [175-177]. EGFR plays a crucial role as a receptor tyrosine kinase involved in regulating various processes essential for the initiation, maintenance, and survival of cancer stem cells [171]. Overexpression of EGFR is a common occurrence in numerous cancer tissues and has been associated with increased aggressiveness and poor clinical outcomes in breast, lung, ovarian, cervical, bladder, esophageal, brain as well as head and neck cancers [175, 178-183]. EGFR also contribute to resistance against different cancer therapies. EGFR acts as a co-transcription factor for genes involved in cell proliferation and angiogenesis when translocated to the nucleus, while also functions as a tyrosine kinase that activating and stabilizing proliferating cell nuclear antigen and DNA-dependent protein kinase. Nuclear localization of EGFR is strongly associated with cancer progression, poor prognosis and increased resistance to radiation, chemotherapy, and anti-EGFR therapies like gefitinib and cetuximab in various cancer types [184]. EGFR tyrosine kinase inhibitors (TKIs) are commonly used

in clinical trials as chemotherapeutic strategies due to their ability to target EGFR and its downstream pathways. These TKI therapies have shown initial efficacy in cancer patients with EGFR mutations or aberrant EGFR activation. However, the development of resistance to TKIs is a common occurrence, leading to tumour recurrence.

Over the past decade, extensive research has been conducted to understand the mechanisms underlying TKI resistance in EGFR-driven cancers. Several mechanisms have been identified as major drivers of resistance to EGFR TKI treatment. One prominent mechanism is the acquisition of a secondary EGFR mutation known as T790M. This mutation alters the conformation of the EGFR protein, reducing its sensitivity to TKIs. Activation of alternative signalling pathways, such as c-Met and Hepatocyte growth factor (HGF), or aberrant activation of AXL receptor tyrosine kinase, can also confer resistance to EGFR TKIs. Moreover, alterations in downstream pathways, such as K-RAS mutations or loss of PTEN, contribute to TKI resistance by bypassing the EGFR signalling cascade. Impairment of the apoptosis pathway mediated by EGFR TKIs, often caused by BCL2-like 11/BIM deletion polymorphism, can also result in resistance. Additionally, histologic transformation, where cancer cells undergo a phenotypic switch, is another mechanism associated with TKI resistance [171, 185].

In the adjuvant setting, endocrine therapy has become the standard of care for early-stage breast cancer with hormone receptor-positive tumours [186]. This therapeutic approach has significantly improved overall survival and disease-free survival in women with early breast cancer [187]. The most commonly used agents in endocrine therapy include selective estrogen receptor modulators (SERMs) such as tamoxifen [188], estrogen synthesis inhibitors like aromatase inhibitors (AIs) including anastrozole, letrozole, and exemestane [189-191], and selective estrogen receptor down-regulators (SERDs) such as fulvestrant [192]. These agents exert their effects by modulating estrogen receptor (ER) signalling. Despite their clinical success, resistance to endocrine therapy remains a significant challenge.

Resistance to endocrine therapy can arise through various mechanisms. Loss of ER-α expression, alterations in the expression of coactivators or coregulators that are critical for ER-mediated gene transcription, and activation of ligand-independent growth factor signalling cascades that lead to kinases and phosphorylation of ER, contributing to resistance. Moreover, altered availability of active tamoxifen metabolites, which is regulated by drug-metabolizing enzymes, as well as dysregulation of the cell cycle and apoptotic machinery can also confer resistance. Additionally, hyperactive receptor tyrosine kinase (RTK) and stress/cell kinase pathways collectively contribute to the development and sustenance of pharmacologic resistance to endocrine therapy [193].

Hypoxia-inducible factor 1 (HIF-1) serves as a transcription factor that regulates oxygen homeostasis. It comprises two subunits: an inducibly-expressed HIF-1α subunit and a constitutively expressed HIF-1β subunit. Under normoxia, HIF-1α undergoes hydroxylation at specific prolyl residues, triggering its immediate ubiquitination and subsequent proteasomal degradation. Conversely, under hypoxia conditions, HIF-1α subunit becomes stabilized and interacts with coactivators such as p300/CBP to modulate its transcriptional activity [194].

HIF-1 governs the expression of target genes involved in enhancing oxygen delivery and mediating adaptive response to oxygen deprivation. These genes often possess specific responsive elements called hypoxia response elements (HREs), which bind to HIF and lead to an adaptive response to hypoxia and tumour progression and metastasis. Given the crucial role of the HIF system and hypoxia in cancer cells, they have been proposed as potential therapeutic targets for cancer treatment [195]. For instance, the HIF-1α/VEGF-A response in T cells directly impacts cancer growth, progression, and vascularization, indicating that therapeutic strategies aimed at targeting HIF-1α or VEGFA signalling in the tumour microenvironment could have broad implications for cancer therapy, including immunotherapy [196]. Obviously, these four signalling pathways all showed strong association with cancer process, further emphasize the potential use of *Limonium Sinense* in cancer treatment.

Limonium Sinense has traditionally been used to replenish blood in the body [7], however, there is no direct evidence to prove the blood-enriching function of *Limonium Sinense* so far. Previous studies have speculated that the blood-enriching effect of *Limonium Sinense* is due to its abundance of vitamins, including vitamin B12, vitamin B2 and vitamin C, as well as trace elements such as iron (Fe), cobalt (Co), nickel (Ni), zinc (Zn) and chromium (Cr) [16]. Among the vitamins found in *Limonium Sinense*, vitamin B12, also known as cobalamin, is a water-soluble vitamin that plays a crucial role in myelin formation, red blood cell maturation, and nucleic acid synthesis [197]. It is usually found in meat, eggs, and dairy products and is essential for erythropoiesis and nerve myelination. Deficiency in vitamin B12 can lead to megaloblastic anaemia due to disrupted DNA synthesis and demyelinated nerves. Megaloblastic changes affect all formed blood elements, with the most significant alterations observed in erythrocytes. The severity of the anaemia corresponds to the extent of morphologic changes in red blood cell [198].

On the other hand, the trace element cobalt (Co) is essential for the synthesis of vitamin B12. Cobalt deficiency manifests in symptoms ranging from reduced growth and weight loss to liver steatosis, anaemia, impaired immune function, impaired reproductive function and even death [199]. The richness of cobalt and vitamin B12 in *Limonium Sinense*, along with other vitamins and trace

elements that play a role in red cell maturation, may collectively contribute to establishing a favourable microenvironment in the body, which can stimulate haematopoiesis.

Our network pharmacology results indicate that peptidyl-tyrosine phosphorylation was top enriched in the biological process analysis. In mammalian cells, tyrosine phosphorylation is a crucial mechanism that regulates signal transduction pathways and key cellular functions [200]. It has been reported that protein tyrosine phosphorylation is involved in inflammation-induced corneal neovascularization, promoting the formation of new blood vessels [201]. These findings suggest that *Limonium Sinense* may possess a potential mechanism for regulating protein tyrosine phosphorylation, which could help to its ability to induce blood vessel formation.

Interestingly, the HIF-1 signalling pathway was also significantly enriched by the *Limonium Sinense* potential targets. Given that HIF-1 signalling plays a key role in angiogenesis and blood vessel formation, we propose that the blood-enriching function of *Limonium Sinense* might through regulating HIF-1 signalling pathway, thereby leading to the production of erythropoietin (EPO) and promoting the blood vessel development.

It is well known that anaemia is a common diagnosis in patients with cancer that may affect both quality of life and survival [202]. Cancer can directly cause or exacerbate anaemia either by suppressing haematopoiesis, cytokine-induced iron sequestration, or reduced red blood cell production [203]. Treatment of anaemia can significantly improve patients' quality of life and potentially enhance clinical outcomes. Therefore, the blood-enriching property of *Limonium Sinense* may help alleviate cancer-related symptoms.

Network pharmacology is recognized as an effective approach for exploring the intricate relationships among components, targets, and diseases [204, 205]. It also provides valuable insights into the mechanisms of action of medications. Our findings indicate a potential link between the biological activity of *Limonium Sinense* and breast cancer. However, it is important to note that this study primarily serves as a theoretical investigation into the anti-cancer effects and the underlying mechanism of action of *Limonium Sinense*, which has great limitations. Further exploration and verification using high-performance liquid chromatography (HPLC) and disease models are required to elucidate the mechanisms underlying these effects and to better understand the blood-enriching and anti-tumour activities of *Limonium Sinense*. Enrichment analyses, such as GO and KEGG analyses, are common bioinformatics tools for interpreting large genomic or transcriptomic datasets by identifying overrepresented terms or pathways. Nevertheless, these methods also have certain limitations. GO terms imply possible functional characteristics of the given gene sets [139]. The effectiveness of GO enrichment analysis relies on gene annotation quality and completeness, impacting results due to poorly or non-annotated genes. GO annotation levels (molecular function,

biological process, cellular component) affect result interpretation, and the choice may influence enriched term identification. The output is typically a list of enriched terms, potentially losing substantial information on term relations. GO terms are usually hierarchically structured, and traditional analysis assumes term independence. However, in reality, terms may be correlated, leading to overrepresentation of certain terms. GO enrichment testing involves multiple hypothesis testing, which may increase the risk of false positives. In practical cases, functional genomic input may naturally form a ranked list. For most applications, setting an arbitrary threshold for the target set is required. The results of the enrichment analysis may often depend on the specific threshold that is set. Tools using the simple hypergeometric distribution require a fixed threshold [206]. Some threshold-free tools have been developed, such as GSEA [81]. GSEA uses a Kolmogorov-Smirnov-like statistic with weighted genes occurrences. It's not specifically for GO enrichment and lacks GO directed acyclic graph (DAG) structure visualization. In addition, GSEA provides estimated, not exact, *p*-values based on permutations, limited by the number of permutations performed.

The functional assignment in KEGG involves linking a set of genes in the genome with a network of interacting molecules in the cell [207]. However, it may not cover all biological processes, and some pathways may be more extensively studied than others, leading to incomplete coverage and a potential oversight of crucial biological information. Regulatory pathways exhibit greater diversity, making it challenging to integrate them into common reference pathway diagrams. For instance, larger pathways may appear enriched simply due to their higher genes count, resulting in biases in favour of well-studied and larger pathways. KEGG enrichment analyses typically assume equal contributions from all genes within a pathway to the biological function, disregarding the varying degrees of importance of different genes. Furthermore, KEGG pathways are curated for specific organisms, leading to variations in the availability and accuracy of pathway information between species. The absence of proper identifiers for functions in regulatory pathways poses a significant problem for KEGG analysis. Some pathways may be better characterized in model organisms than in others. In addition, KEGG pathways are static representations that may not capture dynamic changes or variations in pathway activity under different conditions.

We employed the molecular docking method to explore the interaction between active compounds and potential targets in this analysis. Molecular docking has become a ubiquitous tool in drug discovery and structural biology, enabling the prediction of binding modes and affinities of small molecules (ligands) to target proteins (receptors) [118]. However, it is essential to acknowledge various challenges associated with docking studies, for instance, docking programs rely on scoring functions to evaluate the fitness of ligand-receptor complexes. These scoring functions may not consistently reflect the true binding affinity, leading to potential false positives or false negatives. Common limitations include the oversimplification of complex molecular interactions within the

binding site, neglecting factors such as solvation effects and protein flexibility. Many docking methods assume a rigid protein structure, ignoring the inherent flexibility of both the protein and the ligand, especially concerning protein conformational changes induced by ligand binding, impacting the accuracy of predictions. Most docking studies are typically performed in a vacuum or with simplified solvation models, disregarding the influence of water molecules and other solvent effects. Despite their crucial role in ligand binding, water molecules are often omitted, leading to potential inaccuracies [208]. Additionally, accurately predicting the binding of ligands to metal ions or other cofactors in the binding site poses challenges due to the diversity of metal-ligand interactions. Standard docking simulations frequently neglect induced fit, where ligand binding induces conformational changes in the protein, potentially resulting in inaccurate predictions, especially when substantial structural rearrangements occur upon binding. Docking methods may struggle to accurately predict binding modes and affinities for weak binders or compounds with low affinity, as they may not effectively compete with solvent molecules in the binding site. Furthermore, protein dynamics, encompassing side-chain flexibility and backbone movements, are often disregarded in standard docking simulations, despite their significance in understanding binding events. The accuracy of docking results heavily relies on the selected force field parameters in the simulations, with empirical force fields having limitations in accurately capturing specific interactions or chemical environments [209].

To conclusion, our study predicts that the biological activities of *Limonium Sinense* are strongly associated with breast cancer and multiple cancer-related pathways. The active compounds Apigenin and Luteolin present in *Limonium Sinense* appear to have a crucial role against cancer. GEO datasets validation indicates the involvement of breast cancer pathway upon treatment with Apigenin or Luteolin in human cancer cells. These findings highlight the potential of *Limonium Sinense* as a promising therapeutic agent for the treatment of breast cancer. In a word, the specific regulatory mechanism of *Limonium Sinense* in treating cancers is extremely complex, involving a variety of active ingredients, changes in the microenvironment, immune responses, signalling pathways regulation and other factors, and further exploration is necessary to fully understand these mechanisms, see Chapter 4.

Chapter 4 Integrated analysis reveals effects of bioactive ingredients from *Limonium Sinense* (Girard) Kuntze on hypoxia-inducible factor (HIF) activation

Part of the background of provided in this chapter is from my original draft of the following publication with support from my supervisors:

***Hualong Zhao**, Siyuan Wang, Yilu Zhou, Ayse Ertay, Philip T.F. Williamson, Rob M. Ewing, Xinhui Tang, Jialian Wang, Yihua Wang. Integrated analysis reveals effects of bioactive ingredients from Limonium Sinense (Girard) Kuntze on hypoxia-inducible factor (HIF) activation. Front Plant Sci, 2022. 13: p. 994036. doi: 10.3389/fpls.2022.994036.*

Authors' contributions

*Yihua Wang and Jialian Wang designed the study. **Hualong Zhao**, Siyuan Wang, Yilu Zhou and Ayse Ertay performed the experiments. Yihua Wang, Jialian Wang, Hualong Zhao, Siyuan Wang and Yilu Zhou analysed the data and finalised the figures, and all other authors contributed to data interpretation.*

4.1 Abstract

Limonium Sinense (Girard) Kuntze is a traditional Chinese medicinal herb, showing blood replenishment, anti-tumour, anti-hepatitis, and immunomodulation activities amongst others. However, the mechanism of its pharmacological activities remains largely unknown. Here, we investigated the effects of bioactive ingredients from *Limonium Sinense* using integrated analysis. Water extracts from *Limonium Sinense* (LSW) showed a strong growth inhibitory effect on multiple cells in both 2D and 3D cultures. Global transcriptomic profiling and further connectivity map (CMap) analysis identified several similarly acting therapeutic candidates, including Tubulin inhibitors and hypoxia-inducible factor (HIF) modulators. The effect of LSW on cell cycle was verified with flow cytometry showing a G2/M phase arrest. Integrated analysis suggested a role for gallic acid in mediating HIF activation. Taken together, this study provides novel insights into the bioactive ingredients in *Limonium Sinense*, highlighting the rich natural resource and therapeutic values of herbal plants.

4.2 Introduction

As we described above, *Limonium Sinense* is mainly used for the treatment of fever, hepatitis, haemorrhage, menorrhagia, irregular menstruation, cancer and other disorders [7]. Multiple bioactive ingredients have been identified from *Limonium Sinense*, including polysaccharides, tannins, alkaloids, flavonoids, terpenes, aliphatic compounds, aminoacids, minerals and vitamins [16, 31, 32]. It is reported that the major active constituents found in *Limonium Sinense* are flavonoids, including flavanones, flavonols, flavonol glycosides, flavonol glycoside gallates and flavones [31, 33], while polysaccharides are among the most abundant constituents in its root [21]. Despite these findings, the mechanism of its pharmacological activities remains to be elucidated. Our previous studies in Chapter 3 predicted the biological effects of *Limonium Sinense* using network pharmacology methods and identified its potential use as a candidate for breast cancer treatment. Here, we further investigate the effects of bioactive ingredients from *Limonium Sinense* by integrating bioinformatic analysis and *in vitro* experiments.

4.2.1 Breast cancer

Breast cancer stands as one of the most prevalent malignancies and is the leading cause of mortality in women worldwide [210, 211]. Recent data from the Global Cancer Observation (GCO) (https://gco.iarc.fr), an interactive web-based platform that provides global cancer statistics for cancer control and research, indicates a significant rise in breast cancer incidence. In 2020 alone, there were over 2 million estimated new cases of breast cancer (Figure 4.1A, B), making it the most

frequently diagnosed cancer worldwide, surpassing all other caner types. Additionally, breast cancer was responsible for nearly 700,000 estimated deaths (Figure 4.1C), ranking among the top five leading causes of cancer-related death in women globally. Urgent measures are required to alleviate the suffering of breast cancer patients.

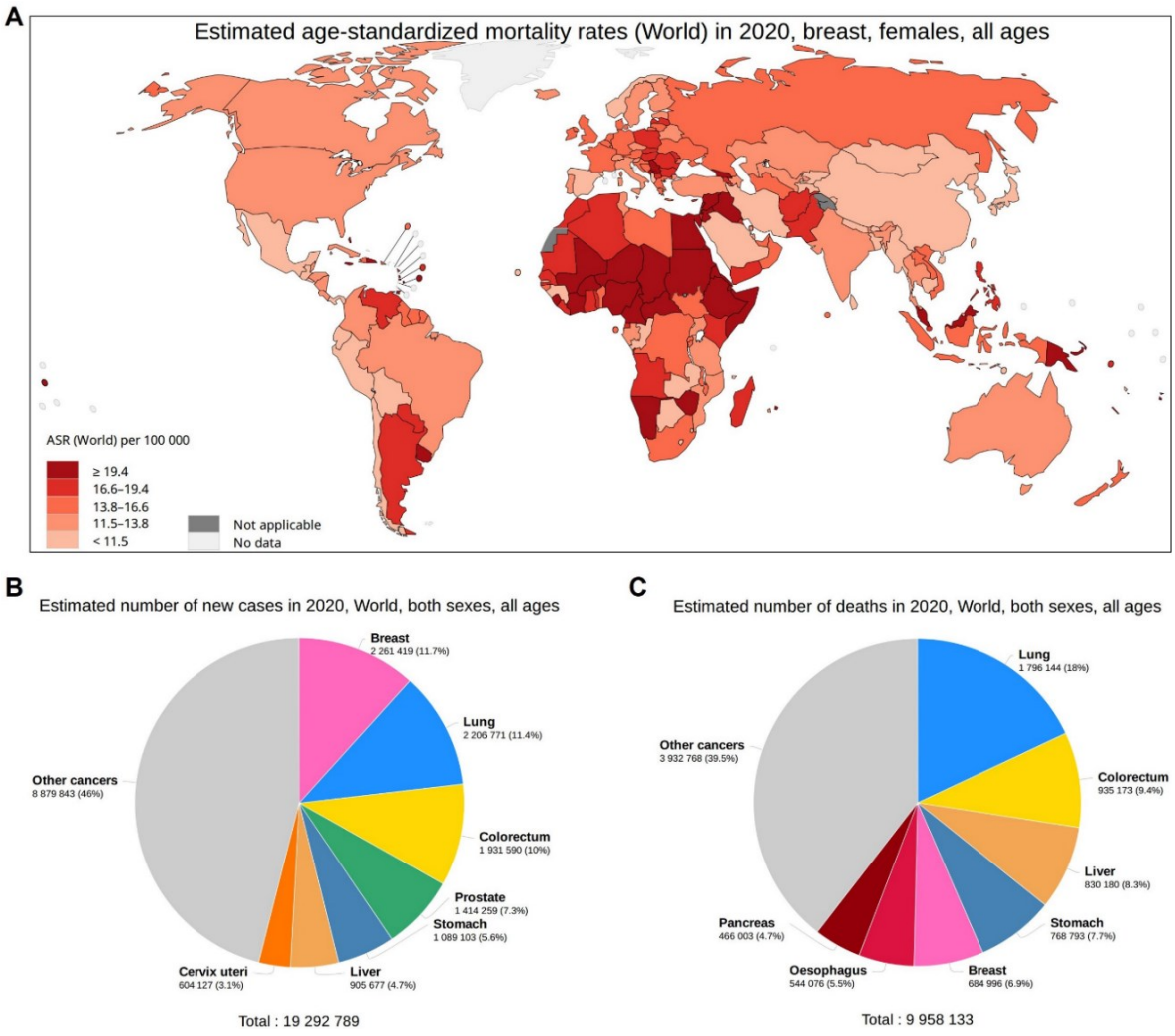


Figure 4.1. Estimated breast cancer data in 2020 from Global Cancer Observation (GCO). (**A**) Graph showing the estimated mortality rates of breast cancer in all ages worldwide, the colour indicates the Age-Standardized Rate (ASR) per 100,000. (**B** and **C**) Pie charts illustrating the estimated number of new cases and death of breast cancer in 2020 all over the world. Data source: GLOBOCAN 2020. Map production: International Agency for Research on cancer (IARC) (http://gco.iarc.fr/today), World Health Organization.

4.2.2 Breast cancer classification

4.2.2.1 Histological subtypes of breast cancer

Breast cancer is a complex disease with diverse biological entities characterized by significant heterogeneity in pathology, genomic alterations, gene expression, and the tumour microenvironment. These factors show a strong relationship with clinical behaviour and treatment response [212]. The diagnosis of human breast cancer involves a broad classification based on histology, distinguishing between *in situ* carcinoma and invasive carcinomas, depending on the extent of malignant cell spread from the breast lobules or ducts into the surrounding stroma [213]. Traditionally, ductal carcinoma *in situ* (DCIS) is the most common form of pre-invasive breast cancer. However, only a subset of 10%-30% of DCIS cases progress to invasive cancer. Invasive carcinomas are further categorized into invasive ductal carcinoma (IDC) and invasive lobular carcinoma (ILC) based on cell morphology (Figure 4.2). The IDC is the most prevalent subtype which accounting for 60%-75% of cases, while ILC comprises 10%-15% of tumours. Apart from ILC, the remaining histological subtypes are rare, constituting only 0.1%-7% of breast cancer cases [214].

4.2.2.2 Clinical subtypes of breast cancer

In addition to morphological classification, breast tumours are clinically classified based on the expression of the estrogen receptor (ER), progesterone receptor (PR), and human epithelial growth factor receptor 2 (HER2/ERBB2), resulting in three main clinical groups: ER+, HER2+, and triple-negative breast cancer (TNBC). ER+ cancers constitute approximately 70% of all breast cancers, defined as ≥ 1% of tumour cells expressing ER. HER2+ tumours can be further subdivided into HER2+ER+ (about 70%) and HER2+ER- (about 30%) subgroups. HER2 status is determined through immunohistochemistry and confirmed using chromogenic or fluorescence *in situ* hybridization to detect gene amplification. Approximately 15% of breast cancers are HER2+, exhibiting significant biological heterogeneity. Recognition is also growing for HER2-low breast cancer, characterized by low (1+) to moderate (2+) HER2 expression detectable by immunohistochemistry without ERBB2 amplification [215]. TNBC, accounting for approximately 15% of breast cancers, lacks expression of ER, PR, and HER2. TNBC comprises multiple subtypes generally characterized by expression of EGFR and cytokeratins CK5 and CK14. These tumours often follow an aggressive clinical course, associated with younger age at diagnosis, and poor prognosis. Furthermore, TNBCs have a higher propensity for early recurrence and metastasis, particularly to the lung and brain, contributing disproportionately to breast cancer mortality [212, 216, 217].

4.2.2.3 Molecular subtypes of breast cancer

Breast cancer can be classified into five major molecular subtypes: Luminal A (ER positive, PR positive and HER2 negative), Luminal B (ER positive and/or PR positive, HER2 positive), HER2 enriched (ER negative, PR negative and HER2 positive), normal-like, and triple-negative (basal-like) (ER negative, PR negative and HER2 negative) subtypes (Figure 4.2) [218].

Luminal A tumours are characterized by the presence of ER and PR and the absence of HER2. They also exhibit low expression of the cell proliferation marker Ki67 which usually less than 20%. Clinically, these tumours are typically low grade, slow growing, and have a favourable prognosis, with a lower incidence of relapse and higher survival rates. Luminal A carcinomas show a high response rate to hormone therapy, such as tamoxifen or aromatase inhibitors, and have a more limited benefit from chemotherapy [219].

Luminal B tumours are characterized by being ER positive and can be PR negative, with a high expression of Ki67 that generally greater than 20%. They are typically of intermediate to high histologic grade. These tumours may benefit from a combination of hormonal therapy and chemotherapy. Luminal B tumours account for 10%-20% of luminal tumours and exhibit a moderately low expression of ER, along with increased expression of proliferation and cell cycle genes. This subtype is associated with a worse prognosis compared to Luminal A tumours [220].

The HER2-positive subgroup constitutes 10%-15% of breast cancers and is characterized by high expression of the HER2 protein, along with the absence of ER and PR. These tumours grow faster compared to luminal subtypes. The HER2-positive subtype is generally more aggressive and rapidly growing. Within this subtype, two subgroups can be distinguished: luminal HER2 (ER+, PR+, HER2+ and Ki67: 15%-30%) and HER2-enriched (HER2+, ER-, PR-, Ki67 > 30%) [221]. Both subtypes have a worse prognosis compared to luminal tumours and require specific drugs directed against the HER2 protein [222].

The normal-like subtype of breast cancer accounts for approximately 7.8% of all cancer cases and exhibits a similar immunohistochemistry status to the Luminal A subtype, as well as a gene expression profile resembling normal breast tissue. Studies have shown that the normal-like signature is significantly less expressed in metastatic tumours compared to primary tumours. Both normal-like and Luminal A signatures exhibit a negative correlation between time to tumour recurrence and the extent of gene expression changes between primary and metastatic disease [223].

TNBC is characterized by ER-negative, PR-negative, and HER2-negative. TNBC constitutes approximately 20% of all breast cancer cases. Within the TNBC subtype, several additional

subgroups have been identified, including basal-like (BL1 and BL2), claudin-low, mesenchymal (MES), luminal androgen receptor (LAR), and immunomodulatory (IM) [224]. Each of these subgroups exhibits unique clinical outcomes, phenotypes, and pharmacological sensitivities. TNBC is known for its aggressive behaviour, early relapse, and a greater tendency to present at advanced stages. It is characterized by a high proliferation rate, alterations in DNA repair genes, and increased genomic instability. Histologically, TNBC is a poorly differentiated, highly proliferative, and heterogeneous neoplasm, encompassing subsets with variable prognosis. Immunohistochemically, TNBC can be further subdivided into basal and non-basal subtypes. The basal TNBC is characterized by the expression of cytokeratins CK5 and CK6, as well as EGFR1, while the non-basal TNBC does not express cytokeratins CK5 and CK6 [225].

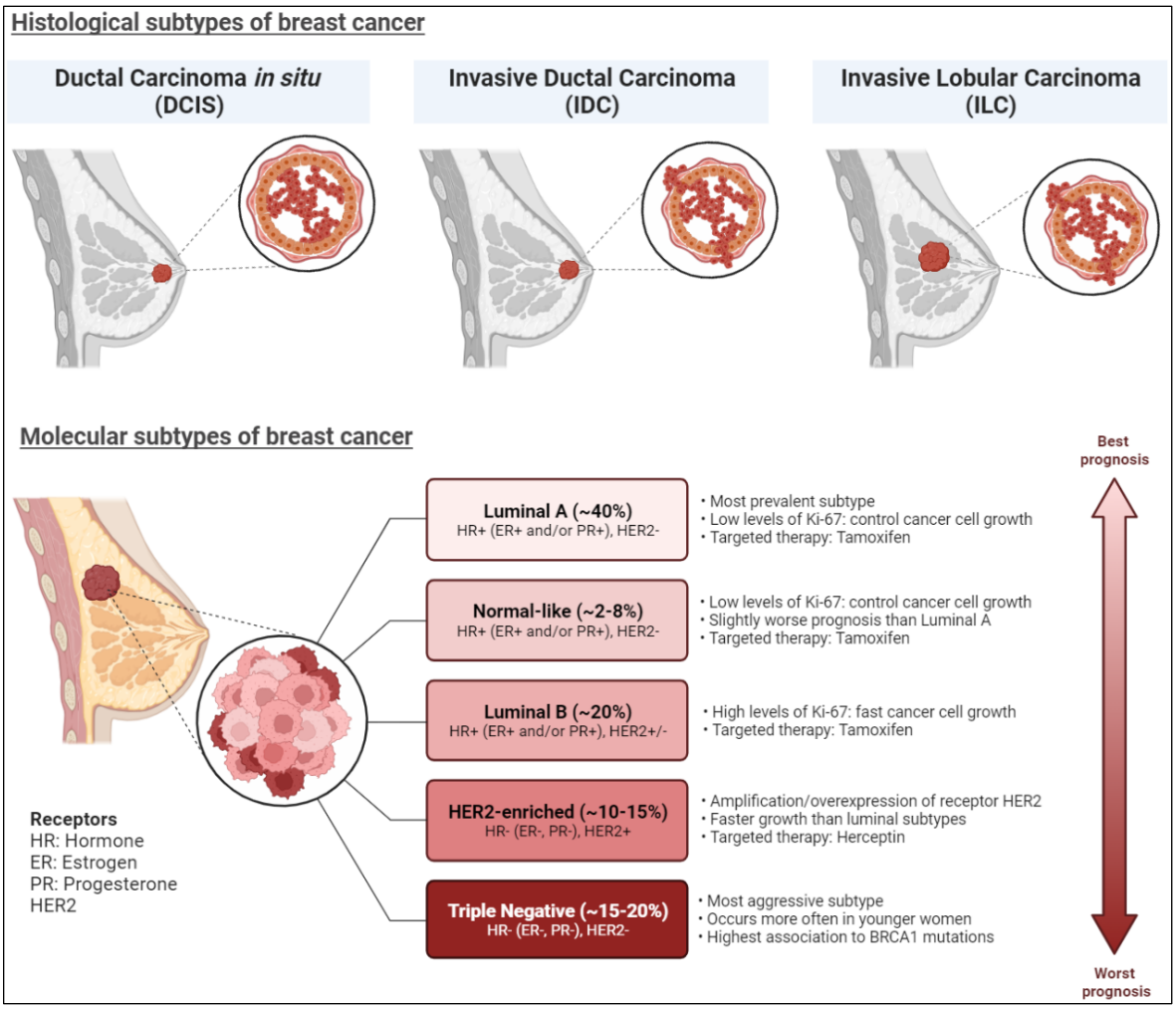


Figure 4.2. Main histological (up) and molecular (down) subtypes of breast cancer. Graph showing the overview of major histological subtypes and molecular subtypes of breast cancer, based on histological features and immunohistochemical expression of estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), and the proliferation marker Ki67. Information collected from [226, 227].

4.2.3 Breast cancer therapies

The currently treatment options for breast cancer include surgery [228], radiation therapy [229, 230], chemotherapy [231], endocrine (hormone) therapy [232, 233], and targeted therapy [234, 235]. These therapeutic approaches are used in various combinations depending on the specific characteristic of the tumour and the individual patient.

Surgery is a primary treatment modality for localized breast cancer [228]. Following surgery, adjuvant therapy is often administered to reduce the risk of recurrence and minimize the changes of metastases [236]. Radiation therapy is commonly employed to target cancer cells that may have been missed during surgery, thus reducing the likelihood of local cancer recurrence.

Radiation therapy involves the targeted exposure of cancer cells to high levels of radiation. Reports have shown that radiation therapy is associated with a decreased risk of breast cancer-related mortality, although this benefit may be offset by increased risk of mortality from vascular causes [237]. When used in combination with chemotherapy, radiation therapy after surgery can help reduce tumour size. However, it is important to note that radiation therapy may have some side effects, including decreased sensation in the breast tissue or under the arm, skin problems in the treated area, and in some cases the skin may become moist and weepy toward the end of treatment [238].

The benefits of chemotherapy are more significant in ER-negative tumours. Chemotherapy is recommended for the majority of TNBC, HER2-positive breast cancers, and in high-risk luminal tumours [228]. In ER-positive tumours, chemotherapy plays a role in inducing ovarian failure, contributing to its therapeutic effects [239].

Endocrine therapy has been extensively studied and is a well-established form of personalized medicine for the treatment of estrogen receptor-positive breast cancer. By lowering estrogen levels and inhibiting cancer growth, endocrine therapy remains a fundamental systemic treatment approach for hormone receptor-positive breast cancer in the adjuvant, metastatic, and sometimes neoadjuvant settings. In recent years, several novel estrogen receptor blockers, such as fulvestrant, as well as agents designed to overcome resistance to endocrine therapy (such as cyclin-dependent kinase 4/6 inhibitors and of mTOR inhibitors) have emerged as valuable treatment options and have gained significant clinical recognition [233].

Targeted therapies for breast cancer involve substances or drugs that interfere with specific molecules responsible for tumour cell proliferation and survival, thereby blocking the growth of cancer cells [234, 235]. The advancements in molecular biology and immunotherapy have led to the development of highly targeted therapies tailored to the specific pathophysiology of different

types of breast cancers. These recent breakthroughs have contributed to more efficient and precise treatment approaches for breast cancer patients. For example, selective estrogen receptor modulators (SERMs) and aromatase inhibitors are targeted therapies used to suppress tumour growth in estrogen-dependent breast cancers. Tamoxifen, the first drug to be approved for estrogen positive metastatic breast cancer, has been shown to reduce recurrences by approximately 40%-50% [240]. In addition, targeted therapies have been approved for VEGF and the drug bevacizumab has demonstrated effectiveness in the treatment of advanced metastatic breast cancer, particularly when used in combination with paclitaxel or docetaxel [241, 242].

4.2.4 Natural products in treating Breast cancer

Natural products have long been recognised as a valuable source of bioactive compounds with therapeutic potential, and they have played a significant role in the discovery of anticancer drugs. Over the past decades, extensive efforts have been made to isolate novel natural products from various sources such as microbes, plants, and marine organisms, in order to assess their anticancer properties and understand their mechanisms of action [243]. Numerous compounds with anticancer potential or unique structural advantages for drug development have been identified, such as irinotecan, vincristine, etoposide, and paclitaxel.

In recent years, natural products have gained considerable attention in breast cancer research due to their advantageous characteristics, including low side effects, low toxicity, and promising efficacy in multitarget therapy [244]. Various classes of natural products, such as alkaloids, flavonoids, terpenoids, and phenylpropanoids, have been extensively investigated for their potential clinical applications in the treatment of breast cancer [244-247]. As examples, noscapine, a phthalideisoquinoline alkaloid derived from *Papaver somniferum* (opium poppy), has demonstrated the ability to suppress the viability of MCF-10F, MCF-7 and MDA-MB-231 cell lines in a dose- and time-dependent manner. Its mechanism of action involves upregulation of Bax protein expression in MCF-10F, MCF-7 and MDA-MB-231 cells, while downregulation of Bcl-2 expression in the MCF-7 and MDA-MB-231 cells [248]. Epigallocatechin-3-gallate (EGCG), a catechin commonly found in green tea, which has shown potential as a chemopreventive dietary agent against breast cancer via influencing cell cycle progression and suppressing various intracellular signalling pathways implicated in breast cancer pathogenesis [249].

4.2.5 Connectivity map (CMap)

CMap is a systematic approach that has been applied in pharmacological research to define drug-disease connections [250]. The concept of CMap is based on the idea of characterizing various

biological states, including physiological conditions, diseases, and experimental perturbations induced by chemicals or genetic manipulations, through genomic signatures. The goal of this tool is to build a comprehensive and publicly accessible database of drug and gene signatures, coupled with the development of pattern-matching tools to identify similarities among these signatures. By leveraging this resource, researchers investigating drug candidates, genes, or disease states can compare their respective signatures to the database, enabling the discovery of unexpected connections, and this approach parallels the process of comparing DNA sequences to the GenBank database to identify similar genes. Examples of the use of CMap include identifying the anthelmintic drug parbendazole as an inducer of osteoclast differentiation [251], discovering celastrol as a leptin sensitizer [252], identifying compounds targeting COX2 and ADRA2A as potential treatments for diabetes [253], finding small molecules that can mitigate skeletal muscular atrophy and spinal muscular atrophy [254], and generating new therapeutic hypotheses for the treatment of inflammatory bowel disease [255] and cancer [256].

4.3 Aims

We have demonstrated the potential effect of *Limonium Sinense* against breast cancer in our previous findings. The aim of this study is to investigate the biological effects of *Limonium Sinense* and its potential mechanisms by using integrated bioinformatic analysis and cellular experiments.

4.4 Methods

4.4.1 Gene Expression Omnibus (GEO) dataset analysis

We screened GEO datasets on human cells treated with herbal extracts/compounds, by searching the keywords "(Traditional Chinese Medicine) AND (herb) AND (herbal) AND (medicinal) AND (medicinal plant)" and publication dates before 18/01/2022 in the National Centre for Biotechnology Information (NCBI) GEO platform. Initially, 673 datasets were identified. Then we only included datasets that met the following criteria: **1)** mRNA expression data; **2)** human sapiens samples; **3)** natural products derived from medicinal plants; **4)** minimum 2 biological replicates. Duplicate datasets were removed. Datasets detected less than 10,000 genes were excluded to balance the number of analysed genes and sample size. Microarray probe IDs were translated to gene symbols according to the GPL annotation files provided in the GEO database. Probes mapped to multiple gene symbols were removed and genes mapped to multiple probe IDs were summarized by calculating the mean. Only genes that are present across all the platforms remained for further analysis. A HIF score for each sample was calculated using Gene Set Variation Analysis (GSVA) to

determine the HIF activity. The compounds that demonstrated an upregulation or downregulation of the HIF score activity were selected for further analysis.

4.4.2 Molecular docking analysis

The crystal structure of *PHD2* (PDB ID: 2G19) was downloaded from Protein Data Bank (https://www.rcsb.org/). Compounds structure of gallic acid was downloaded from PubChem (https://pubchem.ncbi.nlm.nih.gov/) in ".mol2" file. The GOLD (Genetic Optimisation for Ligand Docking) software (*version 2022.3.0*) was used for the molecular docking experiment to predict the conformation. GOLD is an automated ligand docking program utilizes a genetic algorithm (GA) to explore the flexibility of ligand conformations within the context of partial flexible proteins, with a specific focus on ensuring that the ligand displaces loosely bound water upon binding [121]. At the start of a GOLD run, external van der Waals (vdW) energies are cut off when E_{ij} > van der Waals k_{ij} , where k_{ij} is the depth of the vdW well between atoms *i* and *j*. At the end of the run, the cut-off value is FINISH_VDW_LINEAR_CUTOFF. This allows a few bad bumps to be tolerated at the beginning of the run. Similarly, the parameters Hydrogen Bonding and FINAL_VIRTUAL_PT_MATCH_MAX are used to set starting and finishing values of max_distance (the distance between donor hydrogen and fitting point must be less than max_distance for the bond to count towards the fitness score). This allows poor hydrogen bonds to occur at the beginning of a GA run. Briefly, compounds were docked into the binding sites using the GOLD software, docking procedures were performed using the default setting with 100 GA runs of ligands. For each GA run, a maximum of 125,000 operations were performed. The GoldScore was selected for the scoring function. When the top 10 solutions possessed RMSD values within 1.5 Å, docking was terminated. Other parameters were all defaulted. The best conformation was selected and visualized by using Discovery Studio (*version 2021 Client*).

4.5 Results

4.5.1 Bioactive extracts from *Limonium Sinense* show a strong growth inhibitory effect

Traditionally, *Limonium Sinense* has been commonly extracted using boiling water to make an aqueous extract for oral uptake [7, 16], indicating that the active compounds may be present in the water extract of *Limonium Sinense*. On the other hand, our previous network pharmacology findings highlighted the significant role of Apigenin and Luteolin within *Limonium Sinense* in combating breast cancer. Since Apigenin and Luteolin are both poorly soluble in water but well present in the ethanol extract [257, 258], so we utilized both water and ethanol as solvents to

extract compounds from *Limonium Sinense* in this analysis, aiming to capture the potential therapeutic constituents present in both extracts.

To assess the biological effects of extracts from *Limonium Sinense*, multiple cell lines, including an immortalized human breast epithelial cell line MCF10A and seven breast cancer cell lines (BT20, MDA-MB-157, MDA-MB-231, MDA-MB-468, HCC1395, HCC1806 and HCC1937), were treated with *Limonium Sinense* water extracts (LSW) or ethanol extracts (LSE). A cell viability assay was performed to evaluate the impact of the extracts on cell growth. Compared to the control group, both LSW (Figure 4.3A, B) and LSE (Figure 4.3C, D) demonstrated a strong inhibition of growth in all the cell lines treated with LSW and the majority of cell lines treated with LSE in a dose-dependent manner at 24- or 48-hours post-treatment. Obviously, LSW exhibited a more potent growth inhibitory effect on the treated breast cancer cell lines compared to LSE. Among these tested cell lines, the MDA-MB-468 cells displayed the most pronounced inhibition by both LSW and LSE treatments, and thus were selected for subsequent cellular experiments in this study.

4.5.2 Bioactive extracts from *Limonium Sinense* inhibit cell growth in 2D and 3D cultures

Three-dimensional (3D) cell cultures have been recognised to better mimic *in vivo* conditions compared to two-dimensional (2D) monolayer cell cultures [259]. To further confirm the effects of LSW and LSE on cell viability, a 3D mammosphere assay of MDA-MB-468 cells treated with LSW or LSE was performed (Figure 4.4A). Images of spheres were analysed for sphere formation efficiency and sphere volume, and cell viability was determined using a Cell-Titer Glo® assay. The results revealed a significant decrease in cell viability (Figure 4.4B), sphere volume (Figure 4.4C) and sphere formation efficiency (Figure 4.4D) in LSW- and LSE-treated MDA-MB-468 cells in 3D culture. These findings demonstrate the strong growth inhibitory effect of the bioactive extracts from *Limonium Sinense*. Notably, LSW exhibited a stronger effective than LSE on sphere development and cell viability, suggesting that the water extract of *Limonium Sinense* may serve as the primary bioactive component responsible for its pharmacological functions.

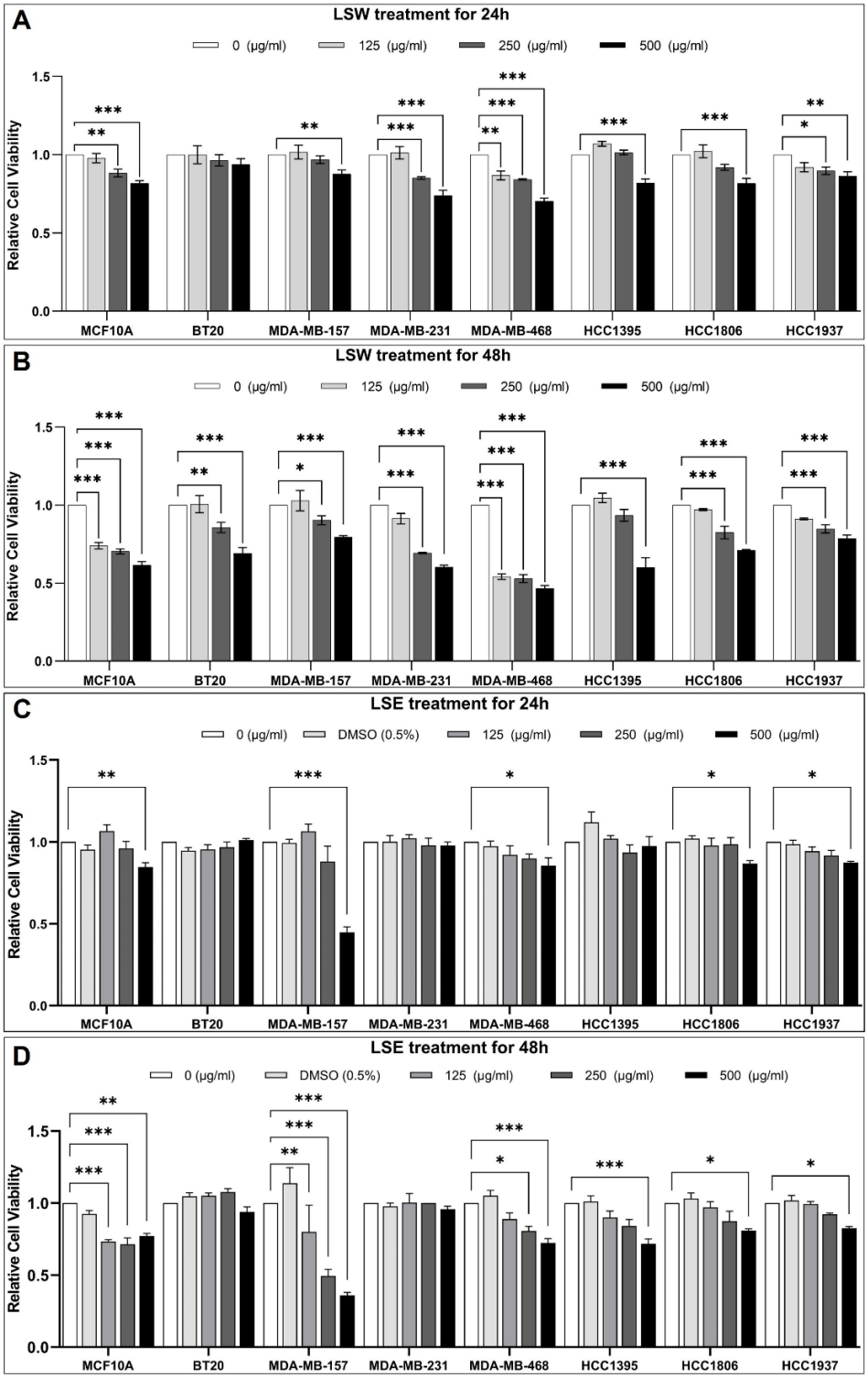


Figure 4.3. Effects of bioactive extracts from *Limonium Sinense* on cell viability. (**A** and **B**) Graphs showing relative cell viability in multiple cell lines treated with *Limonium Sinense* water extract (LSW) at the indicated concentration for 24 (**A**) or 48 hours (**B**). (**C** and **D**) Graphs showing relative cell viability in multiple cell lines treated with *Limonium Sinense* ethanol extract (LSE) at the indicated concentration for 24 (**C**) or 48 hours (**D**). Cell-Titer Glo® assay was performed to measure cell viability. Data are mean ± SD; n = 3 samples per group. ns, not significant; **P* < 0.05; ***P* < 0.01 and ****P* < 0.001 by Dunnett's multiple comparisons test.

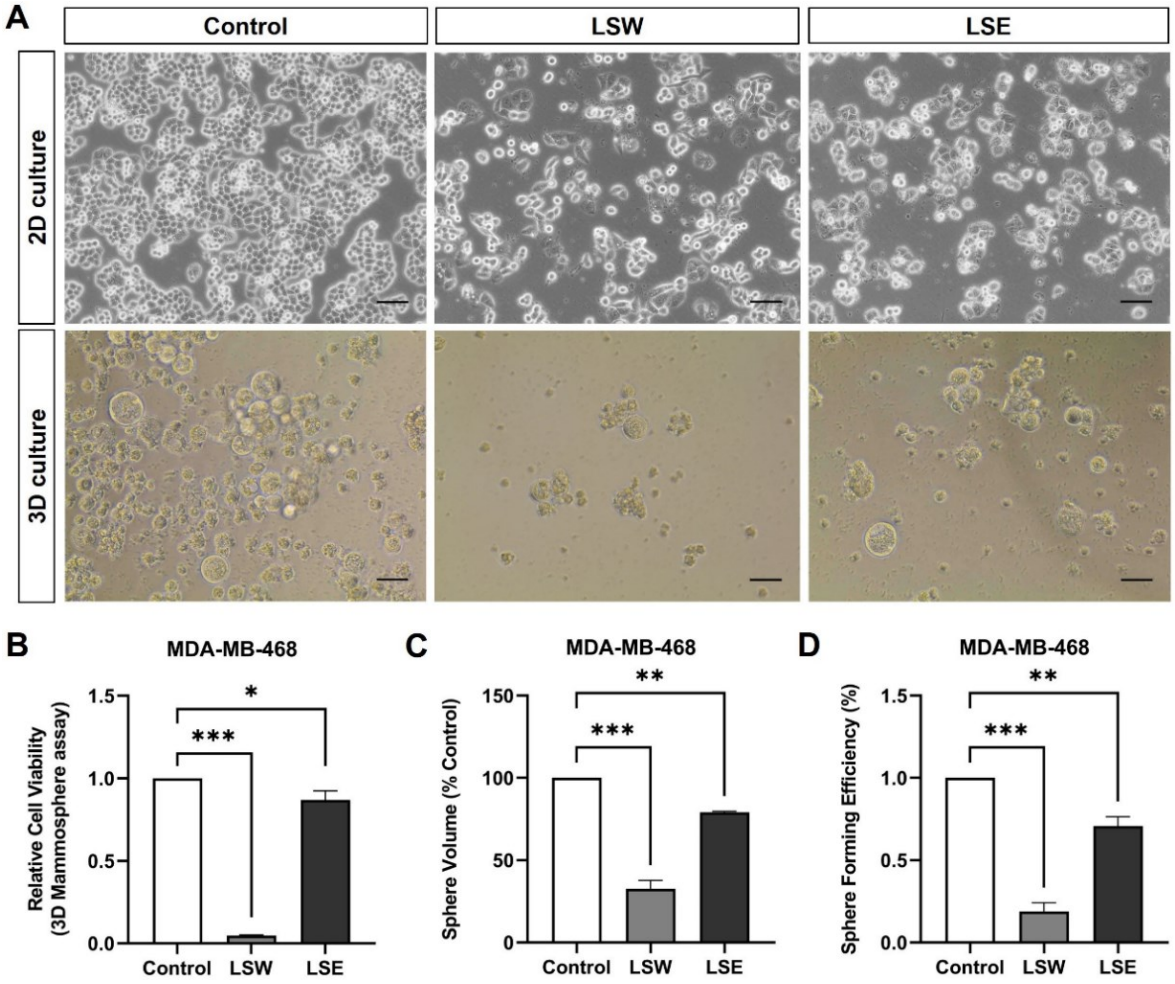


Figure 4.4. Effects of LSW and LSE on MDA-MB-468 cells in 2D and 3D cultures. (**A**) Representative phase contrast microscopy images showing the morphology changes of MDA-MB-468 cells treatment with LSW or LSE in 2D and 3D cultures. Scale bar: 50 µm. (**B**-**D**) Bar plots showing the cell viability (**B**), sphere volume (**C**) and sphere formation efficiency (**D**, Cell-Titer Glo® assay) in MDA-MB-468 cells with the indicated treatment cultured in 3D. Data are mean ± SD. n = 3 samples per group. ***$P < 0.001$; ****$P < 0.0001$ by Dunnett's multiple comparisons test.

4.5.3 Global transcriptomic changes in MDA-MB-468 cells exposed to LSW

To determine how cells respond to bioactive extracts from *Limonium Sinense*, we characterized the global transcriptomic changes in MDA-MB-468 cells exposed to LSW by performing RNA-Seq. t-distributed stochastic neighbour embedding (tSNE) analysis showed a clear separation between control and LSW-treated samples (Figure 4.5A). Genes with adjusted *P* value less than 0.05 and |Log2FoldChange| above 1 were considered as DEGs. A total of 987 DEGs were identified, including 456 upregulated and 531 downregulated genes (Figure 4.5B; Appendix B, Supplementary Table 4.1), and the hierarchical clustering showed that DEGs were grouped into two major clusters (Figure 4.5C).

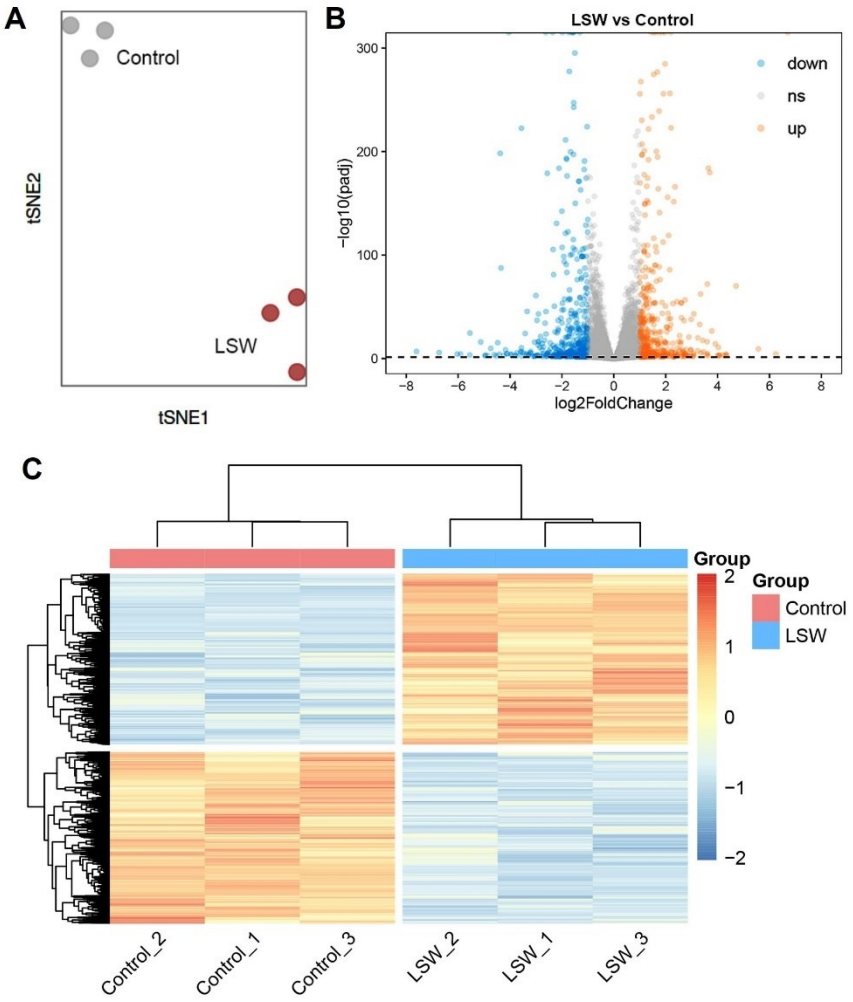
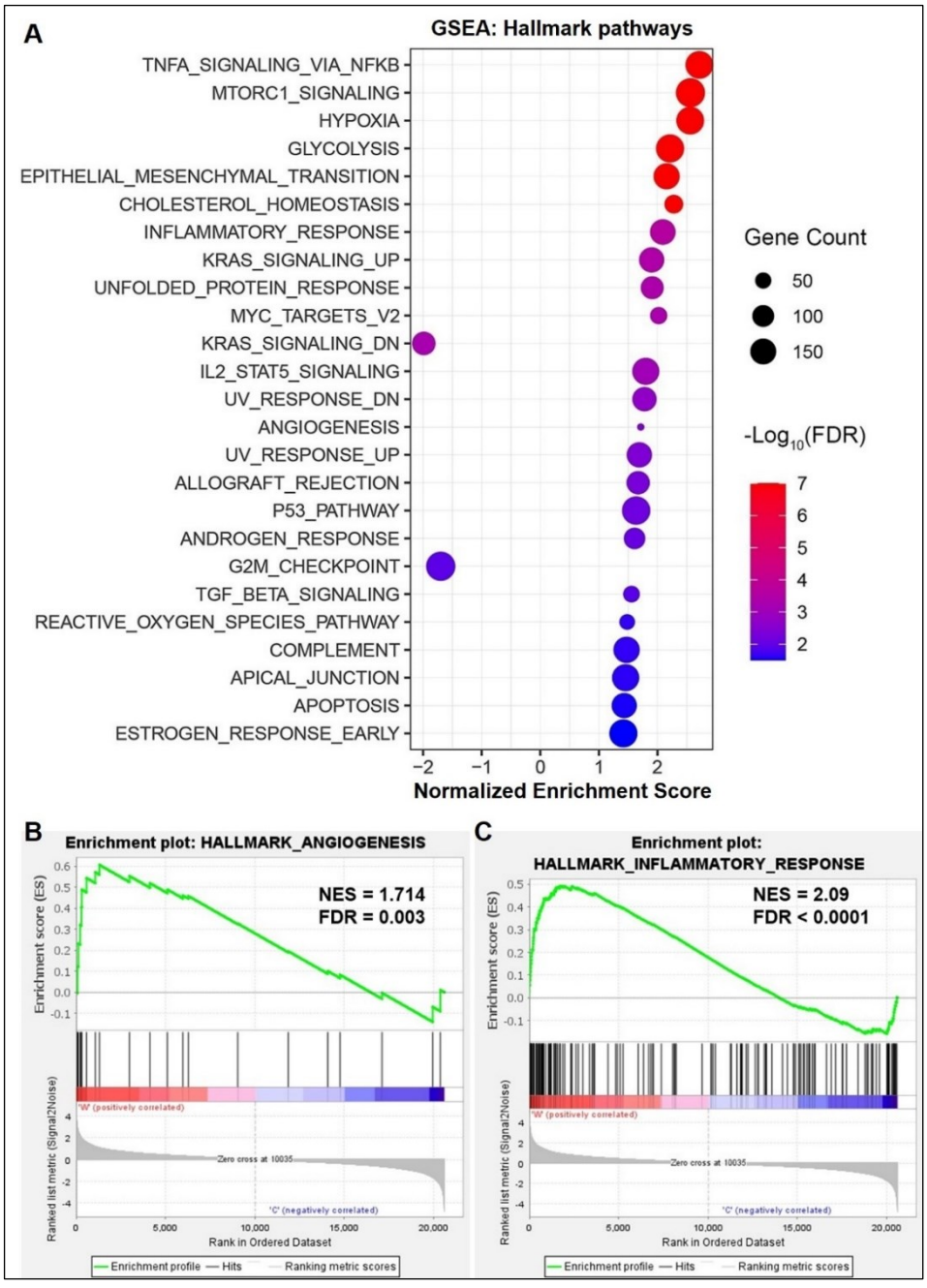


Figure 4.5. Gene expression changes in MDA-MB-468 cells exposed to LSW. (**A**) tSNE plot showing a clear sample separation for control vs. LSW-treated MDA-MB-468 cells (n = 3 samples per group). (**B**) Volcano plot showing up and down-regulated genes in LSW-treated MDA-MB-468 cells. Log2FoldChange in *x*-axis and -Log10(P_{adj}) in *y*-axis. Orange indicates up-regulation and blue down-regulation. (**C**) Heatmap showing differentially expressed genes (DEGs) in LSW-treated MDA-MB-468 cells. Genes with adjusted *P* value less than 0.05 and |Log2FoldChange| above 1 were considered as DEGs. Colour key represents the normalized z-scores. ns: not significant.

4.5.4 GSEA hallmark analysis of MDA-MB-468 cells treated with LSW

We then conducted Hallmark analysis by using GSEA to explore the molecular pathways influenced by the treatment with LSW in MDA-MB-468 cells (Figure 4.6A; Table 4.1). The analysis revealed that several disease-related pathological pathways were identified, including hypoxia (normalized enrichment score, NES = 2.558; FDR < 0.00001) and G2M_checkpoint (NES = -1.699; FDR = 0.011). In addition, we noticed that angiogenesis (Figure 4.6B; NES = 1.71; FDR = 0.009) and inflammatory response (Figure 4.6C; NES = 2.09; FDR < 0.0001) hallmark pathways were also significantly enriched. These findings are consistent with previous reports on the blood enrichment and immunomodulatory functions of *Limonium Sinense*.



(Figure legend is provided on the following page)

Figure 4.6. GSEA analysis of MDA-MB-468 cells upon treatment with LSW. (**A**) Scatter plot showing Gene Set Enrichment Analysis (GSEA) in MDA-MB-468 cells treated with LSW. The sizes of circles represent gene count, and the colours of circles represent the -Log10 of the false discovery rate (FDR) values. Plots showing Angiogenesis (**B**) and inflammatory response (**C**) hallmarks in MDA-MB-468 cell upon treatment with LSW (500 µg/ml), respectively. Normalized Enrichment Score (NES) and false discovery rate (FDR) are indicated.

Table 4.1. GSEA Hallmark analysis of MDA-MB-468 cells upon treatment with LSW

Hallmarks	Gene Count	NES	*P*-value	FDR
TNFA_SIGNALING_VIA_NFKB	177	2.714407	0.0000001	0.0000001
MTORC1_SIGNALING	197	2.56215	0.0000001	0.0000001
HYPOXIA	177	2.55775	0.0000001	0.0000001
CHOLESTEROL_HOMEOSTASIS	69	2.278207	0.0000001	0.0000001
GLYCOLYSIS	182	2.213718	0.0000001	0.0000001
EPITHELIAL_MESENCHYMAL_TRANSITION	153	2.155615	0.0000001	0.0000001
INFLAMMATORY_RESPONSE	145	2.09	0.0000001	2.36E-04
MYC_TARGETS_V2	57	2.02	0.0000001	4.64E-04
UNFOLDED_PROTEIN_RESPONSE	112	1.91	0.0000001	4.12E-04
KRAS_SIGNALING_UP	140	1.9	0.0000001	3.71E-04
IL2_STAT5_SIGNALING	166	1.8	0.0000001	8.46E-04
UV_RESPONSE_DN	129	1.777626	0.0000001	0.001898534
ANGIOGENESIS	20	1.714153	0.009070295	0.003227153
UV_RESPONSE_UP	141	1.68994	0.0000001	0.003885103
ALLOGRAFT_REJECTION	116	1.669914	0.00248139	0.004937911
P53_PATHWAY	186	1.635362	0.0000001	0.006977666
ANDROGEN_RESPONSE	96	1.608708	0.002380953	0.007952092
TGF_BETA_SIGNALING	53	1.556079	0.011135858	0.012136469
REACTIVE_OXYGEN_SPECIES_PATHWAY	46	1.482388	0.037470724	0.023078514
COMPLEMENT	155	1.472136	0.002531646	0.023423934
APICAL_JUNCTION	167	1.456135	0.002570694	0.025044626
APOPTOSIS	139	1.43228	0.009950249	0.029081538
ESTROGEN_RESPONSE_EARLY	182	1.417617	0.008	0.03185364
KRAS_SIGNALING_DN	118	-1.988184	0.0000001	5.31E-04
G2M_CHECKPOINT	199	-1.699333	0.0000001	0.011173217

List of 25 significantly enriched hallmarks (FDR < 0.05) [260] from the molecular signature database (MSigDB) between the control group and the LSW treatment group. "Gene Count" is the number of genes in the gene set after filtering out those genes not in the expression dataset. "NES" is the normalized enrichment score, that is, the enrichment score for the gene set after it has been normalized across analysed gene sets. "*P*-value" is the nominal *P* value, that is, the statistical significance of the enrichment score. The nominal *P* value is not adjusted for gene set size or multiple hypothesis testing. The data list is organized here from maximal to minimal NES.

4.5.5 GO and KEGG pathway enrichment analysis of MDA-MB-468 cells treated with LSW

We next performed the KEGG pathway analysis and GO enrichment analysis. The GO enrichment analysis was further grouped into biological process (BP), cellular component (CC) and molecular function (MF). Top 10 ranked KEGG pathways and Top 5 ranked GO terms of each category enriched by up- and down-regulated DEGs were shown in Table 4.2 and Table 4.3, respectively. Interestingly, the hypoxia-inducible factor (HIF)-1 signalling pathway ($P < 0.0001$) in the KEGG pathway analysis was top enriched by the up-regulated genes in MDA-MB-468 cells treated with LSW, which aligns with the hypoxia hallmark observed in the GSEA analysis. Additionally, we noticed that the Hematopoietic cell lineage was strongly enriched in both the up- and down-regulated genes in MDA-MB-468 cells upon treatment with LSW, indicating the blood replenishing function of *Limonium Sinense* (Figure 4.7).

In the BP category of the GO enrichment analysis, "Positive regulation of angiogenesis" and "Negative regulation of growth" were among the top 10 ranked GO terms, further supporting the blood enrichment effect and cell growth inhibitory effect of *Limonium Sinense*. In the CC category, the enriched terms indicated that the extracellular space was prominently enriched in both the up- and down-regulated genes in MDA-MB-468 cells upon treatment with LSW. Regarding the MF category, the up-regulated genes were associated with iron ion binding, while the down-regulated genes were related to N,N-dimethylaniline monooxygenase activity (Figure 4.7).

Table 4.2. KEGG pathways enriched by up- and down-regulated DEGs of MDA-MB-468 cells upon treatment with LSW (Top 10, respectively)

Term	Gene Count	*P*-value	Fold Enrichment	DEG
HIF-1 signalling pathway	14	3.15E-07	6.15452454	Up
Steroid biosynthesis	7	4.59E-06	14.7708589	Up
Hematopoietic cell lineage	10	1.90E-04	4.85085678	Up
Cytokine-cytokine receptor interaction	16	5.69E-04	2.778762402	Up
Metabolic pathways	45	0.001669	1.557924881	Up
Retinol metabolism	7	0.003774	4.615893405	Up
Biosynthesis of antibiotics	13	0.004155	2.587886329	Up
ECM-receptor interaction	8	0.004252	3.880685424	Up
Drug metabolism - other enzymes	6	0.004281	5.504667911	Up
Complement and coagulation cascades	7	0.005481	4.281408376	Up
Hematopoietic cell lineage	8	0.001586	4.61716587	Down
Cytokine-cytokine receptor interaction	12	0.008467	2.479589078	Down
Mineral absorption	5	0.010763	5.705872595	Down
Amphetamine addiction	5	0.04107	3.803915063	Down
cAMP signalling pathway	9	0.042041	2.282349038	Down
Transcriptional misregulation in cancer	8	0.047135	2.405349884	Down

"**Gene Count**" is the number of genes belonging to an annotation term. "***P*-value**" is the modified Fisher Exact *P*-value, referring to one-tail Fisher Exact probability value used for gene-enrichment analysis. The smaller, the more enriched. "**Fold Enrichment**" is the ratio of the input genes involved in specific annotation term proportion on background associated genes in the same term proportion. "**DEG**" is the differentially expressed gene.

Table 4.3. GO terms enriched by up- and down-regulated DEGs of MDA-MB-468 cells upon treatment with LSW (Top 5 of each category)

Category	Term	Gene Count	*P*-value	Fold Enrichment	DEG
BP	positive regulation of angiogenesis	13	2.656E-06	5.752200264	Up
BP	leukocyte migration	13	4.947E-06	5.422155986	Up
BP	oxidation-reduction process	29	1.631E-05	2.492669943	Up
BP	response to drug	19	3.294E-05	3.18030303	Up
BP	extracellular matrix organization	15	3.32E-05	3.894248609	Up
BP	negative regulation of growth	5	0.0004001	13.76619118	Down
BP	excretion	6	0.0006436	8.48295024	Down
BP	cell activation	4	0.0010098	19.02237327	Down
BP	cholesterol homeostasis	7	0.0013539	5.721573209	Down
BP	neutrophil chemotaxis	7	0.0015895	5.548192202	Down
CC	extracellular space	54	2.825E-07	2.11151306	Up
CC	plasma membrane	112	2.425E-05	1.431472523	Up
CC	cell surface	25	0.0001106	2.42945204	Up
CC	extracellular region	52	0.0001833	1.70115966	Up
CC	integral component of plasma membrane	44	0.0013966	1.63781123	Up
CC	integral component of plasma membrane	56	2.431E-07	2.08448702	Down
CC	extracellular space	52	1.53E-06	2.033308873	Down
CC	extracellular exosome	84	1.588E-05	1.573932301	Down
CC	plasma membrane	110	6.633E-05	1.405910513	Down
CC	cell surface	25	0.0001106	2.42945204	Down
MF	iron ion binding	15	1.81E-06	5.015151515	Up
MF	growth factor activity	14	1.771E-05	4.420763187	Up
MF	structural molecule activity	17	2.883E-05	3.520758189	Up
MF	cytokine activity	12	0.0006878	3.487809917	Up
MF	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	7	0.0008233	6.282137161	Up
MF	N,N-dimethylaniline monooxygenase activity	4	0.0001258	35.38993711	Down
MF	calcium ion binding	26	0.0022705	1.924975659	Down
MF	protein dimerization activity	9	0.0074687	3.18509434	Down
MF	phosphatidylinositol phospholipase C activity	4	0.0137457	7.864430468	Down
MF	glycerol channel activity	3	0.0204893	13.27122642	Down

"**Category**" is the term in the annotation cluster. "**Gene Count**" is the number of genes belonging to an annotation term. "***P*-value**" is the modified Fisher Exact *P*-value, referring to one-tail Fisher Exact probability value used for gene-enrichment analysis. The smaller, the more enriched. "**Fold Enrichment**" is the ratio of the input genes involved in specific annotation term proportion on background associated genes in the same term proportion. "**DEG**" is the differentially expressed gene.

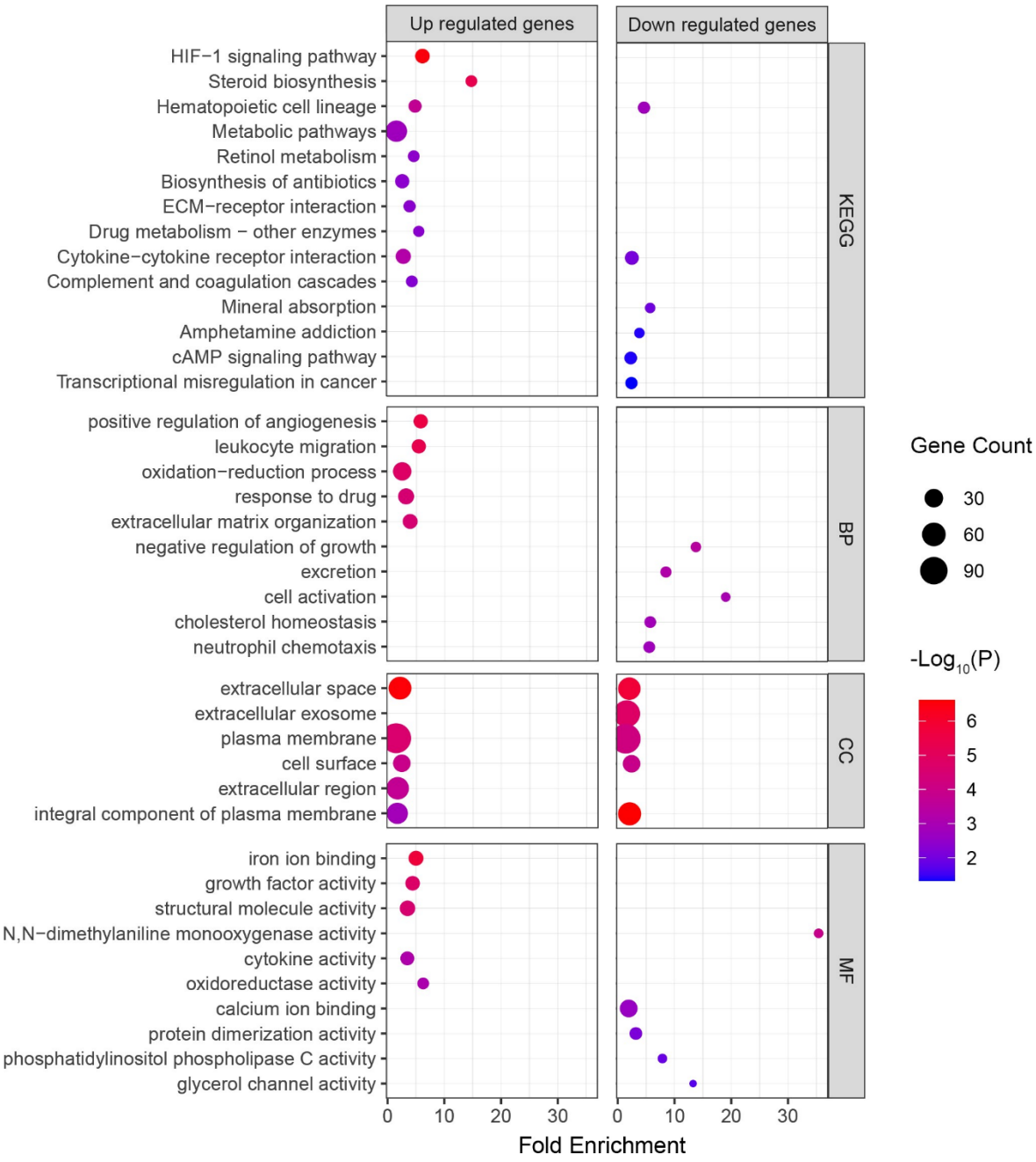


Figure 4.7. GO and KEGG enrichment analysis of MDA-MB-468 cells exposed to LSW. Scatter plot showing enriched Kyoko Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) terms from 3 categories (BP, biological process; CC, cellular component; and MF, molecular function) in MDA-MB-468 cells treated with LSW. The sizes of circles represent gene counts, and the colours of circles represent the -Log10 of the *P*-value.

4.5.6 Connectivity map (CMap) analysis in MDA-MB-468 cells exposed to LSW

We next utilized CMap to explore the small molecular compounds with similar activities to LSW. As mentioned previously, CMap is a systematic approach commonly used in pharmacological research to establish connections between drugs and diseases [250]. In this study, we performed CMap analysis using the top 150 up- and down-regulated genes from MDA-MB-468 cells exposed to LSW and compared them with a database of 2,837 compounds across 9 cell lines [261]. The connectivity score which summarizes the connectivity among signatures across different cell lines, ranged from + 98.17 to - 96.59 (Appendix B, Supplementary Table 4.2). For this analysis, compounds with a connectivity score above 90 were defined as "positive connectivity", while those with a connectivity score less than -90 were considered as "negative connectivity" [262]. In total, we identified 8 compounds with positive connectivity and 7 with negative connectivity (Figure 4.8A). Among those with positive connectivity, two are Tubulin inhibitors (vinorelbine and vincristine) and one HIF modulator (VU-0418946-1). The compound-target network analysis showed distinct clusters among the aforementioned compounds, including Tubulin inhibitors (vinorelbine and vincristine) targeting *TUBB* and HIF modulator (VU-0418946-1) targeting *HIF1A* (Figure 4.8B). The CMap analysis supports the findings from earlier pathway enrichment analysis, highlighting the effects of LSW on G2M_checkpoint in cell cycle and HIF/hypoxia.

The mechanism of actions (MOA) revealed by the CMap analysis indicated that compounds with positive connectivity to LSW were mostly PKC (Protein kinase C) activators and Tubulin inhibitors, whereas compounds with negative connectivity to LSW were primarily HSP (Heat Shock Protein) inhibitors and MEK (Mitogen-Activated Protein kinase) inhibitors (Figure 4.9A). Further examination of the related targets of the identified CMap compounds showed that genes including *PRKCD*, *PRKCE* and *TUBB* were predominantly shared by the positively connected compounds of LSW, while genes such as *MAP2K1* and *MAP2K2* were mainly shared by the negatively connected compounds of LSW (Figure 4.9B).

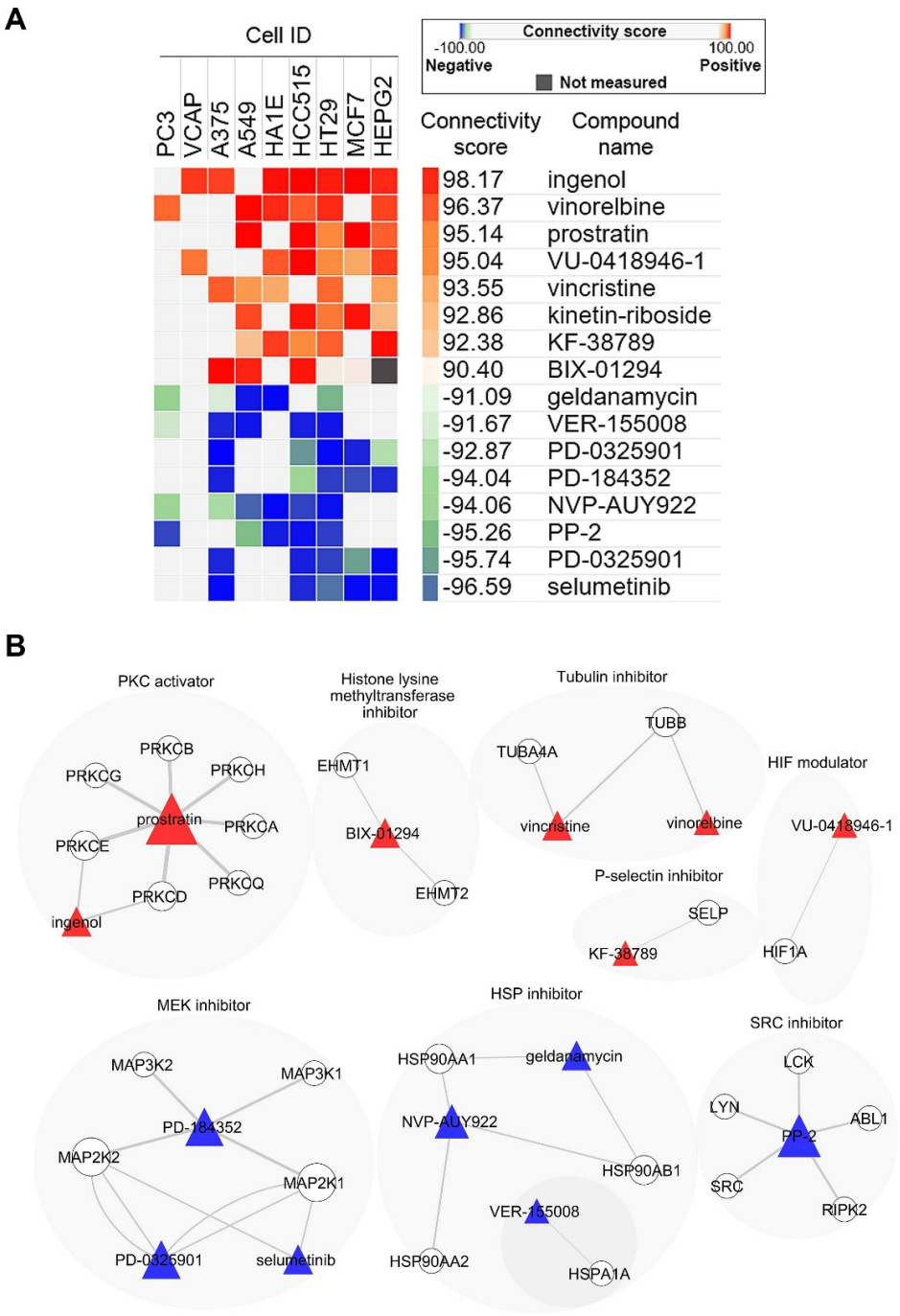


Figure 4.8. CMap analysis in MDA-MB-468 cells exposed to LSW. (**A**) Heatmap showing the connectivity score for the most significant compounds in 9 cell lines. Cell ID, connectivity score (and the colour key) and compound name are indicated. Compounds are considered significantly connected with the reference signature when the connectivity score is above 90 (positive) or below -90 (negative). Compounds are sorted by the decreasing order of their connectivity scores. (**B**) Graph showing the interaction network between compounds and their target genes. The colours and shapes represent the indicated compounds and their target genes. The sizes of the nodes indicated the degrees that the nodes connect to others, and the width of the lines represents the EdgeBetweenness of each gene.

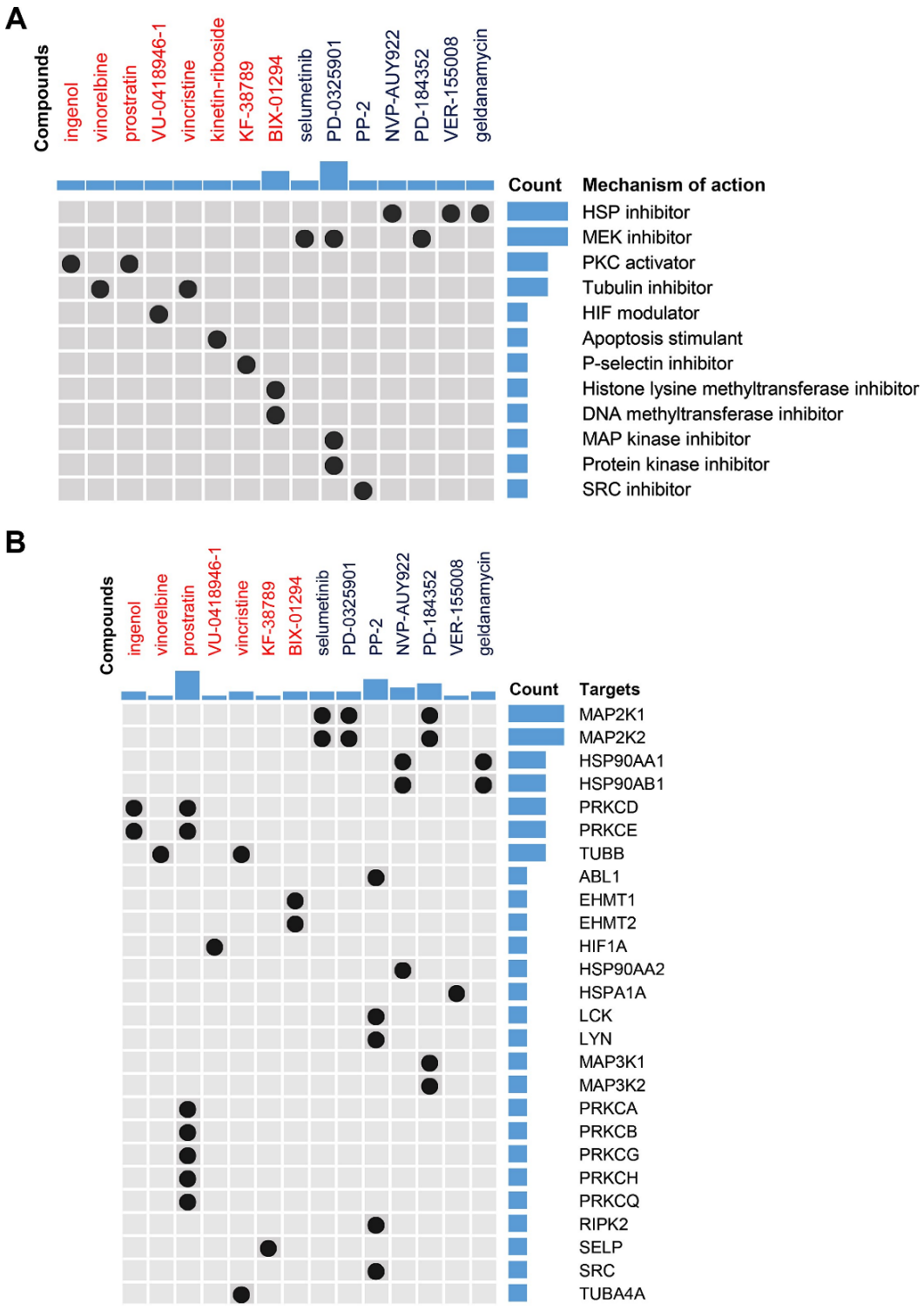


Figure 4.9. Connectivity map (CMap) analysis in MDA-MB-468 cells exposed to bioactive extracts from *Limonium Sinense*. (**A**) Heatmap showing each compound from the CMap that share Mechanism of actions (rows). Sorted by descending number of compounds with shared mechanism of actions. (**B**) Heatmap showing each compound from the CMap that share gene targets in (rows). Sorted by descending number of targets. Orange indicates "positive connectivity", and blue indicates "negative connectivity".

4.5.7 LSW induces G2/M phase arrest in cell cycle

As described earlier, the GSEA analysis identified hallmark “G2M_Checkpoint" negatively enriched in MDA-MB-468 cells upon treatment with LSW (Figure 4.10A). The cell cycle is a complex process involved in cell growth, proliferation, organismal development, DNA damage repair regulation, tissue hyperplasia in response to injury, and diseases such as cancer [263, 264]. Typically, the cell cycle is divided into two main stages, namely interphase and cell division. Interphase consists of G1, S and G2 phases, whereby following the G2 phase the cells enter mitosis (which is the M phase). During G1, individual cells grow, and their cellular content is duplicated. In the S phase, DNA is replicated, followed by further growth and preparation for cellular division in the G2 phase. Finally, in mitosis, the cell divides its contents into two equal daughter cells.

The cell cycle incorporates multiple checkpoints to prevent genomic instability and ensure accurate replication. The G2M checkpoint primarily responses for cell size and DNA replication. Cells with a defective G2M checkpoint, if they enter M phase before repairing their DNA, undergo apoptosis or death after cell division [265, 266]. Moreover, it has been reported that the cellular response to hypoxia and the cell cycle are interconnected through various mechanisms, enabling cells to respond to hypoxia in a manner that ensures survival and minimizes errors during cell division [267].

These findings suggest that the potential mechanism of *Limonium Sinense* inhibiting cell growth may be associated with the cell cycle. To investigate the effect of LSW on the cell cycle, flow cytometry analysis was conducted (Figure 4.10B). The results demonstrated a G2/M phase arrest in the cell cycle (Figure 4.10E; $P < 0.05$), with no significant changes observed in G0/G1 phase (Figure 4.10C) and S phase (Figure 4.10D), indicating that LSW may suppress the DNA replication, leading to G2/M phase arrest and ultimately cell growth inhibition.

The p53 gene is acknowledged for its anti-proliferative effect in various stress responses, including cell cycle arrest and apoptosis [268]. For instance, when cells undergo damage or cell proliferation is abnormal, activation of the p53 gene prompts cell cycle arrest and even cell apoptosis [269]. To better understand how LSW induces cell cycle arrest, we performed a p53 gene knockdown assay to investigate the involvement of p53 gene activation in this process. Surprisingly, our findings revealed that the absence of p53 did not impact cell viability upon LSW treatment (Figure 4.11A), indicating that LSW induces cell cycle arrest through p53-independent way.

Interestingly, subsequent to LSW treatment, there was an observed upregulation of p21 at the gene level (Figure 4.11B), coupled with a significant reduction in the expression of G2/M phase markers, including CDK1, CCNA2 (Cyclin A2) and CCNB2 (Cyclin B2) (Figure 4.11C). p21, a well-recognized cell cycle inhibitor that can arrest cell cycle progression at the G1/S and G2/M transitions by inhibiting

CDK/cyclin complexes [270]. Previous research has suggested that p21 can be activated through both p53-dependent and p53-independent pathways [271, 272]. These results indicate that LSW might potentially induce p21-mediated G2/M phase cell cycle arrest by disrupting the interaction of CDK1/Cyclin A2 and CDK1/Cyclin B2 complexes, thus leading to the inhibition of cell growth. Nevertheless, further validation in future studies remains essential.

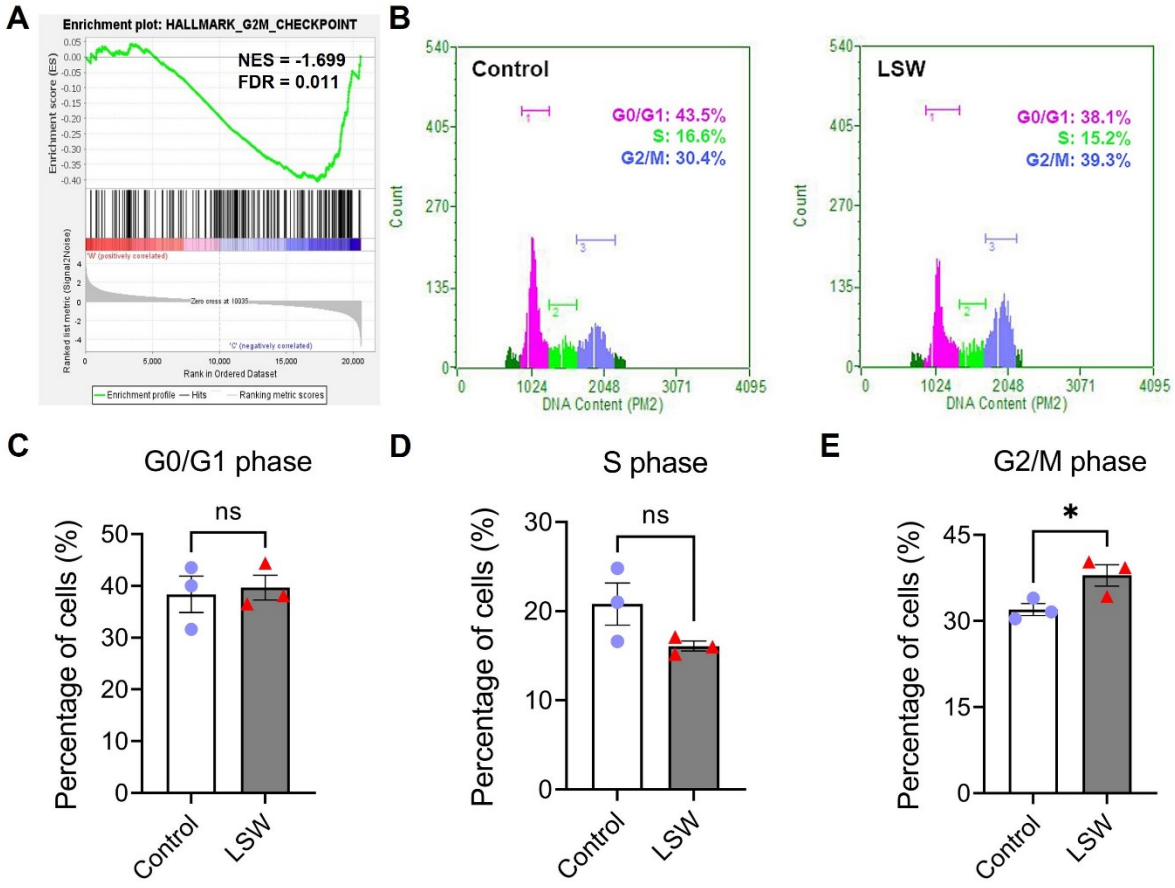


Figure 4.10. Effects of LSW on cell cycle in MDA-MB-468 cells. (**A**) Gene Set Enrichment Analysis (GSEA) plot showing an enrichment of Hallmark_G2/M_checkpoint in MDA-MB-468 cells treated with *Limonium Sinense* water extracts (LSW, 500 μg/ml). Normalized enrichment score (NES) and false discovery rate (FDR) are indicated. (**B**) Representative flow cytometry histograms of percentage of cells in G0/G1, S and G2/M phases of cell cycle from MDA-MB-468 treated with or without LSW (500 μg/ml) for 24 hours. (**C**-**E**) Graphs showing the percentage of cells in G0/G1 (**C**), S (**D**) or G2/M (**E**) phases. Data are mean ± SD. n = 3 samples per group. ns, not significant; *$P < 0.05$ by the Student's *t*-test.

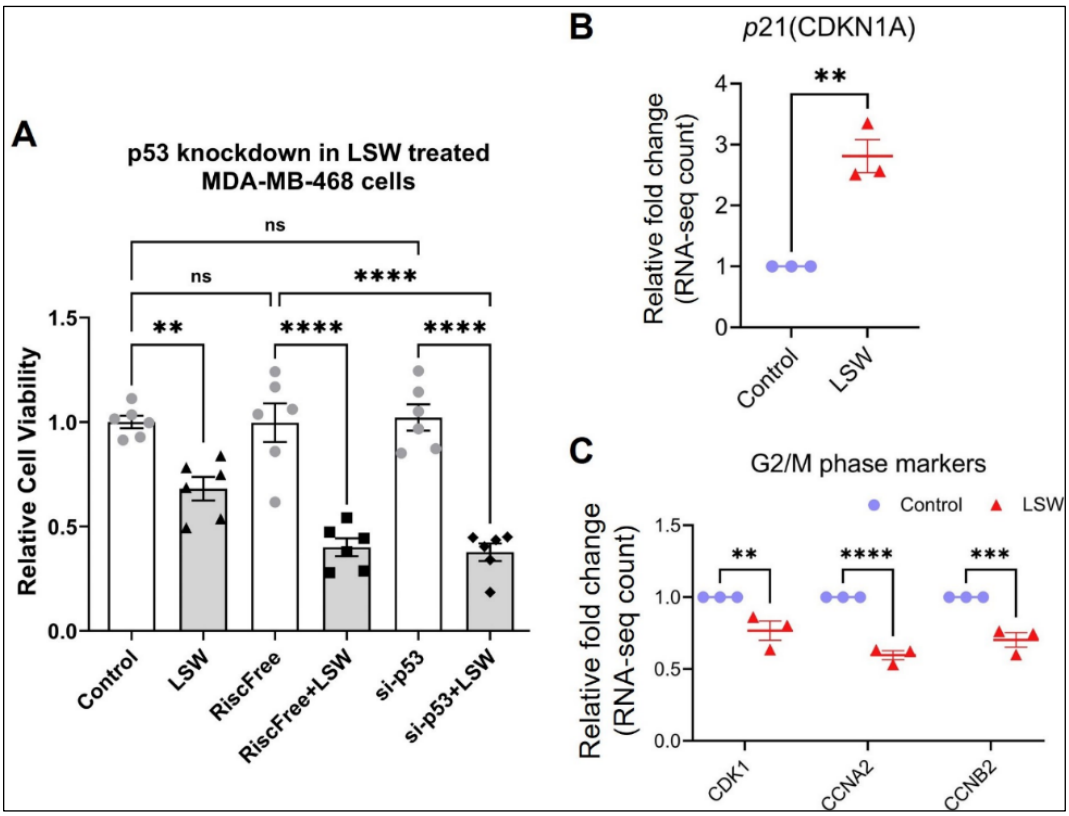


Figure 4.11. LSW-induced cell cycle arrest may through p53-independent p21 activation. (**A**) Bar plot showing knockdown of p53 gene did not affect the inhibitory effect of LSW on MDA-MB-468 cells. Cell-Titer Glo® assay was performed to measure cell viability. Data are mean ± SD; n = 6 samples per group. ns, not significant; ***P* < 0.01, ****P* < 0.001 and *****P* < 0.0001 by the Dunnett's multiple comparisons test. (**B**-**C**) Graphs showing the relative changes of p21, CDK1, CCNA2 (Cyclin A2) and CCNB2 (Cyclin B2) in RNA-seq count levels. Data are mean ± SD; n = 3 samples per group. ns, not significant; ***P* < 0.01, ****P* < 0.001 and *****P* < 0.0001 by the Student's *t*-test and Dunnett's multiple comparisons test.

4.5.8 LSW induces HIF activation

Our analysis strongly supports the involvement of LSW on HIF activation (Figure 4.12A). To evaluate HIF activity, we applied a 15-gene expression signature (*ACOT7*, *ADM*, *ALDOA*, *CDKN3*, *ENO1*, *LDHA*, *MIF*, *MRPS17*, *NDRG1*, *P4HA1*, *PGAM1*, *SLC2A1*, *TPI1*, *TUBB6* and *VEGFA*), which has been validated across multiple populations [84, 273], enabling defining HIF activity by calculating a HIF score for each sample using GSVA [83]. The expression levels of the 15 genes were visualized in a heatmap with upregulations observed in a majority upon LSW treatment (Figure 4.12B), and a significant increase of the HIF score was observed in LSW-treated samples compared to controls (Figure 4.12C; *P* < 0.001).

To validate these findings, the protein levels of HIF-1α was measured in MDA-MB-468 cells upon LSW treatment. The results revealed a significant induction of HIF-1α protein levels in MDA-MB-468 cells in a dose- (Figure 4.12D) and time-dependent manner (Figure 4.12E) upon LSW treatment. Additionally, the immunofluorescence staining of HIF-1α further confirmed a significant activation of HIF-1α following treatment with LSW (Figure 4.12F).

To assess the transcriptional activity of HIF, we employed a hypoxia response elements (HRE) reporter system [274]. Upon hypoxic conditions, HIF-α is upregulated and forms a complex with HIF-1β, which binds to the HRE of the gene promoters, leading to transactivation [275]. DMOG (Dimethyloxalylglycine), a non-specific 2 oxoglutarate (2OG) analogue that can stabilize and activate HIF [276], was used as a positive control. It was found that both DMOG and LSW treatment resulted in a significant increase in the HRE luciferase activity (Figure 4.12G; $P < 0.001$ and $P < 0.01$, respectively), indicating the activation of HRE and subsequent gene transcription induced by HIF-1α upon treatment with LSW.

Furthermore, we also examined the mRNA levels of HIF target genes such as *CA9* and *VEGFA* expressed in MDA-MB-468 cells upon LSW treatment. Our findings revealed that significant upregulation of both *CA9* and *VEGFA* genes upon treatment with LSW (Figure 4.12H; $P < 0.0001$ and $P < 0.05$, respectively). These results further confirm the activation of HIF-1α induced by the treatment of LSW.

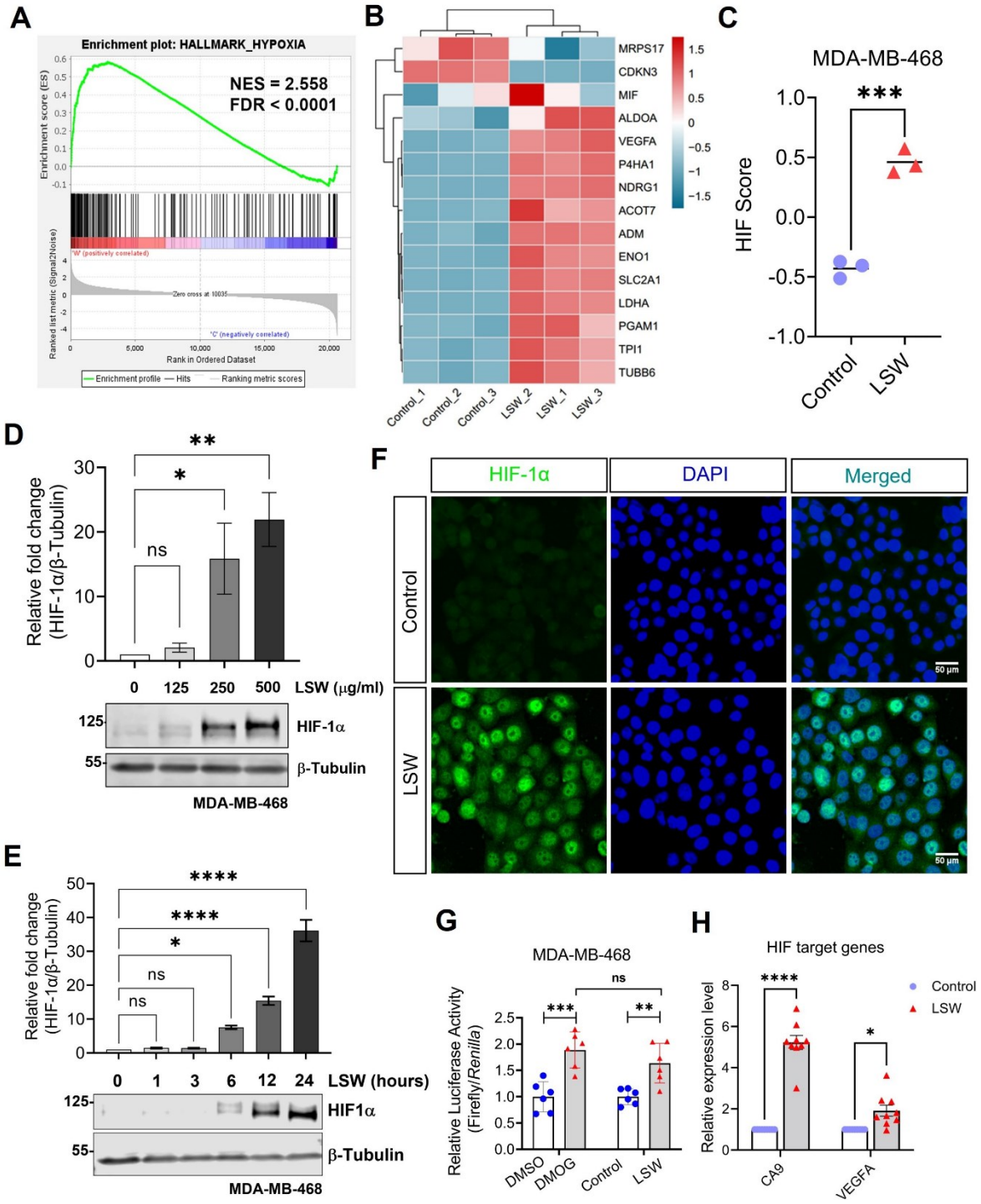


Figure 4.12. Effects of LSW on hypoxia-inducible factor (HIF) activation. (**A**) Gene Set Enrichment Analysis (GSEA) plot showing an enrichment of hallmark Hypoxia in MDA-MB-468 cells treated with LSW (500 μg/ml). Normalized Enrichment Score (NES) and false discovery rate (FDR) are indicated. (**B**) Heatmap showing the expressions of the 15 genes (*ACOT7*, *ADM*, *ALDOA*, *CDKN3*, *ENO1*, *LDHA*, *MIF*, *MRPS17*, *NDRG1*, *P4HA1*, *PGAM1*, *SLC2A1*, *TPI1*, *TUBB6* and *VEGFA*) used to calculate the HIF score in control vs. LSW-treated MDA-MB-468 cells. Red indicates up-regulation and blue down-regulation. (**C**) Graph showing HIF Gene Set Variation Analysis (GSVA) scores from control vs. LSW-treated MDA-MB-468 cells. Data are mean ± SD; n = 3 samples per group. ***P < 0.001 by the

Student's *t*-test. (**D** and **E**) Protein expressions of HIF-1α in MDA-MB-468 cells with the indicated treatment. β-Tubulin was used as a loading control. Graphs showing relative protein levels of HIF-1α. Data are mean ± SD; n = 3 samples per group. ns, not significant; **P* < 0.05; ***P* < 0.01; *****P* < 0.0001 by Dunnett's multiple comparisons test. (**F**) Immunofluorescence staining of HIF-1α (green) in MDA-MB-468 cells with the indicated treatment. 4'6-Diamidino-2-Pheylindole (DAPI) (blue) was used to stain nuclei. Scale bars: 50 µm. (**G**) Graph showing the Hypoxia Response Element (HRE) reporter assay in MDA-MB-468 cells with the indicated treatment. Values represent the relative fold change of Firefly luciferase in relation to *Renilla* luciferase, normalized against control (1.0). Data are mean ± SD; n = 6 samples per group. ns, not significant; ***P* < 0.01; ****P* < 0.001 by Dunnett's multiple comparisons test. (**H**) Graph showing fold change in mRNA levels of HIF-1α target genes (*CA9* and *VEGFA*) in MDA-MB-468 cells with the indicated treatment. β-actin-normalized mRNA levels in control cells were used to set the baseline value at unity. Data are mean ± SD; n = 9 samples per group. **P* < 0.05; *****P* < 0.0001 by the Student's *t*-test.

4.5.9 LSW activates HIF at the protein level

HIF-α is a key regulator of the cellular response to oxygen deprivation, with its activity governed by the Fe(II)- and 2OG-dependent oxygen superfamily. Under normal tissue oxygen levels, HIF-1α expression levels are tightly controlled. Iron- and oxygen dependent hydroxylation occurs at two specific proline residues (P^{402} and P^{564}) within the HIF-1α oxygen-dependent degradation domain (ODD) by prolyl hydroxylase domain (PHD) enzymes (PHD1-3). This modification results in the specific recognition of HIF-1α by the von Hippel-Lindau protein (VHL), making it for ubiquitination and subsequent degradation [277, 278]. Additionally, Factor inhibiting HIF (FIH), another Fe(II)- and 2OG-dependent dioxygenase, which hydroxylates HIF-1α at asparagine (N^{803}) residue, can block the binding of co-activators p300/CBP, leading to the inactivation of the COOH-terminal transactivation domain (C-TAD) and the negative regulation of HIF's transcriptional activity [279].

Mechanistically, we have checked the mRNA expression levels of *HIF1A* and its negative regulators PHDs (encoded by EGLNs) upon LSW treatment (Figure 4.13). We found that the mRNA level of *HIF1A* was not changed, while PHDs were not decreased by LSW treatment, suggesting that LSW treatment induces HIF-1α at the protein level and this is not mediated by down-regulating PHDs.

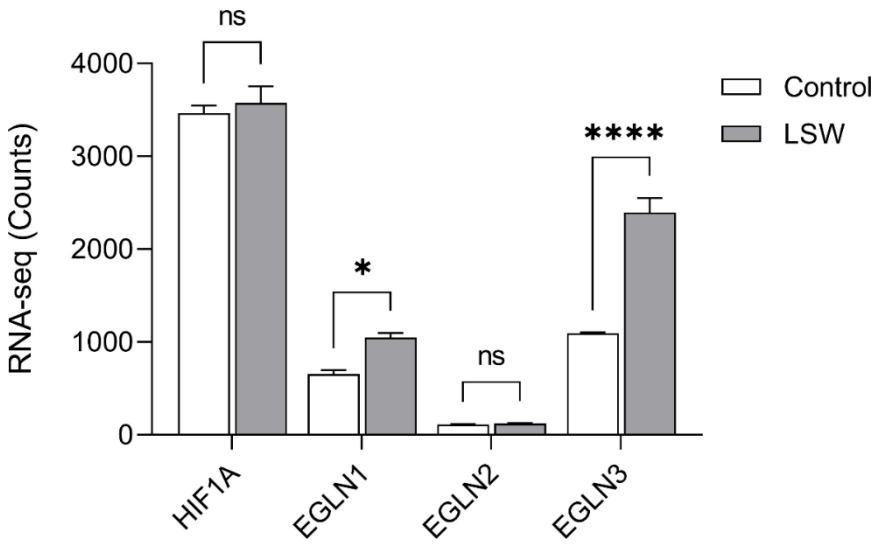


Figure 4.13. Effects of LSW treatment on the mRNA levels of HIF1A and EGLN1-3. Graph showing the mRNA level change in RNA-sequencing data of *HIF-1A*, *EGLN1*, *EGLN2* and *EGLN3*. Data are mean ± SD; n = 3 samples per group. **P* < 0.05; *****P* < 0.0001 by Dunnett's multiple comparisons test. ns, not significant.

4.5.10 Natural compounds/herbal extracts on HIF activity in Gene Expression Omnibus (GEO) dataset

To identify potential bioactive ingredient(s) within *Limonium Sinense* that are responsible for HIF activation, an integrated approach was adopted. We screened GEO datasets on human cells treated with herbal extracts or natural compounds. A comprehensive dataset comprising 871 samples from 36 GEO datasets was compiled (Table 4.4), consisting of 218 control samples and 653 samples treated with various compounds in different cell types. Among these datasets, there are 8 datasets showed an upregulation of the HIF score activity, while 7 datasets demonstrated a downregulation of the HIF score activity. Further details regarding the collection of GEO datasets are provided in Figure 4.14.

Based on the analysis of the obtained GEO datasets, a total of 31 natural compounds or herbal extracts were identified to exhibit the ability to activate HIF, demonstrated by an increase in the value of the HIF score (Figure 4.15A). Among these compounds, 7 were water soluble, and 2 of them were reported to be present in *Limonium Sinense* [25, 32], namely gallic acid and oleanic acid. Gallic acid is soluble in water, while the oleanic acid dissolves easily in ethanol [280]. Additionally, 5 natural compounds or herbal extracts showed an opposite effect on HIF activity, leading to a downregulation of the HIF score (Figure 4.15B). Among these compounds, 3 were water soluble.

However, none of the compounds with downregulated HIF score were found to be present in *Limonium Sinense*.

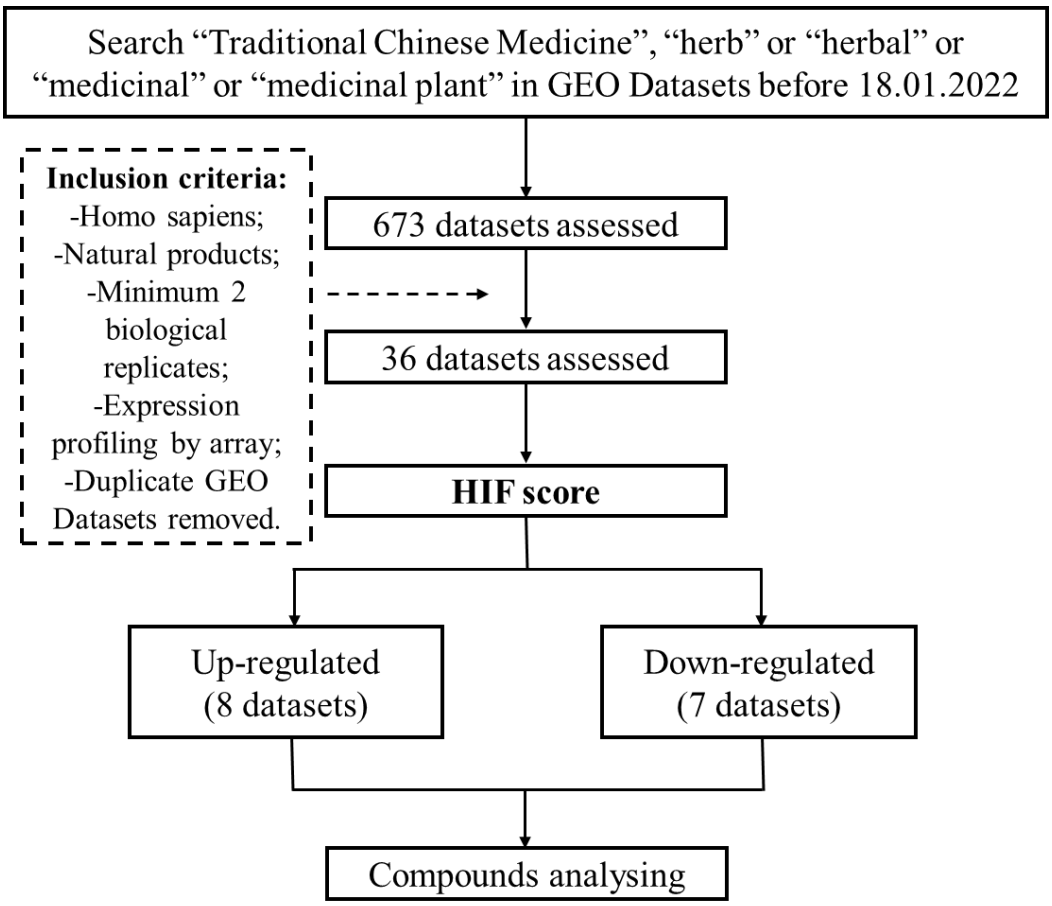


Figure 4.14. Flow chart of Gene Expression Omnibus (GEO) datasets collection. Details of the GEO datasets collection process are provided in the Methods section. Briefly, human cells treated with herbal extracts/ natural compounds were screened by searching for keywords "(Traditional Chinese Medicine) AND (herb) AND (herbal) AND (medicinal) AND (medicinal plant)" before 18/01/2022 on the National Centre for Biotechnology Information (NCBI) GEO platform. The initial datasets were then filtered based on the following criteria: 1) mRNA expression data availability; 2) human sapiens samples; 3) natural products derived from medicinal plants; 4) minimum 2 biological replicates. Duplicate datasets were removed, and datasets with less than 10,000 genes detected were excluded to balance the number of analysed genes and sample size. Microarray probe IDs were translated into gene symbols according to the GPL annotation files provided in the GEO database. Probes mapped to multiple gene symbols were removed, and genes mapped to multiple probe IDs were summarized by calculating the mean expression value. Only genes that are present across all the platforms remained for further analysis. A HIF score for each sample was calculated using Gene Set Variation Analysis (GSVA) to determine the HIF activity. Compounds that demonstrated an upregulation or downregulation of the HIF score activity were selected for further analysis.

Table 4.4. Details of the GEO datasets collected for the integrated analysis

Accession ID	Included sample information	Platform
GSE7848 [281]	Control: DMSO 6h (n=2), 24h (n=4). *Actein*: 20 mg/ml 6h (n=2), 40 mg/ml 6h (n=2), 20 mg/ml 24h (n=3), 40 mg/ml 24h (n=3).	Affymetrix Human Genome U133A 2.0 Array
GSE85871 [282]	Control: DMSO 12h (n=6). *TCM components*: 12h (n=206).	
GSE23610 [283]	Control: DMSO 0.001% 6hr (n=3). *Estradiol*: 0.1 μM 6h (n=3); *Ferulic acid*: 0.1 μM 6h (n=3), 1 μM 6h (n=3), 10 μM 6h (n=3); *Si-Wu-Tang*: 0.0256 mg/ml 6h (n=3), 0.256 mg/ml 6h (n=3), 2.56 mg/ml 6h (n=3).	Affymetrix Human Genome U133 Plus 2.0 Array
GSE6800 [284]	Control: DMSO 24h (n=2). *Cimicifuga*: 24h (n=2); *Estradiol*: 24h (n=2); *Tamoxifen*: 24h (n=2).	
GSE44596 [285]	Control: DMSO 24h (n=3). *SsnB*: 24h (n=3).	
GSE44597 [285]	Control: DMSO 24h (n=2). *SsnB*: 24h (n=2).	
GSE24743 [286]	Control: DMSO 3h (n=2). *Shikonin*: 100 nM 3h (n=2).	
GSE20928 [287]	Un-treated, 12h (n=3). *Calactin*: 0.15 mM 12h (n=3).	
GSE125743 [288, 289]	Un-treated: 24h (n=3), 48h (n=6). *CKI*: 2 mg 24h (n=3), 48h (n=6); *N_2*: 2 mg 24h (n=3), 48h (n=3); *N_3*: 2mg 24h (n=3), 48h (n=3); *OO*: 2mg 24h (n=3), 48h (n=3); *MOO*: 2 mg 24h (n=3), 48h (n=3); *N_Mac*: 2mg 48h (n=3); *N_Nme*: 2 mg 48h (n=3); *N_Omt*: 2 mg 48h (n=3); *N_Tri*: 2 mg 48h (n=3).	HiSeq X Ten
GSE130359 [290, 291]	*A431*: untreated_48h (n=3), vehicle_48h (n=3), *M231*: untreated_48h (n=3). *A431*: *CKI*_48h (n=3), *Doxorubicin*_48h (n=3); *CKI+Doxorubicin* 48h (n=3); *M231*: *CKI*_48h (n=3), *5-FU*_48h (n=3); *CKI+5FU* 48h (n=3).	
GSE78512 [292]	Control: 0h (n=3), 24h (n=3), 48h (n=3). *CKI*: 1mg 24h (n=3), 2mg 24h (n=3), 1mg 48h (n=3), 2mg 48h (n=3); *5-FU*: 24h (n=3), 48h (n=3).	Illumina HiSeq 2500
GSE116121 [293]	Control: 1h (n=3); *F1*: 10 μM 1h (n=2); *Rh1*: 10 μM 1h (n=2); *VEGF*: 2.5 nM 1h (n=2); *VEGF+F1*: 10 μM 1h (n=2); *VEGF+Rh1*: 10 μM 1h (n=2).	

Accession ID	Included sample information	Platform
GSE124715 [294]	*HEP*: untreated_0h (n=3), 24h (n=3), 48h (n=3); *MDA*: untreated_0h (n=3), 24h (n=3), 48h (n=3). *HEP*: VC_24h (n=3), VC_48h (n=3); *CKI*_2 mg_24h (n=3), *CKI*_2 mg_48h (n=3); *5-FU*_24h (n=3), *5-FU*_48h (n=3). *MDA*: VC_24h (n=3), VC_48h (n=3); *CKI*_2 mg_24h (n=3), *CKI*_2 mg_48h (n=3); *5-FU*_24h (n=3), *5-FU*_48h (n=3).	Illumina NextSeq 500
GSE156221 [295]	Control (n=3). *Cotyledon orbiculata extract* (n=3).	
GSE156445 [296]	Control (n=3). *Cissampelos pareira* 1 μg (n=3), 10 μg (n=3), 100 μg (n=3), 500 μg (n=3), 1000 μg (n=3).	Affymetrix Human Transcriptome Array 2.0
GSE95504 [297]	Control: DMSO (n=3). *Baicalein*: 40 μM (n=3), 80 μM (n=3).	
GSE128856 [298]	*BPH1* control (n=2), *WPMY-1* control (n=2). *BPH1 PAO* 250 μg/ml (n=2), *WPMY-1 PAO* 250 μg/ml (n=2).	
GSE112908 [299]	Control: DMSO 0.1% 48h (n=3). *Oridonin*: 10 μM 48h (n=3).	Illumina HiSeq 4000
GSE139929 [300]	Control (n=3). *APBBR1*: (n=3).	
GSE85746 [301]	Untreated: 0h (n=2), 6h (n=3), 12h (n=3), 18h (n=2), 24h (n=3), 48h (n=3). *CCL*: 6h (n=3), 12h (n=3), 18h (n=3), 24h (n=2), 48h (n=2).	Affymetrix Human Gene 2.0 ST Array
GSE110335 [302]	Control: DMSO 6h (n=3). *Glycyrrhetinic acid*: 40 μM 6h (n=3).	
GSE81835 [303]	Untreated (n=3). *Suduxing*: 0.001 μg/ml (n=3), 0.01 μg/ml (n=3).	Agilent-028004 SurePrint G3 Human GE 8x60K Microarray
GSE53415	Control: 2h (n=3), 4h (n=3), 8h (n=3). *Berberine*: 2h (n=3), 4h (n=3), 8h (n=3).	
GSE189697 [304]	Control (n=3). *14,15β-dihydroxyklaineanone*: 24h (n=3).	Illumina NovaSeq 6000
GSE182007	Untreated (n=3), vehicle (n=6). *Trichostatin A*: 0.64 μM (n=3); *Wortmannin*: 1.95 μM (n=6); *W-BR*: 500 μg/ml (n=3), 100 μg/ml (n=3), 20 μg/ml (n=3); *E-BR*: 133 μg/ml (n=3), 26.6 μg/ml (n=3), 5.32 μg/ml (n=3); *Saikosaponin D*: (n=63); *BR-Combination*: (n=9).	MGISEQ-2000RS

Accession ID	Included sample information	Platform
GSE164934 [305]	Control: placebo (n=13). *Herbal preparation*: (n=13).	Affymetrix Human Gene 1.1 ST Array
GSE99820 [306]	Control: vehicle (n=3). *WCE*: (n=3).	Illumina HiScanSQ
GSE100224	Control: vehicle (n=2). *WCE*: (n=2).	Illumina HumanHT-12 V4.0 expression beadchip
GSE86798 [307]	Control: (n=2); *IL-1β*: 1 ng/ml (n=2); *Indomethacin+IL-1β*: (n=2); *Gallic acid+IL-1β*: (n=6); *Piperine+IL-1β*: (n=4); *AVS023+IL-1β*: (n=6).	
GSE61926 [308]	Control: (n=8). *Rikkunshito*: (n=9).	Illumina humanRef-8 v2.0 expression beadchip
GSE42236 [309]	*SHSY5Y*_Control: (n=24); *IMR32*_Control: (n=3). *SHSY5Y_Gelsemium*: (n=24); *IMR32_Gelsemium*: (n=3)	NimbleGen Homo sapiens Expression Array
GSE64111 [310]	Control: DMSO (n=4). *Pachymic acid*: (n=4).	Affymetrix Human Gene Expression Array
GSE3983 [311]	Untreated: (n=4). *Agaricus*: (n=4).	ABI Human Genome Survey Microarray v2.0
GSE24191 [312]	Untreated: (n=8). *Cordyceps sinensis*: (n=4); *LSP*: (n=4); *LSP+Cordyceps sinensis*: (n=4).	MBPL human 30k P7
GSE84074 [313]	Control: (n=5). *(5R)-5-hydroxytriptolide*: (n=5).	Agilent-062918 Human lncRNA array V4.0
GSE103044 [314]	Control: (n=3). *Folate deficiency*: (n=3); *Folate repletion*: (n=3).	Illumina HiSeq 2000

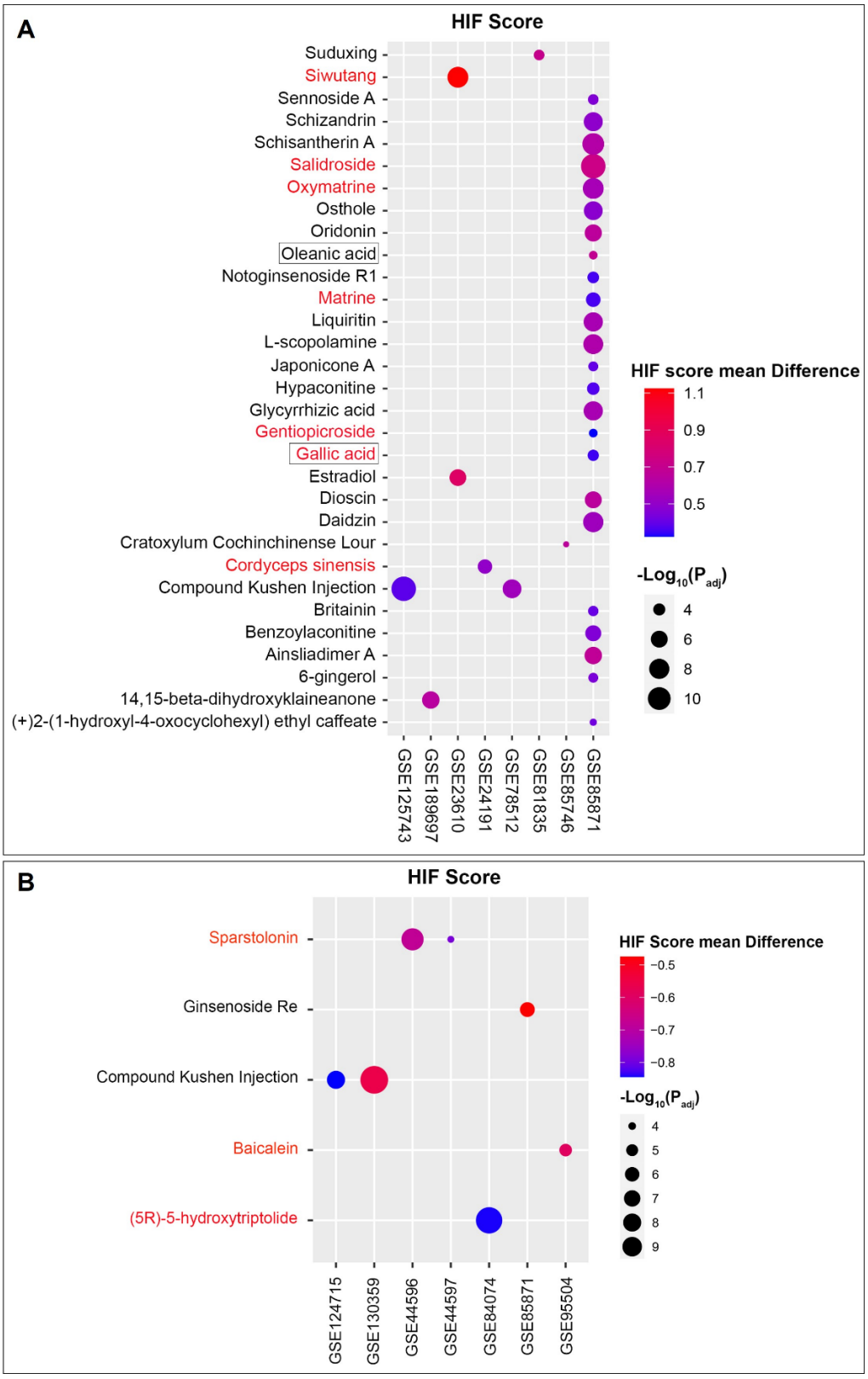


Figure 4.15. Effects of natural compounds/herbal extracts on HIF activity. Scatter plot showing natural compounds or herbal extracts that can upregulate (**A**) or downregulate (**B**) the HIF score. Compounds with red colour are water soluble, and those framed are reported to be present in *Limonium Sinense*. The sizes of circles represent the -Log10 of the P_{adj} values, and the colours of circles represent the HIF score mean difference of each compound compared with control samples.

4.5.11 Gallic acid induces HIF activation

The analysis strongly suggested a potential role for gallic acid with *Limonium Sinense* in mediating HIF activation, given the facts that it is water soluble and has been reported to be present in *Limonium Sinense* [25]. To further investigate this, we examined the HIF score in two additional GEO datasets upon treatment with gallic acid. The datasets included GSE85871, which evaluated gene expression profiles of MCF7 cells treated with traditional Chinese medicine components, and GSE158788, which investigated the gene expression profile analysis of gallic acid-induced cell death processes. The results demonstrated that the HIF score was significantly increased in breast cancer cell line MCF7 (GSE85871; Figure 4.16A) and cervical cancer cell line HeLa (GSE158788; Figure 4.16B) upon gallic acid treatment. These findings further support the role of gallic acid inducing HIF activation. Moreover, these results provide additional evidence linking gallic acid with HIF activation and suggest that gallic acid present in *Limonium Sinense* may contribute to the observed HIF activation.

To confirm the presence of gallic acid in LSW, high-performance liquid chromatography (HPLC) analysis was conducted. The analysis revealed a peak corresponding to gallic acid in both LSW and the gallic acid standard, with a retention time of 2.95 min (Figure 4.16C, D). To validate the effect of gallic acid on HIF, the protein level of HIF-1α was further measured in MDA-MB-468 cells following gallic acid treatment. As shown in Figure 4.16E, the HIF-1α protein level was significantly upregulated in a time-dependent manner upon treatment with gallic acid in MDA-MB-468 cells. Together, these results provide evidence for the involvement of gallic acid within *Limonium Sinense* in mediating HIF activation, at least partially.

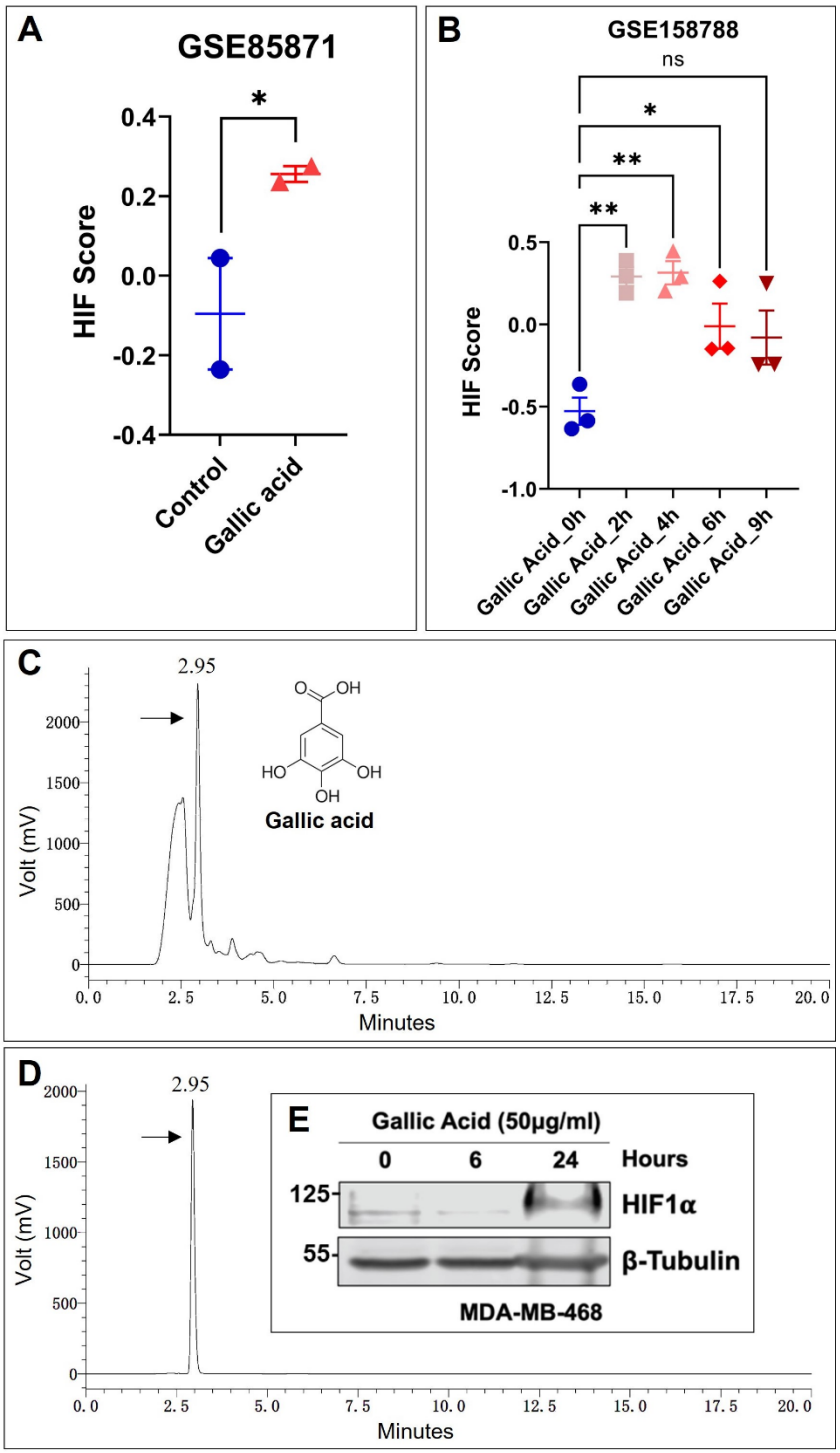
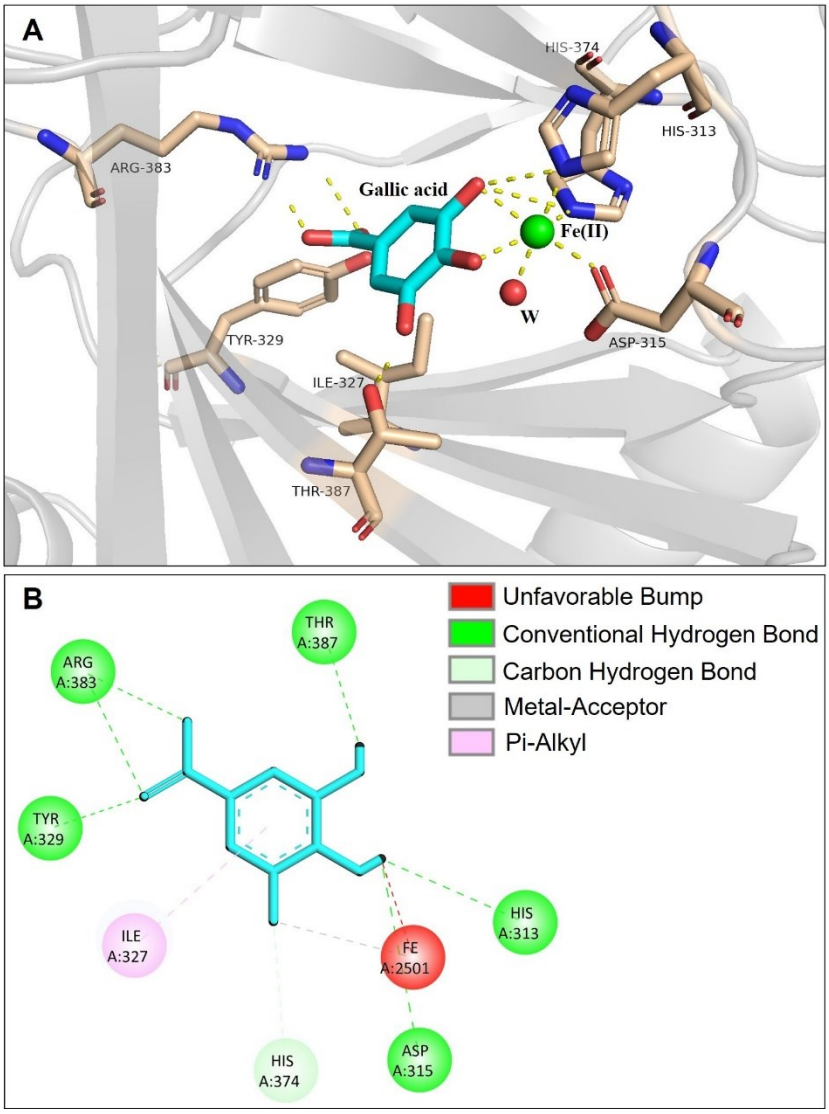


Figure 4.16. Effects of gallic acid on HIF score in GEO datasets. (**A** and **B**) Graphs showing the effect of gallic acid on the HIF score from Gene Expression Omnibus (GEO) dataset GSE85871 (**A**) and GSE158788 (**B**). Data are mean ± SD. ns, not significant; **P* < 0.05; ***P* < 0.01 by the Student's *t*-test and Dunnett's multiple comparisons test, respectively. (**C** and **D**) Chromatograms (HPLC/UV) of *Limonium Sinense* water extracts (LSW) and gallic acid in 271 nm. The retention time (2.95 min) and structure for gallica acid are indicated. (**E**) Protein expressions of HIF-1α in MDA-MB-468 cells treated with or without gallic acid (50 μg/ml) for the indicated time. β-Tubulin was used as a loading control.

4.5.12 Mechanism of gallic acid in activating HIF

It is reported that gallic acid binds to the active site of PHD2 in a specific manner. The phenolate oxygen atoms of gallic acid chelate Fe^{2+}, while the carboxyl group binds to Arg383 site in PHD2 [315]. In our analysis, we examined the X-ray crystallographic structure model of gallic acid binding with PHD2 (Figure 4.17). In the model, we observed that the carboxyl group of gallic acid interacts with Arg383 and forms a conventional hydrogen bond with the Tyr329 site in PHD2. Furthermore, two of the phenolate oxygen atoms of gallic acid chelate Fe^{2+}, while the remaining phenolate oxygen atom forms a conventional hydrogen bond with the Thr387 site in PHD2. In addition, Ile327 site of PHD2 forms a Pi-Alkyl bond with gallic acid. These findings suggest that gallic acid has the potential to inhibit PHD2 activity, leading to the stabilization of HIF-1α. The binding interactions observed in the crystallographic structure model provide insights into the molecular mechanism by which gallic acid may contribute to HIF activation by interfering with PHD2 function.



(**Figure legend is provided on the following page**)

Figure 4.17. Predicted binding pose of gallic acid binding PHD2. (**A**) The 3D graph showing the binding model of gallic acid with PHD2. The key interactions of gallic acid with PHD2 are depicted as yellow dashes. (**B**) The 2D graph showing the specific interactions between gallic acid and the PHD2 protein. Dashes with indicated colours represent the interaction bonds.

4.5.13 LSW interrupts the interaction between HIF and VHL: another potential way of LSW to activate HIF

We also tested the effect of LSW on HIF in a human embryonic kidney cell line HEK293T and it demonstrated a significant upregulation of HIF-1α and HIF-2α protein levels upon LSW treatment in a time-dependent manner (Figure 4.18A).

In addition to PHD, under normal oxygen tension, the protein expression of HIF-1α is also negatively regulated by proteasomal degradation and ubiquitination in a pathway involving von Hippel-Lindau protein (VHL), a recognized component of an E3 ubiquitin protein ligase [316]. Studies have demonstrated that the dissociation of VHL from HIF-1α can inhibit HIF-1α ubiquitination, consequently leading to its accumulation [317, 318]. Therefore, we conducted a Co-immunoprecipitation (Co-IP) assay to validate the impact of LSW on the interaction between VHL and HIF-1α. Our findings revealed that LSW effectively inhibits the binding between VHL and HIF-1α (Figure 4.18B), suggesting a potential additional mechanism by which LSW induces HIF activation. However, the precise mechanism through which LSW disrupts the interaction between VHL and HIF-1α, as well as its potential relationship with the inhibition of PHD activity or other contributing factors, remains unknown and requires further exploration.

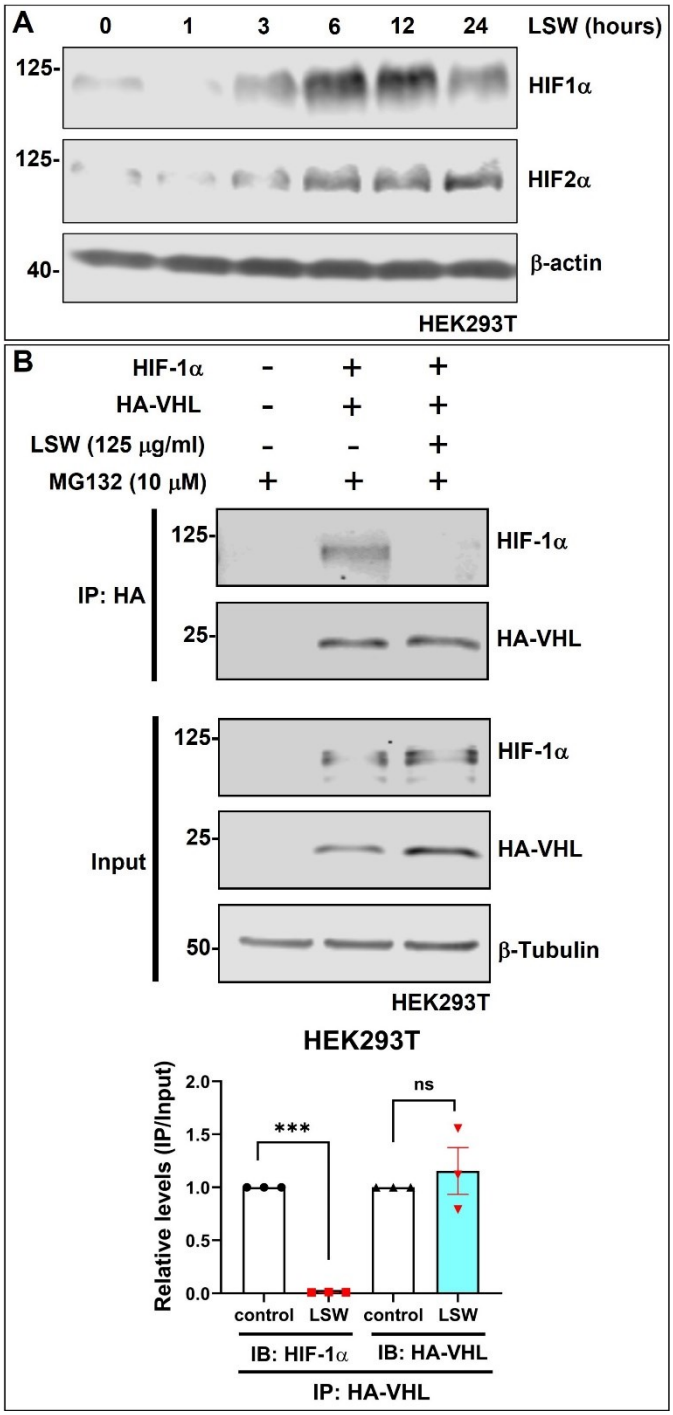


Figure 4.18. LSW inhibits the binding of VHL and HIF-1α. (**A**) Graph showing the protein expression of HIF-1α and HIF-2α in HEK293T cells upon treatment with LSW. β-actin was used as a loading control. (**B**) Co-immunoprecipitation (Co-IP) assay showing the interaction of VHL with HIF-1α was interrupted by LSW treatment. The cell lysates with indicated treatment were immunoprecipitated with anti-HA antibody and the interaction was examined by immunoblot assay. β-Tubulin was used as a loading control. Bar plot showing the relative changes of IP/Input in HEK293T cells upon treatment with LSW. Data are mean ± SD; n = 3 samples per group. **P* < 0.05; ****P* < 0.001 by Dunnett's multiple comparisons test. ns, not significant.

4.6 Discussion

Limonium Sinense has been found to possess multiple pharmacological activities and shows potential as a therapeutic option for the treatment of cancers and other human diseases. The anti-inflammatory effect of *Limonium Sinense* plays a vital role in the treatment of these diseases, as chronic inflammation is closely related to the development of cancer and organ fibrosis. Inflammatory tissues exhibit persistent activation of macrophages and lymphocytes, along with the upregulation of proinflammatory cytokines such as IL-1β, IL-6, IL-17 and TNF-α [319]. Compounds extracted from the root of *Limonium Sinense*, such as polysaccharides and flavonoids, have shown the ability to inhibit tumour growth via enhancing immune function, suggesting that *Limonium Sinense* has the potential to serve as a novel anti-inflammatory drug for the prevention and treatment of cancer and immune-related diseases.

On the other hand, inflammatory cells and epithelial cells can induce the generation of reactive oxygen/nitrogen species (ROS/RNS) inflammation, leading to DNA damage, tissue damage and subsequent accumulation of damaged nucleic acids, proteins, and lipids. In response to this damage, progenitor/stem cells may be activated for tissue regeneration. However, stem cells that are damaged by ROS/RNS can accumulate mutations, potentially leading to the generation of tumour stem cells [320]. In cancer cells, ROS levels are generally elevated compared to normal cells, indicating the critical role of ROS in carcinogenesis and cancer progression. Interestingly, high levels of ROS can also be cytotoxic to cancer cells and induce cell death. This dual role of ROS presents important implications for the development of potential anticancer therapies that aim to modulate cellular redox levels [63].

As previously mentioned, *Limonium Sinense* contains various flavonoids such as isorhamnetin, quercetin, isoquercetin, luteolin and apigenin which exhibit remarkable scavenging activity against oxygen free radical, suggesting that the mechanism of action of *Limonium Sinense* in treating cancer may involve the regulation of redox levels in cells.

Limonium Sinense used in traditional Chinese medicine is often extracted with boiling water to make an aqueous extract for oral uptake [9]. Water extracts from *Limonium Sinense* (LSW) show antiviral, antitumour and immunomodulatory activities in previous studies [24, 25, 27, 28, 321]. In this study, we demonstrated that LSW treatment leads to a strong inhibition of cell growth, potentially by arresting cell cycle at the G2/M phase.

The cell cycle is a tightly regulated process that involves two key checkpoints: the G1/S and G2/M phases. These checkpoints ensure the accurate completion of earlier cellular processes and the faithful transmission of genetic information before cell division. The transitions between these

checkpoints, specifically the G1/S and G2/M transitions, are crucial for controlling cell growth. During these transitions, various cyclins act as positive growth regulators by binding to specific cyclin-dependent kinases (CDKs) [322]. However, when checkpoints are triggered, CDKs are generally inhibited, resulting in cell cycle arrest, which provides time for DNA repair to occur. In cases of severe and irreparable DNA damage, the checkpoints can trigger either cell senescence or apoptosis. Cell cycle checkpoints serve as control mechanisms that stop cell progression at specific stages of the cell cycle. They play a vital role in monitoring DNA and coordinating repair processes. Various stresses, such as nutrient deprivation, mitogenic stimuli, and cytotoxins, can activate checkpoint and induce cell cycle arrest. However, the primary function of checkpoints is to monitor DNA damages and coordinate repair.

Multiple checkpoints have been identified from lower eukaryotes to human, including G1 checkpoint, intra S-phase checkpoint, G2/M checkpoint, and mitotic checkpoint. Among these checkpoints, the G2/M checkpoint is particularly located at the end of G2 phase, which controls the entry into mitosis. This checkpoint prevents cells with damaged DNA from proceeding to mitosis and transmitting the damage to daughter cells. It is especially critical in the presence of lethal DNA lesions such as DNA double-strand breaks (DSBs). Despite its rapid activation compared to the G1/S checkpoint, the G2/M checkpoint also possesses inherent insensitivity [323-325]. Additionally, numerous reports have demonstrated that hypoxia can inhibit cell proliferation by inducing cell cycle arrest [326-328]. Factors such as HIFs and NF-κB control the expression of multiple important components of the cell cycle machinery and control mechanisms, such as p21, p27 and cyclin D1 [267]. These factors play an important role in regulating the cell cycle and can influence cell growth under hypoxic conditions.

Moreover, the CMap analysis revealed Tubulin inhibitors as potential therapeutic candidates with similar actions to LSW. Tubulin inhibitors are chemotherapy drugs that directly interfere with the tubulin system, which is essential for cellular mitosis [329]. Microtubules, composed of α- and β-tubulin heterodimers, play critical roles in many cellular processes such as maintaining cell structure, facilitating cell division, and aiding in intracellular transport [330, 331]. Targeting microtubules has proven to be an attractive strategy for drug discovery due to their involvement in crucial cellular processes and the ability to induce cell cycle arrest in the G2/M phase upon disruption. Microtubule/tubulin inhibitors can be classified into two major categories based on their mechanisms of action. The first category includes agents that promote tubulin polymerization and stabilize microtubule structures, such as paclitaxel, epothilones, discodermolide and taccalonolides. The second category comprises agents that inhibit tubulin polymerization and destabilize microtubule structures, including maytansinoids, auristatins, vinblastine and vincristine [332]. The ability to interact with Tubulin is one of the most important mechanisms of action of natural

compounds or herbal extracts, and it plays a significant role in their effects on cellular processes [333]. Many clinically used tubulin inhibitors are derived from natural products or their synthetic derivatives, such as taxanes (e.g., paclitaxel) and vinca alkaloids (e.g., vinblastine) [334].

Reports have indicated that p21 can arrest cell cycle progression at the G1/S and G2/M transitions by inhibiting CDK4,6/cyclin D and CDK2/cyclin E, respectively [270]. In mammalian cell cycle progression, CDKs and cyclins act as regulatory subunits, and the phosphorylation of retinoblastoma protein by Cyclin/CDK complexes triggers cell cycle progression. p21 disrupts these interactions, thereby inhibiting cell cycle progression [335]. Through our further investigation, we have discovered that the absence of p53 does not affect the ability of LSW to inhibit cell growth, indicating that the growth inhibitory effect of LSW works through a p53-independent pathway. Moreover, we have observed a significant elevation of p21 at the gene level, while several cell cycle regulators including *CDK1*, *CCNA2* (cyclin A2) and *CCNB2* (cyclin B2) [336] were notably downregulated. These findings suggest that LSW may potentially activate p21-induced G2/M phase cell cycle arrest, thus leading to the inhibition of cell growth. Further validation is required in future studies.

The production of red blood cells relies predominantly on the cytokine erythropoietin (EPO) and its transcription factor HIF [337]. Studies have shown that regulating HIF levels provide novel therapeutic strategies for a broad variety of diseases, including anaemia [338-340]. In our findings, LSW can significantly upregulate the expression of HIF-1α protein levels, indicating that the water extract of *Limonium Sinense* can directly interact with hypoxia transcription factors. HIF is a transcription factor that regulates the expression of genes involved in adaptive mechanisms to hypoxia, such as angiogenesis and apoptosis [341]. In addition, the Hematopoietic cell lineage was strongly enriched by the DEGs in MDA-MB-468 cells upon treatment with LSW. Gene ontology enrichment analysis also revealed a strong correlation between LSW treatment and the positive regulation of angiogenesis. This suggests that the mechanism of *Limonium Sinense* in promoting blood circulation and treating blood-related diseases may involve the regulation HIF-1 signalling.

The HIF system plays a crucial role in the body's response to hypoxia, initiating processes such as angiogenesis, anaerobic metabolism promotion, and increased red cell mass in cases of anaemia. Hypoxia is associated with various pathological conditions, including stroke, inflammation, and cancer. These pathological events trigger the activation of repair mechanisms, including angiogenesis, to restore oxygen homeostasis [342]. Given the blood enrichment function of *Limonium Sinense*, it is speculated that the pharmacological effects of *Limonium Sinense* may be associated with the regulation of HIF activity. Additionally, Hypoxia can induce the upregulation of major angiogenic factors, such as VEGF, Angiopoietin-1, TGF-β1 and fibroblast growth factor 2 (FGF-

2), as well as the recruitment of inflammatory and endothelial progenitor cells and the migration of fibroblasts [343]. These factors also act as homing signals for the recruitment of circulating angiogenic cells from distant sites, including bone marrow and vessels from adjacent tissue.

In our findings, LSW significantly enriched the hallmark angiogenesis, and showed a strong relationship with regulation of angiogenesis. Targeting angiogenesis, either through anti-angiogenesis or pro-angiogenesis approaches, has been a therapeutic approach for treating tumours and anaemia-related diseases, and the regulation of HIF-1 has emerged as a crucial factor in these therapeutic strategies. The current results of LSW provide a promising prospect that LSW may regulate HIF-1α expression, thereby activating various HIF target genes that could impact multiple biological processes in the body. However, there are still many questions regarding the intricate relationship between *Limonium Sinense* and hypoxia that need to be explored.

From our integrated analysis, gallic acid was identified as a potentially bioactive ingredient within *Limonium Sinense* in mediating HIF activation. Gallic acid (3,4,5-trihydroxybenzoic acid) is a kind of polyphenol found naturally in various sources, including as a free compound or in combination with other compounds, has been studied for its beneficial properties [344]. It is often found in conjunction with catechins, such as (-)-epigallocatechin gallate (EGCG), (-)-epicatechin gallate (ECG), and (-)-gallocatechin gallate (GCG) [345], which are flavonoids commonly found in green tea. These compounds, including gallic acid, have associated with antioxidant, anti-carcinogenic, and anti-inflammatory activities [346]. Recent research has highlighted the ability of gallic acid and its esters to increase the cellular content of HIF-1α in mammalian cells. For instance, gallic acid and its ester ethyl gallate have been shown to activate the HIF-1α/EPO/VEGFA pathway, leading to reduced infarct volume and improved cerebral ischaemia state in rats [347]. Methyl gallate has demonstrated therapeutic angiogenesis effects by inducing HIF-1α, VEGF, fibroblast growth factor-2 (FGF-2), and miR-146 expression in diabetic hind limb ischemia [348]. n-Propyl gallate (nPG) has exhibited neuroprotective effects in a rat model of moderate forebrain ischemia by upregulating HIF-1α, EPO and VEGF expression [349]. ECG, functioning as a ROS scavenger, has been found to elevate HIF-1α expression neurons subjected to *in vitro* ischemia [350]. In addition, observation revealed that nPG did not inhibit PHD activity up to a concentration of 1.2 mM. Conversely, gallic acid exhibited a dose-dependent inhibition of PHD activity. This is mainly attributed to gallate, the hydrolysed product of nPG, which preserves the polyphenol structure capable of directly inhibiting PHD. The actions of gallate-conjugated catechins in increasing HIF-1α content are mediated by gallate released within cells through hydrolysis. The alteration of the propyl group to other groups may result in compounds with enhanced cell membrane permeability or hydrolysis rate, potentially leading to increased activity in upregulating HIF-1α compared to nPG [315]. These findings suggest a role of gallate compounds in activating HIF, although further studies are needed to elucidate the

precise inhibition mechanism. While it has been suggested that catechins may inhibit HIF-1α hydroxylation, which is a necessary step for its ubiquitination, additional research is required to confirm this possibility.

In this study, we have identified the presence of gallic acid within *Limonium Sinense* and confirmed its role in upregulating HIF-1α expression. In line with our findings, previous research by Tsukiyama and colleagues has reported that gallate is able to inhibit PHD activity, thereby reducing the HIF degradation rate and increasing the protein level of HIF-1α [315]. Taken together, these results provide novel insights into the bioactive ingredients in *Limonium Sinense*, highlighting the rich natural resource and therapeutic values of herbal plants. However, further investigation is needed to fully understand the precise mechanism of *Limonium Sinense* in activating HIF. Given that some compounds in the hydroxybenzoic acid family to which gallic acid belongs have been shown the ability to increase of HIF content, we speculate that the structure of hydroxybenzoic acid may play a role in triggering HIF upregulation, and this effect may be related to the inhibition of PHD2 activity. These aspects will be discussed further in Chapter5.

Chapter 5 Identification of small molecules as potential inhibitors of PHD2 from Hydroxybenzoic acids and their derives using structure-based virtual screening

5.1 Abstract

Prolyl hydroxylase (PHD) enzymes play an important role in cellular responses to hypoxia by regulating the activity of hypoxia-inducible factor-α (HIF-α) transcription factors. Inhibition of PHD enzymes holds promise for the treatment of various diseases, including anaemia, cardiovascular disease, and stroke. In this study, we employed a structure-based virtual screening approach to identify potential inhibitors of human PHD2 from Hydroxybenzoic acids and their derives. As a result, we discovered three potent inhibitors with IC_{50} values of 4.4 nM, 8.6 nM, and 1.5 nM, respectively. Molecular docking studies revealed that these compounds occupy the 2OG binding pocket of PHD2, leading to the inactivation of the enzyme. These findings present novel chemical scaffolds that can serve as starting points for the further development of PHD2 inhibitors, offering potential therapeutic options for the targeted treatment of diseases associated with hypoxia.

5.2 Introduction

In recent years, the discovery of hypoxia-inducible factor-1α (HIF-1α) has greatly contributed to the understanding of many diseases including cancer and cardiovascular diseases. HIF-1α is a key player in the adaptation of tumour cells to hypoxia and is involved in regulating a wide range of biological processes, including cell proliferation, survival, cellular metabolism, angiogenesis, metastasis, cancer stem cell maintenance, and propagation [351]. The critical role of HIF-1α in disease pathogenesis has made it an attractive target for therapeutic interventions. Therefore, modulating the activity of HIF-1α has emerged as a promising approach in the treatment of different diseases.

5.2.1 Hypoxia

Hypoxia refers to a condition characterized by low levels of oxygen that are insufficient to meet the oxygen demands of tissues. It plays an important role in the oxygen supply in the body [352]. The aerobic lifestyle of most animals requires a continuous and adequate oxygen supply, and low

oxygen levels constitute a major environmental threat [353]. Generally, tissue hypoxia can be caused by hypoxemia (low blood oxygen levels), impaired oxygen delivery or impaired tissue oxygen extraction/utilization [354]. Clinical management of tissue hypoxia usually focuses on global hypoxemia and oxygen delivery. A critical question in understanding hypoxia is how mammalian cells detect and respond to changes in oxygen levels to coordinate various biological processes. The most extensively studied mechanism of response to hypoxia involves HIFs. Under low oxygen conditions, HIFs are stabilized and regulate the expression of numerous genes involved in cell survival, angiogenesis, glycolysis, and invasion/metastasis. In addition to the HIF pathway, cells can sense changes in oxygen levels through other stress pathways, as well as alterations in metabolite levels and the generation of ROS by mitochondria. These alternate mechanisms of oxygen sensing contribute to the cellular adaptations in response to hypoxia [355].

5.2.2 Hypoxia in human diseases

Different types of cells response to various levels of hypoxia can range from substantial adaptation to cell death [356, 357]. In response to hypoxia, cellular adaptive mechanisms change in terms of increased ventilation, cardiac output, blood vessel growth, and circulating red blood cell numbers. In addition, ATP consuming reactions are suppressed, and metabolism processes undergo alterations until oxygen homeostasis is restored [355].

In recent years, numerous studies have demonstrated the prominent feature of hypoxia in various pathological conditions, including bacterial infections, inflammation, wounds, cardiovascular defects, and cancer [355]. It is reported that some physiology conditions such as ischemia, haemorrhage, stroke, premature birth, and exhaustive exercise will cause tissue hypoxia [358]. Normally, oxygen is transported by red blood cells to respiring tissues to meet the metabolic demands and prevent microenvironmental toxicity. There is an oxygen delivery system in the body comprising the lungs, heart, blood and vascular system, which work together to maintain the coordination of oxygen supply and demand [359]. When blood oxygen levels are reduced, this coordination will be destroyed. Red blood cells undergo significant changes in their biophysical properties, such as cell morphology, deformability, and adherence under high hypoxic conditions, thus cause various diseases such as anaemia and sick cell disease [360, 361]. Furthermore, damaged tissues may also experience sustained hypoxia due to increased oxygen demand by cells in the wounded areas or infiltrating inflammatory and mesenchymal cells [362]. Acute hypoxia serves as a crucial stimulus for processes necessary for tissue repair, such as wound healing and angiogenesis. However, in some cases of chronic injury, chronic hypoxia, and pathological repair, hypoxia and its associated signalling pathways can drive the progression of fibrosis, leading to excessive scarring and compromised organ function [343].

5.2.3 Hypoxia in cancer

Tumour hypoxia was first described in the 1950s by radiation oncologist as a frequent cause of failure to radiotherapy in solid tumours [363]. It has long been recognised that solid tumours frequently contain regions of profound hypoxia and that tumour hypoxia is associated with adverse prognosis [359]. Numerous studies have demonstrated that hypoxia exits most of the solid tumours caused by the imbalance between oxygen supply and demand of tumour cells [364]. The rapid proliferation of cancer cells in the dynamic and ineffective microvasculature results in a high demand for oxygen supply and consumption. Hypoxia has been associated with poor clinical prognosis in many types of tumours, including breast carcinomas and head and neck squamous cell carcinomas [84, 365]. It has become clear that various cancer-associated processes such as invasion and angiogenesis reply on the activation of HIF target genes [360]. Additionally, multiple oncogenic pathways have been found to be mechanistically linked to the activation of HIF [318, 366, 367]. Recent investigations into HIF hydroxylase pathways have revealed their extraordinarily complexity. When these pathways are activated by oncogenic mutations or even by microenvironmental hypoxia in cancer, they are triggered in their entirety, despite initially evolving for physiological purposes rather than oncogenic ones. Consequently, all components of the pathway become activated in the context of cancer, irrespective of whether they promote oncogenesis, neutral or even restricting the oncogenic process [359]. A systemic view of hypoxia in human tissues and its related diseases was shown in Figure 5.1.

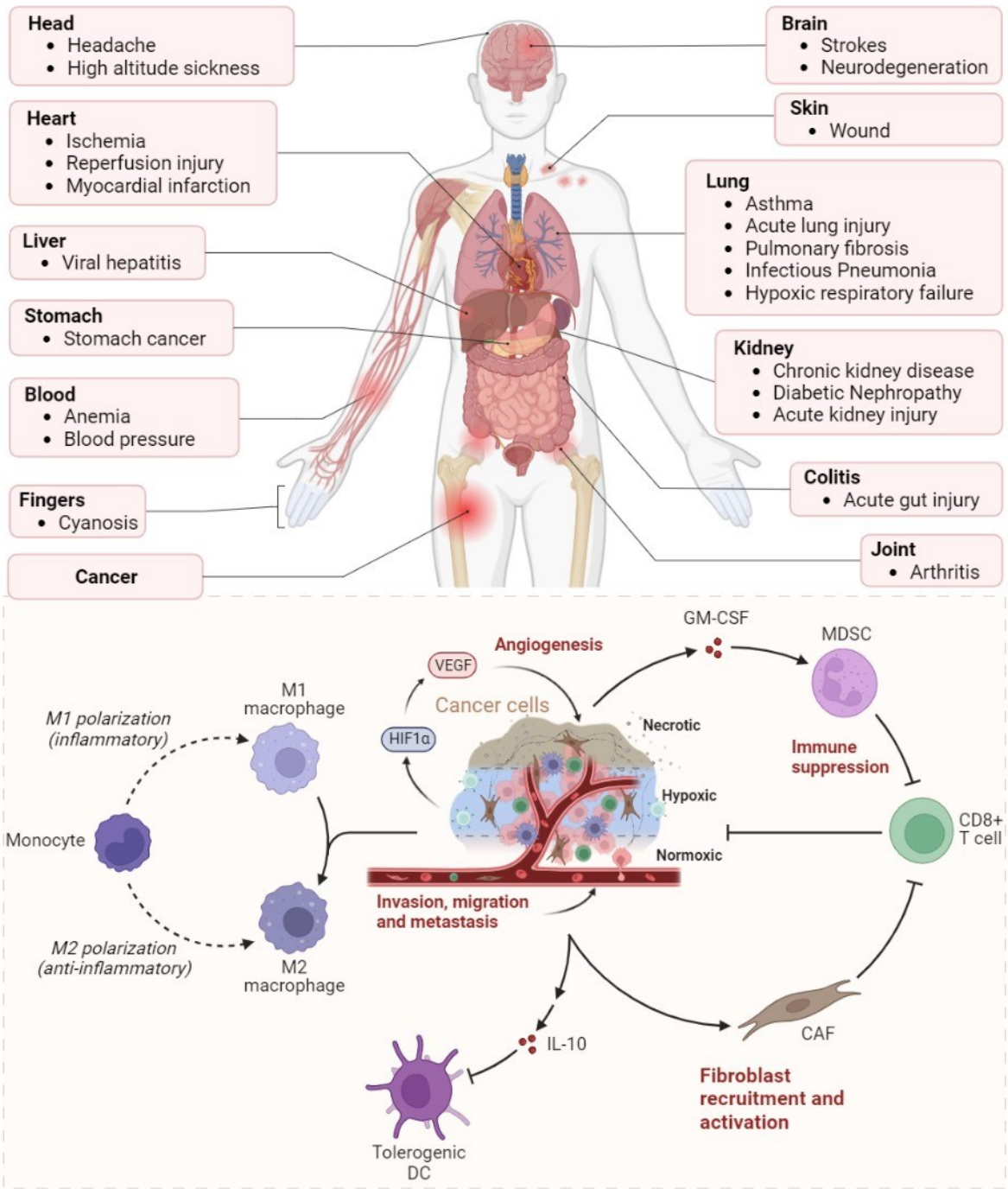


Figure 5.1. Schematic representation of human tissues under hypoxic conditions. Hypoxia drives various mechanisms broadly associated with aggressiveness, invasion, and the acquisition of metastatic properties in cancer. Additionally, oxygen deprivation in the tumour also supports the modulation of inflammation and contributes to the immune suppression. CAF, Cancer-associated fibroblast; GM-CSF, Granulocyte-macrophage colony-stimulating factor; MDSC, Myeloid-derived suppressor cell; HIF-1α, Hypoxia inducible factor 1 alpha; VEGF, Vascular endothelial growth factor. Information collected from [368, 369].

5.2.4 Hypoxia inducible transcription factor (HIF)

Cellular responses to hypoxic conditions are primarily orchestrated by activation of the HIFs [277]. HIFs belong to the family of basic helix-loop-helix (bHLH) proteins, and they are heterodimers composed of an oxygen sensitive α-subunit and an oxygen insensitive aryl hydrocarbon receptor nuclear translocator (ARNT) or β-subunit [370]. The bHLH proteins contain a Per-ARNT-Sim (PAS) domain, which is involved in various functions, including the response to hypoxia [370, 371]. To data, three members of HIFs have been identified, namely HIF-1, HIF-2, and HIF-3, consisting of five subunits: HIF-1α, HIF-2α, HIF-3α, ARNT and ARNT2 [363, 372].

5.2.4.1 HIF-1α

HIF-1 was the first characterized member of HIF family in 1991 and is formed by the dimerization of HIF-1α with ARNT [363, 373] (Figure 5.2A). Structural analysis has revealed that HIF-1α contains four distinct domains: a bHLH domain for DNA binding, a PAS (PAS-A and PAS-B) domain for dimerization and target gene specificity, an oxygen-dependent degradation domain (ODDD) responsible for degradation by the ubiquitin-proteasome pathway, and two transactivation domains located in the C-terminal portion of the protein, namely the NH_2-terminal transactivation domain (N-TAD) and COOH-terminal transactivation domain (C-TAD) [374, 375] (Figure 5.2B). The hypoxia-inducible transactivation capability of HIF-1 is primarily dependent on its HIF-1α subunit since HIF-1α is the oxygen labile subunit of the HIF-1 heterodimer and contains the transactivation domains responsible for HIF-1α's transcriptional activity [375]. These transactivation domains enable HIF-1α to interact with other proteins and activate the expression of target genes involved in the cellular response to hypoxia. It was discovered that HIF-1 binds to an 18-nucleotide fragment of the erythropoietin (EPO) enhancer, resulting in the hypoxic activation of EPO in Hep3B cells [376]. HIF-1α has emerged as a critical regulator of the cellular response to hypoxia due to its ubiquitous expression and its ability to induce the expression of numerous hypoxia-inducible genes [377].

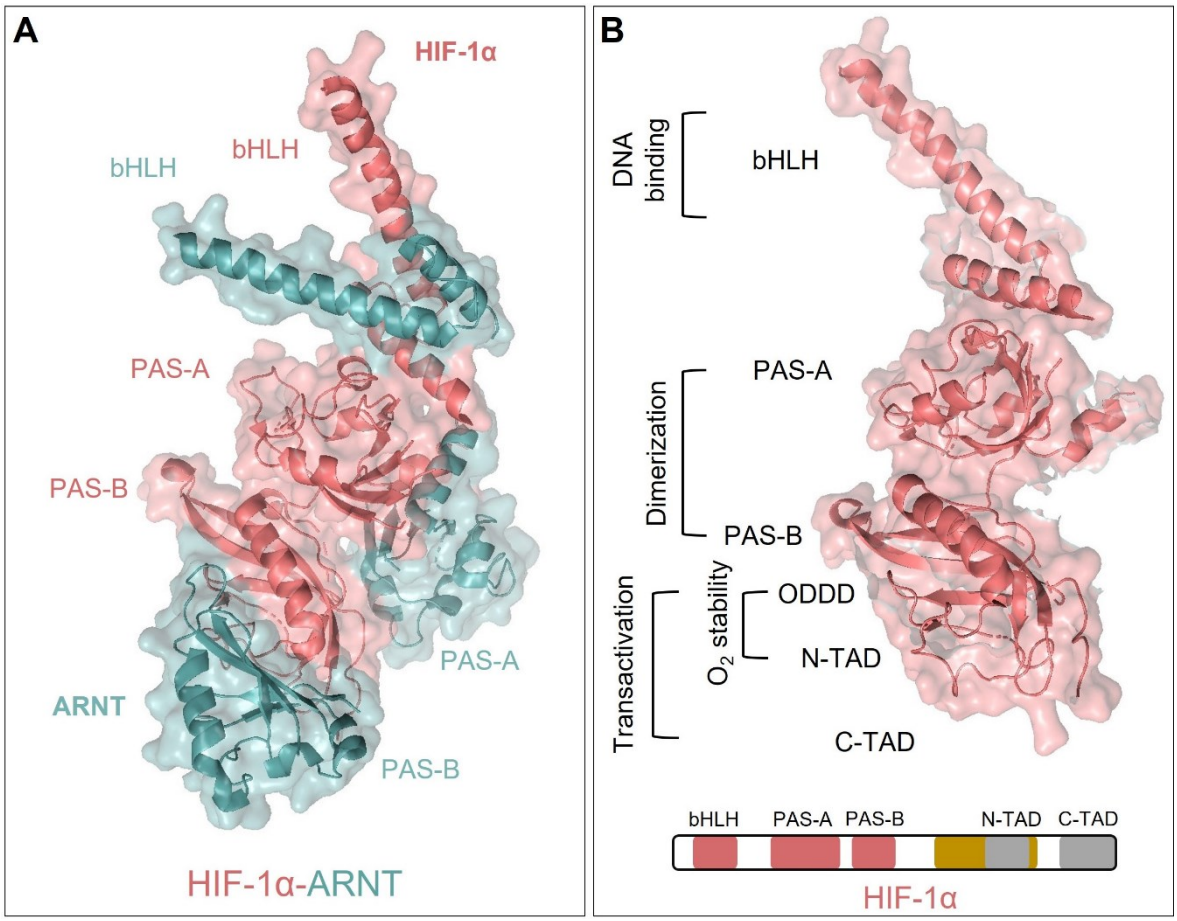


Figure 5.2. Schematic structure of HIF-1. (**A**) Overall structure of HIF-1 heterodimer (PDB ID: 4ZPR), consists of two individual structures labelled as HIF-1α (coloured pink) and ARNT (coloured cyan). bHLH: basic helix-loop-helix domain; PAS: Per/ARNT/Sim domain; ODDD: oxygen dependent degradation domain; N-TAD: NH_2-terminal transactivation domain; C-TAD: C-terminal transactivation domain. (**B**) X-ray crystal structure showing the functional domains of HIF-1α. Information collected from [378].

5.2.4.2 HIF-2α

HIF-2 shares approximately 48% amino-acids sequence homology with HIF-1 [379-381]. It is the second HIF family member and formed by the dimerization of HIF-2α with ARNT (Figure 5.3A). HIF-2α is structurally similar to HIF-1α which also contains bHLH, PAS, ODDD and TAD motifs (Figure 5.3B). HIF-2α mostly expressed in specific cell types including endothelial cells, glial cells, type II pneumocytes, cardiomyocytes, fibroblasts of the kidney, interstitial cells of the pancreas and duodenum, and hepatocytes [382]. Additionally, reports have shown that HIF-2α acts more effectively on EPO gene and genes associated with iron metabolism. On the other hand, another group of genes, such as VEGF and glucose transporter 1 (GLUT-1), are regulated by both HIF-1α and HIF-2α [383, 384].

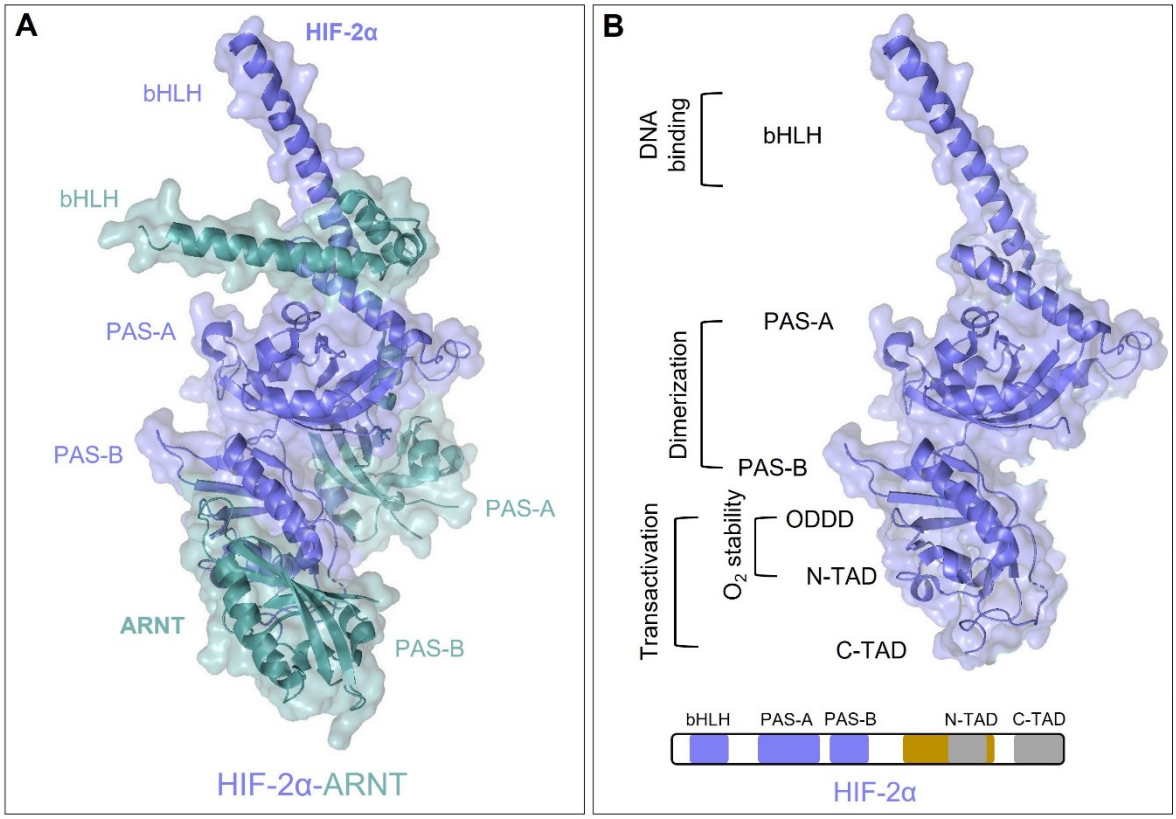


Figure 5.3. Schematic structure of HIF-2. (**A**) Overall structure of HIF-2 heterodimer (PDB ID: 4ZP4), consists of two individual structures labelled as HIF-2α (coloured blue) and ARNT (coloured cyan). bHLH: basic helix-loop-helix domain; PAS: Per/ARNT/Sim domain; ODDD: oxygen dependent degradation domain; N-TAD: NH_2-terminal transactivation domain; C-TAD: C-terminal transactivation domain. (**B**) X-ray crystal structure showing the functional domains of HIF-2α. Information collected from [378].

5.2.4.3 HIF-3α

HIF-3α has a high similarity with human HIF-1α and HIF-2α in the bHLH and PAS domains, and it contains N-TAD but lacks C-TAD transactivation domain [385]. Multiple splice variants of the human HIF-3α have been reported, such as a HIF-3α splice variant with a leucine zipper (LZIP) domain in the place of the C-TAD, which mediates DNA binding and protein-protein interactions [386-388] (Figure 5.4B). Evidence has demonstrated that all human HIF-3α variants are induced by hypoxia and the induction is mediated by HIF-1α but not HIF-2α [386]. HIF-3α dimerizes with ARNT to form HIF-3 (Figure 5.4A). The role of HIF-3 in the hypoxic regulation of target gene expression *in vivo* is not yet well understood. HIF-1α and HIF-2α both activate HRE-dependent gene transcription [277], and similar to HIF-1α and HIF-2α, HIF-3α can dimerize with ARNT and bind to HREs *in vitro* and *in vivo* [363].

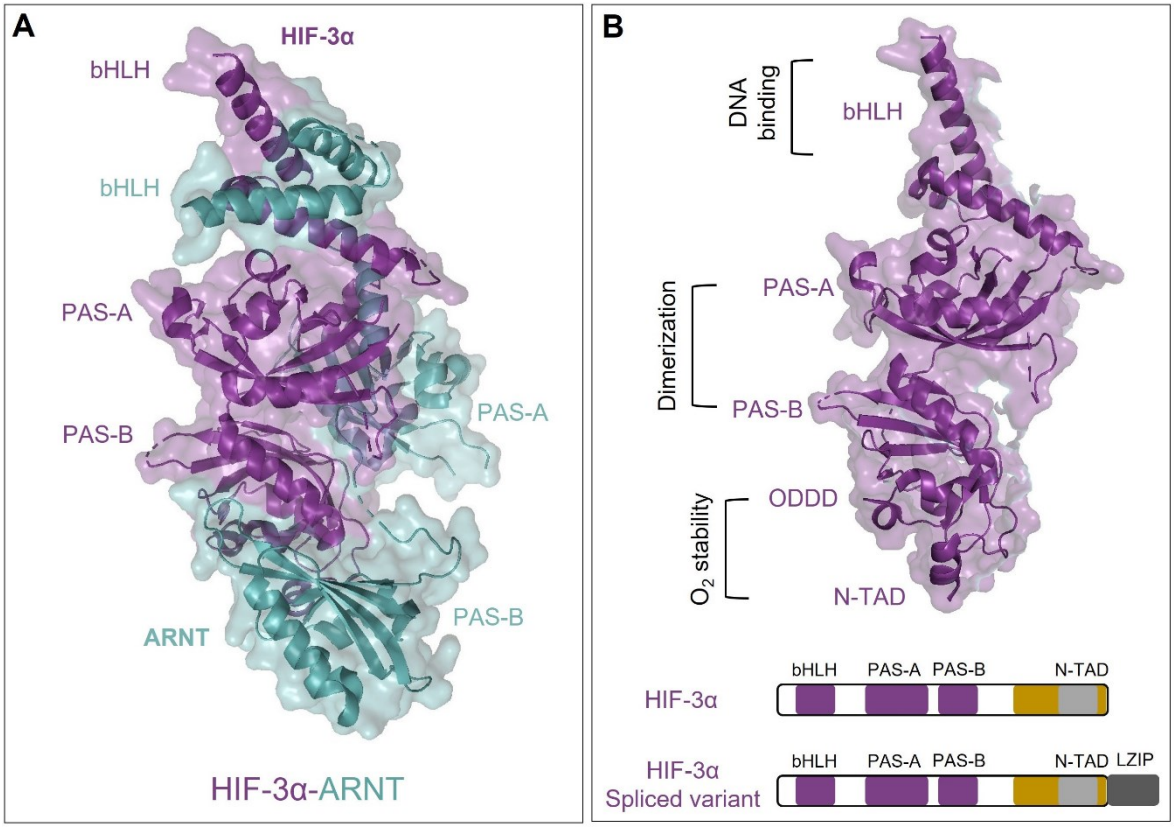


Figure 5.4. Schematic structure of HIF-3. (**A**) Overall structure of HIF-3 heterodimer (PDB ID: 7V7L), consists of two individual structures labelled as HIF-3α (coloured purple) and ARNT (coloured cyan). bHLH: basic helix-loop-helix domain; PAS: Per/ARNT/Sim domain; ODDD: oxygen dependent degradation domain; N-TAD: NH_2-terminal transactivation domain. (**B**) X-ray crystal structure showing the functional domains of HIF-3α. Information collected from [389].

5.2.4.4 ARNT (HIF-β)

ARNT is the general binding partner for all bHLH/PAS family members. It can heterodimerize with the single-minded proteins (SIM1 and SIM2) as well as the alpha subunits of HIF-1, -2 and -3. The structure of ARNT contains bHLH, PAS and transactivation domains, while lacks an ODDD, which is found in some other members of the bHLH/PAS family (Figure 5.5). As a result, ARNT is constitutively expressed in all tissues under aerobic conditions [390]. ARNT2, another HIF-β subunit, can heterodimerize with HIF-α proteins and play a role in mediating hypoxic gene expression in neurons [391]. Notably, the β subunits are constitutively expressed and can form heterodimers with several other members of the bHLH/PAS family to mediate transcriptional responses, whilst the α subunits are specifically involved in the response to hypoxia [359]. Levels of HIF-α, but not HIF-β, are strongly regulated by oxygen availability, as is the transcriptional activity of HIF [277].

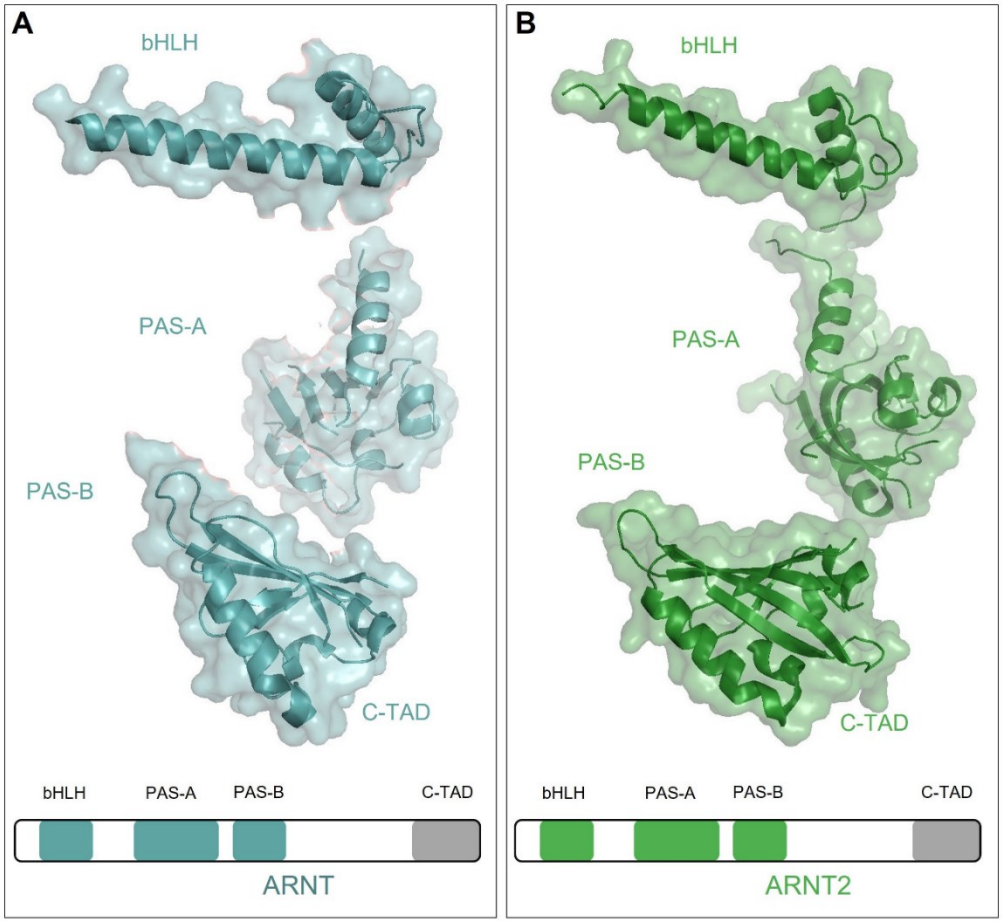
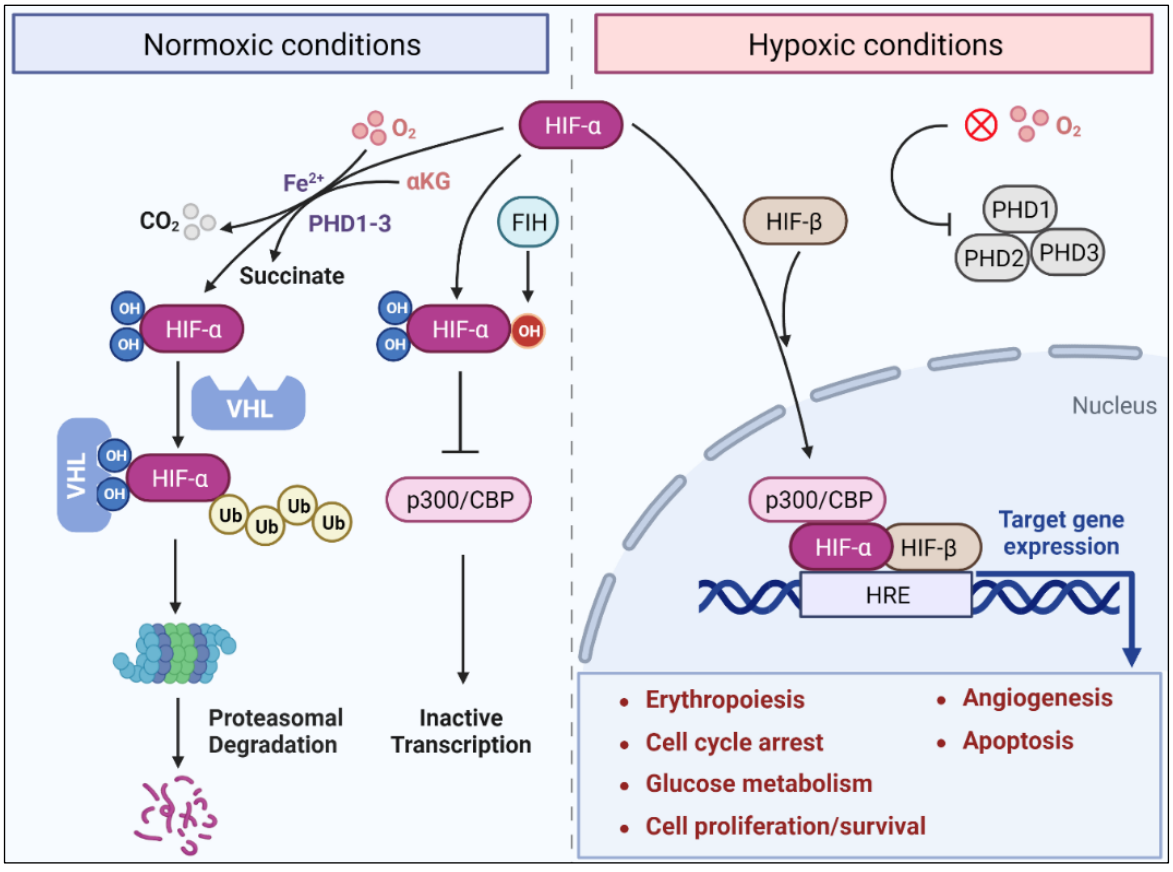


Figure 5.5. Schematic structure of ARNT and ARNT2. Overall structure of (**A**) ARNT (HIF-1β) heterodimer (coloured cyan) (PDB ID: 5SY5) and (**B**) ARNT2 (HIF-2β) (coloured green) (PDB ID: 7XI3). bHLH: basic helix-loop-helix domain; PAS: Per/ARNT/Sim domain; C-TAD: C-terminal transactivation domain. Information collected from [392, 393].

5.2.5 HIF-α signalling pathway

HIF-1 is induced in hypoxic cells and binds to the cis-acting HRE of target genes, such as the human EPO gene, which is required for EPO synthesis [394]. Generally, the bHLH-PAS motifs are essential to allow heterodimer formation between HIF-α and HIF-β subunits and for binding to the HRE-DNA sequence on the target genes [274]. HIF-1 interacts with DNA through its N-terminal bHLH domain and its two PAS domains, and activates transcription through its TADs, namely N-TAD and C-TAD [373]. The C-TAD interacts with co-activators such as p300 and CREP-binding protein (CBP) to modulate gene transcription of HIF-α under hypoxia, while the N-TAD is responsible for stabilizing HIF-α against degradation [395].

HIF-α plays an important role in the adaptation of tumour cells to a hypoxic microenvironment [277]. As described previously, in normoxic conditions, HIF-α is hydroxylated at two proline residues (Pro402 and Pro564 in human HIF-α) within the ODDD by members of the PHD family (PHD1-3), which belong to the Fe(II)- and 2OG-dependent oxygen superfamily and target the HIF proteins for degradation [277, 278]. The hydroxylation reactions utilize O_2 and α-ketoglutarate as substrates and generate CO_2 and succinate as by-products. Hydroxylated HIF-α will be recognised by the VHL E3 ligase, leading to its polyubiquitylation on lysine residues and subsequent proteasomal degradation [396]. FIH is another Fe(II)- and 2OG-dependent dioxygenase that hydroxylates a conserved asparagine residue (Asn803 in human HIF-α) within the HIF-α C-TAD, thus blocks the binding of co-activators p300/CBP and negatively regulates HIF transcriptional activity [279]. Under hypoxic conditions, the hydroxylation reactions are inhibited, resulting in the stabilization of HIF-α. The stabilized HIF-α then translocate to the nucleus, where it forms a heterodimer with HIF-β and interacts with p300/CBP co-activators. The HIF-1 complex, consisting of HIF-1α and HIF-1β, binds to specific DNA sequences known as HREs located in the promoters of target genes, and transcriptionally activates multiple genes involved in various biological processes, including erythropoiesis, angiogenesis, cell cycle arrest, apoptosis, glucose metabolism and cell proliferation and survival [397-403] (Figure 5.6).



(**Figure legend is provided on the following page**)

Figure 5.6. Regulation of HIF-α under normoxic and hypoxic conditions. Scheme showing the major mechanisms of regulation of the stability and functional activity of HIF-α. Details are given in the text. HIF-α, hypoxia-inducible factor-α; αKG, α-ketoglutarate; PHD, prolyl 4 hydroxylase domain; VHL, von Hippel-Lindau; FIH, factor inhibiting HIF; CBP, CREB-binding protein; HIF-β, hypoxia-inducible factor-β; HRE, hypoxia-responsive element, Ub, Ubiquitin.

5.2.6 Targeting HIF-α pathway for therapeutic drugs

HIFs directly or indirectly control the expression of thousands of genes to enable cellular and systemic adaptation to hypoxia [404]. Numerous studies have demonstrated that HIF-1α has highly correlation with tumour metastasis, angiogenesis, poor patient prognosis, and tumour resistance therapy [274, 365], and reports have shown that the expression of both HIF-1α and HIF-2α are commonly increased in multiple human tumours [405]. Therefore, targeting HIF-α pathway has emerged as a promising therapeutic strategy for treating cancer.

In other hand, the stability and transcriptional activity of HIF-α are regulated by various post-transnational modifications and signalling pathways. In non-hypoxic conditions, pathways involving oncogenic activation, growth factors, cytokines and other signalling molecules have considerable contributions to the control of HIF-α protein levels [274] (Figure 5.7). For instance, in oxygen-independent pathways, HIF-α is activated by the phosphatidyl inositol-4,5-bisphosphate-3-kinase (PI3K) or MAPK pathways, and PI3K mediates its effect through its target AKT and the downstream mTOR kinase [406], while on the other hand, Phosphatase and tensin homolog gene (PTEN) attenuates HIF-1α stabilization by inhibiting AKT activation [366]. Protein p53 inactivation also leads to the accumulation and activation of HIF-α target genes [407]. These pathways add complexity to the development of therapeutic strategies targeting HIF-α.

Indeed, the expanding knowledge and understanding of HIF function and regulation in cancer have spurred the discovery and development of novel therapeutic agents targeting the hypoxia signalling pathway. Various approaches have been proposed to directly or indirectly impact HIF signalling [408, 409]. These include: **1)** Inhibition of oncogenic signalling pathways: Agents that target oncogenic signalling pathways such as mTOR (mammalian target of rapamycin) and EGFR inhibitors. **2)** Suppression of HIF-1α protein accumulation: Therapeutic agents that aim to reduce the accumulation of HIF-1α protein, such as topotecan, 2-methoxyestradiol, Hsp90 inhibitors, histone deacetylase inhibitors, thioredoxin inhibitors, nonsteroidal anti-inflammatory drugs (both nonselective and cyclooxygenase-2 selective), ascorbate, 9-betaD-arabinofuranosyl-2-fluoroadenine (Fara-A), flavopiridol, and isoflavonoid genistein. **3)** Inhibition of HIF-1α DNA binding

activity: Agents like polyamides and echinomycin that target the DNA binding activity of HIF-1α. **4)** Suppression of HIF-1α transcriptional activity: Therapeutic approaches that aim to inhibit the transcriptional activity of HIF-1α, including chetomin, proteasome inhibitors, and amphotericin B [410]. These various strategies highlight the diverse range of potential therapeutic targets within the HIF signalling pathway, opening new avenues for the development of targeted therapies against cancer.

Additionally, HIF-1α as a crucial transcriptional regulator, plays a significant role in tumorigenesis by coordinating the activity of various transcription factors and their downstream molecules [411, 412]. It is believed that the overexpression of HIF-1α is strongly associated with poor prognosis in patients with different solid cancers. Hence, targeting HIF-1α has emerged as a novel and effective strategy for cancer therapy, leading to the development of several agents designed to suppress HIF-1α in recent years. These agents include drugs that are currently undergoing clinical trials, as well as compounds that have shown promising results *in vitro* and *in vivo* studies.

Moreover, various compounds have been reported to inhibit HIF-1α activity in direct and indirect ways. Direct HIF-1 inhibitors directly interfere with the transactivation, DNA binding and subsequent transcriptional activity of HIF-1. They can be classified into various groups based on their structural features, such as polyamides, quinols, and naphthoquinone spiroketal analogues, shikonin derivatives, epidithiodiketopiperazines, echinomycin, and bortezomib. On the other hand, indirect HIF-1 inhibitors work by blocking the transcription or translation of HIF-1α or promoting the degradation of HIF-1α protein. These inhibitors belong to different classes, including polyphenols, benzoazaheterocycles, rapamycins, camptothecins, geldanamycins, (aryloxyacetylamino)benzoic acid analogues, 2-methoxyestradiol and analogues, hydroxamic acid compounds, and others [351, 413].

Therefore, targeting HIFs for stabilization or inhibition holds promise as a pharmacological approach. As example the genes targeted by the HIF system make it an appealing pharmacological target, especially via PHD inhibition-mediated HIF upregulation for the treatment of diseases such as anaemia, ischaemic stroke, and wound healing [414]. In addition, several PHD inhibitors including FG4592 (Roxadustat, FibroGen), GSK1278863 (Daprodustat, GlaxoSmithKline), Bay85-3934 (Molidustat, Bayer) and AKB-6548 (Vadadustat, Akebia) are currently in clinical use or trials for the treatment of anaemia in patients with chronic kidney disease (CKD) [415, 416].

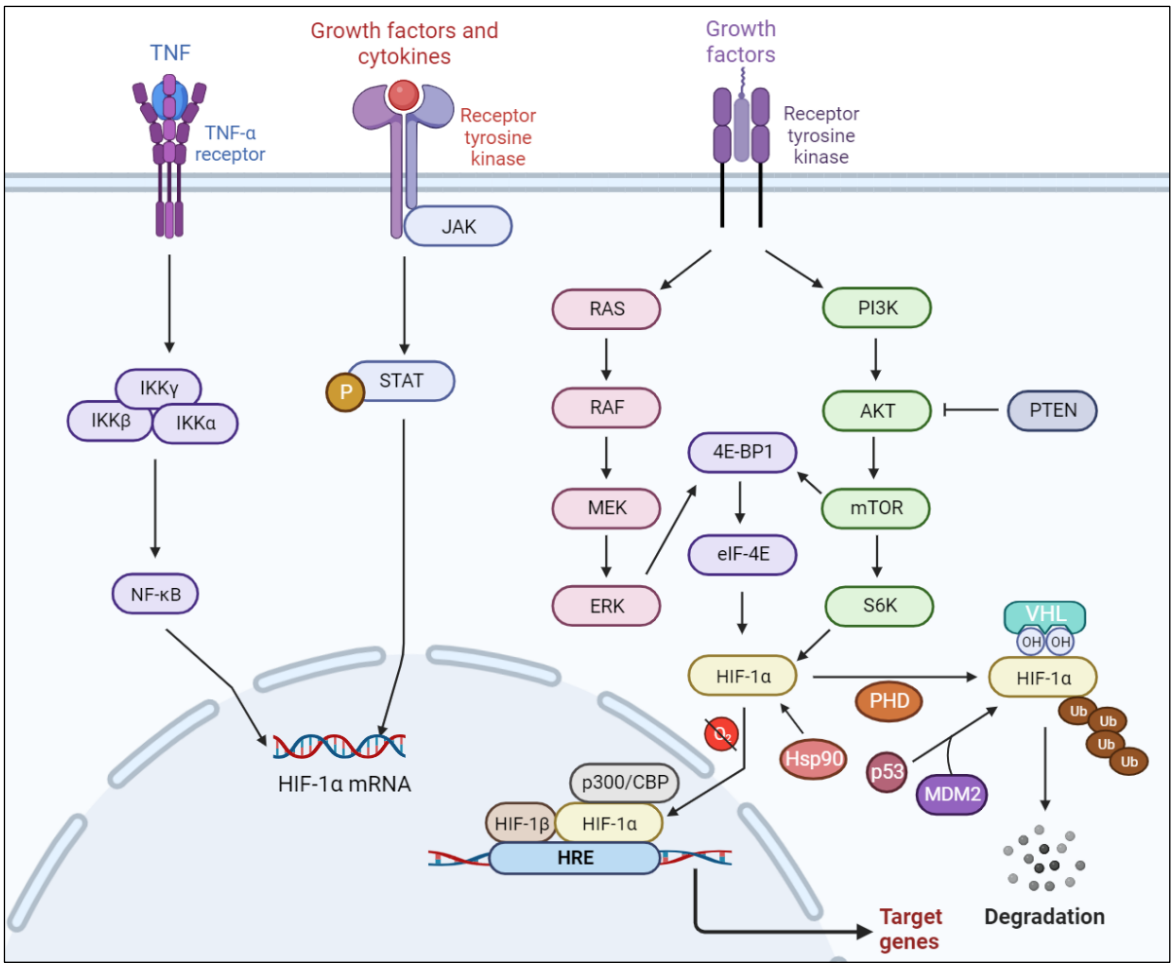


Figure 5.7. Regulation of HIF-α pathway at different levels. HIF-1α undergoes regulation through a complex network of signal transduction pathways in response to growth factors and cytokines. These pathways include NF-κB, RAS-RAF-MEK-ERK, PI3K/AKT/mTOR, and JAK-STAT signalling pathways. Additionally, HSP90 directly interacts with HIF-1α and contributes to its stabilization. The tumour suppressor proto-oncogene p53 (p53) provides a route for HIF-1α degradation, and the murine double minute 2 (MDM2) protein ligase promotes the degradation of HIF-1α by binding to the HIF-1α/p53 complex. CBP, CREB-binding protein; FIH, factor inhibiting HIF; HRE, hypoxia-responsive elements; HSP90, heat shock protein 90; p300, histone acetyltransferase p300; TGF-α, transforming growth factor alpha; Ub, ubiquitin. Information collected from [410].

5.2.7 Herbs in regulating HIF-α pathway

It is known that oxygen was transported by red blood cells in blood vessels, therefore, angiogenesis plays a crucial role for the oxygen delivery and consumption. During the past decades, the regulation of angiogenesis has become a therapeutic target for cancer and ischaemic stroke. Studies have demonstrated that HIF-mediated responses are associated with erythropoiesis and

angiogenesis [417], and HIF-α has been suggested to be an important target in treating cancer and ischaemia stroke duo to its role in regulating the transcriptional activity of downstream genes [342].

Herbal medicines, including herbal formulas, extracts, and chemical ingredients, have been widely used in the treatment of cardiovascular and cerebrovascular diseases for a long history [418]. Recently, there has been increasing attention focused on the herbal medicines or herbal ingredients as potential activators targeting the HIF-1α pathway. For instance, ginsenoside, a major active ingredient of *ginseng*, has been shown to upregulate HIF-1α/VEGF pathway in neural stem cells (NSCs) in oxygen-glucose deprivation/reperfusion (OGD/R) models, and thereby promote the survival, self-renewal and differentiation of NSCs [419]. Total flavonoid extracts from *Caragana sinica* can attenuate neurological deficits, reduce infarct volume and promote angiogenesis in a rat model of transient middle cerebral artery occlusion (tMCAO) via enhancing the expression of CD31, VEGF, Ang-1, HIF-1α, DII4 and Notch1 [420]. Obviously, numerous preclinical studies have provided supportive evidence for the use of herbal medicines as novel anti-angiogenesis or HIF modulator therapies for cancer or other diseases by targeting HIF-1α pathway. However, the underlying mechanisms of these herbal medicines and their complex chemical and pharmacological properties, along with the interactions between the multiple bioactive ingredients are still required to gain further exploration.

5.2.8 Role of PHD in treating anaemia disease

CKD encompasses various disease processes that lead to irreversible damage and dysfunction of the kidneys over a prolonged period, typically months to years. One common feature of CKD is the presence of anaemia, which is often characterized as normocytic, normochromic and hypoproliferative. The development of anaemia in CKD is attributed to multiple factors, including reduced production of EPO by the kidneys, reduced red blood cell survival and iron deficiency [421]. Anaemia in CKD has significant implications for patient outcomes. It is associated with reduced quality of life, increased risk cardiovascular disease, higher rates of hospitalization, cognitive impairment, and increased mortality [422, 423].

In the past few decades, the primary drugs used for the treatment of anaemia were iron supplements and recombinant EPO and its synthetic derivatives, collectively known as erythropoiesis-stimulating agents (ESAs) [424]. These ESAs have been widely used in clinical therapy for renal anaemia, however, their non-oral administration routes often result in low patient compliance. Additionally, some hyporesponsive patients treated with exogenous ESAs have shown increased cardiovascular risks [425-427]. Thus, there is a need for new drug treatment strategies,

such as the development of small-molecule agents, to address these limitations and provide alternative options for anaemia management in CKD patients.

EPO is primarily produced by the kidneys, and it plays a crucial role in stimulating the production of red blood cells in the bone marrow and maintaining haemoglobin homoeostasis. EPO acts as a key regulator of erythropoiesis by inhibiting apoptosis of erythroid progenitor cells and promoting their differentiation [428]. The production of EPO is tightly regulated by HIF-α. As described earlier, when the catalytic activity of PHD is inhibited, HIF-α can escape the hydroxylation, and form a heterodimer with HIF-β, which then translocate into the nucleus, trigger the expression of target genes, including EPO and VEGF [397, 398]. In the context of CKD and anaemia, inhibiting PHD activity will lead to the accumulation of HIF and subsequent stimulation of erythropoiesis even under normoxic conditions. Therefore, pharmacological inhibition of PHD is considered a promising therapeutic strategy for the treatment of anaemia in CKD patients.

There are three main PHD homologs identified in human, namely PHD1, PHD2, and PHD3 [429], while PHD2 is considered a key therapeutic subtype for treating renal anaemia [430]. PHD2 belongs to the 2OG dependent dioxygenase superfamily and possesses a catalytic centre that contains a 2OG molecule, a ferrous iron, and a binding water molecule. Structural studies have provided insights into the interactions within the 2OG binding site. The C1 carboxylate and C2 carbonyl groups of 2OG coordinate with the ferrous iron in a bidentate manner, forming an octahedral complex with residues such as HIS313, HIS374, ASP315, and a water molecule. The C5 carboxylate of 2OG interacts with ARG383 through a salt bridge. Additionally, TYR303 forms a hydrogen bond with the C1 carboxylate of 2OG via an adjacent water molecule [431]. Apart from the 2OG binding site, there is a substrate binding site outside this region. Residues including TRP389, PHE391, TRP258, and ARG322 are part of this region and are believed to interact with the HIF-α substrate through hydrophobic interactions [432]. Based on the understanding of the binding modes within PHD2, a considerable number of PHD2 inhibitors have been developed. These inhibitors target the catalytic centre, aiming to modulate PHD2 activity and stabilize HIF-α, ultimately promoting erythropoiesis and addressing anaemia (Figure 5.8).

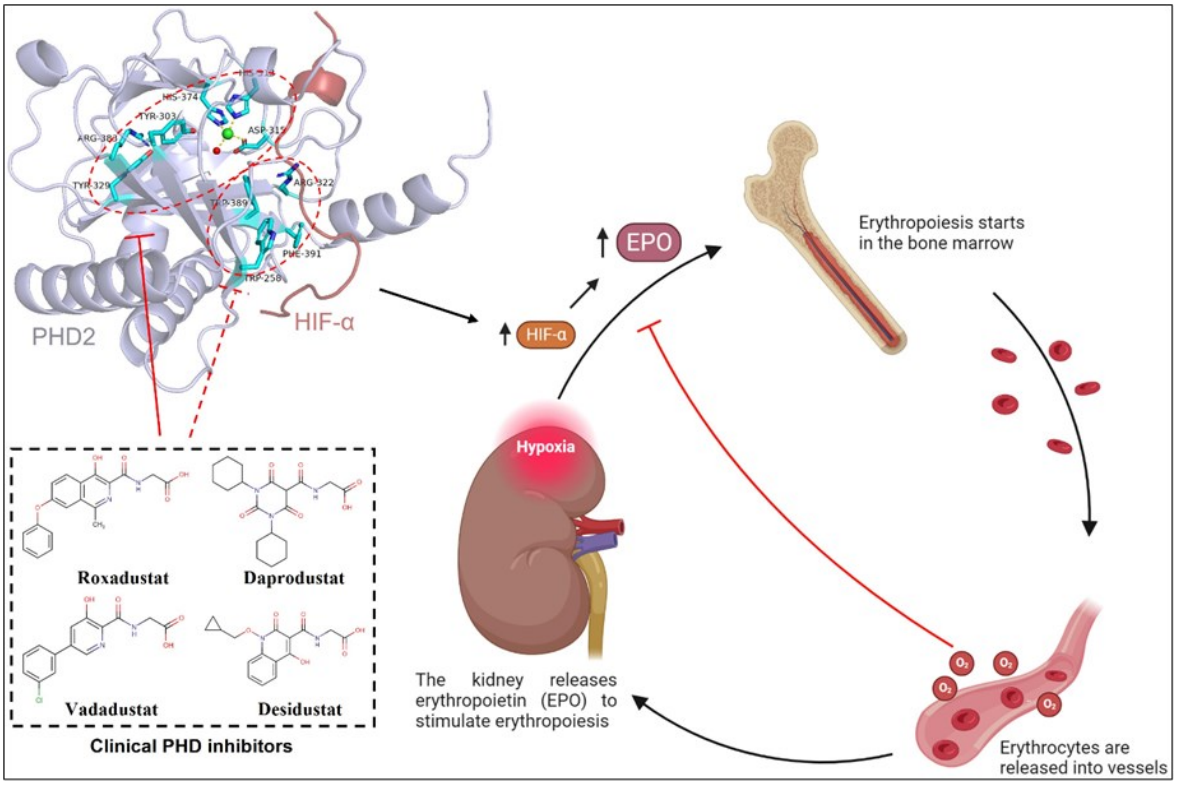


Figure 5.8. PHD2 inhibitors in Erythropoietin (EPO) production. EPO promotes production of mature red blood cells in the bone morrow. More red blood cells in circulation increase, oxygen levels rise, resulting in lower levels of hypoxia-inducible factor (HIF) and suppression of EPO production. In the presence of PHD2 inhibitors, the PHD2 loses its catalytic activity and leads to HIF-α stabilization, which subsequently stimulates EPO production and increases red blood cells count.

5.2.9 Hydroxybenzoic acids

In our previous findings, we have demonstrated the significant effect of gallic acid within *Limonium Sinense* on HIF activation. It was reported that the two phenolate oxygen atoms of gallic acid form a chelate with the active site Fe^{2+} of PHD2, while the carboxyl group of gallic acid interacts strongly through ionic/hydrogen bonding with ARG383. Interestingly, nPG, which has an esterified carboxyl group, is unable to inhibit the hydroxylase [315]. **Since gallic acid belongs to the hydroxybenzoic acid family, these findings suggest us that the hydroxybenzoic acids with similar chemical structure to gallic acid may possess the potential to inhibit PHD2 activity**.

Hydroxybenzoic acid (HBA) and its derivatives belong to the category of phenolic compounds characterized by a general C6-C1 backbone structure, directly derived from benzoic acid (BA), include monohydroxybenzoic acid (MHBA), dihydroxybenzoic acid (DHBA), trihydroxybenzoic acid (THBA), tetrahydroxybenzoic acid (tetraHBA), and pentahydroxybenzoic acid (PentaHBA) (Figure

5.9). The structural variations in hydroxybenzoic acids arise from processes such as aromatic ring hydroxylation and methylation [433]. These compounds exhibit the capacity to mitigate oxidative stress and inflammation by promoting the expression of antioxidant enzymes. Additionally, they possess the ability to inhibit adipocyte lipolysis through the activation of hydroxycarboxylic acid receptors. As a result, they potentially contribute to the improvement of plasma lipid profiles, thereby offering potential benefits for overall health [434].

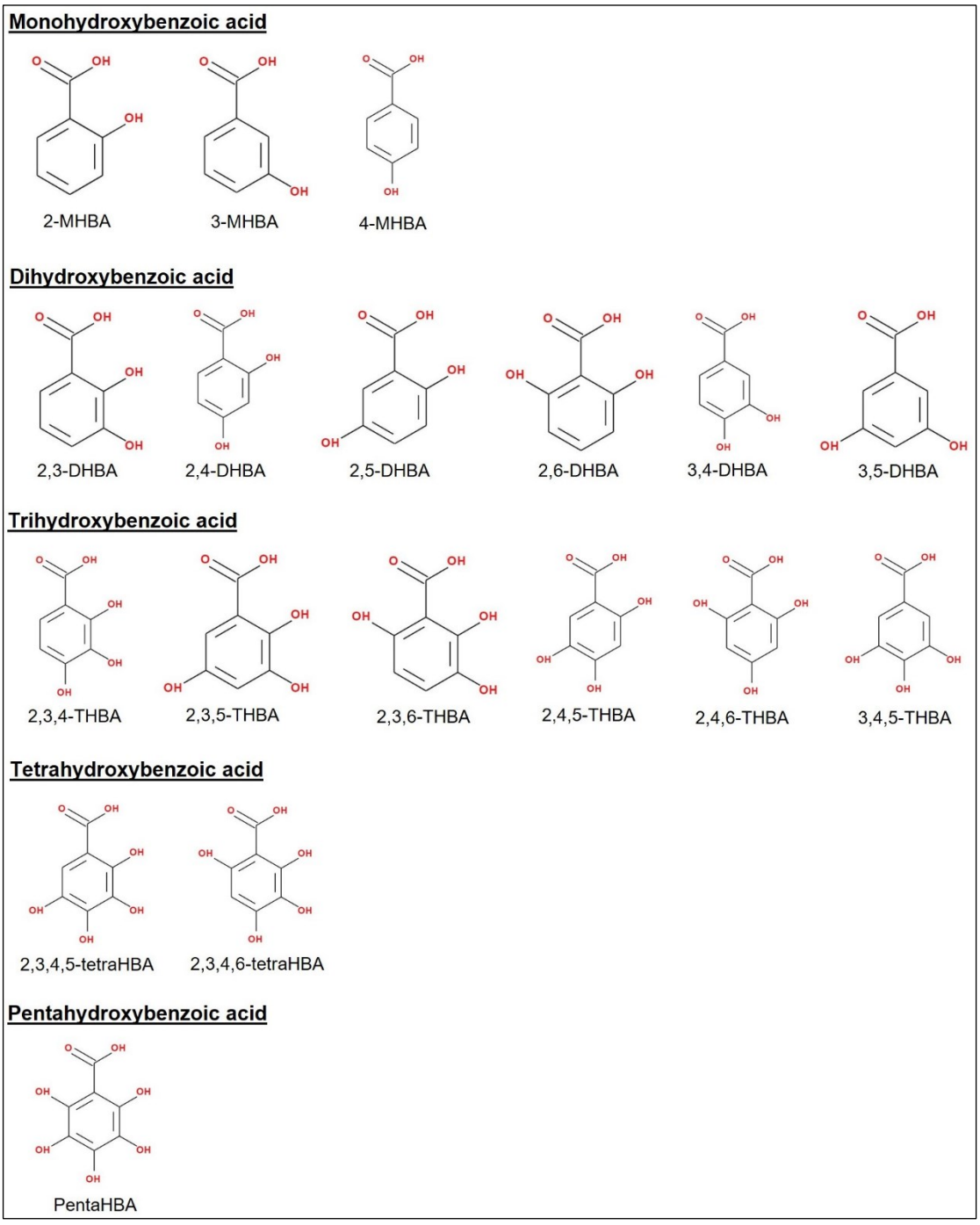


Figure 5.9. Structures of the common natural hydroxybenzoic acids. Graph showing the isomers of monohydroxybenzoic acid (MHBA), dihydroxybenzoic acid (DHBA), trihydroxybenzoic acid (THBA), tetrahydroxybenzoic acid (tetraHBA) and pentahydroxybenzoic acid (PentaHBA). Chemical diagrams drawn using molsoft (https://molsoft.com/mprop/).

Various hydroxybenzoic acids exhibit structural similarities to both 2OG and ascorbate, which are essential reactants in the prolyl-4-hydroxylase reaction. These substances have been identified as competitive inhibitors of prolyl-4-hydroxylase concerning both cosubstrates [435]. The purpose of the present work is to identify novel compounds among hydroxybenzoic acid derivatives that could effectively inhibit the activity of PHD2. Supporting our hypothesis, recent reports have indicated that 3,4-dihydroxybenzoic acid (3,4-DHBA, also known as protocatechuic acid) can stabilize HIF-α by substituting 2OG and ascorbate [436]. Ethyl-3,4-dihydroxybenzoate (EDHB), an esterified analogue of 3,4-DHBA, has demonstrated the ability to elevate serum EPO levels and significantly improve exercise performance in mice subjected to hypoxic conditions [437]. Based on these findings, it is plausible to speculate that hydroxybenzoic acids might possess the potential to effectively inhibit PHD2 activity, holding promise for the treatment of ischemic and vascular diseases such as anaemia.

In this study, we collected hydroxybenzoic acids and their natural derivatives from two public databases, namely the Dictionary of Natural Products (https://dnp.chemnetbase.com) and Coconut (https://coconut.naturalproducts.net). Additionally, we utilized the BindingDB database (http://bdb2.ucsd.edu/bind/index.jsp), a publicly accessible resource that provides information on measured binding affinities between drug-target proteins and small, drug-like molecules. We focused on compounds that exhibit inhibitory effects on PHD2 activity. These collected compounds have been assembled into a compound library to identify new PHD2 inhibitors by using the structure-based virtual screening approach. This approach allows us to explore potential compounds that may interact with PHD2 and modulate its activity, providing new avenues for the development of PHD2 inhibitors with potential therapeutic applications.

5.3 Aims

Based on the previous studies, we have discovered the ability of water extracts from *Limonium Sinense* (LSW) to activate HIF. Subsequent experiments confirmed the presence of gallic acid, a member of the hydroxybenzoic acid family, in LSW, and demonstrated its impact on HIF activation. According to recent findings that some hydroxybenzoic acid derivatives can induce HIF activation via inhibiting PHD2 activity, we therefore intended to explore the potential candidates from hydroxybenzoic acids and compounds that exhibit potent inhibitory effects on PHD2 activity through structure-based virtual screening. The aims of this part of the research are as follows: **1)** Identify new potential PHD2 inhibitors from hydroxybenzoic acid derivates and public databases. **2)** Exploring the potential mechanism of the new PHD2 inhibitor candidates interacting with PHD2.

5.4 Results

5.4.1 Virtual screening

Hydroxybenzoic acids and their natural derivatives were collected from the Dictionary of Natural Products and Coconut databases by searching “hydroxybenzoic acid”, resulting in a total of 881 compounds. Additionally, 266 compounds with inhibitory effects on PHD2 activity, but lacking detailed mechanism research, were obtained from the BindingDB database. Furthermore, as positive controls, four clinical PHD2 inhibitors containing 2OG analogues, namely Roxadustat, Daprodustat, Molidustat and Vadadustat, were also included in the compound collection. Altogether, a compound library was assembled consisting of 1,151 compounds.

A virtual screening was then performed using the compound library. The screening process involved applying filter rules and excluding candidates that had previously reported interaction models with PHD2 from the BindingDB database. As a result, 261 candidates matched the screening criteria (Figure 5.10, Supplementary Table 5.1). The next step involved docking each of these 261 candidates into the crystal structure of PHD2 (PDB: 2G19) using the GOLD software. The docked poses were assessed using GoldScore. Finally, the top 1% ranked candidates based on GoldScore will be selected to do the further analysis.

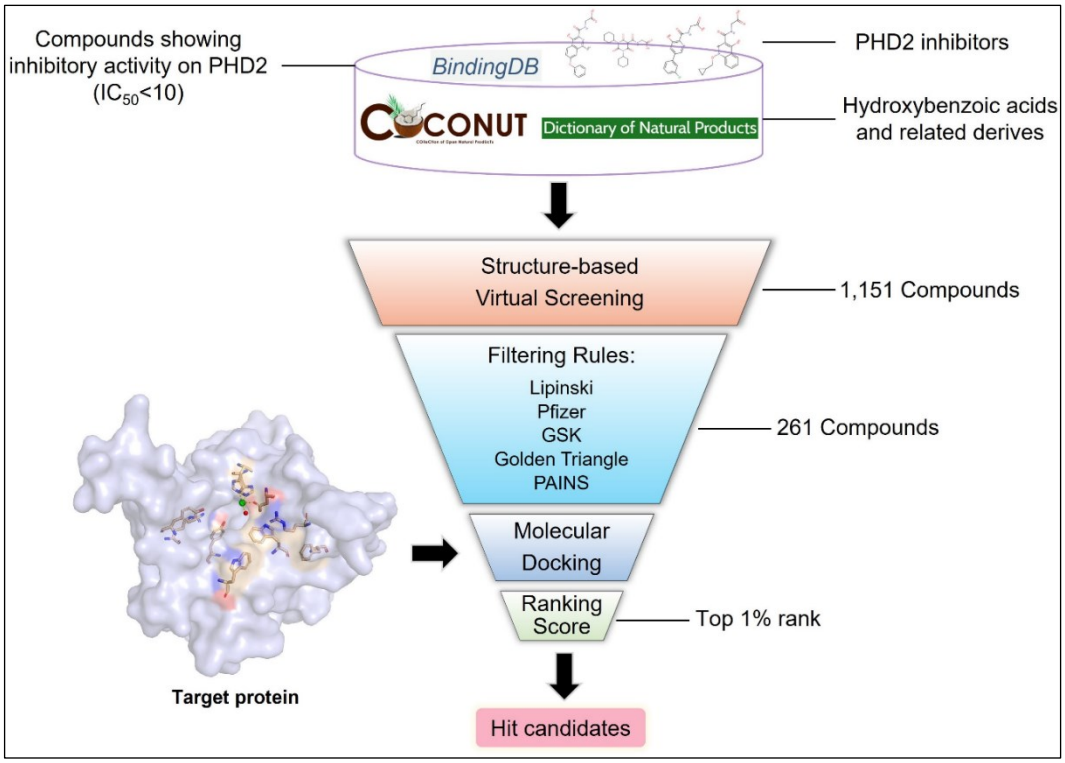


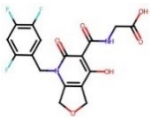
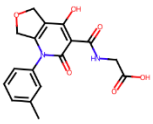
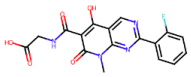
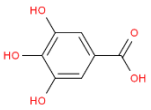
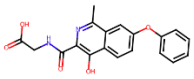
Figure 5.10. The construction of compound library and screening process for PHD2 inhibitors. Graph showing the structure-based virtual screening workflows and filtering rules. Counts of the compounds in each step were indicated.

5.4.2 Molecular docking analysis

Docking studies were performed using the GOLD software to evaluate the binding of the 261 screened compounds with PHD2. To validate the docking models, the native ligands were docked back into the binding pockets. The resulting compounds from the docking screening were ranked based on their GoldScores, which is a scoring function provided by the GOLD software. In addition to the criteria of GoldScore, the docked poses were manually filtered through visual inspection. The focus was on identifying an appropriate pharmacophoric motif responsible for PHD2 binding and eliminating compounds that exhibited detrimental interactions with critical residues of the PHD2 binding pocket or unrealistic docking conformations [438].

Among the 261 compounds that were docked with PHD2, three compounds ranked within the top 1% of the GoldScore, namely 2-[[4-Hydroxy-2-oxo-1-[(2,4,5-trifluorophenyl)methyl]-5,7-dihydrofuro[3,4-b]pyridine-3-carbonyl]amino]acetic acid (short for C1, GoldScore 84.8272), 2-(4-hydroxy-2-oxo-1-(m-tolyl)-1,2,5,7-tetrahydrofuro[3,4-b]pyridine-3-carboxamido)acetic acid (short for C2, GoldScore 84.2385), and 2-(2-(2-Fluorophenyl)-5-hydroxy-8-methyl-7-oxo-7,8-dihydropyrido[2,3-d]pyrimidine-6-carboxamido)acetic acid (short for C3, GoldScore 82.0329). Compared to the reference PHD2 inhibitor Roxadustat (GoldScore 77.898), all three compounds showed higher GoldScores. Furthermore, these compounds also have been reported to possess lower IC_{50} values (4.4 nM, 8.6 nM and 1.5 nM, respectively) compared to the PHD2 inhibitor Roxadustat (which has an IC_{50} value of 9.8 nM, as shown in Table 5.1), indicating that these candidates have the potential to be more effective PHD2 inhibitors. The table also presents the GoldScore of gallic acid binding with PHD2, indicating a lower value compared to other compounds.

Table 5.1. Molecular docking results showing Top 1% ranked compounds plus Gallic acid

Compound ID	ChemPub ID	Name	Chemical Structure	Gold Score	IC50 (nM)
C1	118981731	2-[[4-Hydroxy-2-oxo-1-[(2,4,5-trifluorophenyl)methyl]-5,7-dihydrofuro[3,4-b]pyridine-3-carbonyl]amino]acetic acid		84.8272	4.4
C2	118978236	2-(4-hydroxy-2-oxo-1-(m-tolyl)-1,2,5,7-tetrahydrofuro[3,4-b]pyridine-3-carboxamido)acetic acid		84.2385	8.6
C3	54741100	2-(2-(2-Fluorophenyl)-5-hydroxy-8-methyl-7-oxo-7,8-dihydropyrido[2,3-d]pyrimidine-6-carboxamido)acetic acid		82.0329	1.5
Inhibitor	370	Gallic acid		54.9323	-
Inhibitor	11256664	Roxadustat		77.898	9.8

5.4.3 Binding model analysis

In order to investigate the binding models of the three top-ranked candidates (C1, 2-[[4-Hydroxy-2-oxo-1-[(2,4,5-trifluorophenyl)methyl]-5,7-dihydrofuro[3,4-b]pyridine-3-carbonyl]amino]acetic acid; C2, 2-(4-hydroxy-2-oxo-1-(m-tolyl)-1,2,5,7-tetrahydrofuro[3,4-b]pyridine-3-carboxamido)acetic acid and C3, 2-(2-(2-Fluorophenyl)-5-hydroxy-8-methyl-7-oxo-7,8-dihydropyrido[2,3-d]pyrimidine-6-carboxamido)acetic acid) and Roxadustat (Rox) with PHD2 (PDB ID: 2G19), we examined their individual interaction. The molecular docking results provided valuable information on the interactions between these candidates and the PHD2 protein, as shown in Figure 5.11 and Table 5.2.

For candidate C1, it was observed that the compound formed a conventional hydrogen bond with ARG383 and ARG322. It also established a carbon hydrogen bond with HIS313 and HIS374. Additionally, a pi-sigma bond and a pi-pi stacked bond were formed with MET299 and TYR310, respectively. Furthermore, a halogen (Fluorine) bond was identified between C1 and VAL314 and ASP315 (Figure 5.11A).

Candidate C2 bound toPHD2 through a conventional bond with ARG383 and formed a carbon hydrogen bond with HIS313 and HIS374. A pi-sulfur bond was observed between C2 and MET299, while a pi-pi stacked bond and a pi-alkyl bond were formed with interaction of TYR310 and TRP389, respectively. In addition, C2 exhibited unfavourable acceptor-acceptor bond with TYR303 and ASP315 (Figure 5.11B).

In the case of candidate C3, it formed an unfavourable acceptor-acceptor bond with ASP315 and a conventional hydrogen bond with ARG383 and TYR310. It also established a carbon hydrogen bond with HIS313 and HIS374, as well as a pi-sulfur bond with MET299. Moreover, C3 interacted with ILE256 and TRP389 through a pi-alkyl bond (Figure 5.11C). All the three candidates (C1, C2 and C3) were found to interact with the 2OG binding pockets and exhibit a chelate effect with the iron ion, indicating the potential inhibitory effects of these candidates on PHD2 activity.

In comparison, Roxadustat was observed to bind to ARG383 and TYR303 through conventional hydrogen bond. It also formed a carbon hydrogen bond with HIS313 and HIS374. Additionally, a pi-sulfur bond, a pi-pi stacked bond and a pi-alkyl bond were established with MET299, TYR310 and TRP389, respectively. Obviously, Roxadustat occupies the 2OG binding pocket and demonstrated the inhibitory effect on PHD2 activity (Figure 5.11D).

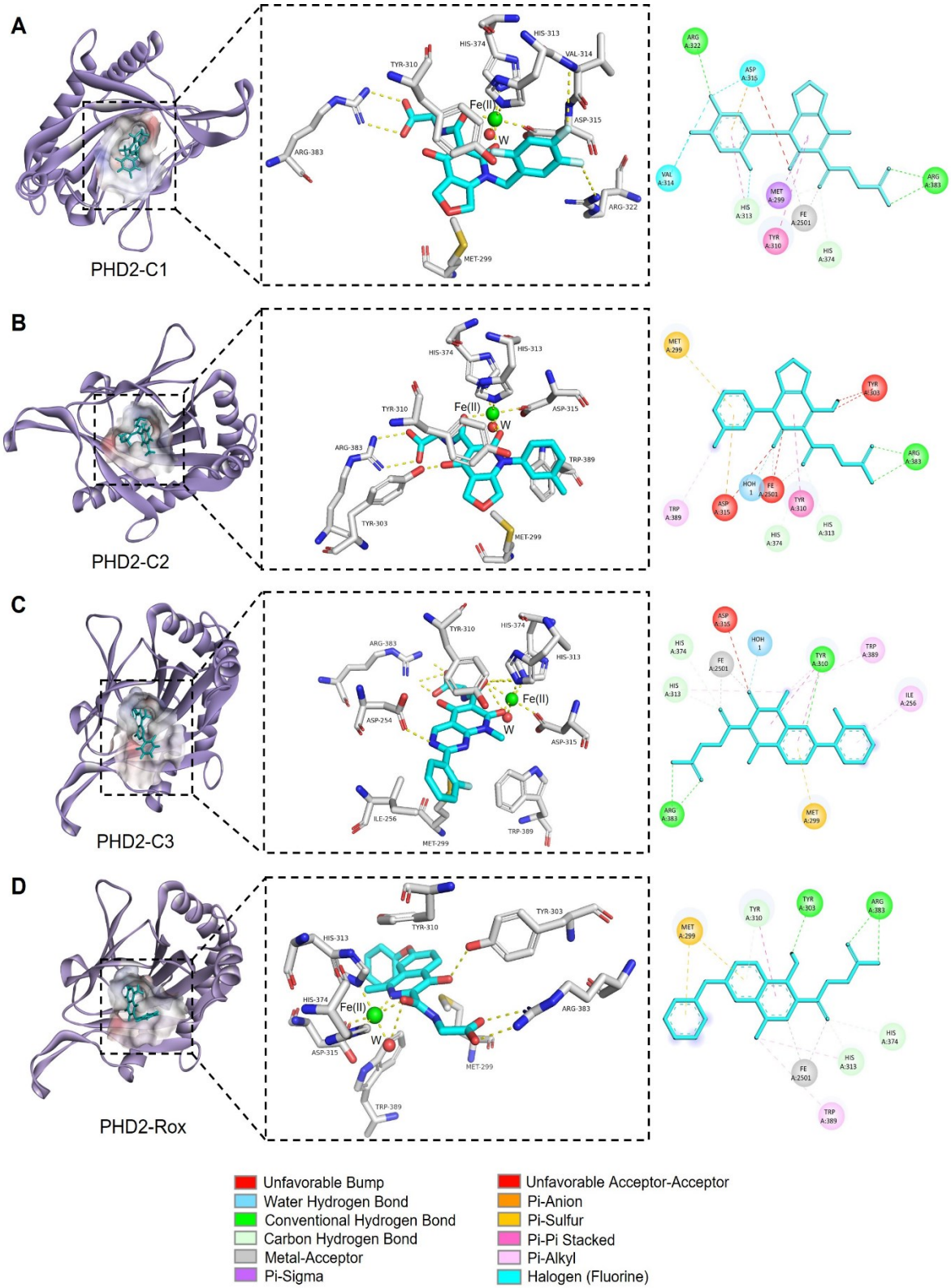


Figure 5.11. Interactions between the three candidate inhibitors and PHD2 protein. Three-dimensional graphs showing the X-ray crystallographic pose of PHD2 docked with candidate C1 (**A**), candidate C2 (**B**), candidate C3 (**C**) and Roxadustat (**D**). Two-dimensional graphs showing the interactions of PHD2 with the three candidates and inhibitor Roxadustat.

Table 5.2. Interactions between PHD2 and identified inhibitors

Protein	Candidate name	Interaction residues	Interaction bonds
PHD2	C1	ARG383, ARG322, HIS313, HIS374, VAL314, ASP315, MET299, TYR310	Conventional Hydrogen Bond, Carbon Hydrogen Bond, Metal-Acceptor, Halogen (Fluorine), Unfavorable Acceptor-Acceptor, Pi-Anion, Pi-Pi Stacked, Pi-Sigma
	C2	TYR303, ASP315, ARG383, HIS313, HIS374, MET299, TYR310, TRP389	Unfavorable Bump, Water Hydrogen Bond, Conventional Hydrogen Bond, Carbon Hydrogen Bond, Metal-Acceptor, Unfavorable Acceptor-Acceptor, Pi-Anion, Pi-Pi Stacked, Pi-Alkyl
	C3	ARG383, TYR310, HIS313, HIS374, ASP315, TRP389, ILE256, MET299	Water Hydrogen Bond, Conventional Hydrogen Bond, Carbon Hydrogen Bond, Metal-Acceptor, Unfavorable Acceptor-Acceptor, Pi-Sulfur, Pi-Pi Stacked, Pi-Alkyl
	Rox	ARG383, TYR303, TYR310, HIS313, HIS374, MET299, TRP389	Conventional Hydrogen Bond, Carbon Hydrogen Bond, Metal-Acceptor, Pi-Donor Hydrogen Bond, Pi-Sulfur, Pi-Pi Stacked, Pi-Alkyl

Overall, we have identified three significant potent PHD2 inhibitors. While additional *in vitro* and *in vivo* experiments are required to confirm the precise mechanisms of action for these novel PHD2 inhibitors, it is hypothesized that they function by occupying the 2OG binding pocket and inhibiting PHD2 activity. Consequently, the accumulation of HIF-α and subsequent production of EPO are promoted, leading to an enhancement in the count of red blood cells.

5.5 Discussion

It is known that hypoxia, or low oxygen levels, triggers the process of angiogenesis, which involves the formation of new blood vessels and sprouting from existing ones to increase the oxygen supply. The expression of various angiogenic markers, including VEGF, TGF-β, PDGF-β, plasminogen activator-1 (PAI-1), EPO and GLUT-1, is regulated by HIF-1α [439].

HIF-1α is known to be involved in all stages of blood vessel formation and is a key mediator of angiogenesis in both physiological and pathophysiological conditions [440, 441]. In addition, the genes targeted by HIF-1α make it an appealing pharmacological target, especially for the treatment of diseases such as anaemia, ischaemic stroke, and wound healing through the upregulation of HIF via inhibition of PHD enzymes [149].

Anaemia is primarily caused by a decrease in EPO production in the kidneys, and it can be partially alleviated by using ESAs. However, randomized controlled trials have demonstrated that targeting normal haemoglobin levels using ESAs can be harmful, raising concerns about the benefits of ESA treatment beyond avoiding blood transfusions [442]. As a result, there is an urgent need for the development of new drugs to effectively treat anaemia.

PHD inhibitors have emerged as a promising category of drugs for the treatment of anaemia by stabilizing HIFs activity, which subsequently increases the production of EPO. Currently, most of the reported PHD2 inhibitors are structurally derived from or mimic the endogenous cofactor 2OG, exhibiting common pharmacophores that are consistent with the binding mode between PHD2 enzyme and 2OG cofactor [443]. The 2OG binding pocket, critical for the catalytic activity of PHD2, comprises a ferrous iron bidentate with HIS313, ASP315 and HIS374, along with a binding water molecule and residues ARG383 and TYR303. Additionally, a substrate binding site, involving residues TRP389, PHE391, TRP258 and ARG322, has identified to potentially interact with the HIF-α substrate through hydrophobic interactions. These two binding sites greatly contribute to the development of drug discovery for anaemia in CKD.

In this study, we have pinpointed three potential PHD inhibitor candidates from hydroxybenzoic acids and their derivatives. All of them demonstrated a robust binding effect with the 2OG pocket of PHD2. Furthermore, when compared to the clinical PHD2 inhibitor Roxadustat used in the study, all three candidates exhibited lower IC_{50} concentrations, indicating their potent inhibitory activity against PHD2. However, further experiments, such as molecular dynamic simulations, cellular experiments, tissue models and X-ray crystallographic analysis, are still necessary to validate the effects and underlying mechanisms of these new PHD2 inhibitor candidates on PHD2.

Chapter 6 Final discussion

6.1 Potential links among *Limonium Sinense*, cell cycle and cancer

In recent years, there has been significant research focusing on the anticancer properties of *Limonium Sinense*. For instance, studies have reported that the root extract of *Limonium Sinense* can effectively inhibit the survival of HepG2 cells in melanoma-bearing mice [444]. Polysaccharides derived from the root of *Limonium Sinense* have demonstrated potent anti-tumour and immunomodulatory activities with low toxicity, effectively suppressing tumour growth in mice [27]. Notably, LSP21, a polysaccharide separated from crude *Limonium Sinense* polysaccharides, has shown remarkable anti-tumour effects by inhibiting cell proliferation and inducing cell death [28]. Our studies have demonstrated that both the water and ethanol extracts of *Limonium Sinense* can significantly inhibit the proliferation of multiple breast cancer cell lines in both 2D and 3D cultures. The mechanism of action may involve the arrestation of G2/M phase in cell cycle.

As mentioned earlier, the cell cycle is a complex process governed by a multiple of regulatory proteins that guide cell through specific events, culminating in mitosis and the production of two daughter cells [445]. Accurate regulation of the cell cycle is vital for cell proliferation, growth, and repair, relying on intricate interactions among proteins, enzymes, cytokines, and cell cycle signalling pathways [268]. Checkpoints play a crucial role in ensuring the fidelity of cell division by monitoring various factors such as growth, DNA replication, DNA damage, and kinetochore binding to the mitotic spindle [446]. Upon DNA damaged, cell cycle checkpoints can be activated at different stages, including G1 phase, S phase, and at the G2/M transition.

Control of cell cycle progression depends on various factors. Among these factors, the cyclin-dependent kinase (CDK) inhibitor p21 (also known as $p21^{WAF1/Cip1}$) is involved in promoting cell cycle arrest in response to diverse stimuli. The p21 gene was the first to be identified as induced by wild-type p53 protein [447]. The expression of the p21 gene can be directly mediated by p53 in a p53-dependent process, or it can be induced by growth factors such as TGF-β or mimosine in a p53-independent manner [448-450].

In the context of cell cycle progression, transitions from G1 to S phase and from G2 to mitosis are regulated by the sequential activation and inactivation of CDK proteins, a family of serine/threonine protein kinases. Cyclins serve as regulatory subunits of CDKs and are degraded or synthesized during the cell cycle. CDK1, CDK2, CDK4 and CDK6 are among the current identified CDKs that playing crucial roles in cell cycle regulation [451, 452]. Cyclins and CDKs are responsible for the regulation of cell cycle progression, and specific Cyclin/CDK complexes can mediate the initiation

and transitions of the cell cycle [453]. Cyclins can bind to different CDKs, forming Cyclin/CDK complexes that regulate cell cycle activities, for instance, Cyclin E/CDK2 complexes in G1-S transition are responsible for licensing DNA origins of replication. CDK2 later binds with Cyclin A during the S phase, while CDK1 participates in the S-G2 and G2-M transitions by sequentially binding to Cyclin A and Cyclin B [454]. The CDK activity is mediated by the inhibitory proteins known as CDK inhibitors. Among these inhibitors, p21 acts as a cell cycle inhibitor and tumour suppressor [455, 456] and has been reported to inhibit Cyclin/CDK complexes at the G1 checkpoint and modulate the transition from the G1 phase to the S phase [268]. The elevated p21 can also cause cell growth arrest at the G2 phase [457]. Studies have demonstrated that p21 can suppress cell cycle progression in G1 and S phases by inhibiting the activity of Cyclin/CDK2, Cyclin/CDK1 and Cyclin/CDK4,6 complexes [458, 459].

On the other hand, p21 is involved in multiple critical pathways that are often dysregulated in cancer, and alterations in p21 expression have been observed in various human cancers [271, 460]. These alterations in p21 expression can have significant implications for cancer susceptibility to chemotherapy, radiation, and targeted therapy [455]. Studies have revealed that p21 has a dual role in tumour cells, as it can either promote or inhibit tumorigenesis. The growth inhibitory activities of p21 primarily involve the inhibition of cyclin-dependent kinases, such as CDK1 and CDK2. In addition, p21 can reduce proliferation independently of CDKs by targeting proteins like proliferating cell nuclear antigen (PCNA), which is required for S phase progression. Furthermore, p21 can enhance the assembly of Cyclin/CDK complexes to promote cell cycle progression at low levels. Moreover, p21 plays a critical role in maintaining tumour stem cell populations and can activate processes such as epithelial-mesenchymal transition (EMT) that promote metastasis. Therefore, p21 serves as a master regulator of the cellular response to stress, exerting its influence on various aspects of tumour development and progression [461].

Overall, p21 plays vital roles in mediating cell cycle control and cancer progression. In our study, we observed that treatment with LSW significantly upregulates the expression of p21 at the gene level. Conversely, the gene expression of several cell cycle regulators such as CDK1, CCNA2 (Cyclin A2) and CCNB2 (Cyclin B2) was significantly downregulated by LSW treatment in MDA-MB-468 cells. Furthermore, we also demonstrated that LSW induces cancer cell growth inhibition through p53-independent pathways. While further evidence is required to confirm these results, our findings suggest that LSW induces G2/M phase arrest possibly by stimulating p21 expression, thereby inhibiting the Cyclin A/CDK1 and Cyclin B/CDK1 complexes, leading to the suppression of the cell cycle checkpoint, ultimately resulting in cell growth inhibition.

6.2 Potential links among *Limonium Sinense*, HIF and anaemia

Our findings highlighted the importance of the HIF-1 signalling pathway in the biological processes mediated by *Limonium Sinense*. HIF-1 acts as a crucial regulator of oxygen homeostasis and plays essential roles in cellular response to hypoxia, such as glycolysis, erythropoiesis, angiogenesis, and vascular remodelling [441].

It is known that HIF involves in a wide range of adaptive responses, primarily focusing on the upregulation of transcriptional cascades that are critical for tissue protection and adaptation. HIF-1α is associated with the upregulation of glycolytic genes such as phosphoglycerate kinase (PGK) and lactate dehydrogenase A (LDHA). These genes facilitate metabolic adaptation to oxygen deprivation and anaerobic ATP synthesis. On the other hand, HIF-2α induces the expression of EPO and VEGF, which are important for improving oxygen supply to hypoxic regions [384]. HIF-1α primarily contributes to acute hypoxia-driven transcriptional responses [462], and undergoes ubiquitination and proteasomal degradation in non-hypoxic cells [317, 374, 463]. However, under hypoxic conditions, HIF-1α escapes hydroxylation and ubiquitination, leading to its accumulation [464, 465].

The stability of HIF-1α is regulated by a variety of interacting proteins, which can influence its degradation in an oxygen-dependent or oxygen-independent manner. Oxygen-dependent degradation is mediated by the prolyl hydroxylase domain proteins PHD1, PHD2 and PHD3 [275]. PHD1 is primarily localized in the nucleus and exhibits high expression levels in the testis, with low expression levels in the heart [466, 467]. PHD2 is predominantly found in the cytoplasm and its inhibition primarily enhances the activity of HIF-1α, rather than HIF-2α. It is moderately expression in brain, kidney, liver, and exhibits high expression levels in the heart [467, 468]. PHD3 is evenly distributed between the cytoplasm and nucleus, with high expression levels in the heart and liver [467]. The activity of HIF hydroxylases is dependent on the activation of PHDs in cells which leads to ubiquitin-mediated proteasomal degradation of prolyl hydroxylated HIF-1α [469].

PHD inhibitors, which stabilize HIF-α and activate the HIF-1 pathway, have been widely used to treat anaemia in CKD patients. Reports have shown that PHD inhibitors can improve iron absorption, downregulate hepcidin, and increase the endogenous production of EPO [470]. Moreover, inhibition of PHDs has been shown in numerous studies to modulate endogenous adaptive programs triggered in response to hypoxia, which has implications in ischemic organ injuries [471], fibrosis [472], inflammation [473], and oxidative stress [474]. These processes have been implicated in the progression of cardiovascular and renal disease [475].

On the other hand, loss of VHL in tumour cells results in increased HIF-1α expression and transcription of downstream target genes such as VEGF [318, 476]. Studies on knockout and knock-in mice have shown the crucial role of HIF-2α [477, 478], PHD2 [479], and VHL [480] in regulating adult erythropoiesis. As *Limonium Sinense* is traditionally used for the treatment of anaemia, the mechanism underlying this effect may involve the induction of HIF activation and elevation of EPO levels.

Our results indicate that the HIF-1 signalling pathway was significantly enriched both in the network pharmacology analysis and the RNA-sequencing analysis of LSW-treated MDA-MB-468 cells. Furthermore, *in vitro* experiments validated the upregulating effect of LSW on HIF activation. We demonstrated that LSW treatment induces HIF-1α at the protein level and this is not mediated by down-regulating PHDs. In line with our findings, Tsukiyama and colleagues reported that gallate is able to inhibit PHD activity, thereby reducing the HIF degradation rate and increasing the protein level of HIF-1α [315].

Overall, these findings provide novel insights for the research on *Limonium Sinense* in the treatment of anaemia. Based on our findings, Figure 6.1 illustrates the potential mechanism of *Limonium Sinense* in inducing the cell cycle arrest and inhibiting cancer cell growth.

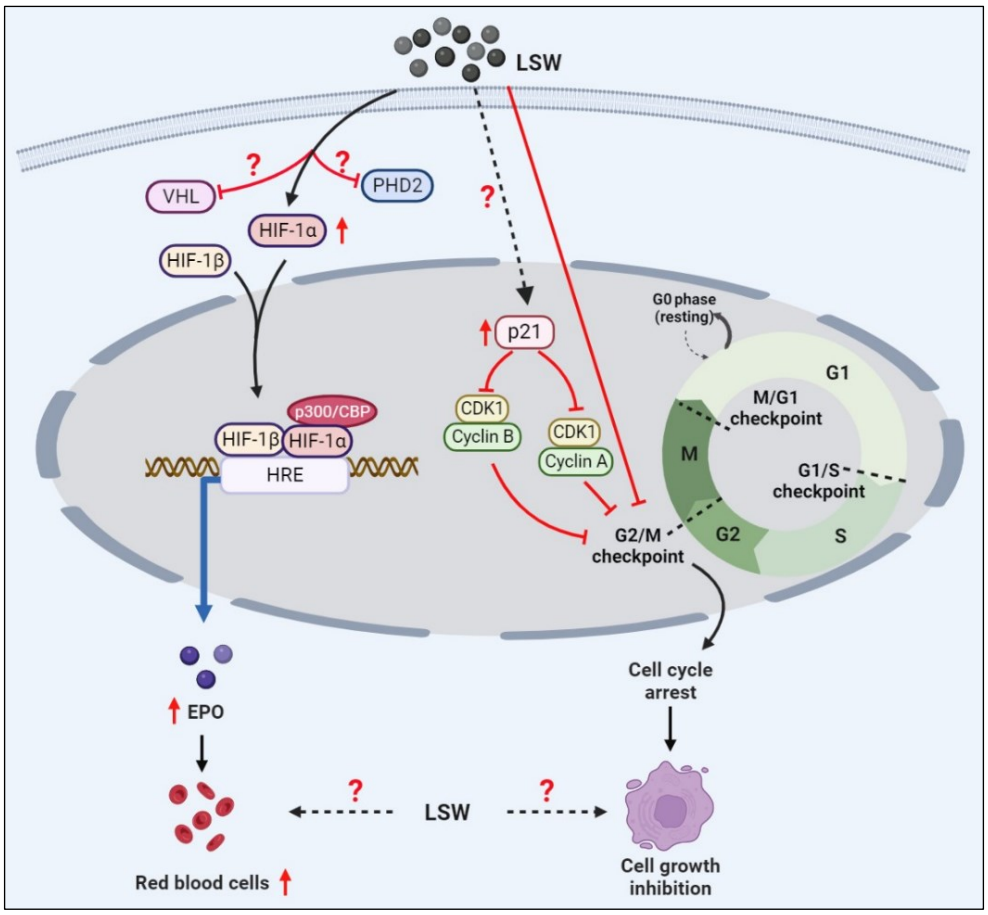


Figure 6.1. Potential mechanism for *Limonium Sinense* water extract (LSW) in activating HIF and inducing cell cycle arrest. The potential mechanism by which LSW induces G2/M phase cell cycle arrest may involve the activation of p21, thus inhibiting the activity of Cyclin A/CDK1 and Cyclin B/CDK1 complexes and leading to the cell cycle arrest at the G2/M checkpoint. However, additional experiments are necessary to address the following aspects: 1) Confirm whether LSW-induced cell cycle arrest is mediated by p21 activation, this will help establish the involvement of p21 as a key factor in the cell cycle arrest observed after LSW treatment. 2) Validate the exact role of p21 in LSW-induced cell growth inhibition due to its dual role in tumour cells, this will contribute to provide a clearer understanding of its mechanism of action in the context of tumour cells. It is also important to investigate whether there are alternative mechanisms underlying the LSW-induced cell growth inhibition. On the other hand, LSW induces the upregulation of HIF-1α in protein levels, which can promote the production of EPO and enhance the cell counts of red blood cells. Nevertheless, the effects of LSW on EPO and red blood cell counts still required validation. It is also important to investigate if LSW has alternative mechanisms for enriching red blood cells. Furthermore, the specific mechanism by which LSW activates HIF, such as whether it inhibits PHD2 or VHL activities, remains unclear and warrants further investigation. CDK1, Cyclin Dependent Kinase 1; CDK2, Cyclin Dependent Kinase 2; CDK4, Cyclin Dependent Kinase 4; CDK6, Cyclin Dependent Kinase 6; EPO, Erythropoietin; HIF-1α, Hypoxia-inducible factor 1 alpha; HIF-1β, Hypoxia-inducible factor 1 beta; PHD2, proline-hydroxylase 2; VHL, von Hippel-Lindau.

6.3 Other factors implicated in the effects of *Limonium Sinense*

Obviously, HIF plays a central role in connecting the cell cycle arrest, anticancer effects, and anti-anaemia effects of *Limonium Sinense*. It has been reported that the induction of HIF-1α leads to G1 phase cell cycle arrest in multiple cell types, including various cancer cell lines [326, 481-483]. HIFs can alter cell cycle progression through transcriptional targets such as Cyclin D1 and indirect modulation of CDK inhibitors p21 and p27 [484-487]. In addition, HIF-1α serves as a key transcription factor in cancer progression and target therapy in cancer. Overexpression of HIF-1α supports cancer progression through various mechanisms including angiogenesis, cell proliferation and survival, metabolism reprogramming, invasion and metastasis, and cancer stem cell maintenance [488, 489]. Therefore, targeting HIF-1α presents a promising approach in the treatment of cancers.

The anti-inflammatory and antioxidant activities of *Limonium Sinense* are believed to play important roles in its biological effects. As previously described, *Limonium Sinense* exhibits significant immunomodulatory effects by increasing the production of IFN-γ and IL-2 while decreasing the production of IL-4 [30]. Additionally, *Limonium Sinense* shows remarkable scavenging abilities against ROS, particularly DPPH, ABTS radicals and hydroxyl radicals [21-23]. Both inflammation and oxidative stress are closely associated with HIF-1 signalling. Immune cells extravasate from the rich bloodstream to the site of inflammation will cause hypoxia and subsequently active HIF-1 signalling [490, 491]. Reports have demonstrated the involvement of HIF-1 in both intrinsic and extrinsic activation of tumour-associated inflammatory signalling [492, 493]. Hypoxic and necrotic regions within tumours can generate proinflammatory mediators that recruit immune cells, resulting in immune response suppression at the tumour site and promoting tumour cell proliferation, angiogenesis, and metastasis [493, 494]. On the other hand, both inflammatory cells and hypoxia can induce the generation of ROS, while sustained oxidative stress can contribute to chronic inflammation, which is implicated in various chronic diseases, including cancer, cardiovascular, neurological, and pulmonary diseases [495-497]. It is reported that oxidative stress can influence the expression and activity of HIF-1α through signalling molecules such as PHDs [498], PI3K [499], and microRNA [500]. Additionally, oxidative stress can also participate in inflammation mediated by HIF-1 through interacting with NF-κB and TGF-β signalling pathways [474]. Overall, the connections among inflammation, oxidative stress, cell cycle regulation, anaemia and cancer progression are intricately intertwined through the HIF-1 signalling pathway, suggesting a promising approach for investigating the biological effects of *Limonium Sinense*.

6.4 Future work

As discussed earlier, future research will primarily concentrate on the following areas to enhance our understanding of the biological effects of *Limonium Sinense*.

6.4.1 Validating the mechanism of LSW induces G2/M cell cycle arrest

As we observed that the water extract of *Limonium Sinense* (LSW) has been shown to have a significant impact on inducing G2/M phase arrest in the cell cycle, it is important to examine the expression of cell cycle checkpoint-related factors to verify the underlying mechanism, such as p21, cyclins and CDKs. Our analysis revealed notable alterations in the gene expression levels of p21, CDK1, CCNA2 (Cyclin A2) and CCNB2 (Cyclin B2) through RNA-seq count levels following LSW treatment. To validate these findings, it is necessary to further examine the changes in protein levels of these factors.

6.4.2 Validating the effects of LSW on angiogenesis, EPO and red blood counts

Based on our observations, there was a notable enrichment of angiogenesis in both GSEA Hallmark analysis and KEGG pathway analysis, and the expression of angiogenesis-related genes was significantly altered by LSW treatment. However, further validation is required to confirm the specific effects of LSW on angiogenesis. Additionally, considering the association of *Limonium Sinense* with anaemia treatment, it is important to investigate the effects of LSW on EPO and red blood cell counts through *in vitro* and *in vivo* experiments. This will help determine the potential impact of LSW on promoting red blood cell production and addressing anaemia.

6.4.3 Validating the mechanism of LSW on HIF activation

In our current results, we have demonstrated that gallic acid, one of the active compounds present in LSW, exhibits the ability to activate HIF. The potential mechanism behind this could involve the inhibition of PHD2 activity. To confirm this result, further validation of the effect of LSW on PHD2 activity will be conducted. Furthermore, we have observed that LSW can inhibit the binding between VHL and HIF-1α. Therefore, it is crucial to investigate the effect of LSW on VHL and explore the underlying mechanism involved in this interaction.

6.4.4 Isolation and identification of active compounds from LSW

Further investigations will prioritize the isolation and identification of specific components within LSW. The identification of the precise component or components responsible for mediating HIF

activation or inducing cell growth inhibition is crucial. Isolation experiments on LSW can offer valuable insights and potentially lead to the discovery of new candidate drug for the treatment of human diseases.

6.4.5 Validating the effectiveness of PHD2 inhibitors identified through virtual screening

In this study, we have identified three potential PHD2 inhibitors using a structural-based virtual screening method. However, further experiments are essential to validate the effectiveness of these newly discovered inhibitors on PHD2 activity. These validations will involve conducting *in vitro* and *in vivo* experiments. Additionally, it is crucial to investigate the underlying mechanisms of these new PHD2 inhibitor candidates and confirm their impact on PHD2 activity. More importantly, it is crucial to assess the effects of these PHD2 inhibitors on EPO levels and red blood cell counts to confirm their potential application in treating anaemia disease. These investigations will help determine the efficacy and suitability of these inhibitors for managing anaemia diseases.

Appendix A R Scripts

A.1 R Scripts for chapter 3

Raw data were imported into RStudio (*version 4.2.0*), Microsoft Windows (*version 11*) and R scripts were run.

setwd("C:/Users/hz2u19/Downloads").

A.1.1 R codes for obtaining GSE119552 and GSE120550 rawdata

```
library(GEOquery)

library(limma)

#GSE119552

gse119552 <- getGEO("GSE119552", GSEMatrix =TRUE, AnnotGPL=FALSE)

if (length(gse119552) > 1) idx <- grep("GPL16699", attr(gse119552, "names")) else idx <- 1

gse119552 <- gse119552[[idx]]

gse119552_rawdata<- cbind(fData(gse119552), exprs(gse119552))

gse119552_rawdata <- aggregate(x=gse119552_rawdata[,2:(ncol(gse119552_rawdata))],

by=list(gse119552_rawdata $Gene_Symbol), FUN = median)

gse119552_rawdata <- na.omit(gse119552_rawdata)

rownames(gse119552_rawdata) <- gse119552_rawdata $Group.1

gse119552_rawdata <- gse119552_rawdata [ ,-1]

write.csv(gse119552_rawdata, "GSE119552_rawdata.csv")


#GSE120550

gse120550 <- getGEO("GSE120550", GSEMatrix =TRUE, AnnotGPL=FALSE)

if (length(gse120550) > 1) idx <- grep("GPL17692", attr(gse119552, "names")) else idx <- 1
```

```
gse120550 <- gse120550[[idx]]

gse120550_rawdata<- cbind(fData(gse120550), exprs(gse120550))

gse120550_rawdata <- aggregate(x=gse120550_rawdata[,2:(ncol(gse120550_rawdata))],

by=list(gse120550_rawdata $Gene_Symbol), FUN = median)

gse120550_rawdata <- na.omit(gse120550_rawdata)

rownames(gse120550_rawdata) <- gse120550_rawdata $Group.1

gse120550_rawdata <- gse120550_rawdata [ ,-1]

write.csv(gse120550_rawdata, "GSE120550_rawdata.csv")
```

A.1.2 R codes for Figure 3.5

```
library(ggplot2)

library(RColorBrewer)

library(gridExtra)

library(dplyr)

#Top 20 KEGG pathways (Figure 3.5A)

Top20_kegg<- read.csv("Top20_KEGG_david.csv", header = T)

ggplot(Top20_kegg, aes(x = Fold.Enrichment, y = reorder(Name, -PValue),

                       color = -log10(PValue), size = Count))+ geom_point()+

  scale_color_gradient(low = "deepskyblue4", high = "red3")+

  theme_bw()+

  labs(x="Fold Enrichment", y="", title="KEGG enrichment analysis")+

  theme(axis.text.x = element_text(angle = 270, hjust = 0.1, vjust = 0.5))

#Top 5 GO_BP, GO_CC and GO_MF (Figure 3.5B)

top5_Go<- read.csv("Top5_GO_david.csv", header = T)
```

```
ggplot(top5_Go, aes(x = Fold.Enrichment, y = reorder(Pathway, -PValue),
                    color = -log10(PValue), size = Count))+
 geom_point()+
 scale_color_gradient(low = "blue", high = "red")+
 theme(plot.subtitle = element_text(size=15, angle=0))+
 theme(axis.title.y = element_text(size=15, angle=0)) +
 theme(axis.title.x = element_text(size=15, angle=0)) +
 theme(legend.position="top") +
 theme_bw()+
 theme(panel.grid.major = element_line(color = "gray95",
                                       size = 0.01, linetype = 1)) +
 labs(x="Fold Enrichment", y="", title="") +
 facet_grid(cols=vars(Category), scales = "free_y",space = "free_x")
```

A.1.3 R codes for Figure 3.8

```
library(disgenet2r)
disgenet_api_key <- get_disgenet_api_key(email = "hz2u19@soton.ac.uk",
 password = "ZhaoHL1988")
Sys.setenv(DISGENET_API_KEY= disgenet_api_key)

#Figure 3.8A
res_enrich <- disease_enrichment(entities = c("PGD","AKR1C1","AKR1C2","AKR1C3","AMY1A",
                                              "ANTXR2","CA14","CA12","CA3","CA1","CA2","CA4","CA5A",
                                              "CA5B","CA6","CA7","CA9","CTRC","HSD17B14","POLA1","POLB",
```

"DPP4","ELAVL3","ERN1","CES1","CES2","FTO","FUT4","FUT7",

"GABRB1","GLRA3","GLRB","GPR35","HIF1AN","HMGB1","HSPA1A",

"IGFBP1","IGFBP2","IGFBP3","IGFBP4","IGFBP5","IGFBP6","KDM2A",

"KDM3A","KDM4D","KDM4E","KDM5C","SELL","SELE","SELP",

"NEU3","NR4A1","DDO","P4HA1","P4HTM","SERPINE1","PARP15",

"PGAM1","POLH","POLI","RARB","RARG","SLC22A6","SRD5A2",

"CXCL12","ST3GAL3","SMAD3","TOP1","TPMT","TTR",

"TYR","SLC7A11","ACR","ATF1","GLB1",

"CACNA1D","COMT","CYP1A2","DBF4","EIF2AK3",

"ELANE","GABRA2","GABRA6","GABRB2",

"GABRG2","GRM6","GSTM2","KAT2A","KDM6A",

"PRKCH","PRKCZ","LNPEP","ALOX15","ALOX5",

"LOXL2","NCOA1","NEK6","P2RY12","PLA2G5",

"PABPC1","PHLPP2","PIN4","PLAA","ALPG",

"RARA","SLC22A2","SLC29A1","SENP7","STAT1",

"TLR1","TLR2","TDP1","MPG","APP","ABCG2",

"ACHE","AHR","AKR1B10","ALDH2","AKR1B1",

"AR","MAOA","CALM1","CBR1","CBS","CCR4",

"CDK5R1","CDK6","CFTR","CYP1A1","CYP19A1",

"CYP1B1","CYP2C8","CYP2C19","CYP3A4","CREB1",

"CSNK2A1","DAPK1","ODC1","HSD17B1","HSD17B2",

"HSD17B3","ELAVL1","ERAP1","ESRRB","ESR1",

"ESR2","FLT3","GRK6","GSK3B","GSTO1","IL2",

"KCND3","RPS6KA3","SYK","GLO1","ALOX12",

"ABCB1","MIF","ABCC1","NMUR2","NOX4","NUAK1",

"PDE4D","P4HB","MPO","PTGS2","PKN1","PGF",

"PON1","ACP1","ALPI","PTPRS","ST6GAL1",

"TERT","TNKS","TNKS2","NAE1","VEGFA","XDH",

"MAOB","KLK2","PDK4","PTPN1","TAS2R31",

"CD38","CSNK2A2","CSNK2B","QDPR","MMP12",

"MMP2","MMP9","PARP1","SNCA","MAPT","GNG2",

"GNB1","MYLK","NQO1","TOP2A","YWHAG",

"STK17B","ALPL","SFN","GLA","AKR1A1","AKR1C4",

"GFER","BCL2A1","BACE1","CA13","CAPN1","CCR2",

"HSPE1","HSPD1","BCHE","LIG1","DNM1","EP300",

"ERCC1","ESRRG","EZR","FOS","FOXO1","FYN",

"HDAC11","HDAC10","HDAC2","HDAC4","HDAC1",

"HDAC5","HDAC3","HDAC7","HDAC6","HDAC9",

"HDAC8","HTT","HPSE","JUN","KCNK2","LHCGR",

"MMP1","MYOC","IKBKG","NFE2L2","NFKB1","NQO2",

"NR0B2","PTGDR","PTGER2","PTGER4","PYGL",

"PTP","SLC22A1","SLC6A5","SCNN1A","SPR",

"STAT6","YARS1","TAAR1","TRPM2","TRPM5",

"TRPV1","TH","ERCC4","AKT1","ALK","AURKB",

"BRAF","CDK1","CYP2C9","CXCR1","DRD4","EGFR",

"PTK2","IGF1R","CAMK2B","MET","MMP13","MMP3",

"NEK2","PIK3R1","PLA2G1B","PIM1","F2","AXL",

"AVPR2","KDR","LRP6","MCL1","SLC22A12",

"NDUFAB1","NDUFAF1","DUSP3","ABCC2","NDUFA1",

"NDUFA2","NDUFA3","NDUFA4","NDUFA5","NDUFA6",

"NDUFA7","NDUFA8","NDUFA9","NDUFA10","NDUFA11",

"NDUFA12","NDUFA13","NDUFB1","NDUFB2","NDUFB3",

"NDUFB5","NDUFB4","NDUFB6","NDUFB7","NDUFB9",

"NDUFB8","NDUFB10","NDUFB11","NDUFC1","NDUFC2",

"NDUFAF2","NDUFAF3","NDUFAF4","NDUFS1","NDUFS2",

"NDUFS3","NDUFS5","NDUFS4","NDUFS6","NDUFS7",

"NDUFS8","NDUFV1","NDUFV2","NDUFV3","NR0B1",

"MT-ND1","MT-ND2","MT-ND3","MT-ND4L","MT-ND4",

"MT-ND5","MT-ND6","NDUFA4L2","PTPN7","SLC22A3",

"SHBG","ABCC8","ADCY1","ADRB1","ADRB2","ADRB3",

"BDKRB2","CISD1","CNR2","CYP24A1","DHCR7",

"DRD3","DRD2","GCGR","HIF1A","IMPDH1","IMPDH2",

"CAMK2A","KLF5","LTB4R","METAP2","IGF2R","MTNR1A",

"MTNR1B","NOD2","NR1H4","ADCYAP1R1","PASK","PDK3",

"PTGER1","PTGS1","PTGES","SLC28A3","SENP1","TAOK3",

"TUBB1","RELA","F3","TLR9","TNFRSF1A","APEX1",

"CHEK1","TCF4","RAF1","ACVRL1","IL6","MAP3K7",

"MAP2K6","PDGFRA","SQLE","LDHA","LDHB","BCL2L1",

"ALB","DAO","SRC","LCK","MVD","FASN","TYMS",

"CDK2","ADRA2A","ADRA2C","ADRA2B","ADRA1A",

"DYRK1A","CLK1","DYRK1B","CRHR1","PLA2G4B",

"GRM2","KDM4A","STS","ALDH5A1","ABAT","CNR1",

"KDM6B","KDM4C","BCL2","ADRA1D","OPRM1",

"IDO1","RXRB","RXRG","RXRA","FOLH1","HTR3A",

"ADAMTS5","CDC7","CCNE1","PTK2B","TYMP",

"ERBB2","GRIA1","GRIA4","GRIA3","GRIA2",

"KCNMA1","CDK5","CCNB3","CCNB1","CCNB2",

"NISCH","CTSL","IKBKB","MPI","CXCR2","GPR84",

"NOS2","UPP1","MB","SLC9A1","PIK3CG","ADAM17",

"MMP14","VCAM1","MMP7","MMP8","PLEC","AHCY",

"CSNK1A1","CSNK1D","PRKDC","ADORA1","ADORA2A",

"ARG1","ESRRA","PFKFB3","TBXAS1","PLK1",

"ADORA3","HTR2C","PLG","GRM5","PPARG","SLC5A2",

"PLA2G2A","PLA2G10","KLK1","CHRNA7","KIT",

"FGFR1","INSR","EDNRA","SIRT2","PTGER3",

"PIK3CB","PIK3CA","ADCY5","VCP","WEE1",

"HNF4A","CDK4","CDK3","AURKA","KCNH2","FFAR1",

"CTSB","ATP4B","ATP4A","PNP","SAE1","UBA2",

"MTOR","PDK1","SGK1","ANPEP","EPHA2","YES1",

"BLK","CSK","EPHB2","BMX","LYN","EPHA5",

"EPHA4","TXK","FGR","EPHA6","PTK6","EPHB3",

"EPHA3","BTK","TYRO3","COQ8B","EPHA1","NR3C1",

"ROCK1","PRKACA","PDE4A","PDE4C","BMP1","ABL1",

"EPHA8","EPHA7","EPHB1","ROCK2","RPS6KB1",

"THRA","THRB","CLK3","DYRK2","MELK","NEK1",

"CTSD","MAP4K4","DNMT1","HPGDS","GABRA1",

```
                    "PRMT1","CCNA1","CCNA2","MPL","GUSB","NTRK1",
                    "OPRK1","ECE1","HSP90AA1","GLI2","CHEK2","GRK2","SLC5A1",
                    "TRPM8","CFD","TNF","AGTR1","PNMT","TRAP1","HSP90B1",
                    "HSP90AB1","AVPR1A","ADAM10","TSPO","PDE7A","MMP16","AKT2",
                    "MAP2K1","MAPK14","HPRT1","PLA2G7","GRIN1","GRIN2B",
                    "GRIN2A","PPARA","MME","CELA1","SIRT1","MTAP"),
          database = "CURATED")
plot(res_enrich, class = "Enrichment", count = 3, cutoff = 0.01, nchars=70, limit = 20)


#Figure 3.8B
res_enrich2 <- disease_enrichment(entities = c("AKT1","EGFR","SRC","ESR1", "GSK3B","PTGS2"),
                                  database = "CURATED")
plot(res_enrich2, class = "Enrichment", count = 3, cutoff = 0.01, nchars =70, limit = 20)
```

A.1.4 R codes for Figure 3.11

```
library(EnhancedVolcano)
#Figure 3.11A
gse119552_vol<- read.csv("gse119552_DEGs_from_GEO2.csv", header = T)
EnhancedVolcano(gse119552_vol, lab = NA, selectLab = NULL,
    subtitle = "DEGs of GSE119552", subtitleLabSize = 18,
    x = 'logFC', y = 'P.Value', ylim = c(0, 12), pCutoff = 0.05, FCcutoff = 1, title = "")

#Figure 3.11B
gse120550_vol<- read.csv("gse120550_DEGs_from_GEO2.csv", header = T)
```

```
EnhancedVolcano(gse120550_vol, lab = NA,
        selectLab = NULL, subtitle = "DEGs of GSE120550",
        subtitleLabSize = 18, x = 'logFC', y = 'P.Value',
        ylim = c(0, 10), xlim = c(-4, 3), pCutoff = 0.05, FCcutoff = 1, title = "")


#Figure 3.11C
gse212598_a2780_vol<- read.csv("gse212598_A2780_DEGs_from_GEO2.csv", header = T)
EnhancedVolcano(gse212598_a2780_vol, lab = NA, selectLab = NULL,
        subtitle = "DEGs of GSE212598(A2780)",
        subtitleLabSize = 18, x = 'log2FoldChange', y = 'pvalue',
        ylim = c(0, 300), xlim = c(-10, 5), pCutoff = 0.05,
        FCcutoff = 1, title = "")


#Figure 3.11D
gse212598_es2_vol<- read.csv("gse212598_es2_DEGs_from_GEO2.csv", header = T)
EnhancedVolcano(gse212598_es2_vol, lab = NA, selectLab = NULL,
        subtitle = "DEGs of GSE212598(ES2)", subtitleLabSize = 18,
        x = 'log2FoldChange', y = 'pvalue',
        ylim = c(0, 400), xlim = c(-4, 3), pCutoff = 0.05, FCcutoff = 1, title = "")
```

A.1.5 R codes for Figure 3.12

```
library(ggplot2)
library(RColorBrewer)
library(gridExtra)
```

Appendix A

```
library(dplyr)

library(pheatmap)

#Figure 3.12A

gse119552_kegg<- read.csv("gse119552_kegg.csv", header = T)

ggplot(gse119552_kegg, aes(x = DEG, y = reorder(Name, -PValue),

                           color = -log10(PValue), size = Count))+

 geom_point()+

 scale_color_gradient(low = "deepskyblue4", high = "red3")+

 theme_bw()+

 labs(x="", y="", title="")+

 theme(axis.text.x = element_text(angle = 270, hjust = 0.1, vjust = 0.5))+

 facet_grid(cols = vars(), scales = "free_x", space = "free_x")


#Figure 3.12B

gse120550_kegg<- read.csv("gse120550_kegg.csv", header = T)

ggplot(gse120550_kegg, aes(x = DEG, y = reorder(Name, -PValue),

                           color = -log10(PValue), size = Count))+

 geom_point()+

 scale_color_gradient(low = "deepskyblue4", high = "red3")+

 theme_bw()+

 labs(x="", y="", title="")+

 theme(axis.text.x = element_text(angle = 270, hjust = 0.1, vjust = 0.5))+

 facet_grid(cols = vars(), scales = "free_x", space = "free_x")
```

```
#Figure 3.12C

gse119552_breast_genes<- read.csv("gse119552_breastcancer_genes_david.csv",
                          row.names = 1)
pheatmap(gse119552_breast_genes,
       scale = 'row', treeheight_row = 20, treeheight_col = 20,
       annotation_row = NA, annotation_names_row = FALSE,
       cluster_rows = T, cluster_cols = T,
       color = colorRampPalette(c("deepskyblue4","white", "red3"))(500),
       border_color = "grey90", angle_col = 45, fontsize_col = 10)


#Figure 3.12D

gse119552_pic<- read.csv("gse119552_pathwaysincancer_genes_david.csv", row.names = 1)
pheatmap(gse119552_pic, scale = 'row',
       treeheight_row = 20, treeheight_col = 20,
       annotation_row = NA, annotation_names_row = FALSE,
       cluster_rows = T, cluster_cols = T,
       color = colorRampPalette(c("deepskyblue4","white", "red3"))(500),
       border_color = "grey90", angle_col = 45, fontsize_col = 10)


#Figure 3.12E

gse120550_pic<- read.csv("gse120550_pathwaysincancer.csv", row.names = 1)
pheatmap(gse120550_pic, scale = 'row', treeheight_row = 20,
       treeheight_col = 20, annotation_row = NA, annotation_names_row = FALSE,
       cluster_rows = T, cluster_cols = T,
```

```
    color = colorRampPalette(c("deepskyblue4","white", "red3"))(500),
    border_color = "grey90", angle_col = 45, fontsize_col = 10)
```

A.1.6 R codes for Figure 3.13

```
library(ggplot2)
library(RColorBrewer)
library(gridExtra)
library(dplyr)
library(pheatmap)
#Figure 3.13A
gse212598_a2780_kegg<- read.csv("gse212598_a2780_kegg.csv", header = T)
ggplot(gse212598_a2780_kegg, aes(x = DEG, y = reorder(Name, -PValue),
                    color = -log10(PValue), size = Count))+
 geom_point()+
 scale_color_gradient(low = "deepskyblue4", high = "red3")+
 theme_bw()+
 labs(x="", y="", title="")+
 theme(axis.text.x = element_text(angle = 270, hjust = 0.1, vjust = 0.5))+
 facet_grid(cols = vars(), scales = "free_x", space = "free_x")

#Figure 3.13B
gse212598_es2_kegg<- read.csv("gse212598_es2_kegg.csv", header = T)
ggplot(gse212598_es2_kegg, aes(x = DEG, y = reorder(Name, -PValue),
                      color = -log10(PValue), size = Count))+
```

```
 geom_point()+
 scale_color_gradient(low = "deepskyblue4", high = "red3")+
 theme_bw()+
 labs(x="", y="", title="")+
 theme(axis.text.x = element_text(angle = 270, hjust = 0.1, vjust = 0.5))+
 facet_grid(cols = vars(), scales = "free_x", space = "free_x")


#Figure 3.13C
gse212598_a2780_breast_genes<- read.csv("gse212598_a2780_breastcancer_genes2.csv",
                        row.names = 1)
pheatmap(gse212598_a2780_breast_genes, scale = 'row',
       treeheight_row = 20, treeheight_col = 20, annotation_row = NA,
       annotation_names_row = FALSE, cluster_rows = T, cluster_cols = T,
       color = colorRampPalette(c("deepskyblue4","white", "red3"))(500),
       border_color = "grey90", angle_col = 45, fontsize_col = 10)


#Figure 3.13D
gse212598_a2780_pic<- read.csv("gse212598_a2780_pathwaysincancer.csv", row.names = 1)
pheatmap(gse212598_a2780_pic, scale = 'row',
       treeheight_row = 20, treeheight_col = 20, annotation_row = NA,
       annotation_names_row = FALSE, cluster_rows = T, cluster_cols = T,
       color = colorRampPalette(c("deepskyblue4","white", "red3"))(500),
       border_color = "grey90",angle_col = 45, fontsize_col = 10)
```

Appendix A

#Figure 3.13E

```
gse212598_es2_breast_genes<- read.csv("gse212598_es2_breastcancer.csv", row.names = 1)

pheatmap(gse212598_es2_breast_genes, scale = 'row', treeheight_row = 20,

    treeheight_col = 20, annotation_row = NA, clustering_method = "ward.D2",

    #clustering_distance_rows = "minkowski", clustering_distance_cols = "euclidean",

    annotation_names_row = FALSE, cluster_rows = T, cluster_cols = T,

    color = colorRampPalette(c("deepskyblue4","white", "red3"))(500),

    border_color = "grey90", angle_col = 45, fontsize_col = 10)
```

A.2 R Scripts for chapter 4

A.2.1 R codes for obtaining LSW-treated DEGs from rawdata

```
RNA_seq <- read.table('LSW_rawcount.txt', header = T, sep = '\t')

expr = RNA_seq[, c(1, 7:12)]

colnames(expr) <- c("Gene_ID", 'C1', 'C2', 'C3', 'W1', 'W2', 'W3')

expr <- expr[c(2:67152), ]

#Prepare the row annotation from
#https://github.com/theislab/scvelo_notebooks/blob/master/data/biomart/mart_export#_human.txt.

annotation <- read.table('mart_export_human.txt', header = T, sep = '\t')

colnames(annotation) <- c("Gene_ID", "Gene_name")

expr <- merge(expr, annotation, by='Gene_ID')

expr <- expr[ , c(8, 2:7)]

#Remove the duplicated rows

expr1 <- aggregate(x=expr, by=list(expr$Gene_name), FUN=median)
```

```
rownames(expr1) <- expr1$Group.1

expr1 <- expr1[ , -1]

expr1 <- na.omit(expr1)

write.csv(expr1, "rsedata.csv")

rsedata <- read.csv("rsedata.csv", row.names = 1)

condition_rna_W <- factor(c(rep('C', length(rsedata[1,])/2), rep('W', length(rsedata[1,])/2)))

condition_rna_W

coldata_W <- data.frame(row.names=colnames(rsedata), condition_rna_W)

coldata_W

dds_W <- DESeqDataSetFromMatrix(countData=rsedata, colData=coldata_W,
design=~condition_rna_W)

nrow(dds_W)

dds_W <- dds_W[rowSums(counts(dds_W))>10, ]

dds_W <- DESeq(dds_W)

resultsNames(dds_W)

res <- lfcShrink(dds_W, coef="condition_rna_W_W_vs_C", type="apeglm")

res <- as.data.frame(res)

raw_count <- counts(dds_W, normalized=F)

raw_count <- as.data.frame(raw_count)

normalizedCounts2 <- counts(dds_W, normalized=T)

res <- data.frame(res, normalizedCounts2)

res <- data.frame(res,raw_count)

write.csv(res, "LSW_vs_Control_DESeq2.csv")
```

A.2.2 R codes for Figure 4.5

```
#Figure 4.5A

library(ggplot2)

library(RColorBrewer)

library(dplyr)

vdata<- read.csv("DEseq2-DEG-WvsC.csv", row.names = 1)

vdata_gene_type<- vdata%>%mutate(gene_type= case_when(

                        log2FoldChange>=1 & padj<= 0.05 ~ "up",

                        log2FoldChange<=-1 & padj<= 0.05 ~ "down", TRUE ~ "ns"))

cols<- c("up"="#ff8027", "down"="#008dd9", "ns"="grey")

ggplot(data = vdata_gene_type, aes(x = log2FoldChange, y = -log10(padj))) +

        geom_point(aes(colour = gene_type), alpha=0.3, shape=16, size=1.5) +

        geom_hline(yintercept=-log10(0.05), linetype="dashed") +

        scale_colour_manual(values=cols) +

        scale_x_continuous(breaks=c(seq(-10, 10, 2)), limits = c(-8, 8)) +

        scale_y_continuous(breaks=c(seq(0, 300, 100)), limits = c(0, 300)) +

        labs(title="LSW vs Control", x="Log2FoldChange",

             y="-Log10(padj)", colour="Expression\nchange") +

        theme_bw() +

        theme(panel.border=element_rect(colour="black", fill=NA, size=0.5),

            panel.grid.minor=element_blank(), panel.grid.major=element_blank())



#Figure 4.5A

library(pheatmap)
```

```
library(RColorBrewer)

w_data<- read.csv("wdata.csv", row.names = 1)

w_data<- wdata[rowSums(wdata[, -1])>10, ]

col_name2<- data.frame(Group=rep(c("Control", "LSW"), c(3,3)))

col.names(col_name2)<- colnames(w_data)

my_colour2 = list(Group = c(Control = 'lightcoral', LSW = 'steelblue1'))

pheatmap(w_ data, treeheight_row = 40, treeheight_col = 40, scale = 'row',

      annotation_col = col_name2, annotation_colors = my_colour2,

      annotation_row = NA, annotation_names_row = FALSE,

      cluster_rows = T, cluster_cols = T, fontsize_col = 15)
```

A.2.3 R codes for Figure 4.6

```
library(ggplot2)

library(RColorBrewer)

wgsea<- read.csv("LSW_ gsea.csv", header = TRUE)

ggplot(wgsea, aes(x = NES, y = reorder(ï..Hallmarks, -FDR),

            color = -log10(FDR), size = Gene Count))+

            geom_point()+

            scale_color_gradient(low = "blue", high = "red")+

            theme(plot.subtitle = element_text(size=11, angle=0))+

            theme(axis.title.y = element_text(size=11, angle=0)) +

            theme(axis.title.x = element_text(size=11, angle=0)) +

            theme(legend.position="top") +

            theme_bw()+
```

```
        theme(panel.grid.major = element_line(color = "grey80",
        size = 0.05, linetype = 1)) +
        labs(x="Normalized Enrichment Score", y="", title="GSEA:Hallmark pathways")
```

A.2.4 R codes for Figure 4.7

```
library(ggplot2)
library(RColorBrewer)
library(dplyr)
library(gridExtra)
w_go_kegg<- read.csv("lsw_go_kegg.csv", header = T)
ggplot(w_go_kegg, aes(x = Fold.Enrichment, y = reorder(Term, -PValue),
        color = -log10(PValue), size = Count))+
        geom_point()+
        scale_color_gradient(low = "blue", high = "red")+
        theme(plot.subtitle = element_text(size=11, angle=0))+
        theme(axis.title.y = element_text(size=11, angle=0)) +
        theme(axis.title.x = element_text(size=11, angle=0)) +
        theme(legend.position="top") +
        theme_bw()+
        theme(panel.grid.major = element_line(color = "grey80",
        size = 0.05, linetype = 1)) +
        labs(x="Fold Enrichment", y="", title="") +
        facet_grid(vars(ï..Category), vars(group), scales = "free_y",
              space = "free_y")
```

A.2.5 R codes for Figure 4.10

```
#Figure 4.10B

library(pheatmap)

library(RColorBrewer)

w_data<- read.csv("wdata.csv", row.names = 1)

w_data<- wdata[rowSums(wdata[, -1])>10, ]

pheatmap(wdata[c('VEGFA','SLC2A1','PGAM1','ENO1','LDHA','TPI1','P4HA1',

    'MRPS17','CDKN3','ADM','NDRG1','TUBB6','ALDOA','MIF','ACOT7'),],

    scale = 'row', treeheight_row = 40, treeheight_col = 40,

    annotation_row = NA, annotation_names_row = FALSE,

    cluster_rows = T, cluster_cols = T,

    color = colorRampPalette(c("blue","white", "red"))(500),

    border_color = "grey90", angle_col = 45, fontsize_col = 10)


#Figure 4.10C

library(GSVA)

w_data<- read.csv("wdata.csv", row.names = 1)

w_data<- wdata[rowSums(wdata[, -1])>10, ]

h<-list(c('VEGFA','SLC2A1','PGAM1','ENO1','LDHA','TPI1','P4HA1','MRPS17',

    'CDKN3','ADM','NDRG1','TUBB6','ALDOA','MIF','ACOT7'))

gsva_w <- gsva(as.matrix(wdata), h, method='gsva')

write.csv(gsva_w, "lsw_ hif_score.csv")
```

A.2.6 R codes for Figure 4.12

```
#Figure 4.12A

library(ggplot2)

library(RColorBrewer)

herb_up_data<- read.csv("herb up regulated hif score.csv", header = T)

ggplot(herb_up_data, aes(x = GEO.dataset, y = ï..Name,

                color = HIF.score.mean.difference, size = -log(Pvalue)))+

                geom_point()+

                scale_color_gradient(low = "blue", high = "red")+

                theme(axis.text.x = element_text(size = 10, angle=90,

                    hjust = 1), axis.title.y = element_text(size=10))+

                labs(x=" ", y=" ", title="HIF score")



#Figure 4.12B

library(ggplot2)

library(RColorBrewer)

herb_down_data<- read.csv("herb down regulated hif score.csv", header = T)

ggplot(herb_down_data, aes(x = GEO.dataset, y = ï..Name,

                    color = HIF.score.mean.difference, size = -log(Pvalue)))+

    geom_point()+

    scale_color_gradient(low = "blue", high = "red")+

    theme(axis.text.x = element_text(size = 10, angle=90, hjust = 1),

        axis.title.y = element_text(size=10))+

    labs(x=" ", y=" ", title="HIF score")
```

Appendix B Supplementary Tables

B.1 Supplementary Tables for chapter 3

Supplementary Table 3.1. Potential targets of *Limonium Sinense*

Target Name	Human Symbol
CA2	carbonic anhydrase 2
CA7	carbonic anhydrase 7
CA1	carbonic anhydrase 1
CA3	carbonic anhydrase 3
CA6	carbonic anhydrase 6
CA12	carbonic anhydrase 12
CA14	carbonic anhydrase 14
CA9	carbonic anhydrase 9
FUT7	fucosyltransferase 7
CA4	carbonic anhydrase 4
CA5B	carbonic anhydrase 5B
CA5A	carbonic anhydrase 5A
CA13	carbonic anhydrase 13
SQLE	squalene epoxidase
LDHA	lactate dehydrogenase A
LDHB	lactate dehydrogenase B
TTR	transthyretin
IGF1R	insulin like growth factor 1 receptor
ALK	ALK receptor tyrosine kinase
SERPINE1	serpin family E member 1
ESR2	estrogen receptor 2
BCL2L1	BCL2 like 1
GPR35	G protein-coupled receptor 35
COMT	catechol-O-methyltransferase
TPMT	thiopurine S-methyltransferase
AKR1C3	aldo-keto reductase family 1 member C3
ALB	albumin
TYR	tyrosinase
DAO	D-amino acid oxidase
AURKB	aurora kinase B
SRC	SRC proto-oncogene, non-receptor tyrosine kinase
PTK2	protein tyrosine kinase 2
KDR	kinase insert domain receptor
MET	MET proto-oncogene, receptor tyrosine kinase
NEK2	NIMA related kinase 2
AXL	AXL receptor tyrosine kinase
SRD5A2	steroid 5 alpha-reductase 2
AKR1C2	aldo-keto reductase family 1 member C2
AKR1C1	aldo-keto reductase family 1 member C1

Target Name	Human Symbol
FYN	FYN proto-oncogene, Src family tyrosine kinase
LCK	LCK proto-oncogene, Src family tyrosine kinase
CNR2	cannabinoid receptor 2
POLA1	DNA polymerase alpha 1, catalytic subunit
POLB	DNA polymerase beta
ESR1	estrogen receptor 1
ERN1	endoplasmic reticulum to nucleus signalling 1
PARP1	poly(ADP-ribose polymerase 1)
MVD	mevalonate diphosphate decarboxylase
TUBB1	tubulin beta 1 class VI
FASN	fatty acid synthase
TYMS	thymidylate synthetase
MCL1	MCL1 apoptosis regulator, BCL2 family member
MAOB	monoamine oxidase B
CDK2	cyclin dependent kinase 2
ELANE	elastase, neutrophil expressed
MAOA	monoamine oxidase A
ADRA2A	adrenoceptor alpha 2A
ADRA2C	adrenoceptor alpha 2C
ADRA2B	adrenoceptor alpha 2B
ADRB2	adrenoceptor beta 2
ADRB1	adrenoceptor beta 1
DRD2	dopamine receptor D2
ADRA1A	adrenoceptor alpha 1A
RARG	retinoic acid receptor gamma
RARB	retinoic acid receptor beta
DYRK1A	dual specificity tyrosine phosphorylation regulated kinase 1A
CLK1	CDC like kinase 1
DYRK1B	dual specificity tyrosine phosphorylation regulated kinase 1B
CRHR1	corticotropin releasing hormone receptor 1
DRD3	dopamine receptor D3
PTPN1	protein tyrosine phosphatase non-receptor type 1
PTGS2	prostaglandin-endoperoxide synthase 2
ACHE	acetylcholinesterase (Cartwright blood group)
PLA2G4B	phospholipase A2 group IVB
GRM2	glutamate metabotropic receptor 2
KDM4E	lysine demethylase 4E
EGFR	epidermal growth factor receptor
KDM4A	lysine demethylase 4A
DPP4	dipeptidyl peptidase 4
STS	steroid sulfatase
ALDH5A1	aldehyde dehydrogenase 5 family member A1
ABAT	4-aminobutyrate aminotransferase
CNR1	cannabinoid receptor 1
KDM3A	lysine demethylase 3A
KDM6B	lysine demethylase 6B
FTO	FTO alpha-ketoglutarate dependent dioxygenase

Target Name	Human Symbol
KDM4C	lysine demethylase 4C
BCL2	BCL2 apoptosis regulator
ADRA1D	adrenoceptor alpha 1D
OPRM1	opioid receptor mu 1
HSD17B1	hydroxysteroid 17-beta dehydrogenase 1
SLC29A1	solute carrier family 29 member 1 (Augustine blood group)
IDO1	indoleamine 2,3-dioxygenase 1
CYP19A1	cytochrome P450 family 19 subfamily A member 1
PTGS1	prostaglandin-endoperoxide synthase 1
RXRB	retinoid X receptor beta
RXRG	retinoid X receptor gamma
RXRA	retinoid X receptor alpha
FOLH1	folate hydrolase 1
HTR3A	5-hydroxytryptamine receptor 3A
ADAMTS5	ADAM metallopeptidase with thrombospondin type 1 motif 5
SNCA	synuclein alpha
MMP9	matrix metallopeptidase 9
XDH	xanthine dehydrogenase
TLR9	toll like receptor 9
MMP2	matrix metallopeptidase 2
DBF4	DBF4 zinc finger
CDC7	cell division cycle 7
CCNE1	cyclin E1
PTK2B	protein tyrosine kinase 2 beta
TYMP	thymidine phosphorylase
ERBB2	erb-b2 receptor tyrosine kinase 2
GRIA1	glutamate ionotropic receptor AMPA type subunit 1
GRIA4	glutamate ionotropic receptor AMPA type subunit 4
GRIA3	glutamate ionotropic receptor AMPA type subunit 3
GRIA2	glutamate ionotropic receptor AMPA type subunit 2
ALPL	alkaline phosphatase, biomineralization associated
KCNMA1	potassium calcium-activated channel subfamily M alpha 1
BRAF	B-Raf proto-oncogene, serine/threonine kinase
CSNK2A1	casein kinase 2 alpha 1
CDK5R1	cyclin dependent kinase 5 regulatory subunit 1
CDK5	cyclin dependent kinase 5
CCNB3	cyclin B3
CDK1	cyclin dependent kinase 1
CCNB1	cyclin B1
CCNB2	cyclin B2
GSK3B	glycogen synthase kinase 3 beta
NISCH	nischarin
CISD1	CDGSH iron sulfur domain 1
CTSL	cathepsin L
TOP1	DNA topoisomerase I
IKBKB	inhibitor of nuclear factor kappa B kinase subunit beta
MIF	macrophage migration inhibitory factor

Target Name	Human Symbol
CHEK1	checkpoint kinase 1
RPS6KA3	ribosomal protein S6 kinase A3
MPI	mannose phosphate isomerase
CXCR2	C-X-C motif chemokine receptor 2
GPR84	G protein-coupled receptor 84
NOS2	nitric oxide synthase 2
UPP1	uridine phosphorylase 1
PLAA	phospholipase A2 activating protein
MB	myoglobin
SLC9A1	solute carrier family 9 member A1
ALOX15	arachidonate 15-lipoxygenase
PIK3CG	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit gamma
MMP1	matrix metallopeptidase 1
ADAM17	ADAM metallopeptidase domain 17
MMP14	matrix metallopeptidase 14
MMP13	matrix metallopeptidase 13
VCAM1	vascular cell adhesion molecule 1
MMP7	matrix metallopeptidase 7
MMP12	matrix metallopeptidase 12
MMP8	matrix metallopeptidase 8
PLEC	plectin
AHCY	adenosylhomocysteinase
CSNK1A1	casein kinase 1 alpha 1
CSNK1D	casein kinase 1 delta
PRKDC	protein kinase, DNA-activated, catalytic subunit
NOX4	NADPH oxidase 4
AKR1B1	aldo-keto reductase family 1 member B
FLT3	fms related receptor tyrosine kinase 3
ADORA1	adenosine A1 receptor
CDK6	cyclin dependent kinase 6
ADORA2A	adenosine A2a receptor
SYK	spleen associated tyrosine kinase
ABCC1	ATP binding cassette subfamily C member 1
CFTR	CF transmembrane conductance regulator
CYP1B1	cytochrome P450 family 1 subfamily B member 1
ABCG2	ATP binding cassette subfamily G member 2 (Junior blood group)
AKR1B10	aldo-keto reductase family 1 member B10
TNKS2	tankyrase 2
TNKS	tankyrase
ALOX5	arachidonate 5-lipoxygenase
ABCB1	ATP binding cassette subfamily B member 1
ALOX12	arachidonate 12-lipoxygenase, 12S type
PTPRS	protein tyrosine phosphatase receptor type S
GLO1	glyoxalase I
APP	amyloid beta precursor protein
CD38	CD38 molecule
ARG1	arginase 1

Target Name	Human Symbol
ESRRA	estrogen related receptor alpha
PFKFB3	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3
AMY1A	amylase alpha 1A
GRK6	G protein-coupled receptor kinase 6
HSD17B2	hydroxysteroid 17-beta dehydrogenase 2
AHR	aryl hydrocarbon receptor
CBR1	carbonyl reductase 1
AR	androgen receptor
TERT	telomerase reverse transcriptase
PIM1	Pim-1 proto-oncogene, serine/threonine kinase
NAE1	NEDD8 activating enzyme E1 subunit 1
TBXAS1	thromboxane A synthase 1
PLK1	polo like kinase 1
BCHE	butyrylcholinesterase
ADORA3	adenosine A3 receptor
HTR2C	5-hydroxytryptamine receptor 2C
DAPK1	death associated protein kinase 1
MPG	N-methylpurine DNA glycosylase
SLC22A12	solute carrier family 22 member 12
ST6GAL1	ST6 beta-galactoside alpha-2,6-sialyltransferase 1
F2	coagulation factor II, thrombin
PLG	plasminogen
AVPR2	arginine vasopressin receptor 2
DRD4	dopamine receptor D4
MPO	myeloperoxidase
PIK3R1	phosphoinositide-3-kinase regulatory subunit 1
PYGL	glycogen phosphorylase L
MMP3	matrix metallopeptidase 3
PKN1	protein kinase N1
CXCR1	C-X-C motif chemokine receptor 1
CAMK2B	calcium/calmodulin dependent protein kinase II beta
AKT1	AKT serine/threonine kinase 1
NEK6	NIMA related kinase 6
SHBG	sex hormone binding globulin
TAS2R31	taste 2 receptor member 31
PLA2G1B	phospholipase A2 group IB
GRM5	glutamate metabotropic receptor 5
CES1	carboxylesterase 1
PPARG	peroxisome proliferator activated receptor gamma
CES2	carboxylesterase 2
SLC5A2	solute carrier family 5 member 2
PLA2G2A	phospholipase A2 group IIA
PLA2G5	phospholipase A2 group V
PLA2G10	phospholipase A2 group X
BACE1	beta-secretase 1
KLK1	kallikrein 1
KLK2	kallikrein related peptidase 2

Target Name	Human Symbol
CHRNA7	cholinergic receptor nicotinic alpha 7 subunit
KIT	KIT proto-oncogene, receptor tyrosine kinase
FGFR1	fibroblast growth factor receptor 1
NQO2	N-ribosyldihydronicotinamide:quinone reductase 2
HSD17B14	hydroxysteroid 17-beta dehydrogenase 14
INSR	insulin receptor
ESRRB	estrogen related receptor beta
EDNRA	endothelin receptor type A
SIRT2	sirtuin 2
IGFBP3	insulin like growth factor binding protein 3
PTGER1	prostaglandin E receptor 1
PTGER4	prostaglandin E receptor 4
PTGER2	prostaglandin E receptor 2
PTGER3	prostaglandin E receptor 3
PIK3CB	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit beta
CYP2C9	cytochrome P450 family 2 subfamily C member 9
CYP3A4	cytochrome P450 family 3 subfamily A member 4
PIK3CA	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha
F3	coagulation factor III, tissue factor
ADCY5	adenylate cyclase 5
PGF	placental growth factor
VEGFA	vascular endothelial growth factor A
YWHAG	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein gamma
VCP	valosin containing protein
WEE1	WEE1 G2 checkpoint kinase
HNF4A	hepatocyte nuclear factor 4 alpha
CDK4	cyclin dependent kinase 4
CDK3	cyclin dependent kinase 3
AURKA	aurora kinase A
NUAK1	NUAK family kinase 1
AKR1C4	aldo-keto reductase family 1 member C4
AKR1A1	aldo-keto reductase family 1 member A1
MAPT	microtubule associated protein tau
KCNH2	potassium voltage-gated channel subfamily H member 2
ST3GAL3	ST3 beta-galactoside alpha-2,3-sialyltransferase 3
FUT4	fucosyltransferase 4
FFAR1	free fatty acid receptor 1
HIF1A	hypoxia inducible factor 1 subunit alpha
CTSB	cathepsin B
ODC1	ornithine decarboxylase 1
ATP4B	ATPase H+/K+ transporting subunit beta
ATP4A	ATPase H+/K+ transporting subunit alpha
PNP	purine nucleoside phosphorylase
SAE1	SUMO1 activating enzyme subunit 1
UBA2	ubiquitin like modifier activating enzyme 2
HSD17B3	hydroxysteroid 17-beta dehydrogenase 3

Target Name	Human Symbol
PGD	phosphogluconate dehydrogenase
CBS	cystathionine beta-synthase
QDPR	quinoid dihydropteridine reductase
ELAVL3	ELAV like RNA binding protein 3
PGAM1	phosphoglycerate mutase 1
MTNR1A	melatonin receptor 1A
MTNR1B	melatonin receptor 1B
MTOR	mechanistic target of rapamycin kinase
CYP24A1	cytochrome P450 family 24 subfamily A member 1
PDK1	pyruvate dehydrogenase kinase 1
SGK1	serum/glucocorticoid regulated kinase 1
DUSP3	dual specificity phosphatase 3
ANPEP	alanyl aminopeptidase, membrane
EPHA2	EPH receptor A2
YES1	YES proto-oncogene 1, Src family tyrosine kinase
BLK	BLK proto-oncogene, Src family tyrosine kinase
CSK	C-terminal Src kinase
EPHB2	EPH receptor B2
BMX	BMX non-receptor tyrosine kinase
LYN	LYN proto-oncogene, Src family tyrosine kinase
EPHA5	EPH receptor A5
EPHA4	EPH receptor A4
TXK	TXK tyrosine kinase
FGR	FGR proto-oncogene, Src family tyrosine kinase
EPHA6	EPH receptor A6
PTK6	protein tyrosine kinase 6
EPHB3	EPH receptor B3
EPHA3	EPH receptor A3
BTK	Bruton tyrosine kinase
TYRO3	TYRO3 protein tyrosine kinase
COQ8B	coenzyme Q8B
EPHA1	EPH receptor A1
CAPN1	calpain 1
GCGR	glucagon receptor
NR3C1	nuclear receptor subfamily 3 group C member 1
ROCK1	Rho associated coiled-coil containing protein kinase 1
PRKACA	protein kinase cAMP-activated catalytic subunit alpha
DNM1	dynamin 1
PDE4A	phosphodiesterase 4A
PDE4D	phosphodiesterase 4D
PDE4C	phosphodiesterase 4C
BMP1	bone morphogenetic protein 1
ABL1	ABL proto-oncogene 1, non-receptor tyrosine kinase
EPHA8	EPH receptor A8
EPHA7	EPH receptor A7
EPHB1	EPH receptor B1
ROCK2	Rho associated coiled-coil containing protein kinase 2

Target Name	Human Symbol
RPS6KB1	ribosomal protein S6 kinase B1
HDAC2	histone deacetylase 2
HDAC8	histone deacetylase 8
THRA	thyroid hormone receptor alpha
THRB	thyroid hormone receptor beta
CLK3	CDC like kinase 3
DYRK2	dual specificity tyrosine phosphorylation regulated kinase 2
MELK	maternal embryonic leucine zipper kinase
NEK1	NIMA related kinase 1
HDAC3	histone deacetylase 3
HDAC6	histone deacetylase 6
HDAC1	histone deacetylase 1
CTSD	cathepsin D
TOP2A	DNA topoisomerase II alpha
MYLK	myosin light chain kinase
APEX1	apurinic/apyrimidinic endodeoxyribonuclease 1
MAP4K4	mitogen-activated protein kinase kinase kinase kinase 4
DNMT1	DNA methyltransferase 1
STAT1	signal transducer and activator of transcription 1
HPGDS	hematopoietic prostaglandin D synthase
GABRA1	gamma-aminobutyric acid type A receptor subunit alpha1
GABRB2	gamma-aminobutyric acid type A receptor subunit beta2
GABRG2	gamma-aminobutyric acid type A receptor subunit gamma2
PRMT1	protein arginine methyltransferase 1
CCNA1	cyclin A1
CCNA2	cyclin A2
MPL	MPL proto-oncogene, thrombopoietin receptor
GUSB	glucuronidase beta
NTRK1	neurotrophic receptor tyrosine kinase 1
OPRK1	opioid receptor kappa 1
ECE1	endothelin converting enzyme 1
HSP90AA1	heat shock protein 90 alpha family class A member 1
GLI2	GLI family zinc finger 2
ALDH2	aldehyde dehydrogenase 2 family member
PRKCZ	protein kinase C zeta
CHEK2	checkpoint kinase 2
GRK2	G protein-coupled receptor kinase 2
SLC5A1	solute carrier family 5 member 1
TRPM8	transient receptor potential cation channel subfamily M member 8
CFD	complement factor D
TNF	tumour necrosis factor
AGTR1	angiotensin II receptor type 1
PNMT	phenylethanolamine N-methyltransferase
TRAP1	TNF receptor associated protein 1
HSP90B1	heat shock protein 90 beta family member 1
HSP90AB1	heat shock protein 90 alpha family class B member 1
AVPR1A	arginine vasopressin receptor 1A

Target Name	Human Symbol
NR1H4	nuclear receptor subfamily 1 group H member 4
ADAM10	ADAM metallopeptidase domain 10
TSPO	translocator protein
PDE7A	phosphodiesterase 7A
MMP16	matrix metallopeptidase 16
AKT2	AKT serine/threonine kinase 2
MAP2K1	mitogen-activated protein kinase kinase 1
MAPK14	mitogen-activated protein kinase 14
HPRT1	hypoxanthine phosphoribosyltransferase 1
MT-ND4	NADH dehydrogenase subunit 4
LTB4R	leukotriene B4 receptor
PLA2G7	phospholipase A2 group VII
GRIN1	glutamate ionotropic receptor NMDA type subunit 1
GRIN2B	glutamate ionotropic receptor NMDA type subunit 2B
GRIN2A	glutamate ionotropic receptor NMDA type subunit 2A
PPARA	peroxisome proliferator activated receptor alpha
MME	membrane metalloendopeptidase
CYP2C19	cytochrome P450 family 2 subfamily C member 19
CELA1	chymotrypsin like elastase 1
SIRT1	sirtuin 1
TARS1	threonyl-tRNA synthetase 1
MTAP	methylthioadenosine phosphorylase

Supplementary Table 3.2. Details of the information of Compound-Target network

Name	BetweennessCentrality	ClosenessCentrality	Degree
Homoeriodictyol	0.115118719	0.409137056	107
Ethyl gallate	0.181664419	0.409969481	106
Apigenin	0.064189127	0.407482305	104
Naringenin	0.107552895	0.407482305	104
Luteolin	0.045552788	0.407482305	104
Kaempferol	0.044978625	0.407482305	104
Quercetin	0.044246566	0.407482305	104
Morin	0.044468377	0.407482305	104
N-trans-feruloyltyramine	0.18709518	0.406659939	104
Isorhamnetin	0.045692973	0.407482305	104
Eriodictyol	0.090700351	0.407482305	103
Isodihydrosyringetin	0.1346498	0.407482305	103
N-trans-caffeoyltyramine	0.207430517	0.405840886	101
Gallic acid	0.186174914	0.405025126	100
Catechin	0.023366618	0.353198948	27
CDK1	0.009686887	0.442371021	18
CA7	0.013714931	0.470245041	14
CA6	0.013714931	0.470245041	14
CA12	0.013714931	0.470245041	14
CA9	0.013714931	0.470245041	14
CA2	0.008678562	0.429179979	13
CA1	0.008678562	0.429179979	13
CA3	0.008678562	0.429179979	13
CA4	0.008678562	0.429179979	13
CA5A	0.012490399	0.464821223	13
ESR2	0.008678562	0.429179979	13
CDK2	0.007866911	0.441401972	13
CYP19A1	0.014774535	0.47579693	13
MMP9	0.014524474	0.466975666	13
MMP2	0.014524474	0.466975666	13
MMP13	0.014524474	0.466975666	13
CA13	0.010675279	0.459521095	12
SRC	0.007276179	0.424657534	12
KDR	0.007276179	0.424657534	12
ACHE	0.013969415	0.463751438	12
HSD17B1	0.008842277	0.435675676	12
MMP12	0.008088617	0.432867884	12
ADORA1	0.011060787	0.432867884	12
CYP1B1	0.009881921	0.412487206	12
IGF1R	0.005383345	0.409137056	11
MET	0.005922948	0.412487206	11
ESR1	0.00732935	0.427359491	11
CDK5R1	0.009379092	0.434735707	11
CDK5	0.009379092	0.434735707	11
AKR1C3	0.004212245	0.393170732	10

Name	BetweennessCentrality	ClosenessCentrality	Degree
TYR	0.009770043	0.420229406	10
MAOA	0.004102502	0.392405063	10
CCNB1	0.004949468	0.411644535	10
SYK	0.004406588	0.392405063	10
ABCC1	0.002422373	0.364046974	10
ABCG2	0.002422373	0.364046974	10
APP	0.002422373	0.364046974	10
CA14	0.004816172	0.390125847	9
ALK	0.004816172	0.390125847	9
EGFR	0.007213005	0.39626352	9
ALOX15	0.005410676	0.40019861	9
NOX4	0.001459876	0.351351351	9
AKR1B1	0.001459876	0.351351351	9
ABCB1	0.001789081	0.355066079	9
ESRRA	0.005264963	0.389371981	9
HSD17B2	0.003033334	0.381267739	9
MMP3	0.003033334	0.381267739	9
PLA2G1B	0.001958493	0.360787825	9
BACE1	0.001958493	0.360787825	9
CA5B	0.008153008	0.426455026	8
TTR	0.002119274	0.354441513	8
AURKB	0.002119274	0.354441513	8
PTK2	0.002119274	0.354441513	8
NEK2	0.002119274	0.354441513	8
AXL	0.002119274	0.354441513	8
PARP1	0.001858199	0.357586513	8
MAOB	0.012326764	0.456398641	8
TOP1	0.003249841	0.393170732	8
ADORA2A	0.002656803	0.370064279	8
ALOX12	0.003296817	0.367365542	8
TERT	0.00153762	0.354441513	8
AKT1	9.29E-04	0.34123624	8
INSR	0.005767894	0.408308004	8
GPR35	9.95E-04	0.32951758	7
BCL2	0.008698275	0.432867884	7
XDH	7.99E-04	0.330598852	7
CSNK2A1	7.99E-04	0.330598852	7
CCNB3	7.99E-04	0.330598852	7
CCNB2	7.99E-04	0.330598852	7
GSK3B	7.99E-04	0.330598852	7
TNKS2	0.002314882	0.340084388	7
ALOX5	0.00128048	0.333885667	7
ADORA3	0.006643958	0.420229406	7
F2	4.03E-04	0.326315789	7
PYGL	7.55E-04	0.331687243	7
FUT7	0.003453343	0.379830349	6
SERPINE1	0.003487746	0.387872955	6

Name	BetweennessCentrality	ClosenessCentrality	Degree
AKR1C2	7.60E-04	0.325788197	6
AKR1C1	7.60E-04	0.325788197	6
POLB	0.003487746	0.387872955	6
DYRK1A	0.003487746	0.387872955	6
KDM4E	8.37E-04	0.328979592	6
CCNE1	0.001979375	0.372114497	6
FLT3	2.71E-05	0.29918337	6
CDK6	2.71E-05	0.29918337	6
AKR1B10	2.71E-05	0.29918337	6
TNKS	0.001939247	0.339511373	6
PTPRS	2.71E-05	0.29918337	6
GLO1	2.71E-05	0.29918337	6
CD38	2.71E-05	0.29918337	6
ARG1	2.71E-05	0.29918337	6
AHR	2.71E-05	0.29918337	6
PIM1	2.71E-05	0.29918337	6
PLK1	2.71E-05	0.29918337	6
DAPK1	2.71E-05	0.29918337	6
AVPR2	2.71E-05	0.29918337	6
DRD4	2.71E-05	0.29918337	6
MPO	2.71E-05	0.29918337	6
PIK3R1	2.71E-05	0.29918337	6
PKN1	2.71E-05	0.29918337	6
CXCR1	2.71E-05	0.29918337	6
CAMK2B	2.71E-05	0.29918337	6
NEK6	2.71E-05	0.29918337	6
TAS2R31	0.00502513	0.367365542	6
BCL2L1	0.004958267	0.401794616	5
CLK1	0.004958267	0.401794616	5
DYRK1B	0.004958267	0.401794616	5
PTGS2	0.00205564	0.361434978	5
PTGS1	0.001875729	0.360787825	5
RXRA	0.001875729	0.360787825	5
CHEK1	0.004460082	0.399405352	5
MMP7	0.005458386	0.405840886	5
MMP8	0.005458386	0.405840886	5
CBR1	8.19E-04	0.354441513	5
MPG	2.07E-05	0.298739807	5
SLC22A12	2.07E-05	0.298739807	5
PLG	5.10E-04	0.339511373	5
KLK2	7.62E-04	0.327908869	5
NUAK1	8.60E-06	0.295671313	5
AKR1C4	8.60E-06	0.295671313	5
AKR1A1	8.60E-06	0.295671313	5
MAPT	5.72E-04	0.336675021	5
SQLE	0.002039062	0.367365542	4
LCK	0.00104956	0.357586513	4

Name	BetweennessCentrality	ClosenessCentrality	Degree
CNR2	0.004462202	0.376283847	4
ERN1	0.003432674	0.379830349	4
TYMS	0.004462202	0.376283847	4
GRM2	0.00133406	0.350130321	4
PIK3CG	3.56E-04	0.321628093	4
BCHE	9.67E-04	0.355692851	4
SHBG	3.73E-04	0.325262308	4
GRM5	3.73E-04	0.325262308	4
CES1	3.73E-04	0.325262308	4
PPARG	3.73E-04	0.325262308	4
CES2	3.73E-04	0.325262308	4
SLC5A2	3.73E-04	0.325262308	4
PLA2G5	3.73E-04	0.325262308	4
PLA2G10	3.73E-04	0.325262308	4
KLK1	3.73E-04	0.325262308	4
FGFR1	3.73E-04	0.325262308	4
EDNRA	3.73E-04	0.325262308	4
ST3GAL3	4.60E-04	0.319081552	4
FUT4	4.60E-04	0.319081552	4
MCL1	0.001946629	0.362084456	3
DRD2	0.00267688	0.353198948	3
STS	8.44E-04	0.337803856	3
SNCA	5.47E-04	0.333333333	3
TYMP	0.001695667	0.361434978	3
ALPL	7.67E-04	0.327908869	3
CTSL	0.001185292	0.352580927	3
MMP1	0.001800302	0.348917749	3
ADAM17	0.001800302	0.348917749	3
MMP14	0.001646822	0.360787825	3
CFTR	4.71E-06	0.294806145	3
PFKFB3	4.71E-06	0.294806145	3
AMY1A	4.71E-06	0.294806145	3
GRK6	4.71E-06	0.294806145	3
AR	8.29E-04	0.334995844	3
PLA2G2A	1.90E-04	0.31657502	3
CHRNA7	1.39E-04	0.31410756	3
KIT	1.90E-04	0.31657502	3
NQO2	0.001159606	0.338940286	3
HSD17B14	0.001159606	0.338940286	3
ESRRB	0.001264615	0.346517627	3
PGF	3.05E-04	0.312645462	3
VEGFA	3.05E-04	0.312645462	3
YWHAG	2.12E-04	0.304610733	3
VCP	0.001264615	0.346517627	3
WEE1	0.001264615	0.346517627	3
AURKA	1.39E-04	0.31410756	3
FFAR1	1.72E-04	0.31657502	3

Name	BetweennessCentrality	ClosenessCentrality	Degree
ODC1	1.72E-04	0.31657502	3
PGD	2.77E-04	0.308339709	3
DUSP3	0.001448465	0.351965066	3
DNM1	0.001448465	0.351965066	3
BMP1	0.002132355	0.356953056	3
RPS6KB1	0.001448465	0.351965066	3
TOP2A	7.90E-07	0.291817524	3
MYLK	7.90E-07	0.291817524	3
APEX1	7.90E-07	0.291817524	3
LDHA	3.53E-04	0.315089914	2
LDHB	3.53E-04	0.315089914	2
COMT	3.53E-04	0.315089914	2
TPMT	3.53E-04	0.315089914	2
POLA1	3.53E-04	0.315089914	2
TUBB1	3.53E-04	0.315089914	2
ELANE	0.001430162	0.32951758	2
DRD3	8.79E-04	0.326315789	2
PTPN1	8.79E-04	0.326315789	2
CNR1	3.53E-04	0.315089914	2
SLC29A1	3.53E-04	0.315089914	2
HTR3A	8.79E-04	0.326315789	2
TLR9	8.68E-04	0.327376117	2
CDC7	3.06E-04	0.323175621	2
BRAF	5.64E-04	0.323694779	2
MIF	5.64E-04	0.323694779	2
RPS6KA3	8.68E-04	0.327376117	2
UPP1	8.68E-04	0.327376117	2
SIRT2	2.76E-05	0.301421092	2
IGFBP3	2.76E-05	0.301421092	2
PTGER4	2.76E-05	0.301421092	2
PTGER2	2.76E-05	0.301421092	2
PTGER3	2.76E-05	0.301421092	2
PIK3CB	1.03E-04	0.310238645	2
CDK4	3.47E-04	0.322142286	2
CDK3	3.47E-04	0.322142286	2
KCNH2	5.94E-05	0.305071915	2
HIF1A	5.94E-05	0.305071915	2
CTSB	5.94E-05	0.305071915	2
PNP	4.04E-05	0.304610733	2
SAE1	4.04E-05	0.304610733	2
UBA2	4.04E-05	0.304610733	2
MTNR1A	3.68E-04	0.31410756	2
MTOR	3.68E-04	0.31410756	2
PDK1	3.68E-04	0.31410756	2
SGK1	6.64E-04	0.328443358	2
ANPEP	3.68E-04	0.31410756	2
EPHA2	3.68E-04	0.31410756	2

Name	BetweennessCentrality	ClosenessCentrality	Degree
EPHB2	3.68E-04	0.31410756	2
EPHA5	3.68E-04	0.31410756	2
EPHA4	3.68E-04	0.31410756	2
EPHB3	3.68E-04	0.31410756	2
PDE4D	3.68E-04	0.31410756	2
PDE4C	3.68E-04	0.31410756	2
ABL1	3.68E-04	0.31410756	2
EPHA8	3.68E-04	0.31410756	2
EPHA7	3.68E-04	0.31410756	2
ROCK2	3.68E-04	0.31410756	2
THRA	3.68E-04	0.31410756	2
THRB	3.68E-04	0.31410756	2
HDAC1	3.68E-04	0.31410756	2
DNMT1	7.24E-05	0.305534496	2
STAT1	7.24E-05	0.305534496	2
HPGDS	4.17E-04	0.324738114	2
GABRA1	7.24E-05	0.305534496	2
GABRB2	7.24E-05	0.305534496	2
GABRG2	7.24E-05	0.305534496	2
ECE1	7.24E-05	0.305534496	2
ADAM10	6.48E-04	0.327376117	2
ALB	0	0.288475304	1
DAO	0	0.288475304	1
SRD5A2	0	0.288475304	1
FYN	0	0.288475304	1
MVD	0	0.288475304	1
FASN	0	0.288475304	1
ADRA2A	0	0.288475304	1
ADRA2C	0	0.288475304	1
ADRA2B	0	0.288475304	1
ADRB2	0	0.288475304	1
ADRB1	0	0.288475304	1
ADRA1A	0	0.288475304	1
RARG	0	0.288475304	1
RARB	0	0.288475304	1
CRHR1	0	0.288475304	1
PLA2G4B	0	0.288475304	1
KDM4A	0	0.288475304	1
DPP4	0	0.288475304	1
ALDH5A1	0	0.288475304	1
ABAT	0	0.288475304	1
KDM3A	0	0.288475304	1
KDM6B	0	0.288475304	1
FTO	0	0.288475304	1
KDM4C	0	0.288475304	1
ADRA1D	0	0.288475304	1
OPRM1	0	0.288475304	1

Name	BetweennessCentrality	ClosenessCentrality	Degree
IDO1	0	0.288475304	1
RXRB	0	0.288475304	1
RXRG	0	0.288475304	1
FOLH1	0	0.288475304	1
ADAMTS5	0	0.290974729	1
DBF4	0	0.290974729	1
PTK2B	0	0.290974729	1
ERBB2	0	0.290974729	1
GRIA1	0	0.290974729	1
GRIA4	0	0.290974729	1
GRIA3	0	0.290974729	1
GRIA2	0	0.290974729	1
KCNMA1	0	0.290974729	1
NISCH	0	0.290974729	1
CISD1	0	0.290974729	1
IKBKB	0	0.290974729	1
MPI	0	0.290974729	1
CXCR2	0	0.290974729	1
GPR84	0	0.290974729	1
NOS2	0	0.290974729	1
PLAA	0	0.290974729	1
MB	0	0.290974729	1
SLC9A1	0	0.290974729	1
VCAM1	0	0.290974729	1
PLEC	0	0.290974729	1
AHCY	0	0.290974729	1
CSNK1A1	0	0.290974729	1
CSNK1D	0	0.290974729	1
PRKDC	0	0.290974729	1
NAE1	0	0.289719626	1
TBXAS1	0	0.289719626	1
HTR2C	0	0.289719626	1
ST6GAL1	0	0.289719626	1
PTGER1	0	0.289719626	1
CYP2C9	0	0.289719626	1
CYP3A4	0	0.289719626	1
PIK3CA	0	0.289719626	1
F3	0	0.289719626	1
ADCY5	0	0.289719626	1
HNF4A	0	0.289719626	1
ATP4B	0	0.289719626	1
ATP4A	0	0.289719626	1
HSD17B3	0	0.289719626	1
CBS	0	0.26117952	1
QDPR	0	0.26117952	1
ELAVL3	0	0.26117952	1
PGAM1	0	0.26117952	1

Name	BetweennessCentrality	ClosenessCentrality	Degree
MTNR1B	0	0.288888889	1
CYP24A1	0	0.288888889	1
YES1	0	0.288888889	1
BLK	0	0.288888889	1
CSK	0	0.288888889	1
BMX	0	0.288888889	1
LYN	0	0.288888889	1
TXK	0	0.288888889	1
FGR	0	0.288888889	1
EPHA6	0	0.288888889	1
PTK6	0	0.288888889	1
EPHA3	0	0.288888889	1
BTK	0	0.288888889	1
TYRO3	0	0.288888889	1
COQ8B	0	0.288888889	1
EPHA1	0	0.288888889	1
CAPN1	0	0.288888889	1
GCGR	0	0.288888889	1
NR3C1	0	0.288888889	1
ROCK1	0	0.288888889	1
PRKACA	0	0.288888889	1
PDE4A	0	0.288888889	1
EPHB1	0	0.288888889	1
HDAC2	0	0.288888889	1
HDAC8	0	0.288888889	1
CLK3	0	0.288888889	1
DYRK2	0	0.288888889	1
MELK	0	0.288888889	1
NEK1	0	0.288888889	1
HDAC3	0	0.288888889	1
HDAC6	0	0.288888889	1
CTSD	0	0.288888889	1
MAP4K4	0	0.290555155	1
PRMT1	0	0.290555155	1
CCNA1	0	0.290555155	1
CCNA2	0	0.290555155	1
MPL	0	0.290555155	1
GUSB	0	0.290555155	1
NTRK1	0	0.290555155	1
OPRK1	0	0.290555155	1
HSP90AA1	0	0.289303661	1
GLI2	0	0.289303661	1
ALDH2	0	0.289303661	1
PRKCZ	0	0.289303661	1
CHEK2	0	0.289303661	1
GRK2	0	0.289303661	1
SLC5A1	0	0.289303661	1

Name	BetweennessCentrality	ClosenessCentrality	Degree
TRPM8	0	0.289303661	1
CFD	0	0.289303661	1
TNF	0	0.289303661	1
AGTR1	0	0.289303661	1
PNMT	0	0.289303661	1
TRAP1	0	0.289303661	1
HSP90B1	0	0.289303661	1
HSP90AB1	0	0.289303661	1
AVPR1A	0	0.289303661	1
NR1H4	0	0.289303661	1
TSPO	0	0.289303661	1
PDE7A	0	0.289303661	1
MMP16	0	0.289303661	1
AKT2	0	0.289303661	1
MAP2K1	0	0.289303661	1
MAPK14	0	0.289719626	1
HPRT1	0	0.289719626	1
MT-ND4	0	0.289719626	1
LTB4R	0	0.289719626	1
PLA2G7	0	0.289719626	1
GRIN1	0	0.289719626	1
GRIN2B	0	0.289719626	1
GRIN2A	0	0.289719626	1
PPARA	0	0.289719626	1
MME	0	0.289719626	1
CYP2C19	0	0.289719626	1
CELA1	0	0.289719626	1
SIRT1	0	0.289719626	1
TARS	0	0.289719626	1
MTAP	0	0.289719626	1

Supplementary Table 3.3. Details of the information of Target-Pathway network

Name	BetweennessCentrality	ClosenessCentrality	Degree
Pathways in cancer	0.543088	0.496598639	68
Neuroactive ligand-receptor interaction	0.408604	0.384210526	51
cAMP signalling pathway	0.253227	0.403314917	34
Prostate cancer	0.04807	0.32735426	24
Endocrine resistance	0.034182	0.325892857	23
Progesterone-mediated oocyte maturation	0.108048	0.325892857	23
HIF-1 signalling pathway	0.070353	0.324444444	22
EGFR tyrosine kinase inhibitor resistance	0.043115	0.321585903	20
Central carbon metabolism in cancer	0.041998	0.320175439	19
Nitrogen metabolism	1	1	12
PIK3CB	0.032721	0.42074928	8
AKT2	0.032721	0.42074928	8
AKT1	0.032721	0.42074928	8
MAP2K1	0.032721	0.42074928	8
PIK3CA	0.032721	0.42074928	8
PIK3R1	0.032721	0.42074928	8
IGF1R	0.010537	0.373401535	6
EGFR	0.007358	0.362282878	6
ERBB2	0.007358	0.362282878	6
BRAF	0.020596	0.4	6
MTOR	0.007358	0.362282878	6
BCL2	0.005343	0.358722359	5
PRKACA	0.012885	0.387267905	4
ADCY5	0.012885	0.387267905	4
RPS6KB1	0.003945	0.353510896	4
GSK3B	0.001969	0.343529412	3
HSP90AB1	0.004022	0.351807229	3
EDNRA	0.040155	0.435820896	3
HSP90AA1	0.004022	0.351807229	3
MMP9	0.001371	0.341920375	3
MET	0.002463	0.346793349	3
CAMK2B	0.008907	0.379220779	3
PTGER2	0.040155	0.435820896	3
PTGER3	0.040155	0.435820896	3
HIF1A	0.00314	0.346793349	3
VEGFA	0.003075	0.350119904	3
CDK2	0.004022	0.351807229	3
FGFR1	0.002178	0.345153664	3
FLT3	0.001201	0.34032634	2
IKBKB	7.26E-04	0.337182448	2
MMP2	5.03E-04	0.337182448	2
F2	0.030823	0.418338109	2

Name	BetweennessCentrality	ClosenessCentrality	Degree
CCNA2	0.002805	0.346793349	2
CCNA1	0.002805	0.346793349	2
AR	7.26E-04	0.337182448	2
CCNE1	7.26E-04	0.337182448	2
KIT	0.001201	0.34032634	2
AGTR1	0.030823	0.418338109	2
PTGER4	0.030823	0.418338109	2
ROCK1	0.003837	0.365914787	2
ROCK2	0.003837	0.365914787	2
PTGER1	0.030823	0.418338109	2
HSP90B1	7.26E-04	0.337182448	2
NTRK1	0.001201	0.34032634	2
NOS2	0.001796	0.343529412	2
ESR1	5.03E-04	0.337182448	2
ESR2	5.03E-04	0.337182448	2
PTK2	5.03E-04	0.337182448	2
CDK4	5.03E-04	0.337182448	2
BCL2L1	0.001056	0.3387471	2
ADORA1	0.005496	0.335632184	2
ADORA2A	0.005496	0.335632184	2
GRIA1	0.005496	0.335632184	2
GRIA2	0.005496	0.335632184	2
ADRB1	0.005496	0.335632184	2
ADRB2	0.005496	0.335632184	2
CRHR1	0.005496	0.335632184	2
GRIN2A	0.005496	0.335632184	2
DRD2	0.005496	0.335632184	2
GRIA3	0.005496	0.335632184	2
GRIA4	0.005496	0.335632184	2
GRIN2B	0.005496	0.335632184	2
GRIN1	0.005496	0.335632184	2
SRC	1.12E-04	0.252158895	2
MAPK14	3.65E-04	0.257495591	2
LDHB	1.43E-04	0.251290878	2
LDHA	1.43E-04	0.251290878	2
PDK1	1.43E-04	0.251290878	2
ALK	0	0.332574032	1
GLI2	0	0.332574032	1
PIM1	0	0.332574032	1
DAPK1	0	0.332574032	1
MMP1	0	0.332574032	1
PGF	0	0.332574032	1
RARB	0	0.332574032	1
PPARG	0	0.332574032	1
HDAC2	0	0.332574032	1
HDAC1	0	0.332574032	1
PTGS2	0	0.332574032	1

Name	BetweennessCentrality	ClosenessCentrality	Degree
RXRB	0	0.332574032	1
RXRA	0	0.332574032	1
TERT	0	0.332574032	1
ABL1	0	0.332574032	1
RXRG	0	0.332574032	1
STAT1	0	0.332574032	1
CDK6	0	0.332574032	1
GABRB2	0	0.278095238	1
THRB	0	0.278095238	1
THRA	0	0.278095238	1
HTR2C	0	0.278095238	1
ADRA1D	0	0.278095238	1
NR3C1	0	0.278095238	1
ADRA1A	0	0.278095238	1
LTB4R	0	0.278095238	1
GRM2	0	0.278095238	1
GRM5	0	0.278095238	1
ADORA3	0	0.278095238	1
TSPO	0	0.278095238	1
GPR35	0	0.278095238	1
AVPR2	0	0.278095238	1
AVPR1A	0	0.278095238	1
OPRM1	0	0.278095238	1
ADRA2C	0	0.278095238	1
ADRA2B	0	0.278095238	1
GABRG2	0	0.278095238	1
ADRA2A	0	0.278095238	1
MTNR1A	0	0.278095238	1
MTNR1B	0	0.278095238	1
CHRNA7	0	0.278095238	1
PLG	0	0.278095238	1
CNR2	0	0.278095238	1
CNR1	0	0.278095238	1
DRD3	0	0.278095238	1
DRD4	0	0.278095238	1
GABRA1	0	0.278095238	1
GCGR	0	0.278095238	1
OPRK1	0	0.278095238	1
CA12	0	0.52173913	1
CA1	0	0.52173913	1
CA5B	0	0.52173913	1
CA3	0	0.52173913	1
CA2	0	0.52173913	1
CA5A	0	0.52173913	1
CA4	0	0.52173913	1
CA7	0	0.52173913	1
CA6	0	0.52173913	1

Name	BetweennessCentrality	ClosenessCentrality	Degree
CA9	0	0.52173913	1
CA14	0	0.52173913	1
CA13	0	0.52173913	1
SRD5A2	0	0.247038917	1
MMP3	0	0.247038917	1
SLC9A1	0	0.287968442	1
PDE4A	0	0.287968442	1
PDE4D	0	0.287968442	1
PDE4C	0	0.287968442	1
PPARA	0	0.287968442	1
CFTR	0	0.287968442	1
PLK1	0	0.246205734	1
AURKA	0	0.246205734	1
CCNB3	0	0.246205734	1
CCNB2	0	0.246205734	1
RPS6KA3	0	0.246205734	1
CCNB1	0	0.246205734	1
CDK1	0	0.246205734	1
PGAM1	0	0.242928453	1
AXL	0	0.243739566	1
KDR	0	0.243739566	1
PFKFB3	0	0.245378151	1
INSR	0	0.245378151	1
SERPINE1	0	0.245378151	1

B.2 Supplementary Tables for chapter 4

Supplementary Table 4.1. DEGs in MDA-MB-468 cells upon treatment with LSW

Gene name	log2FoldChange	*P*-value	P_{adj}	change
CYP1A1	6.699422657	0	0	up
KRT34	6.24663334	9.35E-06	4.73E-05	up
ACOXL	5.571067951	5.23E-11	4.84E-10	up
INHBA	4.712051977	1.06E-72	8.82E-71	up
SIRPB1	4.349979499	2.38E-05	0.000113	up
PRSS3	4.336691576	0.000715	0.002629	up
IL1RL1	4.304478971	3.48E-05	0.00016	up
IL6	4.219275172	0.000166	0.000685	up
UGT1A7	4.089285071	0.000344	0.001342	up
AL596223.1	4.063720583	0.000214	0.000867	up
HKDC1	4.031815239	1.27E-12	1.35E-11	up
AOX1	4.023467802	0.000292	0.001151	up
RORA	4.012730865	0.000575	0.002149	up
GDF15	3.914069277	4.97E-16	6.66E-15	up
NMRAL2P	3.776172949	8.71E-29	2.20E-27	up
LIF	3.700985545	5.14E-183	1.69E-180	up
FOSL1	3.647284682	3.35E-187	1.16E-184	up
AL022724.3	3.631166885	0.00072	0.002647	up
KRT6A	3.607801156	1.29E-74	1.10E-72	up
SYT6	3.606998982	2.17E-06	1.20E-05	up
LINC01564	3.596913144	3.61E-18	5.61E-17	up
AC004264.1	3.589832134	3.39E-05	0.000157	up
ROS1	3.457404611	5.95E-08	3.99E-07	up
PTPRH	3.428046667	1.56E-19	2.63E-18	up
AC089983.1	3.361746226	7.78E-07	4.55E-06	up
MAP1B	3.34488567	7.11E-65	5.02E-63	up
AC009093.4	3.309972742	3.45E-35	1.09E-33	up
MRPS24	3.238015904	5.23E-22	9.89E-21	up
AL033397.1	3.23548182	3.91E-09	3.00E-08	up
MYH15	3.090684547	8.65E-05	0.000374	up
GPAT3	3.086611982	1.66E-45	7.40E-44	up
TIMP4	3.070679538	1.69E-06	9.46E-06	up
ARHGAP9	3.060502711	0.002772	0.009001	up
SPOCD1	3.026852229	0.002852	0.009234	up
CLMP	2.982638881	3.45E-40	1.29E-38	up
AC009229.1	2.977539748	4.31E-10	3.64E-09	up
AC019117.2	2.957869191	9.34E-10	7.63E-09	up
AL121956.1	2.942593645	0.000604	0.002248	up
PDE2A	2.917087217	0.002967	0.009567	up
SP8	2.896716951	8.28E-23	1.62E-21	up
KIF1A	2.863640507	0.000258	0.001029	up
NECAB2	2.824982876	0.000437	0.001676	up

Gene name	log2FoldChange	*P*-value	P_{adj}	change
SERPINE1	2.794005022	2.37E-10	2.05E-09	up
TMEM200B	2.773609043	1.65E-05	7.99E-05	up
RAB39B	2.773354446	2.32E-06	1.27E-05	up
LINC00707	2.756452586	0.003804	0.012003	up
ALOXE3	2.746016458	3.10E-05	0.000144	up
LUCAT1	2.744798541	4.27E-59	2.67E-57	up
RPS6KA2-IT1	2.708993894	6.73E-05	0.000298	up
MYO7B	2.707851327	2.69E-05	0.000127	up
TUBB2B	2.706553135	6.91E-39	2.46E-37	up
GPR3	2.699133526	2.02E-06	1.12E-05	up
NT5DC4	2.689011623	0.003864	0.012179	up
CCN3	2.659685129	2.58E-08	1.81E-07	up
KRT16P2	2.657284272	0.000706	0.002598	up
SOCS2	2.648246603	3.30E-10	2.82E-09	up
AC009229.2	2.636071324	5.37E-07	3.21E-06	up
KLHDC7B	2.610244959	1.14E-41	4.50E-40	up
AC034213.1	2.604691366	0.002964	0.009559	up
PPFIA4	2.600673766	3.06E-56	1.80E-54	up
GDA	2.596708106	3.15E-56	1.85E-54	up
UGT1A1	2.585967957	2.78E-18	4.34E-17	up
AREG	2.579961668	2.08E-16	2.87E-15	up
AL109615.2	2.540623222	5.73E-10	4.79E-09	up
HTR1D	2.539560776	4.26E-24	9.03E-23	up
HES7	2.501474708	4.34E-18	6.70E-17	up
PCOLCE2	2.463045259	1.44E-05	7.08E-05	up
HTR7	2.460733384	5.18E-12	5.21E-11	up
GLP2R	2.452816406	0.00068	0.00251	up
AC011455.6	2.449059867	2.72E-05	0.000128	up
FLG	2.434456853	2.00E-57	1.20E-55	up
MYOSLID	2.430657344	0.002694	0.008766	up
PLAT	2.429988016	0.000764	0.002798	up
TEX37	2.382066828	9.85E-07	5.68E-06	up
TMEM156	2.359050502	4.40E-169	1.17E-166	up
GCNT3	2.348403989	1.99E-07	1.25E-06	up
DLL4	2.345920367	0.001598	0.005455	up
AL359834.1	2.339476807	0.00383	0.012079	up
INHBA-AS1	2.33415031	0.006964	0.020643	up
IRS2	2.312583852	1.24E-154	2.98E-152	up
AC080128.3	2.30561723	0.007326	0.021578	up
RMDN2-AS1	2.284760776	0.006688	0.019906	up
XDH	2.28387324	1.14E-16	1.61E-15	up
SFTA1P	2.280623848	1.52E-55	8.64E-54	up
IGFBP1	2.258994485	0.000428	0.001645	up
FGL1	2.25179982	2.82E-10	2.43E-09	up
MARCO	2.232551504	8.54E-08	5.62E-07	up
SLC35G2	2.221450708	1.11E-07	7.20E-07	up

Gene name	log2FoldChange	*P*-value	P_{adj}	change
LAMB3	2.207177745	2.36E-226	1.31E-223	up
IL7R	2.195636739	8.69E-35	2.73E-33	up
CYP1B1	2.190554812	0	0	up
UGT1A6	2.187132102	5.95E-41	2.28E-39	up
OLFML3	2.16446615	8.99E-260	6.71E-257	up
CA9	2.15568013	1.64E-88	1.74E-86	up
SLC16A1-AS1	2.149887059	6.53E-31	1.79E-29	up
AC041005.1	2.139122331	0.002612	0.008531	up
ANKRD29	2.124671302	3.33E-06	1.79E-05	up
TGFBI	2.114098687	6.87E-122	1.20E-119	up
FYN	2.111138364	1.13E-40	4.33E-39	up
AL136169.1	2.108840867	0.000469	0.001787	up
NAT8L	2.093046121	5.32E-46	2.40E-44	up
AKAP12	2.076283266	1.27E-05	6.28E-05	up
IL1R2	2.075752202	2.24E-16	3.09E-15	up
TIPARP	2.074235527	5.69E-159	1.42E-156	up
EPGN	2.070871998	6.57E-05	0.000291	up
ECM2	2.051581298	1.92E-06	1.07E-05	up
GJB2	2.05016666	2.57E-32	7.58E-31	up
CLCF1	2.044977809	7.62E-28	1.86E-26	up
BX640514.2	2.042509449	2.24E-10	1.94E-09	up
ADGRF4	2.037977737	9.38E-33	2.79E-31	up
FOXI3	2.010379996	0.001018	0.003623	up
ANKRD33B	1.993481946	4.01E-53	2.18E-51	up
KCNE5	1.992961587	0.005306	0.016132	up
IL6R	1.982937707	1.95E-05	9.38E-05	up
PNMA2	1.982268658	2.08E-17	3.08E-16	up
GAL	1.980363353	2.98E-17	4.35E-16	up
STC1	1.976475285	0	0	up
ARTN	1.974065104	1.91E-288	1.77E-285	up
LAMC3	1.962012312	4.35E-08	2.97E-07	up
HIF1A-AS3	1.952005173	1.11E-18	1.77E-17	up
CYP27B1	1.94514256	3.60E-05	0.000166	up
LINP1	1.94143729	9.31E-05	0.000401	up
TNFRSF11B	1.923889167	2.69E-09	2.10E-08	up
LURAP1L-AS1	1.920990851	0.001478	0.005081	up
FYB1	1.92025941	0.007203	0.021276	up
AP003733.3	1.91681813	3.28E-06	1.77E-05	up
F3	1.914020522	2.01E-259	1.44E-256	up
NCR3LG1	1.910397026	1.32E-08	9.57E-08	up
PFKFB4	1.901277395	1.88E-115	3.08E-113	up
SLC34A3	1.900970532	0.010196	0.028853	up
AC005336.1	1.88256598	2.35E-09	1.85E-08	up
AC133644.3	1.870352419	0.000732	0.002689	up
UPP1	1.868235753	0	0	up
KLF4	1.865454731	5.25E-46	2.38E-44	up

Gene name	log2FoldChange	*P*-value	P_{adj}	change
RIMKLA	1.862682677	4.59E-16	6.17E-15	up
INA	1.858940964	8.67E-57	5.18E-55	up
NRAD1	1.842177477	5.27E-05	0.000237	up
ASB2	1.840690704	0.004504	0.013956	up
CR2	1.838511952	0.006766	0.020113	up
RAB3B	1.833068339	2.00E-10	1.75E-09	up
PANX2	1.817891241	6.75E-28	1.66E-26	up
ANGPT4	1.799082219	0.002384	0.00785	up
FHL1	1.793673436	9.89E-22	1.84E-20	up
PYGL	1.784774016	7.06E-79	6.49E-77	up
NEXN	1.764794853	5.70E-17	8.22E-16	up
LINC01895	1.76424036	4.56E-144	9.94E-142	up
AL133517.1	1.757840505	0.006939	0.020582	up
SUGT1P4-STRA6LP	1.754362856	6.98E-05	0.000308	up
NDRG1	1.75189545	0	0	up
MAFF	1.751501769	2.51E-91	2.80E-89	up
BNIP3	1.747380018	1.32E-242	8.23E-240	up
INHBE	1.742856062	2.18E-05	0.000104	up
CYP1B1-AS1	1.74219382	1.46E-27	3.52E-26	up
NIBAN1	1.713832109	4.39E-167	1.15E-164	up
PADI1	1.706295581	1.38E-11	1.34E-10	up
FOSB	1.703165514	1.60E-16	2.24E-15	up
SSTR1	1.70163873	2.72E-05	0.000128	up
LINC02615	1.692983115	9.98E-15	1.23E-13	up
CXCL8	1.688548443	2.28E-16	3.14E-15	up
LIPG	1.687465486	3.55E-198	1.38E-195	up
NPBWR1	1.684012282	1.82E-07	1.15E-06	up
PDE4B	1.678417905	2.95E-40	1.10E-38	up
TSC22D1	1.677264992	1.48E-280	1.25E-277	up
HBEGF	1.661628634	2.67E-19	4.44E-18	up
SERPINE2	1.659966515	1.70E-08	1.22E-07	up
SLC7A11	1.645919738	6.47E-106	9.10E-104	up
FSCN1	1.645053025	2.33E-77	2.10E-75	up
TNFRSF21	1.640114045	1.80E-203	7.95E-201	up
SPRY4-AS1	1.638414523	5.95E-05	0.000265	up
GNG4	1.625454434	0.001772	0.005998	up
EGR3	1.617081985	0.000291	0.001149	up
LAMC2	1.616114434	8.01E-93	8.98E-91	up
STC2	1.615932277	0	0	up
KRT6B	1.612274697	5.87E-111	9.04E-109	up
SLC2A6	1.611120995	9.65E-05	0.000414	up
SLCO4A1	1.609733293	4.05E-26	9.22E-25	up
TMEM45B	1.609181783	7.36E-20	1.26E-18	up
SNHG12	1.603278259	4.67E-56	2.72E-54	up
MAGEC1	1.596614541	0.000276	0.001094	up
SPATS2L	1.58549359	1.07E-17	1.60E-16	up

Gene name	log2FoldChange	*P*-value	P_{adj}	change
SRPX2	1.584994764	7.97E-09	5.93E-08	up
S100P	1.578419294	6.32E-93	7.13E-91	up
LINC02065	1.578034782	3.37E-05	0.000156	up
ITGA2	1.576450919	1.56E-56	9.24E-55	up
UPK1A	1.565593844	0.00319	0.010212	up
MMP24	1.559995393	9.29E-06	4.70E-05	up
VIM	1.559090939	1.95E-40	7.34E-39	up
AC245100.3	1.557471453	1.23E-05	6.09E-05	up
ITGA5	1.556308494	2.26E-28	5.64E-27	up
BMP6	1.548555473	5.56E-06	2.91E-05	up
SH3RF3-AS1	1.548018228	1.30E-06	7.40E-06	up
AC027373.1	1.546874096	2.51E-05	0.000119	up
GATA6-AS1	1.546745216	0.001562	0.005344	up
MVD	1.545934844	0	0	up
INPP4B	1.543929749	5.83E-35	1.84E-33	up
AL136981.3	1.54348364	0.003462	0.01102	up
PDK1	1.542105367	4.25E-119	7.23E-117	up
LOXL2	1.539663944	3.22E-278	2.60E-275	up
HOXA9	1.536702132	0.008768	0.025248	up
SIRPB2	1.534498125	0.00648	0.019364	up
SLC7A5P1	1.533292923	2.01E-07	1.26E-06	up
ELOVL6	1.530348145	6.97E-56	4.00E-54	up
KCNQ5	1.52822879	3.54E-08	2.45E-07	up
ENO2	1.526911823	1.39E-68	1.05E-66	up
SLAMF7	1.525764454	0.004626	0.014294	up
AC108925.1	1.522509449	1.71E-19	2.88E-18	up
CYP4F11	1.517573721	7.91E-40	2.91E-38	up
ETV5	1.516685659	1.70E-86	1.72E-84	up
TRABD2B	1.502772289	4.31E-05	0.000197	up
BATF3	1.496102778	5.36E-09	4.05E-08	up
OTUB2	1.494504243	3.14E-07	1.93E-06	up
HDAC9	1.488024077	2.47E-07	1.54E-06	up
PLK3	1.48706729	4.04E-43	1.68E-41	up
DUSP6	1.471871592	5.34E-19	8.73E-18	up
GPRC5A	1.470575857	5.91E-237	3.58E-234	up
FUT11	1.470490079	2.18E-38	7.68E-37	up
CLIP4	1.468643126	5.27E-41	2.03E-39	up
ELOVL4	1.465947385	4.70E-27	1.11E-25	up
ERFL	1.459953672	1.55E-05	7.54E-05	up
AC009495.3	1.45505773	0.016338	0.043552	up
CTH	1.446377658	3.17E-72	2.61E-70	up
FADS3	1.445003094	2.08E-88	2.20E-86	up
DLGAP1-AS2	1.439321909	2.85E-17	4.17E-16	up
GCLM	1.439225657	0	0	up
AC133785.1	1.438069098	0.001465	0.005037	up
SPAG4	1.436062363	9.43E-14	1.08E-12	up

Gene name	log2FoldChange	*P*-value	P_{adj}	change
CTNNAL1	1.429259108	3.71E-40	1.38E-38	up
RASGEF1A	1.426333817	0.002719	0.008842	up
HERC5	1.424705212	1.72E-05	8.35E-05	up
WNT5B	1.419305446	0.004628	0.014301	up
NUAK1	1.416060405	1.88E-20	3.30E-19	up
CELF5	1.408506267	8.39E-07	4.89E-06	up
AL357079.1	1.407034925	0.001814	0.00613	up
ARAP3	1.406674898	1.00E-10	9.06E-10	up
ADM	1.405775626	2.23E-55	1.26E-53	up
CYP26A1	1.399436043	0.001928	0.006476	up
RDH5	1.398418805	0.012669	0.034869	up
AL161756.1	1.397883544	0.006521	0.019464	up
ITGB3	1.387474225	4.78E-07	2.88E-06	up
PICART1	1.382074975	0.001019	0.003623	up
MFNG	1.376311335	1.33E-07	8.58E-07	up
HMOX1	1.368923498	2.29E-72	1.90E-70	up
L1CAM	1.358450852	1.14E-10	1.02E-09	up
PIM1	1.35599638	8.39E-68	6.26E-66	up
KLHL2P1	1.35437611	0.000212	0.000861	up
CDKN1A	1.354333438	5.77E-44	2.45E-42	up
SLCO2A1	1.350463396	1.51E-38	5.35E-37	up
MIR210HG	1.349599486	1.17E-06	6.65E-06	up
ETV4	1.345621069	3.57E-60	2.28E-58	up
ANKLE1	1.341788054	3.31E-28	8.21E-27	up
AL354872.2	1.340751966	0.007364	0.021677	up
ANK1	1.340195007	0.003923	0.012349	up
LINC01285	1.34010371	7.38E-11	6.71E-10	up
PRR7-AS1	1.337770565	6.40E-06	3.33E-05	up
HMGCS1	1.337636614	2.63E-225	1.38E-222	up
ERO1A	1.335246261	2.88E-195	1.05E-192	up
HRH1	1.333910361	4.87E-14	5.70E-13	up
TRIM16L	1.333290316	1.30E-153	3.05E-151	up
FADS1	1.332225495	3.28E-88	3.42E-86	up
AC018978.1	1.326939114	0.000149	0.00062	up
FBXO27	1.325731578	1.00E-06	5.78E-06	up
GJA3	1.325561882	1.61E-09	1.29E-08	up
LINC01589	1.319902304	0.0002	0.000815	up
PHLDB2	1.31900423	4.00E-74	3.39E-72	up
PDK4	1.31258444	0.002449	0.008031	up
SLC16A3	1.311003181	1.16E-102	1.52E-100	up
GADD45A	1.308232713	1.73E-177	5.09E-175	up
CNTD2	1.306799377	9.68E-24	2.01E-22	up
RCAN2	1.306784772	4.26E-22	8.07E-21	up
RGS20	1.306098922	8.14E-06	4.15E-05	up
LONRF1	1.304414415	1.04E-13	1.18E-12	up
IL18	1.302686003	2.06E-51	1.06E-49	up

Gene name	log2FoldChange	*P*-value	P_{adj}	change
LAMA3	1.302550027	1.44E-13	1.62E-12	up
COL12A1	1.302258203	7.41E-71	5.87E-69	up
NEDD9	1.300789243	1.75E-24	3.78E-23	up
ABLIM3	1.299800456	0.002418	0.007949	up
ERRFI1	1.298377282	2.70E-112	4.33E-110	up
ST3GAL3	1.294961397	1.27E-08	9.19E-08	up
ENOX1	1.293782284	2.30E-15	2.96E-14	up
ATF3	1.293110948	5.90E-30	1.56E-28	up
APLN	1.291662673	0.003171	0.010161	up
SORCS2	1.285451615	1.64E-09	1.31E-08	up
ULBP1	1.284337549	3.16E-18	4.92E-17	up
CAV1	1.281222903	1.20E-155	2.94E-153	up
KCTD15	1.278642886	5.62E-69	4.31E-67	up
TNFSF9	1.278276341	3.10E-22	5.93E-21	up
PLEKHG1	1.276970706	3.04E-06	1.65E-05	up
GATA6	1.274798596	3.40E-27	8.06E-26	up
BVES	1.27353659	7.47E-17	1.06E-15	up
AL161431.1	1.271638553	0.001184	0.00415	up
CCDC71L	1.271029376	1.27E-169	3.42E-167	up
TOX2	1.267177492	1.11E-13	1.25E-12	up
FBXW10	1.266315837	0.001096	0.003868	up
SLC2A14	1.260030113	0.000191	0.000781	up
IGFBP6	1.259337532	8.60E-09	6.37E-08	up
SERPINB5	1.256905917	2.26E-20	3.94E-19	up
JCAD	1.255938442	9.56E-05	0.000411	up
PNPLA3	1.255597665	6.10E-14	7.07E-13	up
CLGN	1.252792282	7.70E-15	9.53E-14	up
CRISPLD2	1.251235746	8.93E-59	5.50E-57	up
ANGPTL4	1.247134629	0.00072	0.002647	up
CARMIL2	1.245744572	0.000288	0.001137	up
IGFL1	1.244884505	1.30E-06	7.42E-06	up
ATP2B4	1.244576026	6.54E-56	3.77E-54	up
COL4A1	1.244158125	0.000262	0.001041	up
KRT16P6	1.243277525	1.98E-11	1.90E-10	up
GAREM2	1.243116008	5.48E-06	2.87E-05	up
SBK2	1.242155598	0.009755	0.027738	up
JAG1	1.238295634	1.98E-54	1.09E-52	up
GFPT2	1.237011793	2.09E-20	3.66E-19	up
WFDC3	1.233284064	0.001862	0.006276	up
MGLL	1.232976765	3.33E-63	2.26E-61	up
PFKFB3	1.22588708	8.36E-91	9.21E-89	up
UCKL1-AS1	1.225777677	6.82E-06	3.53E-05	up
AC008063.1	1.224200123	0.000163	0.000676	up
LINC02492	1.223607603	7.34E-06	3.78E-05	up
AC245052.4	1.222992447	0.000493	0.001869	up
PRELID3A	1.222872865	3.03E-39	1.09E-37	up

Gene name	log2FoldChange	*P*-value	P_{adj}	change
SAMD4A	1.220467116	1.34E-31	3.81E-30	up
LINC00431	1.219865659	0.001701	0.005781	up
SAMD12-AS1	1.218172533	0.000116	0.000492	up
CASP16P	1.21756584	0.010342	0.029202	up
GRIK2	1.215103285	1.06E-12	1.13E-11	up
P4HA1	1.214151262	8.41E-82	8.08E-80	up
ACSL4	1.20890708	7.54E-61	4.89E-59	up
AC025442.2	1.208162074	0.005375	0.016321	up
VEGFA	1.202424823	1.30E-141	2.71E-139	up
OSBPL6	1.19773254	1.48E-40	5.61E-39	up
HK2	1.19763505	1.39E-152	3.17E-150	up
PALD1	1.193986533	7.54E-12	7.49E-11	up
IDI2-AS1	1.191798799	0.000423	0.001625	up
TCAF2	1.19065659	1.33E-10	1.18E-09	up
KRT17	1.189969712	2.56E-152	5.77E-150	up
SGTB	1.18964059	2.81E-21	5.12E-20	up
MORN4	1.187330685	1.01E-14	1.25E-13	up
AKNA	1.184907739	1.07E-12	1.14E-11	up
CHAC1	1.184653929	1.77E-36	5.86E-35	up
CCNB1IP1	1.178192341	5.94E-107	8.47E-105	up
MYO1B	1.175490546	4.27E-204	1.93E-201	up
MSMO1	1.172876118	8.10E-181	2.58E-178	up
ME1	1.172188085	2.33E-105	3.23E-103	up
FTH1P2	1.170922959	0.000208	0.000842	up
SLC24A4	1.170891124	0.008423	0.024355	up
ANXA3	1.170786931	1.56E-106	2.21E-104	up
MYO7A	1.170439143	0.000352	0.001369	up
MXRA7	1.169753918	2.95E-17	4.31E-16	up
AMIGO2	1.165082666	7.06E-27	1.66E-25	up
GJB3	1.163952279	2.79E-62	1.86E-60	up
ANKZF1	1.162027388	8.44E-71	6.65E-69	up
MILR1	1.161996369	0.011072	0.031028	up
SPRY2	1.160959993	5.54E-53	2.98E-51	up
ESPNP	1.159828369	0.00427	0.013317	up
CALB2	1.158689901	1.19E-08	8.69E-08	up
TMEFF1	1.157997876	0.011183	0.031306	up
CSTA	1.15425135	1.93E-06	1.08E-05	up
KLK10	1.154121922	7.02E-112	1.11E-109	up
P4HA2	1.152772552	7.67E-44	3.24E-42	up
TMEM97	1.151723586	1.32E-70	1.04E-68	up
ARNTL2	1.151508277	6.36E-23	1.26E-21	up
RELB	1.150466888	4.95E-38	1.73E-36	up
AP3B2	1.149279371	0.001082	0.003827	up
ACTBL2	1.143956205	0.007973	0.02322	up
GPX3	1.143851005	0.004553	0.0141	up
MANCR	1.140270636	0.001559	0.005333	up

Gene name	log2FoldChange	*P*-value	P_{adj}	change
CACNB1	1.140042207	2.28E-16	3.14E-15	up
AL590004.3	1.136923741	2.83E-05	0.000133	up
CD177	1.13423548	0.006981	0.020686	up
CUEDC1	1.133564922	4.10E-15	5.18E-14	up
AC009236.1	1.130946954	0.003608	0.01144	up
LY6K	1.13071092	8.84E-37	2.96E-35	up
DUSP4	1.130547262	1.20E-66	8.72E-65	up
RFLNB	1.128558496	0.000117	0.000495	up
DNAH11	1.127567542	4.29E-37	1.46E-35	up
GLI3	1.124567063	3.11E-23	6.25E-22	up
IZUMO1	1.124359571	0.002142	0.00712	up
CYP51A1	1.121747078	1.07E-24	2.34E-23	up
AL513534.2	1.121134031	0.008031	0.023372	up
PHLDA1	1.120208388	1.72E-41	6.77E-40	up
CD55	1.117298793	5.38E-48	2.58E-46	up
SCEL	1.115595584	0.017259	0.045607	up
OSMR	1.114809278	1.14E-87	1.18E-85	up
BMPR1B	1.114012657	0.000367	0.001424	up
KCNK1	1.113562254	1.77E-42	7.21E-41	up
ACP7	1.112441378	3.13E-09	2.42E-08	up
DAPK1	1.112216914	9.46E-35	2.97E-33	up
TRIM36	1.112172274	0.000938	0.003373	up
SPRY4	1.111049244	6.93E-22	1.30E-20	up
PROC	1.108423608	3.40E-07	2.09E-06	up
RAB3IL1	1.104127749	2.27E-21	4.15E-20	up
GOLGA8A	1.101932956	5.62E-23	1.12E-21	up
KIF21B	1.101880725	1.08E-39	3.96E-38	up
HELT	1.096428026	0.014249	0.038639	up
ACAT2	1.092989533	1.93E-200	7.79E-198	up
SLC7A5	1.089947015	1.35E-210	6.41E-208	up
PKD1P6	1.085179158	8.23E-14	9.44E-13	up
STK39	1.084737986	9.34E-21	1.66E-19	up
FKBP9P1	1.0826967	8.91E-05	0.000385	up
AC011447.7	1.080327486	9.56E-05	0.000411	up
STARD4	1.080049797	5.51E-29	1.40E-27	up
PKD1	1.079405809	2.83E-47	1.33E-45	up
ALDH1A3	1.078742707	1.11E-233	6.50E-231	up
SLC16A13	1.075758232	1.64E-27	3.94E-26	up
SQLE	1.075083676	6.92E-200	2.74E-197	up
CDA	1.07305637	0.004458	0.01383	up
AXL	1.071802098	0.003885	0.012241	up
FAM24B	1.071292459	1.38E-05	6.77E-05	up
SMIM2-AS1	1.070662243	1.20E-05	5.96E-05	up
AC027031.2	1.066338075	2.95E-11	2.79E-10	up
NR4A1	1.064407904	7.86E-16	1.05E-14	up
CCT6P1	1.061003352	3.62E-11	3.39E-10	up

Gene name	log2FoldChange	*P*-value	P_{adj}	change
LPIN1	1.053904658	1.84E-64	1.28E-62	up
PAX9	1.05346803	0.000474	0.001805	up
RHBDL3	1.053236127	3.73E-06	2.00E-05	up
KIF3C	1.053182333	6.92E-31	1.90E-29	up
TLCD3A	1.052184521	4.58E-35	1.45E-33	up
CYP4F3	1.051671149	4.04E-12	4.10E-11	up
PLIN2	1.0493054	1.13E-96	1.33E-94	up
AC145098.2	1.049100754	5.92E-05	0.000264	up
COL23A1	1.045423684	6.27E-05	0.000279	up
HHEX	1.044594416	1.84E-08	1.31E-07	up
LARP6	1.043902689	7.58E-22	1.42E-20	up
TFRC	1.04322364	3.51E-271	2.73E-268	up
ZNF404	1.041159289	3.06E-11	2.89E-10	up
SMIM10L2B	1.041134213	0.012537	0.034568	up
ALKBH6	1.04099238	1.37E-05	6.72E-05	up
TNFRSF10A	1.040733548	3.06E-24	6.54E-23	up
ADCY7	1.040078754	4.43E-18	6.84E-17	up
MXI1	1.038813272	2.42E-48	1.17E-46	up
NMNAT2	1.038694677	4.24E-06	2.25E-05	up
HSD17B7	1.034882099	8.51E-40	3.12E-38	up
SPHK1	1.033790862	1.37E-29	3.58E-28	up
C5AR1	1.032517907	0.000542	0.002036	up
NPAS2	1.031721477	7.31E-39	2.60E-37	up
NPIPA9	1.024043426	0.012998	0.035648	up
ZNF829	1.022782006	1.03E-22	2.02E-21	up
SLC1A1	1.017938414	5.89E-37	2.00E-35	up
SDAD1P1	1.01732383	6.04E-05	0.000269	up
ZFP28	1.015408075	1.05E-26	2.46E-25	up
HS6ST2	1.014927823	3.00E-21	5.44E-20	up
SLC1A4	1.01434689	2.01E-42	8.14E-41	up
SUN3	1.009878257	4.31E-37	1.47E-35	up
SC5D	1.009683243	6.09E-173	1.66E-170	up
DHCR7	1.009589996	2.67E-259	1.85E-256	up
SEC14L2	1.009114356	9.28E-28	2.25E-26	up
VGF	1.007169297	7.87E-06	4.02E-05	up
KCNG3	1.006440056	0.000677	0.0025	up
EGLN3	1.006249822	9.35E-76	8.10E-74	up
ALDH1L2	1.005853647	5.51E-81	5.18E-79	up
RASSF8	1.005704639	1.83E-44	7.91E-43	up
LINC01714	1.00557743	0.012163	0.033662	up
LINC01619	1.004571455	0.012811	0.035209	up
GJB5	1.002600701	8.48E-24	1.77E-22	up
SLC26A6	1.002307163	3.76E-42	1.50E-40	up
STAMBPL1	1.001013067	7.70E-18	1.17E-16	up
ARID3B	1.00090675	3.51E-09	2.71E-08	up
NPHP1	1.000106582	5.83E-07	3.47E-06	up

Gene name	log2FoldChange	*P*-value	P_{adj}	change
FLVCR2	-1.000198251	3.00E-09	2.33E-08	down
SELENOW	-1.000334839	2.28E-137	4.66E-135	down
H4C9	-1.000839142	2.47E-11	2.35E-10	down
S100A7A	-1.00250957	3.05E-14	3.63E-13	down
RAP1GAP	-1.003315696	4.94E-17	7.14E-16	down
SFXN2	-1.003801309	7.32E-11	6.66E-10	down
MAP2K6	-1.006702413	0.000643	0.002384	down
TSC22D3	-1.007597238	4.21E-125	7.56E-123	down
SLC15A2	-1.008886591	0.009729	0.027672	down
PRADC1	-1.011427037	7.80E-19	1.26E-17	down
FAM167B	-1.011507639	0.000113	0.000479	down
SCNN1B	-1.012076052	0.002479	0.008122	down
MYORG	-1.015706241	1.87E-31	5.31E-30	down
TYMSOS	-1.015971979	6.44E-08	4.31E-07	down
LINC01132	-1.016140655	1.02E-06	5.87E-06	down
ADIRF-AS1	-1.016167124	2.14E-12	2.22E-11	down
SESN3	-1.016437647	5.19E-08	3.50E-07	down
LYRM9	-1.016978887	0.009148	0.026218	down
VASN	-1.018586001	1.45E-20	2.57E-19	down
TNS1	-1.020631595	2.26E-31	6.33E-30	down
PPP1R3C	-1.021693655	3.32E-40	1.24E-38	down
AC092164.1	-1.023701432	0.005532	0.01675	down
CX3CL1	-1.02430934	6.86E-66	4.93E-64	down
FMO4	-1.025353725	2.45E-05	0.000116	down
ARHGEF39	-1.026012873	1.06E-27	2.57E-26	down
ERMAP	-1.02867922	2.23E-25	4.95E-24	down
CD59	-1.029215628	1.77E-227	1.01E-224	down
PLAU	-1.029297998	2.81E-28	6.98E-27	down
CERS4	-1.029883458	1.22E-76	1.09E-74	down
AC103957.2	-1.030498975	1.26E-08	9.15E-08	down
SLC4A8	-1.031653586	1.37E-08	9.94E-08	down
C1GALT1C1L	-1.032601692	0.000847	0.003068	down
AC080080.1	-1.037057512	0.012127	0.033577	down
AC099343.4	-1.037369401	0.000218	0.000879	down
TOLLIP-AS1	-1.039368096	0.003609	0.011443	down
SYNC	-1.042362904	0.003643	0.011541	down
DCXR	-1.044444511	1.04E-54	5.77E-53	down
IL17RC	-1.044456688	1.13E-30	3.08E-29	down
SEPSECS-AS1	-1.044487333	0.004267	0.01331	down
SIDT1	-1.04546262	0.000149	0.000619	down
MYCN	-1.047327914	9.06E-09	6.70E-08	down
MZB1	-1.048032098	0.010099	0.028602	down
TNS4	-1.051286566	2.70E-23	5.44E-22	down
IL1R1	-1.051661082	8.34E-33	2.49E-31	down
LINC01182	-1.0535845	2.05E-13	2.28E-12	down
CLIC5	-1.054283205	4.08E-13	4.45E-12	down

Gene name	log2FoldChange	*P*-value	P_{adj}	change
MYLIP	-1.055561909	3.90E-32	1.14E-30	down
UNC5B-AS1	-1.057764785	0.000462	0.001763	down
ZNF233	-1.058286372	0.000125	0.000527	down
GJC3	-1.058491561	0.003407	0.010858	down
PDCD4	-1.058629688	1.25E-101	1.61E-99	down
CD6	-1.059049016	0.00568	0.01716	down
VAV1	-1.059258336	9.38E-09	6.92E-08	down
MISP3	-1.059834148	3.04E-09	2.36E-08	down
AP002761.4	-1.059839623	5.74E-18	8.81E-17	down
GPD1L	-1.06131484	4.06E-71	3.26E-69	down
LINC01089	-1.0621871	0.000729	0.00268	down
SV2C	-1.065247993	0.001025	0.003643	down
AC111200.2	-1.065657729	0.002206	0.007313	down
MME	-1.066210604	1.82E-177	5.27E-175	down
AC106795.2	-1.066878177	0.001024	0.003638	down
C1orf115	-1.067478992	3.26E-48	1.56E-46	down
LBH	-1.068793133	2.91E-31	8.10E-30	down
TMC1	-1.069275897	8.27E-09	6.14E-08	down
SLC25A20	-1.069650204	9.75E-23	1.91E-21	down
PRRT3	-1.07503002	1.28E-05	6.33E-05	down
PBX1	-1.075183396	1.29E-83	1.27E-81	down
AKR1D1	-1.076947816	0.008189	0.023768	down
KYNU	-1.077208021	8.00E-132	1.55E-129	down
PLAAT3	-1.079622857	3.40E-70	2.63E-68	down
P2RY2	-1.081627369	2.08E-27	4.97E-26	down
EPHX2	-1.082998103	4.72E-08	3.20E-07	down
ANKRD1	-1.083731883	7.92E-06	4.04E-05	down
AC104454.2	-1.085235417	0.000552	0.002073	down
CCR3	-1.088455429	1.90E-06	1.06E-05	down
AP003307.2	-1.089128068	0.008326	0.024146	down
ANXA8L1	-1.092425998	2.19E-63	1.49E-61	down
ZSCAN16-AS1	-1.094704819	0.000139	0.00058	down
LINC00920	-1.09528891	2.35E-06	1.29E-05	down
MEGF10	-1.101586033	0.000184	0.000754	down
AC093904.2	-1.102769149	0.000205	0.000832	down
GALNT12	-1.104204555	8.52E-06	4.33E-05	down
LIMK2	-1.109056895	1.53E-109	2.27E-107	down
JMJD1C-AS1	-1.111191845	0.000762	0.002789	down
SEPTIN1	-1.111744192	0.006206	0.018607	down
SLC9A3R2	-1.114282139	4.58E-90	4.91E-88	down
WNT10A	-1.11569939	3.07E-11	2.89E-10	down
MTUS2	-1.116936334	3.42E-09	2.64E-08	down
AP001636.1	-1.11817467	0.010142	0.028716	down
SCUBE3	-1.120655354	5.71E-06	2.99E-05	down
DEPTOR	-1.125115218	6.69E-21	1.20E-19	down
ACSS1	-1.125902162	3.44E-12	3.51E-11	down

Gene name	log2FoldChange	*P*-value	P_{adj}	change
ATP2A1	-1.128789688	0.00444	0.013786	down
ARHGAP25	-1.131041844	1.78E-05	8.63E-05	down
H2BC15	-1.131209979	0.001841	0.006214	down
STARD8	-1.131767621	1.90E-18	3.02E-17	down
LY6E	-1.132026687	4.66E-194	1.68E-191	down
GYG2	-1.132504509	6.20E-43	2.57E-41	down
NREP	-1.134109221	3.07E-08	2.14E-07	down
SLC6A12	-1.134612136	0.00575	0.017356	down
TMEM169	-1.135754414	0.008186	0.023764	down
AL121820.2	-1.135754474	0.002336	0.007705	down
TRIM6	-1.1359424	4.94E-16	6.63E-15	down
RAB26	-1.137304548	2.45E-06	1.34E-05	down
PIK3IP1	-1.13919715	1.35E-13	1.52E-12	down
TENT5B	-1.139243905	1.86E-33	5.61E-32	down
C11orf71	-1.139544214	6.52E-06	3.39E-05	down
HOXC-AS1	-1.140882118	0.001756	0.005949	down
SYT12	-1.141626697	6.89E-186	2.30E-183	down
LFNG	-1.141859175	4.54E-08	3.09E-07	down
AEBP1	-1.141871796	0.008881	0.025542	down
MFSD3	-1.145316659	3.08E-48	1.48E-46	down
PDZD2	-1.14593478	2.67E-20	4.64E-19	down
AL359513.1	-1.146916847	0.00161	0.005493	down
AC083841.1	-1.148477074	2.46E-08	1.73E-07	down
AC074212.1	-1.148662615	0.001292	0.004497	down
GGT6	-1.155329998	3.06E-59	1.91E-57	down
SYNPO	-1.1576244	2.06E-35	6.62E-34	down
C12orf60	-1.160099412	6.87E-05	0.000303	down
H4C5	-1.161368097	0.005088	0.015542	down
C6orf141	-1.163422594	6.37E-07	3.77E-06	down
CLDN10	-1.168386724	1.74E-101	2.22E-99	down
CCDC121	-1.168580189	1.86E-07	1.18E-06	down
DOK7	-1.173623662	0.00524	0.015954	down
SNHG19	-1.174223641	3.06E-15	3.89E-14	down
CXCL17	-1.174919752	0.006559	0.019566	down
AL356215.1	-1.178202969	4.76E-06	2.51E-05	down
ARHGEF38	-1.180467361	1.83E-13	2.04E-12	down
TMEM270	-1.181831294	0.000841	0.003048	down
LINC01671	-1.181862574	1.65E-10	1.46E-09	down
PIF1	-1.183353004	8.53E-27	1.99E-25	down
OLFM4	-1.18434826	7.49E-14	8.63E-13	down
AC018629.1	-1.18623	1.30E-05	6.40E-05	down
SUN2	-1.195627853	1.16E-108	1.71E-106	down
SLC46A3	-1.198899627	2.68E-27	6.38E-26	down
LINC01487	-1.199873154	0.009247	0.026476	down
SOX18	-1.20421041	7.96E-08	5.27E-07	down
KLRF1	-1.204573475	1.10E-06	6.29E-06	down

Gene name	log2FoldChange	*P*-value	P_{adj}	change
H2AW	-1.207328269	3.57E-10	3.04E-09	down
TNNC1	-1.207891001	8.51E-06	4.33E-05	down
LINC01270	-1.208115718	3.36E-08	2.33E-07	down
RORC	-1.209211905	3.00E-17	4.38E-16	down
CHRNA3	-1.211643293	2.55E-07	1.58E-06	down
LINC01740	-1.216383062	0.00076	0.002782	down
KIAA1217	-1.219214644	2.03E-101	2.57E-99	down
SYTL3	-1.219714037	3.82E-12	3.89E-11	down
DSC1	-1.220998278	0.000606	0.002257	down
CKAP2L	-1.222484625	1.48E-54	8.22E-53	down
TMEM139	-1.223267008	0.000196	0.000798	down
AF274858.1	-1.224355565	0.002111	0.00703	down
CMAHP	-1.227730644	1.52E-101	1.95E-99	down
CLEC2D	-1.228212224	4.00E-07	2.44E-06	down
VPS37D	-1.228985793	4.74E-05	0.000215	down
PSRC1	-1.230218316	3.46E-101	4.30E-99	down
SOWAHA	-1.23134831	3.54E-06	1.90E-05	down
AC093575.1	-1.231808901	5.69E-05	0.000255	down
MAGI3	-1.233869637	4.02E-81	3.80E-79	down
RGL1	-1.23642465	6.58E-20	1.13E-18	down
GTF2IP7	-1.236552555	0.001325	0.0046	down
LEMD1	-1.242570324	8.73E-11	7.90E-10	down
TTC36	-1.24992265	0.009013	0.025875	down
AC084018.1	-1.25000677	0.008557	0.024712	down
MN1	-1.250632804	1.30E-14	1.59E-13	down
TSSK3	-1.25735544	0.010572	0.029769	down
C20orf204	-1.258918292	4.00E-07	2.43E-06	down
SNHG21	-1.259565819	1.64E-05	7.97E-05	down
RERG	-1.260218673	1.96E-79	1.82E-77	down
H3C6	-1.261878076	5.74E-05	0.000256	down
AQP1	-1.264588144	8.71E-08	5.73E-07	down
AP002381.2	-1.264659554	9.73E-07	5.62E-06	down
MMAA	-1.264822956	2.13E-11	2.04E-10	down
LNX1	-1.26561565	3.93E-09	3.02E-08	down
DCHS1	-1.268682936	0.004109	0.012873	down
PLEKHG4B	-1.269700884	3.59E-23	7.19E-22	down
PPFIBP2	-1.271976698	5.48E-32	1.58E-30	down
SCPEP1	-1.278746144	4.96E-166	1.28E-163	down
IKZF2	-1.280031093	6.52E-19	1.06E-17	down
AC004938.2	-1.282168673	0.000214	0.000868	down
LINC01508	-1.283255276	3.19E-05	0.000148	down
PLCL1	-1.283895528	0.000348	0.001358	down
AC010173.1	-1.28497617	0.000736	0.002701	down
STOML1	-1.285043382	2.91E-10	2.50E-09	down
SNCG	-1.287715628	1.93E-10	1.69E-09	down
CHADL	-1.291905954	1.28E-05	6.33E-05	down

Gene name	log2FoldChange	*P*-value	P_{adj}	change
CHI3L1	-1.296772895	2.97E-07	1.83E-06	down
H3C8	-1.298490979	0.000237	0.000953	down
CRACR2B	-1.29936622	5.48E-105	7.49E-103	down
FAM102B	-1.302912122	6.09E-61	3.98E-59	down
ZNF703	-1.303734877	9.27E-62	6.10E-60	down
TNFSF10	-1.304585638	9.69E-87	9.84E-85	down
SVOPL	-1.306244422	1.80E-06	1.01E-05	down
AC108134.2	-1.309622544	0.0131	0.035887	down
AZGP1	-1.311583381	0	0	down
GPR55	-1.311715526	0.004617	0.01427	down
PLCH1	-1.312632234	6.25E-17	8.97E-16	down
RHOU	-1.313181737	1.58E-06	8.87E-06	down
LINC00592	-1.317065564	0.007272	0.021451	down
AC021066.1	-1.320042459	9.84E-13	1.05E-11	down
STX17-AS1	-1.321100347	0.001705	0.005793	down
SPTSSB	-1.323648232	2.75E-17	4.03E-16	down
IL17RE	-1.32383621	5.11E-30	1.36E-28	down
MSLN	-1.324875396	0.001217	0.00426	down
CYP2T1P	-1.32647339	2.93E-05	0.000137	down
LINC01311	-1.327248591	3.11E-05	0.000145	down
TEX9	-1.328451206	2.37E-29	6.17E-28	down
AC008514.1	-1.328754727	0.000303	0.001192	down
SEMA3C	-1.329716104	1.65E-48	7.98E-47	down
KLHL30	-1.330086065	0.000326	0.001278	down
AC000068.1	-1.334564491	0.000213	0.000863	down
SPDEF	-1.335609203	1.79E-174	5.12E-172	down
AC005523.1	-1.335764111	0.008945	0.025695	down
ABCA1	-1.337746086	2.00E-05	9.62E-05	down
LY6E-DT	-1.339157715	5.30E-18	8.14E-17	down
CNIH3	-1.341600003	6.23E-05	0.000277	down
AL022345.4	-1.342434412	0.003514	0.011166	down
SNORD3B-1	-1.34415027	0.000997	0.003557	down
HCAR2	-1.347032865	9.85E-17	1.39E-15	down
NEK10	-1.347500439	1.56E-06	8.80E-06	down
AC007325.4	-1.349899259	0.002167	0.007199	down
MT1F	-1.350125495	3.23E-16	4.36E-15	down
ATP6V0A4	-1.352211687	2.12E-174	5.97E-172	down
BNIPL	-1.352397367	8.35E-05	0.000363	down
NINJ2-AS1	-1.353784118	0.00035	0.001365	down
PTGIS	-1.360851189	0.000106	0.000453	down
GADD45G	-1.361139297	6.33E-09	4.75E-08	down
TNS2	-1.361306555	2.92E-31	8.12E-30	down
AXIN2	-1.361941098	1.50E-31	4.25E-30	down
CMYA5	-1.362391476	6.08E-24	1.28E-22	down
AC011330.1	-1.364044301	0.006835	0.020298	down
CDH18	-1.365994329	0.003938	0.012393	down

Gene name	log2FoldChange	*P*-value	P_{adj}	change
RGCC	-1.368438105	1.48E-15	1.94E-14	down
SLC7A2	-1.368881636	1.18E-08	8.63E-08	down
SCGB2A2	-1.369991057	0.00019	0.000779	down
LDHD	-1.370821329	9.54E-09	7.03E-08	down
FBXO43	-1.375518877	2.39E-16	3.27E-15	down
NGFR	-1.382650126	0.001137	0.004002	down
CISH	-1.382802736	2.91E-12	3.00E-11	down
PIGR	-1.383884191	0.004428	0.013754	down
ABCC6P1	-1.385423025	0.002997	0.00966	down
MAGEA4	-1.386249384	1.74E-05	8.41E-05	down
LINC01630	-1.388209222	0.001704	0.005791	down
ERP27	-1.391267536	1.18E-16	1.67E-15	down
FUZ	-1.395178859	0.008624	0.024882	down
CAPN13	-1.396846198	3.50E-36	1.15E-34	down
HOXA-AS2	-1.403264024	4.89E-13	5.31E-12	down
PLB1	-1.405790226	0.000601	0.002239	down
AC103740.2	-1.407368617	0.000924	0.003323	down
CPAMD8	-1.408815992	6.06E-79	5.60E-77	down
ADORA2A-AS1	-1.410505922	0.002834	0.009181	down
KRT4	-1.41539518	5.42E-99	6.65E-97	down
LRRC17	-1.416193404	0.007792	0.022759	down
RBBP8NL	-1.417523941	2.00E-27	4.78E-26	down
ANKRD30B	-1.418083206	0.000738	0.002708	down
AP003392.5	-1.419210374	0.005682	0.017163	down
MAOA	-1.42938191	0.000958	0.003436	down
EDN2	-1.429996353	5.40E-56	3.13E-54	down
THRB-AS1	-1.432720582	0.003437	0.010949	down
GAMT	-1.433067062	8.23E-20	1.40E-18	down
FAM49A	-1.43345572	1.45E-40	5.49E-39	down
PLA2G4F	-1.43510851	0.000867	0.003133	down
MUC5B	-1.436717261	1.18E-08	8.64E-08	down
CHST8	-1.441988921	0.01766	0.04654	down
AL512599.1	-1.44239403	1.37E-30	3.71E-29	down
AL031847.1	-1.451275701	0.006933	0.020566	down
C9orf152	-1.452474412	2.50E-44	1.08E-42	down
RASSF10	-1.454665653	4.64E-36	1.52E-34	down
ART3	-1.456815845	2.68E-05	0.000126	down
LRIG2-DT	-1.457194499	0.004841	0.014872	down
H2AJ	-1.481449433	4.98E-153	1.15E-150	down
UPK2	-1.484156317	2.64E-15	3.39E-14	down
LINC01208	-1.489417606	1.02E-05	5.12E-05	down
PDZK1IP1	-1.491937622	1.35E-45	6.04E-44	down
DLX3	-1.492821782	2.44E-46	1.13E-44	down
CITED4	-1.499175661	6.49E-299	6.29E-296	down
ARHGAP5-AS1	-1.499299297	4.84E-05	0.000219	down
CD14	-1.49934141	6.02E-108	8.72E-106	down

Gene name	log2FoldChange	*P*-value	P_{adj}	change
AC005077.4	-1.500034892	3.09E-07	1.90E-06	down
PKP4-AS1	-1.504088013	0.004803	0.014774	down
APBA1	-1.51125433	0.000964	0.003451	down
SLURP1	-1.513979842	3.23E-14	3.83E-13	down
CD82	-1.514495504	1.10E-304	1.13E-301	down
GJA5	-1.526421925	3.98E-44	1.70E-42	down
LINC00456	-1.528090001	0.000423	0.001625	down
LYPD2	-1.529808081	0.006489	0.019384	down
SCIN	-1.530737234	1.09E-07	7.08E-07	down
TTC39A-AS1	-1.531192553	0.000297	0.001168	down
VGLL1	-1.543528946	2.50E-246	1.62E-243	down
CYP4X1	-1.54521199	1.04E-250	6.96E-248	down
CPHL1P	-1.545639562	3.63E-05	0.000167	down
H2BC8	-1.547934597	6.27E-05	0.000279	down
LINC02635	-1.548005129	5.94E-05	0.000265	down
CD24	-1.552633329	0	0	down
OBP2B	-1.554436342	5.79E-07	3.45E-06	down
S100A8	-1.556273651	1.67E-200	6.89E-198	down
GLYATL2	-1.55923879	8.53E-110	1.27E-107	down
CP	-1.559611254	7.33E-118	1.23E-115	down
TLR5	-1.562540185	4.38E-20	7.56E-19	down
TGFB3	-1.563421062	7.86E-12	7.79E-11	down
H1-0	-1.567053527	0	0	down
AL451165.2	-1.568880765	0.00028	0.001108	down
HLF	-1.576023988	2.36E-11	2.25E-10	down
CD24P4	-1.592629893	1.85E-07	1.17E-06	down
SLC6A14	-1.599282287	1.60E-08	1.15E-07	down
NEURL3	-1.599516732	3.67E-31	1.02E-29	down
H2BC21	-1.601738795	1.46E-05	7.18E-05	down
LY6D	-1.611797617	1.11E-103	1.48E-101	down
H2AC8	-1.613247374	0.000577	0.002157	down
CAB39L	-1.614420946	8.09E-26	1.83E-24	down
RNF152	-1.614670204	1.77E-15	2.31E-14	down
LINC01235	-1.614774543	8.74E-32	2.51E-30	down
ASCL2	-1.615112284	1.11E-13	1.25E-12	down
MPP1	-1.618388779	4.86E-05	0.00022	down
CSF3R	-1.625172559	1.23E-42	5.02E-41	down
SPNS3	-1.625818288	2.64E-09	2.06E-08	down
CLDN16	-1.625941275	0.002205	0.007309	down
KIF20A	-1.635907223	0	0	down
KRT14	-1.639992065	9.20E-49	4.48E-47	down
MIR99AHG	-1.645188817	1.99E-07	1.25E-06	down
UPB1	-1.647773348	0.000609	0.002267	down
CCL17	-1.648973311	0.0079	0.023032	down
MUC15	-1.665065822	1.43E-30	3.88E-29	down
SMPD1	-1.665283833	7.64E-46	3.44E-44	down

Gene name	log2FoldChange	*P*-value	P_{adj}	change
LINC02331	-1.665448806	5.58E-10	4.66E-09	down
SLC4A11	-1.666734798	4.33E-203	1.86E-200	down
AC026689.1	-1.668904587	1.94E-05	9.34E-05	down
PROM1	-1.674538033	1.19E-59	7.57E-58	down
SELENBP1	-1.67537298	8.65E-103	1.14E-100	down
VNN2	-1.690167771	0.004672	0.014419	down
LINC01003	-1.692689959	1.92E-07	1.21E-06	down
ID2	-1.698905389	7.60E-09	5.66E-08	down
AC093001.1	-1.699801707	2.11E-09	1.67E-08	down
ACPP	-1.702634762	2.64E-29	6.84E-28	down
SLC1A2	-1.702909507	3.13E-16	4.23E-15	down
AGR2	-1.706582294	7.24E-69	5.51E-67	down
APCDD1	-1.707094777	2.97E-26	6.79E-25	down
H4C8	-1.711870692	1.04E-08	7.63E-08	down
IGFBP5	-1.713050231	4.22E-281	3.72E-278	down
H2BC12	-1.717371052	1.57E-60	1.01E-58	down
IFITM2	-1.732293891	4.21E-39	1.51E-37	down
NPY1R	-1.732510566	4.17E-05	0.000191	down
PLAAT2	-1.736286529	2.03E-23	4.13E-22	down
ST3GAL5	-1.736895003	1.24E-15	1.64E-14	down
CLU	-1.741210744	0	0	down
SEMA3E	-1.747031602	4.43E-59	2.76E-57	down
AKAP6	-1.750326116	5.64E-34	1.73E-32	down
H1-2	-1.755253029	3.92E-23	7.84E-22	down
ITGAL	-1.7609337	0.000181	0.000745	down
LIPK	-1.767998322	1.06E-07	6.90E-07	down
ARHGEF6	-1.768008842	1.92E-125	3.49E-123	down
TIPARP-AS1	-1.769233934	0.00588	0.017721	down
AC144450.1	-1.780734147	7.64E-06	3.92E-05	down
RASGRP2	-1.781808043	9.00E-07	5.23E-06	down
MSMB	-1.784499403	1.02E-06	5.87E-06	down
SLC30A2	-1.786118726	6.37E-07	3.77E-06	down
SLC5A1	-1.793379195	1.97E-30	5.31E-29	down
AMOT	-1.800677316	1.30E-179	4.07E-177	down
AL109976.1	-1.80853765	0.000166	0.000688	down
AQP2	-1.810055741	0.005257	0.016	down
TMPRSS4	-1.810094809	5.98E-86	5.98E-84	down
AL137077.2	-1.813699175	0.001773	0.006	down
TIMP3	-1.8175653	1.01E-195	3.77E-193	down
AC007906.2	-1.820429057	4.24E-22	8.05E-21	down
SSPO	-1.826166526	1.66E-07	1.06E-06	down
CDKN2C	-1.830121567	7.96E-197	3.03E-194	down
CRIP2	-1.831505984	1.12E-93	1.27E-91	down
UPK3B	-1.856289688	8.80E-23	1.72E-21	down
PLAAT4	-1.859271223	1.25E-214	6.06E-212	down
CCR1	-1.859365489	0.000465	0.001772	down

Gene name	log2FoldChange	*P*-value	P_{adj}	change
MXD3	-1.860280852	5.75E-124	1.02E-121	down
PNMT	-1.863949655	9.77E-05	0.000419	down
CAPN9	-1.867702162	0.003055	0.00984	down
AC141930.1	-1.870802675	2.10E-05	0.000101	down
BLNK	-1.871546526	4.70E-15	5.89E-14	down
EWSAT1	-1.871549911	0.003049	0.009821	down
H2BC9	-1.898511335	1.77E-08	1.27E-07	down
KIT	-1.899193937	5.88E-60	3.75E-58	down
LRRC26	-1.902169207	1.12E-06	6.43E-06	down
EFEMP1	-1.911549584	2.11E-96	2.48E-94	down
AZGP1P1	-1.913207389	1.12E-13	1.26E-12	down
LINC02028	-1.914749303	3.07E-05	0.000143	down
CCL22	-1.917217859	6.15E-17	8.83E-16	down
PCDH8	-1.917739914	7.60E-06	3.90E-05	down
ATP13A5	-1.930363439	2.15E-101	2.69E-99	down
GBP2	-1.930570365	1.21E-49	6.01E-48	down
ANXA2R	-1.95869466	0.000128	0.000538	down
AP001207.3	-1.972245616	6.25E-11	5.73E-10	down
ABCG1	-1.979985515	6.38E-56	3.69E-54	down
MGP	-1.984185956	0	0	down
AC007325.2	-1.987595467	3.43E-15	4.35E-14	down
PDE1C	-1.988502155	0.0002	0.000814	down
ALDH1A1	-1.988899349	2.66E-110	4.00E-108	down
AL603840.1	-1.993009765	0.009408	0.026878	down
ESPN	-1.995119135	3.07E-49	1.52E-47	down
AC046168.1	-1.99544307	0.000854	0.003089	down
CAPS	-1.996711012	2.09E-18	3.30E-17	down
ZBTB7C	-1.996860322	1.27E-36	4.22E-35	down
LRRC55	-1.997385043	7.95E-06	4.06E-05	down
CREB3L4	-1.999390004	1.88E-145	4.14E-143	down
PALM3	-2.001991278	6.13E-17	8.80E-16	down
FAM110D	-2.010890905	0.000656	0.002428	down
ELF5	-2.024478109	0.000596	0.002223	down
DHRS2	-2.039442696	2.74E-41	1.06E-39	down
ISLR	-2.045718004	2.72E-08	1.90E-07	down
FAT2	-2.047613681	0.000572	0.002141	down
RUNDC3A	-2.055032417	0.000533	0.002007	down
CSF2RB	-2.057237327	9.46E-07	5.47E-06	down
AP000880.1	-2.072104606	0.008496	0.024555	down
TRIM63	-2.077500251	0.002436	0.007996	down
H2BC4	-2.091451187	2.47E-09	1.93E-08	down
ADAMTS1	-2.093316648	0.000208	0.000845	down
MYCNOS	-2.097087156	5.08E-09	3.84E-08	down
CCDC160	-2.105258544	6.01E-06	3.14E-05	down
MOGAT2	-2.108267219	1.43E-06	8.08E-06	down
MAB21L4	-2.112630608	3.76E-187	1.28E-184	down

Gene name	log2FoldChange	*P*-value	P_{adj}	change
AC108479.1	-2.132379544	0.007697	0.02251	down
AC124947.2	-2.134264195	0.000241	0.000965	down
BCAS1	-2.134370413	2.66E-06	1.45E-05	down
HSD17B2	-2.137647496	3.42E-29	8.83E-28	down
MT1X	-2.143984941	4.21E-39	1.51E-37	down
PLEKHS1	-2.145600289	5.61E-31	1.54E-29	down
ANTXRL	-2.155989562	0.000444	0.001699	down
PLEKHD1	-2.166343454	9.04E-14	1.03E-12	down
H2BC11	-2.16735401	3.38E-10	2.89E-09	down
EXTL1	-2.172785955	1.96E-05	9.41E-05	down
PSCA	-2.180411103	6.05E-27	1.42E-25	down
MIR3150BHG	-2.196545663	0.00431	0.013427	down
PPP1R1B	-2.204666143	1.19E-133	2.38E-131	down
MT2P1	-2.214722616	0.000997	0.003556	down
LGALS9	-2.220643344	4.67E-07	2.82E-06	down
NRG2	-2.225487991	9.86E-08	6.44E-07	down
AC097369.2	-2.231940923	0.003688	0.011667	down
COLCA2	-2.252240726	2.28E-12	2.36E-11	down
NDRG2	-2.260537823	9.34E-28	2.26E-26	down
GRAP2	-2.260962119	0.000393	0.001519	down
INMT	-2.284534945	0.0073	0.021511	down
H3C10	-2.287038921	8.82E-12	8.71E-11	down
UPP2	-2.292337024	0.001589	0.005428	down
ATP6V1B1	-2.29302894	5.45E-53	2.94E-51	down
H2BC17	-2.293422133	0.000429	0.001647	down
DIRAS2	-2.296683604	1.99E-05	9.58E-05	down
SCARA5	-2.31567635	0.000124	0.000525	down
H2AC6	-2.319641151	8.27E-84	8.23E-82	down
AC110619.1	-2.333105739	1.80E-14	2.18E-13	down
METTL7A	-2.339181361	3.19E-108	4.65E-106	down
LINC02747	-2.343365865	0.001055	0.003742	down
STAC2	-2.364586304	0	0	down
GPR12	-2.378481907	0.002135	0.007099	down
HSPD1P6	-2.383898149	0.0028	0.009085	down
LIPJ	-2.411178647	0.005474	0.016595	down
AC017002.4	-2.444195975	6.73E-12	6.72E-11	down
LINC00578	-2.449629816	4.97E-17	7.18E-16	down
FYB2	-2.450431117	4.49E-07	2.72E-06	down
SNCAIP	-2.509981822	1.66E-08	1.19E-07	down
SPINK8	-2.549418318	0.000244	0.000979	down
AL031668.2	-2.550759408	3.88E-06	2.07E-05	down
BBOX1	-2.553829394	5.95E-07	3.53E-06	down
KIAA1324	-2.56486416	2.76E-182	8.93E-180	down
PRODH	-2.570920196	3.74E-06	2.00E-05	down
PXMP4	-2.571460622	0.002053	0.006856	down
THRSP	-2.60630849	4.29E-05	0.000196	down

Gene name	log2FoldChange	*P*-value	P_{adj}	change
AP001636.3	-2.611809576	9.22E-26	2.08E-24	down
TRIL	-2.615561487	8.71E-48	4.13E-46	down
MT2A	-2.617445631	0	0	down
MAPK4	-2.626903737	0.004212	0.013153	down
H2BC5	-2.648156646	2.45E-36	8.07E-35	down
LGR6	-2.668626496	1.10E-19	1.86E-18	down
CALML5	-2.706027572	2.81E-307	3.03E-304	down
CALML3-AS1	-2.707937076	0.000114	0.000482	down
CYP4Z2P	-2.713415626	5.39E-13	5.83E-12	down
FMO3	-2.718533708	2.60E-06	1.42E-05	down
COL14A1	-2.75572346	2.53E-16	3.46E-15	down
LINC00892	-2.762424992	4.47E-10	3.77E-09	down
LINC02568	-2.771980838	0.004467	0.013853	down
SLC12A3	-2.817545501	2.74E-23	5.51E-22	down
SYN3	-2.849692599	1.46E-12	1.53E-11	down
TFAP2B	-2.872263583	0.000107	0.000458	down
AC015910.1	-2.877216668	2.23E-18	3.52E-17	down
MT1E	-2.904058677	1.94E-45	8.58E-44	down
DIRAS3	-2.914930697	7.24E-15	8.97E-14	down
SLC23A1	-2.916291408	3.07E-18	4.78E-17	down
FMO2	-2.935284444	0.000244	0.000978	down
FATE1	-2.952999375	0.000225	0.000908	down
CRYAB	-2.959655095	3.53E-39	1.27E-37	down
FMO6P	-3.001428738	1.95E-63	1.34E-61	down
AC011374.1	-3.007545323	0.000163	0.000676	down
LINC02240	-3.040648783	0.001768	0.005986	down
CLCA2	-3.052353912	9.48E-18	1.43E-16	down
CAV3	-3.054649264	0.00184	0.00621	down
ACSM3	-3.102774007	0.002639	0.008609	down
AL355538.2	-3.200450981	1.29E-08	9.32E-08	down
TACR1	-3.206519714	1.61E-06	9.06E-06	down
KRT79	-3.212389512	3.36E-07	2.06E-06	down
CREB3L1	-3.28876833	0.000205	0.000832	down
SOSTDC1	-3.294594514	5.44E-05	0.000244	down
C8orf31	-3.340839539	6.17E-48	2.95E-46	down
GABRP	-3.363062262	2.70E-05	0.000127	down
LINC02303	-3.372477464	2.56E-06	1.40E-05	down
TBATA	-3.415811983	2.69E-26	6.17E-25	down
AP001636.2	-3.50338006	3.86E-07	2.36E-06	down
AC053513.1	-3.54314975	0.001131	0.003985	down
SELENOP	-3.55731002	5.28E-226	2.85E-223	down
COLCA1	-3.558516715	4.00E-12	4.07E-11	down
SPARC	-3.573048344	1.11E-08	8.15E-08	down
DAPL1	-3.591654153	1.05E-07	6.82E-07	down
AC108063.2	-3.738635931	5.93E-07	3.53E-06	down
AC245297.1	-3.753880399	0.000967	0.003462	down

Gene name	log2FoldChange	*P*-value	P_{adj}	change
AQP5	-3.803032065	0.000215	0.000869	down
PAMR1	-3.953059084	8.47E-06	4.31E-05	down
LCP1	-3.957043816	7.34E-05	0.000323	down
CYP4B1	-4.047128506	0	0	down
EDAR	-4.110473593	4.57E-06	2.42E-05	down
AC124067.4	-4.124784237	3.22E-06	1.74E-05	down
LGALS12	-4.175418621	4.61E-10	3.89E-09	down
ASCL4	-4.239557706	0.0002	0.000813	down
CYP4Z1	-4.346341173	2.97E-90	3.21E-88	down
CRISP3	-4.379074645	1.30E-201	5.49E-199	down
PIP	-4.412494574	1.67E-05	8.09E-05	down
FAIM2	-4.472749138	3.30E-17	4.80E-16	down
LY6L	-4.525138061	0.000933	0.003356	down
FAM178B	-4.610819647	4.92E-05	0.000222	down
AC025154.2	-4.743509439	1.79E-11	1.72E-10	down
AC008870.1	-4.867989142	0.000377	0.001462	down
AL049536.1	-4.947367783	0.000275	0.001089	down
AL355538.1	-4.962498565	0.000413	0.001591	down
POU2AF1	-5.12743238	7.31E-18	1.11E-16	down
CNGA2	-5.544268112	1.10E-26	2.56E-25	down
RHAG	-5.550192486	8.50E-05	0.000369	down
PSAPL1	-5.908266464	1.00E-05	5.04E-05	down
H2AC13	-6.0366813	1.00E-05	5.05E-05	down
CHRM1	-6.731773564	3.42E-07	2.10E-06	down
CD36	-7.60484205	2.11E-08	1.49E-07	down

List of 987 differentially expressed genes (DEGs) ($P_{adj} < 0.05$ & |Log2FoldChange| > 1) identified by DESeq2 between the control group and the LSW treatment group. "Log2FoldChange" is the log-ratio of gene expression values in control and LSW treatment group. "*P*-value" is the result of Wald test for estimating which gene would be statistically significant. "P_{adj}" is the adjusted P-value by the Benjamini-Hochberg method. "change" is the direction of the differentially expressed gene. The data list is organized here from maximal to minimal Log2FoldChange.

Supplementary Table 4.2. CMap analysis in MDA-MB-468 cells upon treatment with LSW

Compound Id	Compound Name	Connectivity Score
BRD-A52650764	ingenol	98.17
BRD-K10916986	vinorelbine	96.37
BRD-K91145395	prostratin	95.14
BRD-K44432556	VU-0418946-1	95.04
BRD-A76528577	vincristine	93.55
BRD-K94325918	kinetin-riboside	92.86
BRD-A54927599	KF-38789	92.38
BRD-K26818574	BIX-01294	90.4
BRD-A19500257	geldanamycin	-91.09
BRD-K32330832	VER-155008	-91.67
BRD-A50454580	PD-0325901	-92.87
BRD-K05104363	PD-184352	-94.04
BRD-K41859756	NVP-AUY922	-94.06
BRD-K95785537	PP-2	-95.26
BRD-K49865102	PD-0325901	-95.74
BRD-K57080016	selumetinib	-96.59

"Id" indicates the Broad Institute compound identifier. "Name" indicates the vendor compound name. "Connectivity score" is the “median tau score” in the CMap platform which summarizes the connectivity between a Query signature and a drug across multiple cell types. The drugs in the table are sorted by decreasing order of connectivity score.

B.3 Supplementary Tables for chapter 5

Supplementary Table 5.1. Compounds in the library conform the filtering rules

ChemPub ID	Name	MW	LogP	nHA	nHD	TPSA	nRot
118981731	2-[[4-Hydroxy-2-oxo-1-[(2,4,5-trifluorophenyl)methyl]-5,7-dihydrofuro[3,4-b]pyridine-3-carbonyl]amino]acetic acid	398	0.222	8	3	117.86	6
118978236	2-(4-hydroxy-2-oxo-1-(m- tolyl)-1,2,5,7-tetrahydrofuro[3,4-b]pyridine- 3-carboxamido)acetic acid	344	-0.259	8	3	117.86	5
54741100	2-(2-(2-Fluorophenyl)-5-hydroxy-8-methyl-7-oxo-7,8-dihydropyrido[2,3-d]pyrimidine-6-carboxamido)acetic acid	372	0.845	9	3	134.41	5
71133989	2-[[1-[(4-Chlorophenyl)methyl]-4-hydroxy-2-oxo-5,7-dihydrofuro[3,4-b]pyridine-3-carbonyl]amino]acetic acid	379	1.066	8	3	117.86	6
117601276	6-(4-(4-cyano-2-methylphenyl)-5-hydroxy-1H-pyrazol-1-yl)-N-(cyclobutylmethyl)nicotinamide	387	3.445	7	2	103.57	6
118978254	2-(4-hydroxy-2-oxo-1-phenyl- 1,2,5,7-tetrahydrofuro[3,4-b] pyridine-3-carboxamido) acetic acid	330	0.08	8	3	117.86	5
118978243	2-(1-(4-chlorophenyl)-4- hydroxy-2-oxo-1,2,5,7- tetrahydrofuro[3,4-b]pyridine- 3-carboxamido)acetic acid	365	0.858	8	3	117.86	5
117601519	4-(5-hydroxy-1-(5-(3-methoxyazetidine-1-carbonyl)pyridin-2-yl)-1H-pyrazol-4-yl)-3-methylbenzonitrile	389	1.636	8	1	104.01	5
118978237	2-[[4-Hydroxy-2-oxo-1-[4-(trifluoromethyl)phenyl]-5,7-dihydrofuro[3,4-b]pyridine-3-carbonyl]amino]acetic acid	398	0.915	8	3	117.86	6
44218981	2-(4-Hydroxy-5-([(1-methyl-1-phenylethyl)amino]carbonyl)pyrimidin-2-yl)isonicotinic acid	378	2.053	8	3	125.04	6
118981498	2-[[1-[(2,4-Difluorophenyl)methyl]-4-hydroxy-2-oxo-5,7-dihydrofuro[3,4-b]pyridine-3-carbonyl]amino]acetic acid	380	0.17	8	3	117.86	6
121245745	2-[[1-[(4-chlorophenyl)methyl]-4-hydroxy-2-oxo-7,8-dihydro-5H-pyrano[4,3-b]pyridine-3-carbonyl]amino]acetic acid	393	1.292	8	3	117.86	6

ChemPub ID	Name	MW	LogP	nHA	nHD	TPSA	nRot
117600853	6-(4-(4-cyano-2-methylphenyl)-5-hydroxy-1H-pyrazol-1-yl)-N,N-dimethylnicotinamide	347	1.643	7	1	94.78	4
117601032	6-(4-(4-cyanophenyl)-5-hydroxy-1H-pyrazol-1-yl)-N-(3-methoxy-2-methylpropyl)nicotinamide	391	1.838	8	2	112.8	8
117600775	6-(4-(4-cyano-2-methylphenyl)-5-hydroxy-1H-pyrazol-1-yl)-N-ethylnicotinamide	347	2.162	7	2	103.57	5
117601165	6-(4-(4-cyano-2-methylphenyl)-5-hydroxy-1H-pyrazol-1-yl)-N-methylnicotinamide	333	1.851	7	2	103.57	4
44231076	N-benzhydryl-6-oxo-2-pyridazin-3-yl-1H-pyrimidine-5-carboxamide	383	1.953	7	2	100.63	6
44231440	6-oxo-2-pyrazol-1-yl-N-[[4-(trifluoromethoxy)phenyl]methyl]-1H-pyrimidine-5-carboxamide	379	2.328	8	2	101.9	7
54749369	2-(7-Hydroxy-4-methyl-5-oxo-4,5-dihydrothieno[3,2-b]pyridine-6-carboxamido)acetic Acid	282	0.088	7	3	108.63	4
44232167	1-[6-oxo-5-(2-phenylpropan-2-ylcarbamoyl)-1H-pyrimidin-2-yl]pyrazole-4-carboxylic acid	367	1.589	9	3	129.97	6
53493908	N-(6-oxo-2-pyrazol-1-yl-1H-pyrimidin-5-yl)-2-(4-phenylphenyl)acetamide	371	2.617	7	2	92.67	6
117601031	6-(4-(4-(cyanomethyl)phenyl)-5-hydroxy-1H-pyrazol-1-yl)-N-(3-methoxypropyl)nicotinamide	391	1.448	8	2	112.8	9
53493245	2-naphthalen-2-yl-N-(6-oxo-2-pyrazol-1-yl-1H-pyrimidin-5-yl)acetamide	345	2.07	7	2	92.67	5
44231442	N-benzhydryl-4-hydroxy-2-(1H-pyrazol-1-yl)pyrimidine-5-carboxamide	371	2.585	7	2	92.67	6
117601435	(S)-6-(4-(4-cyanophenyl)-5-hydroxy-1H-pyrazol-1-yl)-N-(4-methoxybutan-2-yl)nicotinamide	391	1.822	8	2	112.8	8
117601149	6-(4-(4-cyanophenyl)-5-hydroxy-1H-pyrazol-1-yl)-N-(3-methoxybutyl)nicotinamide	391	1.808	8	2	112.8	8
117601355	2-fluoro-4-(5-hydroxy-1-(5-(pyrrolidine-1-carbonyl)pyridin-2-yl)-1H-pyrazol-4-yl)-3-methylbenzonitrile	391	2.567	7	1	94.78	4
117600583	6-(5-hydroxy-4-(2-methoxypyridin-4-yl)-1H-pyrazol-1-yl)-N-(3-methoxybutyl)nicotinamide	397	1.667	9	2	111.13	9
44231439	6-oxo-N-[(4-phenylphenyl)methyl]-2-pyrazol-1-yl-1H-pyrimidine-5-carboxamide	371	2.74	7	2	92.67	6
57878392	N-[(1S)-1-(4-methoxyphenyl)ethyl]-6-oxo-2-pyrazol-1-yl-1H-pyrimidine-5-carboxamide	339	1.438	8	2	101.9	6
121246848	2-[[10-Hydroxy-12-oxo-2-(1,3-thiazol-5-yl)-4-oxa-1-	387	0.63	9	3	130.75	5

ChemPub ID	Name	MW	LogP	nHA	nHD	TPSA	nRot
	azatricyclo[7.3.1.05,13]trideca-5,7,9(13),10-tetraene-11-carbonyl]amino]acetic acid						
44231907	N-[(4-chlorophenyl)methyl]-6-oxo-2-pyrazol-1-yl-1H-pyrimidine-5-carboxamide	330	2.684	7	2	92.93	5
44231560	6-oxo-N-[(1S)-1-phenylethyl]-2-pyrazol-1-yl-1H-pyrimidine-5-carboxamide	309	1.319	7	2	92.67	5
44231792	N-[(1S)-1-(4-fluorophenyl)ethyl]-6-oxo-2-pyrazol-1-yl-1H-pyrimidine-5-carboxamide	327	1.397	7	2	92.67	5
44231673	6-oxo-N-[(1R)-1-phenylethyl]-2-pyrazol-1-yl-1H-pyrimidine-5-carboxamide	309	1.317	7	2	92.67	5
117601525	(R)-6-(4-(4-cyanophenyl)-5-hydroxy-1H-pyrazol-1-yl)-N-(4-methoxybutan-2-yl)nicotinamide	391	1.955	8	2	112.8	8
135895253	1-(4-oxo-6-o-tolyl-3,4-dihydro-quinazolin-2-yl)-1H-pyrazole-4-carboxylic acid	346	3.391	7	2	100.87	3
44230798	4-Hydroxy-N-(1-methyl-1-phenylethyl)-2-(1H-pyrazol-1-yl)pyrimidine-5-carboxamide	323	1.786	7	2	92.67	5
45359386	4-[2,4-dihydroxy-6-(2-oxopropyl)phenyl]-6,8-dihydroxy-7-methoxy-3-methyl-3,4-dihydro-1H-2-benzopyran-1-one	388	2.103	8	4	133.52	4
15816603	4-(4-hydroxybenzoyloxy)-3-(trimethylazaniumyl)butanoate	281	-2.252	6	1	86.66	7
60208858	5-carbamoylpentan-2-yl 4-chloro-3,5-dihydroxybenzoate	301	1.798	6	4	109.85	7
2748027	methyl 2,6-dibromo-3,5-dihydroxybenzoate	324	2.835	4	2	66.76	2
50742287	(phenylformamido)methyl 4-hydroxybenzoate	271	2.418	5	2	75.63	6
56677797	3-(3-hydroxybenzoyloxy)-8-methyl-8-azabicyclo[3.2.1]octane-2-carboxylic acid	305	1.452	6	2	87.07	4
375472	5-[(2,5-dihydroxyphenyl)methylidene]amino-2-hydroxybenzoic acid	273	2.982	6	4	110.35	3
162821769	Monocillin Vii	332	2.407	6	2	93.06	0
162822164	2,3-Dihydroxybenzoic Acid 3-O-Beta-D-Xyloside	286	-0.143	8	5	136.68	3
24066839	5-chloro-4,6,8-trihydroxy-7-methoxy-3-methyl-3,4-dihydro-1H-2-benzopyran-1-one	274	1.807	6	3	96.22	1
162823885	3-(8,9-dihydroxy-3,7,9-trimethyldeca-2,6-dien-1-yl)-4-hydroxybenzoic acid	348	3.67	5	4	97.99	8

ChemPub ID	Name	MW	LogP	nHA	nHD	TPSA	nRot
162839444	6,8-dihydroxy-3-[2-(4-hydroxyphenyl)ethyl]-3,4-dihydro-1H-2-benzopyran-1-one	300	3.267	5	3	86.99	3
162843848	Colletomellein A	294	3.179	5	3	86.99	4
162846882	4-hydroxy-3-methylbut-2-en-1-yl 2-hydroxybenzoate	222	2.704	4	2	66.76	5
25156984	Euparvic Acid	280	2.568	6	4	111.13	4
150898	2,4-dihydroxy-6-(2-oxopropyl)benzoic acid	210	1.001	5	3	94.83	3
162853734	Threo-L-3-[(2,4-Dihydroxy-6-Methylbenzoyl)Oxy]-2-Hydroxybutanoic Acid	270	0.702	7	4	124.29	5
139585373	Myxotrichin D	274	1.454	6	2	96.97	1
139587117	3,4,5-trihydroxypentan-2-yl 2,4-dihydroxy-6-methylbenzoate	286	0.226	7	5	127.45	6
162856902	(-)-(7S,8R)-8-Hydroxysydowic Acid	280	2.394	5	3	86.99	2
75614512	4-hydroxy-3-(7-hydroxy-3,7-dimethyl-4-oxooct-5-en-1-yl)-5-(4-hydroxy-3-methylbut-2-en-1-yl)benzoic acid	390	2.946	6	4	115.06	10
162858467	Rel-(2R,3R,4R)-2,3,4-Trihydroxypentyl 2,4-Dihydroxy-6-Methylbenzoate	286	0.333	7	5	127.45	6
68181280	Monaschromone	208	2.795	4	2	66.76	0
23727686	5-[(1-carboxyeth-1-en-1-yl)oxy]-2-hydroxybenzoic acid	224	1.62	6	3	104.06	4
101297713	3,6,8-trihydroxy-3,4,5,7-tetramethyl-3,4-dihydro-1H-2-benzopyran-1-one	252	2.34	5	3	86.99	0
125416272	2-(3-hydroxybenzoyloxy)prop-2-enoic acid	208	1.02	5	2	83.83	4
25131105	propan-2-yl 2,4-dihydroxy-6-methylbenzoate	210	2.797	4	2	66.76	3
78103806	(3,4,5-trihydroxyoxolan-3-yl)methyl 4-hydroxybenzoate	270	-0.006	7	4	116.45	4
91721009	propan-2-yl 3-formyl-2,4-dihydroxy-6-methylbenzoate	238	3.576	5	2	83.83	4
73228838	Sch725680	386	3.115	7	3	113.29	4
162867919	4-acetyl-6,8-dihydroxy-3-methoxy-5-methyl-3,4-dihydro-1H-2-benzopyran-1-one	266	1.368	6	2	93.06	2
73718904	Monocillin Iii	332	2.527	6	2	96.36	0
37269	Asperentin	292	2.937	5	2	75.99	2
60147886	methyl 2,4,6-trihydroxy-3-(3-methylbut-2-en-1-yl)benzoate	252	3.863	5	3	86.99	4
75228160	3R,5S-Sonnerlactone	266	2.445	5	3	86.99	0
129990444	2-(1,2-dihydroxypropyl)-6-hydroxybenzoic acid	212	-0.537	5	4	97.99	3
10987936	3,4,6,8-tetrahydroxy-3-methyl-3,4-dihydro-1H-2-benzopyran-1-one	226	0.475	6	4	107.22	0
162881997	2-benzyl-4,6-dihydroxybenzoic acid	244	3.176	4	3	77.76	3

ChemPub ID	Name	MW	LogP	nHA	nHD	TPSA	nRot
162883120	4-hydroxy-3,5-bis(3-hydroxy-3-methylbutyl)benzoic acid	310	2.442	5	4	97.99	7
578033	methyl 5-ethenyl-4-[2-(4-hydroxybenzoyloxy)ethyl]pyridine-3-carboxylate	327	2.874	6	1	85.72	8
85100035	4-Chloro-6-Hydroxymellein	228	2.51	4	2	66.76	0
135563785	Apah	210	-0.585	5	5	95.58	5
162886113	(-)-(7R,10S)-10-Hydroxysydowic Acid	280	2.19	5	3	86.99	2
73069410	Orsellide C	310	1.519	7	2	102.29	4
74030051	Monocillin I	330	2.382	6	2	96.36	0
162887908	4-hydroxy-2-methylbut-2-en-1-yl 4-hydroxybenzoate	222	2.05	4	2	66.76	5
72996259	Orsellide E	278	2.14	6	2	93.06	3
9920917	3-[(1-carboxyeth-1-en-1-yl)oxy]-4-hydroxybenzoic acid	224	0.988	6	3	104.06	4
14729105	3-(2,4-dihydroxy-6-methylbenzoyloxy)-2-hydroxy-2-methylbutanoic acid	284	1.583	7	4	124.29	5
78144561	Penicimarin A	224	1.162	5	3	86.99	1
162895392	5-(6,8-dihydroxy-5-methyl-1-oxo-3,4-dihydro-1H-2-benzopyran-3-yl)pentyl methyl carbonate	338	3.608	7	2	102.29	8
14598767	5'-Hydroxyasperentin	308	1.767	6	3	96.22	2
9794294	Acetophthalidin	208	1.798	5	3	90.9	1
162896226	Threo-L-2-[(2,4-Dihydroxy-6-Methylbenzoyl)Oxy]-3-Hydroxybutanoic Acid	270	0.898	7	4	124.29	5
4545936	Jaeskeanin	374	3.777	5	2	79.29	4
14394294	3-(3,4-dihydroxy-3-methylbut-1-yn-1-yl)-4-hydroxybenzoic acid	236	1.572	5	4	97.99	2
162898826	2-(2-hydroxyethyl)-3-(hydroxymethyl)-4-methylcyclopent-3-en-1-yl 4-hydroxybenzoate	292	1.906	5	3	86.99	6
132526850	1-Hydroxyboivinianic Acid	236	1.997	5	2	83.83	2
3320872	Pseudolaroside A	300	-0.706	8	5	136.68	4
162899860	(4S)-(+)-Ascochin	234	2.321	5	2	87.74	1
15726674	2-Formyl-3,5-Dihydroxy-4-Hydroxymethylbenzoic Acid	212	0.565	6	4	115.06	3
78138652	5,7-dihydroxy-3-(1-hydroxyethyl)-4-methoxy-1,3-dihydro-2-benzofuran-1-one	240	1.072	6	4	103.29	2
121433876	3-(4-hydroxy-3-methoxyphenyl)prop-2-en-1-yl 4-hydroxybenzoate	300	3.167	5	2	75.99	6
86136225	5,7-dihydroxy-4-(methoxymethyl)-6-methyl-1,3-dihydro-2-benzofuran-1-one	224	1.383	5	3	83.06	2
78088469	Hydroxypropan-2', 3'-Diol Orsellinate	242	0.706	6	4	107.22	5

ChemPub ID	Name	MW	LogP	nHA	nHD	TPSA	nRot
73016244	Pochonin A	366	3.388	6	2	96.36	0
102411352	4-hydroxy-3,5-bis(2-hydroxy-3-methylbut-3-en-1-yl)benzoic acid	306	2.148	5	4	97.99	7
162908811	6,8-dihydroxy-3-[(3-hydroxy-6-methyloxan-2-yl)methyl]-3,4-dihydro-1H-2-benzopyran-1-one	308	1.851	6	3	96.22	2
162909406	butan-2-yl 2,4-dihydroxy-6-(6-oxooct-1-en-1-yl)benzoate	334	3.75	5	2	83.83	10
24123447	6,8-dihydroxy-7-methoxy-3-methyl-4-[(6-methyl-2-oxo-2H-pyran-4-yl)oxy]-3,4-dihydro-1H-2-benzopyran-1-one	348	2.147	8	2	115.43	3
162912199	Soudanone E	308	3.558	5	3	86.99	6
44715330	4-acetyl-6,8-dihydroxy-5-methyl-3,4-dihydro-1H-2-benzopyran-1-one	236	1.61	5	2	83.83	1
74218984	5'-Hydroxymonocillin Iii	348	1.248	7	2	113.43	0
73240802	7,8,16,18-tetrahydroxy-12-methyl-3,13-dioxatricyclo[13.4.0.0^{2,4}]nonadeca-1(15),9,16,18-tetraen-14-one	350	1.64	7	4	119.75	0
56671199	Radicicol D	382	2.57	7	3	113.29	0
12315415	Sclerotinin B	238	2.197	5	3	86.99	0
441562	Dianthramine	289	3.189	7	5	127.09	4
11608479	Phomopsilactone	248	2.693	5	2	87.74	1
89726	1,2-dimethyl 4-hydroxybenzene-1,2-dicarboxylate	210	1.469	5	1	72.83	4
90887469	methyl 3,6-dihydroxy-2-[2-(2-hydroxyphenyl)ethynyl]benzoate	284	3.972	5	3	86.99	2
60028445	Pochonin O	384	2.366	7	3	116.59	0
162934290	Soudanone D	306	3.153	5	2	83.83	6
129861862	methyl 3-(2-hydroperoxy-3-methylbut-3-en-1-yl)-2,4,6-trihydroxybenzoate	284	2.194	7	4	116.45	6
3849318	2,3-dihydroxypropyl 2-hydroxybenzoate	212	1.212	5	3	86.99	5
25111594	4-Hydroxy-3-(2-Hydroxy-3-Methyl-3-Butenyl)-5-(3-Methyl-2-Butenyl)Benzoic +	290	3.675	4	3	77.76	6
162938665	10,12-dihydroxy-3-(4-hydroxypentyl)-3,4,5,6,7,8-hexahydro-1H-2-benzoxecin-1-one	322	3.624	5	3	86.99	4
85313102	Scorzocreticin	286	3.221	5	2	75.99	2
162946132	3-formyl-2a-hydroxy-6,7b-dimethyl-1H,2H,2aH,4aH,5H,6H,7H,7aH,7bH-cyclobuta[e]inden-2-yl 2,4-dihydroxy-6-methylbenzoate	386	3.597	6	3	104.06	4
14378899	6,9,12-trihydroxy-1,10,10-trimethyl-5-(propan-2-yl)-8-oxatricyclo[7.4.0.0^{2,7}]trideca-2(7),3,5-triene-3-carboxylic acid	350	3.339	6	4	107.22	2

ChemPub ID	Name	MW	LogP	nHA	nHD	TPSA	nRot
129831560	6,8-dihydroxy-3-methoxy-3-methyl-3,4-dihydro-1H-2-benzopyran-1-one	224	1.674	5	2	75.99	1
591722	methyl 2,4-dihydroxy-3,5,6-trimethylbenzoate	210	2.867	4	2	66.76	2
15720287	ethyl 2-ethyl-4,6-dihydroxybenzoate	210	2.965	4	2	66.76	4
162949882	5-Chloro-4,6-Dihydroxymellein	244	1.9	5	3	86.99	0
162950473	3,4,5-trihydroxyoxan-2-yl 2,5-dihydroxybenzoate	286	-0.365	8	5	136.68	3
12223651	1,2-dimethyl 3-hydroxybenzene-1,2-dicarboxylate	210	2.009	5	1	72.83	4
131831648	(3,4,5,6-tetrahydroxyoxan-2-yl)methyl 4-hydroxybenzoate	300	-0.746	8	5	136.68	4
162955766	methyl 4-hydroxy-2-[4-(pyridin-2-yl)buta-1,3-dien-1-yl]benzoate	281	2.878	4	1	59.42	5
53462156	methyl 3-(3-hydroxybenzoyloxy)-8-methyl-8-azabicyclo[3.2.1]octane-2-carboxylate	319	1.869	6	1	76.07	5
162958043	7-[(2,3-dihydroxy-2-methylbutanoyl)oxy]methyl-hexahydro-1H-pyrrolizin-1-yl 2-hydroxybenzoate	393	2.025	8	3	116.53	8
56681397	Ampelomin G	296	1.448	6	4	107.22	3
74193127	(3R,4S)-3,4-Dihydro-4,5,8-Trihydroxy-3-Methylisocoumarin	210	1.007	5	3	86.99	0
4014619	5-[amino(carboxy)methyl]-2-hydroxybenzoic acid	211	-1.145	6	5	120.85	3
15558543	Agrimonolide	314	3.645	5	2	75.99	4
90966548	2-(2,6-dihydroxy-4-methylbenzoyl)-6-hydroxybenzoic acid	288	3.97	6	4	115.06	3
76185781	Pochonin P	382	2.143	7	2	113.43	0
85100613	4-Bromo-6-Hydroxymellein	272	2.599	4	2	66.76	0
162975813	7-hydroxy-hexahydro-1H-pyrrolizin-1-yl 2-hydroxybenzoate	263	1.772	5	2	70	3
162976273	3-hydroxy-4-(2-hydroxy-3-methoxy-3-oxopropyl)benzoic acid	240	0.842	6	3	104.06	5
162979277	(7-hydroxy-hexahydro-1H-pyrrolizin-1-yl)methyl 2-hydroxybenzoate	277	2.341	5	2	70	4
191188	6,8-dihydroxy-3-[(4-hydroxy-6-methyloxan-2-yl)methyl]-3,4-dihydro-1H-2-benzopyran-1-one	308	1.724	6	3	96.22	2
162980233	Soudanone F	308	3.759	5	3	86.99	7
78138653	3,6,8-trihydroxy-5-methoxy-3-methyl-3,4-dihydro-1H-2-benzopyran-1-one	240	1.494	6	3	96.22	1
49871294	(+-)-Matteucen B	302	2.681	6	3	96.22	2
158899	2-benzamido-4-hydroxybenzoic acid	257	2.576	5	3	86.63	4
74956704	7-hydroxy-7-methyl-6-oxo-3-(prop-1-en-1-yl)-7,8-dihydro-6H-isochromen-8-yl 2,4-dihydroxy-6-methylbenzoate	384	3.291	7	3	113.29	4

ChemPub ID	Name	MW	LogP	nHA	nHD	TPSA	nRot
139586551	Stoloniferol A	252	2.614	5	2	75.99	1
9971390	2-(2,4-dihydroxybenzamido)-4-hydroxybenzoic acid	289	2.305	7	5	127.09	4
71436686	methyl 3-hydroxy-2,4-bis(methylamino)benzoate	210	1.875	5	3	70.59	4
162991744	6-hydroxy-2a,7a-dimethyl-5-(propan-2-yl)-decahydroazuleno[5,6-b]oxiren-5-yl 4-hydroxybenzoate	374	3.567	5	2	79.29	4
583884	5-(2-amino-2-carboxyethyl)-2-hydroxybenzoic acid	225	-1.166	6	5	120.85	4
162994033	4',6,8-trihydroxy-6'-methyl-1,4-dihydrospiro[2-benzopyran-3,2'-oxan]-1-one	280	1.493	6	3	96.22	0
13846908	Methyl 4-Hydroxy-3-(3'-Methyl-2'-Hydroxy-3'-Butenyl)Benzoate	236	2.259	4	2	66.76	5
13855157	Specionin	394	2.17	8	2	106.98	8
15559745	Illudoic Acid	278	2.733	5	3	86.99	1
101805479	4,6-Dihydroxy-3,9-Dehydromellein	208	0.963	5	3	90.9	0
75059807	(3R*,4S*)-6,8-Dihydroxy-3,4,7-Trimethylisocoumarin	222	2.872	4	2	66.76	0
261618	Cochlearine	261	2.33	4	1	49.77	3
5033	Radicicol	364	3.338	6	2	96.36	0
74015882	Cephalanone F	288	3.878	6	4	115.06	3
163002891	8-methyl-8-azabicyclo[3.2.1]oct-6-en-3-yl 4-hydroxybenzoate	259	2.208	4	1	49.77	3
163003173	Chaetophenol G	362	3.028	7	4	120.36	3
71308285	6,8-dihydroxy-3-(3-hydroxy-4-methoxyphenyl)-3,4-dihydro-1H-2-benzopyran-1-one	302	2.676	6	3	96.22	2
73001684	Orsellide A	326	1.212	8	3	122.52	4
11184361	Hostmaniene	254	1.456	5	3	86.99	5
72978900	Globosumone B	268	1.128	6	3	104.06	6
9799439	Monorden B	368	3.502	6	2	96.36	0
73718905	Pochonin B	382	2.143	7	2	113.43	0
14131421	3-ethyl-5,7-dihydroxy-3,6-dimethyl-1,3-dihydro-2-benzofuran-1-one	222	2.84	4	2	66.76	1
13963332	(3,5-dihydroxy-4-methylphenyl)methyl 3-formyl-2,4-dihydroxy-6-methylbenzoate	332	3.759	7	4	124.29	5
57509387	4,6,8-trihydroxy-3-methyl-3,4-dihydro-1H-2-benzopyran-1-one	210	0.845	5	3	86.99	0
14807796	6,8-dihydroxy-3,5-dimethyl-3,4-dihydro-1H-2-benzopyran-1-one	208	2.912	4	2	66.76	0
85395150	Kuhistanicaol H	390	3.177	6	3	99.52	4
102411350	ethyl 4-hydroxy-3-(2-hydroxy-3-methylbut-3-en-1-yl)benzoate	250	2.724	4	2	66.76	6

ChemPub ID	Name	MW	LogP	nHA	nHD	TPSA	nRot
163020419	15G256K	298	1.146	7	4	124.29	7
85221882	(R,S)-5,7-Dihydroxy-3-(1-Hydroxyethyl)Phthalide	210	0.896	5	4	94.06	1
10333412	6,8-dihydroxy-3-(4-hydroxyphenyl)-3,4-dihydro-1H-2-benzopyran-1-one	272	2.778	5	3	86.99	1
163022961	Soudanone G	322	3.66	6	2	93.06	7
163022969	4-(2-carboxy-2-hydroxyethyl)-3-hydroxybenzoic acid	226	0.354	6	4	115.06	4
21126437	3-Hydroxybenzoic Acid-3-O-Sulphate	218	-0.272	6	2	100.9	3
45359989	Montagnetol	272	-0.081	7	5	127.45	6
163030696	Exserolide D	294	2.714	6	2	85.22	3
163032099	Soudanone B	322	3.674	6	3	104.06	7
85187905	6,6a-bis(hydroxymethyl)-2H,3H,3aH,4H,6aH-cyclopenta[b]furan-4-yl 4-hydroxybenzoate	306	0.695	6	3	96.22	5
163036410	4-(hydroxymethyl)-5-oxo-7-oxabicyclo[4.1.0]hept-3-en-2-yl 2-hydroxy-6-methylbenzoate	290	1.408	6	2	96.36	4
75614550	4-hydroxy-3-(7-hydroxy-3,7-dimethyl-4-oxooct-5-en-1-yl)-5-(3-methylbut-2-en-1-yl)benzoic acid	374	3.978	5	3	94.83	9
2826719	Olivetolic Acid	224	3.616	4	3	77.76	5
75051822	Globosumone C	284	0.448	7	4	124.29	6
83668779	Ethyl 24-Dihydroxy-56-Dimethylbenzoate	210	2.998	4	2	66.76	3
91427654	3,5-dihydroxy-4-(5-methoxy-5-oxopent-1-en-1-yl)benzoic acid	266	2.297	6	3	104.06	6
163042897	5-Hydroxyl-6-O-Methylasperentin	322	2.993	6	2	85.22	3
18690341	4-Hydroxybenzoic Acid-4-O-Sulphate	218	-0.312	6	2	100.9	3
21187450	(3,4,5,6-Tetrahydroxytetrahydro-2H-Pyran-2-Yl)Methyl 2-Hydroxybenzoate	300	-0.377	8	5	136.68	4
163048040	Soudanone C	308	3.14	6	3	104.06	7
86221420	3,6,8-trihydroxy-3-methyl-3,4-dihydro-1H-2-benzopyran-1-one	210	1.245	5	3	86.99	0
879332	2-[3-(furan-2-yl)prop-2-enamido]-5-hydroxybenzoic acid	273	2.43	6	3	99.77	5
14190390	3,5,7-trihydroxy-4,6-dimethyl-1,3-dihydro-2-benzofuran-1-one	210	1.831	5	4	94.06	0
9840632	17,19-dihydroxy-4-methyl-3,7-dioxatricyclo[13.4.0.0^{6},^{8}]nonadeca-1(15),16,18-triene-2,13-dione	334	2.522	6	2	96.36	0
163050744	6,8-dihydroxy-3-(1-hydroxyethyl)-7-methyl-3,4-dihydro-1H-2-benzopyran-1-one	238	1.896	5	3	86.99	1
102516032	Rubramin	224	1.768	6	2	100.9	4
591773	Methyl Haematomate	210	2.81	5	2	83.83	3

ChemPub ID	Name	MW	LogP	nHA	nHD	TPSA	nRot
163052289	3,5,7-trihydroxy-6-methyl-4-(3-methylbut-2-en-1-yl)-1,3-dihydro-2-benzofuran-1-one	264	3.928	5	4	94.06	2
163052365	6,8-dihydroxy-3-(2-hydroxy-4-oxopentyl)-3,4-dihydro-1H-2-benzopyran-1-one	280	0.905	6	3	104.06	4
163054597	3-(2,4-dihydroxypentyl)-6,8-dihydroxy-3,4-dihydro-1H-2-benzopyran-1-one	282	1.011	6	4	107.22	4
3022325	2,3-dihydroxypropyl 4-hydroxybenzoate	212	0.671	5	3	86.99	5
163056432	10,13-dihydroxy-11,12-dimethoxy-4-propyl-3,7-dioxatricyclo[7.4.0.0^{2},^{6}]trideca-1(9),10,12-trien-8-one	324	2.719	7	2	94.45	4
89041760	4-3-[(dihydroxyamino)oxy]propoxy-2-hydroxybenzoic acid	259	1.07	8	4	119.69	7
11206471	3-chloro-4,6-dihydroxy-2-methylbenzoic acid	202	2.706	4	3	77.76	1
75966406	(R)-3,4-Dihydro-4,6,8-Trihydroxy-4,5-Dimethyl-3-Methyleneisochromen-1-One	236	1.546	5	3	86.99	0
23757117	Lignicol	240	0.821	6	3	96.22	1
45782752	methyl 4-[(4,6-dihydroxy-5-methoxy-2,5-dimethyl-3-oxocyclohex-1-en-1-yl)oxy]-2-hydroxy-3,6-dimethylbenzoate	380	2.491	8	2	119.36	5
614234	6,8-dihydroxy-7-methoxy-3-methyl-3,4-dihydro-1H-2-benzopyran-1-one	224	2.331	5	2	75.99	1
44715229	Roseopurpurin B	380	2.219	8	2	119.36	5
163063260	10'-Oxorelgro	320	3.281	5	2	83.83	4
163063310	Orsellide A1	340	1.775	8	3	122.52	5
123740797	2-(2,4-Dihydroxy-6-Methylbenzoyl)-Glycerol	242	0.211	6	4	107.22	5
44715229	methyl 2-[(2,5-dihydroxy-6-methoxy-3,6-dimethyl-4-oxocyclohex-2-en-1-yl)oxy]-4-hydroxy-3,6-dimethylbenzoate	380	2.219	8	2	119.36	5
163063667	Periplanetin D	224	1.182	5	3	86.99	1
163063934	(-)-(7R,10R)-Iso-10-Hydroxysydowic Acid	280	2.35	5	3	86.99	3
139585494	4,6,8-trihydroxy-3-methoxy-3,7-dimethyl-3,4-dihydro-1H-2-benzopyran-1-one	254	1.336	6	3	96.22	1
139587775	6,8-dihydroxy-3-methoxy-3,7-dimethyl-3,4-dihydro-1H-2-benzopyran-1,4-dione	252	1.547	6	2	93.06	1
163064201	7, 9-Dihydroxy-10-Methyl-2H, 4Ah, 6H, 10Bh-Pyrano[5,6-C][2]Ben-Zopyran-2, 6-Dione	262	1.519	6	2	96.97	0

ChemPub ID	Name	MW	LogP	nHA	nHD	TPSA	nRot
163064257	(1'S,5'S,6'S)-5',6'-Dihydroxy-2',6'-Dimethyl-3'-(2''-Oxopentyl)Cyclohex-2'-En-1'-Yl 2,4-Dihydroxy-6-Methylbenzoate	392	2.952	7	4	124.29	6
163064350	Dothideomynone A	252	2.007	5	3	86.99	1
139587759	Hypoxyphenone	210	2.226	5	2	83.83	3
163065739	Colletomellein B	294	3.079	5	3	86.99	4
163065763	4-Dehydro-Dihydromelleolide	388	3.373	6	4	107.22	4
163065906	Lycopodiellactone	344	2.792	7	3	121.11	2
163065969	Pyrenomycin	238	1.932	5	3	86.99	2
163069091	3-ethyl-6,8-dihydroxy-7-(hydroxymethyl)-3,4-dihydro-1H-2-benzopyran-1-one	238	1.886	5	3	86.99	2
163071099	Nigrospoxydon C	294	-0.09	7	3	121.13	4
78200715	Kuhistaferone	390	2.934	6	2	93.06	7
3940691	ethyl 3-formyl-2,4-dihydroxy-6-methylbenzoate	224	3.37	5	2	83.83	4
75311343	4,6,8-trihydroxy-3-phenyl-3,4-dihydro-1H-2-benzopyran-1-one	272	2.021	5	3	86.99	1
72961102	Globosumone A	250	2.409	5	2	83.83	5
73030948	Orsellide B	326	1.153	8	3	122.52	4
73154815	(5Z)-12-Chloro-13,15-Dihydroxy-4,7,8,9-Tetrahydro-2-Benzoxacyclotridecine-1,10(3H,11H)-Dione	324	3.741	5	2	83.83	0
11075034	Herbaric Acid	224	1.116	6	4	111.13	2
129990580	2,4-dihydroxy-6-(methoxycarbonyl)benzoic acid	212	1.364	6	3	104.06	3
85197254	2-(2-hydroxyethyl)-3,4-bis(hydroxymethyl)cyclopent-3-en-1-yl 4-hydroxybenzoate	308	0.525	6	4	107.22	7
71753478	5-[2-(2-carboxyphenyl)acetamido]-2-hydroxybenzoic acid	315	2.88	7	4	123.93	6
56667780	3-(4-hydroxybenzoyloxy)-8-methyl-8-azabicyclo[3.2.1]octane-2-carboxylic acid	305	1.414	6	2	87.07	4
56677454	Pestaphthalide B	224	1.023	5	4	94.06	1
91638211	5-[2-(4,5-dimethoxy-3-oxo-1,3-dihydro-2-benzofuran-1-yl)acetamido]-2-hydroxybenzoic acid	387	2.309	9	3	131.39	7
91638315	4-[2-(4,5-dimethoxy-3-oxo-1,3-dihydro-2-benzofuran-1-yl)acetamido]-2-hydroxybenzoic acid	387	2.327	9	3	131.39	7
84368	3-bromo-4-hydroxybenzoic acid	216	2.298	3	2	57.53	1
163098223	3-hydroxy-5-(4-phenoxybutoxy)benzoic acid	302	3.754	5	2	75.99	8
101304431	Decarboxydihydrocitrinone	222	2.868	4	2	66.76	0

ChemPub ID	Name	MW	LogP	nHA	nHD	TPSA	nRot
85370952	5-hydroxy-2-[3-(4-hydroxy-3,5-dimethoxyphenyl)prop-2-enamido]benzoic acid	359	2.318	8	4	125.32	7
76857	3,5-dibromo-4-hydroxybenzoic acid	294	2.846	3	2	57.53	1
11256664	Roxadustat	352	3.599	7	3	108.75	6
91617630	Daprodustat	393	1.653	9	3	130.63	6
23634441	Vadadustat	306	2.772	6	3	99.52	5
75593290	Desidustat	332	1.334	8	3	117.86	7

Bibliography

1. Koutroumpa, K., et al., *An expanded molecular phylogeny of Plumbaginaceae, with emphasis on Limonium (sea lavenders): Taxonomic implications and biogeographic considerations.* Ecol Evol, 2018. **8**(24): p. 12397-12424.

2. Lee, Y.G., et al., *Limonium tetragonum Promotes Running Endurance in Mice through Mitochondrial Biogenesis and Oxidative Fiber Formation.* Nutrients, 2022. **14**(19).

3. Hamadou, M.H., et al., *Apigenin rich-Limonium duriusculum (de Girard) Kuntze promotes apoptosis in HCT116 cancer cells.* Nat Prod Res, 2021. **35**(17): p. 2910-2914.

4. Blainski, A., et al., *Antibacterial activity of Limonium brasiliense (Baicuru) against multidrug-resistant bacteria using a statistical mixture design.* J Ethnopharmacol, 2017. **198**: p. 313-323.

5. Murray, A.P., et al., *Antioxidant metabolites from Limonium brasiliense (Boiss.) Kuntze.* Z Naturforsch C J Biosci, 2004. **59**(7-8): p. 477-80.

6. Yamashiro, S., et al., *Cardioprotective effects of extracts from Psidium guajava L and Limonium wrightii, Okinawan medicinal plants, against ischemia-reperfusion injury in perfused rat hearts.* Pharmacology, 2003. **67**(3): p. 128-35.

7. Dong, B., *Research on the conservation of Limonium sinense in the coast of Jiangsu.* Chin Wild Plant Res, 2005. **24**(06): p. 28-30.

8. Li, H., *Plumbaginaceae.* Flora of Taiwan, 1998. **4**: p. 79-82.

9. Chaung, S.S., et al., *The hepatoprotective effects of Limonium sinense against carbon tetrachloride and beta-D-galactosamine intoxication in rats.* Phytother Res, 2003. **17**(7): p. 784-91.

10. College, J.N.M., *Chinese Materia Medica Dictionary*. 1977, Shanghai Press of Science and Technology Shanghai.

11. Chang, Z., D. Jia, and J. Bare, *Chinese Materia Medica*. 2015: People's Medical Publishing House.

12. Department, X.A.L.D.H., *Chinese Herbal Medicine Handbook in Xinjiang*. 1970: Xinjiang People's Publishing House.

13. Ding, G., et al., *Analysis of genetic variability and population structure of the endemic medicinal Limonium sinense using molecular markers.* Gene, 2013. **520**(2): p. 189-93.

14. Nybom, H., *Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants.* Mol Ecol, 2004. **13**(5): p. 1143-55.

15. Avila-Carrasco, L., et al., *Natural Plants Compounds as Modulators of Epithelial-to-Mesenchymal Transition.* Front Pharmacol, 2019. **10**: p. 715.

16. Xiao ZF, Z.L., Cai XF, Zhang JB., *Effect of Limonium Sinense (Girard) O. Kuntze on Hemorrhagic Anaemia-An Analysis of the principal consituents of Limonium Sinense O. Kuntze.* Journal of Shantou University (Natural Science), 1991(01): p. 78-83.

17. Li, R., Z. Jia, and M.A. Trush, *Defining ROS in Biology and Medicine.* React Oxyg Species (Apex), 2016. **1**(1): p. 9-21.

18. Angelova, P.R. and A.Y. Abramov, *Functional role of mitochondrial reactive oxygen species in physiology.* Free Radic Biol Med, 2016. **100**: p. 81-85.

19. Huang, G., X. Mei, and J. Hu, *The Antioxidant Activities of Natural Polysaccharides.* Curr Drug Targets, 2017. **18**(11): p. 1296-1300.

20. Pizzino, G., et al., *Oxidative Stress: Harms and Benefits for Human Health.* Oxid Med Cell Longev, 2017. **2017**: p. 8416763.

21. Tang, X., et al., *Optimization of extraction process and investigation of antioxidant effect of polysaccharides from the root of Limonium sinense Kuntze.* Pharmacogn Mag, 2011. **7**(27): p. 186-92.

22. Dong MM, Y.F., Liu J, et al., *Comparison of antioxidant and anticancer activity among five kinds of flavonoids in Limonium sinense Kuntze.* Jiangsu Agricultural Sciences, 2015. **43**(02): p. 297-299.

23. Chen B, L.J., Guo Q, Chen Y, *Radical scavenging activity of polyphenol extracts from Limonium sinense.* Journal of Plant Resources and Environment, 2011. **20**(03): p. 36-42.

24. Kuo, Y.C., et al., *Samarangenin B from Limonium sinense suppresses herpes simplex virus type 1 replication in Vero cells by regulation of viral macromolecular synthesis.* Antimicrob Agents Chemother, 2002. **46**(9): p. 2854-64.

25. Hsu, W.C., et al., *Limonium sinense and gallic acid suppress hepatitis C virus infection by blocking early viral entry.* Antiviral Res, 2015. **118**: p. 139-47.

26. Tang, X.H., et al., *Mitochondrial modulation is involved in the hepatoprotection of Limonium sinense extract against liver damage in mice.* Journal of Ethnopharmacology, 2008. **120**(3).

27. Tang, X.H., et al., *Antitumor and immunomodulatory activity of polysaccharides from the root of Limonium sinense Kuntze.* Int J Biol Macromol, 2012. **51**(5): p. 1134-9.

28. Tang, X.H., et al., *Isolation and identification of anti-tumor polysaccharide LSP21 from Limonium sinense (Girard) Kuntze.* Int J Biol Macromol, 2014. **70**: p. 138-42.

29. Yan W, X.Y., Lin Y, Chen BH, *Study of Limonium sinense Polyphenols Inducing Apoptosis in HL-60 Human Leukemia Cells.* Journal of Fujian Normal University (Natural Science Edition), 2014. **30**(02): p. 119-124.

30. Tang, X.-H., et al., *Antitumor and immunomodulatory activity of polysaccharides from the root of Limonium sinense Kuntze.* International Journal of Biological Macromolecules, 2012. **51**(5).

31. Lin, L.C. and C.J. Chou, *Flavonoids and phenolics from Limonium sinense.* Planta Med, 2000. **66**(4): p. 382-3.

32. Liu, X., *Chemical constituents of Limonium sinense.* Chinese Traditional and Herbal Drugs., 2011. **42**(02): p. 230-233.

33. Lin, L.C., Y.C. Kuo, and C.J. Chou, *Anti-herpes simplex virus type-1 flavonoids and a new flavanone from the root of Limonium sinense.* Planta Med, 2000. **66**(4): p. 333-6.

34. Fan, Y., *Simultaneous determination of isoquercetin,morin,quercetin,luteolin and apigenin in Limonium sinense (Girard) Kuntze by RP-HPLC.* Chinese Journal of Pharmaceutical Analysis, 2014. **34**(04): p. 632-635.

35. Hongzhu, G. and Y. Jiurong, *Studies on the Chemical Constituents of Chinese Sealavender (Limonium sinense).* Chinese Traditional and Herbal Drugs, 1994.

36. Cassidy, A. and A.M. Minihane, *The role of metabolism (and the microbiome) in defining the clinical efficacy of dietary flavonoids.* Am J Clin Nutr, 2017. **105**(1): p. 10-22.

37. Wasser, S.P., *Medicinal mushrooms as a source of antitumor and immunomodulating polysaccharides.* Appl Microbiol Biotechnol, 2002. **60**(3): p. 258-74.

38. Zhang, R., et al., *Astragalus polysaccharides attenuate pulmonary fibrosis by inhibiting the epithelial-mesenchymal transition and NF-kappaB pathway activation.* Int J Mol Med, 2020. **46**(1): p. 331-339.

39. Saha, B.C. and F.M. Racine, *Biotechnological production of mannitol and its applications.* Appl Microbiol Biotechnol, 2011. **89**(4): p. 879-91.

40. Kim, H., J.Y. Seo, and K.H. Kim, *Effects of mannitol and dimethylthiourea on helicobacter pylori-induced IL-8 production in gastric epithelial cells.* Pharmacology, 1999. **59**(4): p. 201-11.

41. Schreibman, D.L., et al., *Mannitol and Hypertonic Saline Reduce Swelling and Modulate Inflammatory Markers in a Rat Model of Intracerebral Hemorrhage.* Neurocrit Care, 2018. **29**(2): p. 253-263.

42. Rane, R., et al., *Marine bromopyrrole alkaloids: synthesis and diverse medicinal applications.* Curr Top Med Chem, 2014. **14**(2): p. 253-73.

43. Park, J.B. and N. Schoene, *N-Caffeoyltyramine arrests growth of U937 and Jurkat cells by inhibiting protein tyrosine phosphorylation and inducing caspase-3.* Cancer Lett, 2003. **202**(2): p. 161-71.

44. Jiang, Y., L. Yu, and M.H. Wang, *N-trans-feruloyltyramine inhibits LPS-induced NO and PGE2 production in RAW 264.7 macrophages: Involvement of AP-1 and MAP kinase signalling pathways.* Chem Biol Interact, 2015. **235**: p. 56-62.

45. Tarnawski, M., et al., *HPLC determination of phenolic acids and antioxidant activity in concentrated peat extract--a natural immunomodulator.* J Pharm Biomed Anal, 2006. **41**(1): p. 182-8.

46. Fernandes, F.H. and H.R. Salgado, *Gallic Acid: Review of the Methods of Determination and Quantification.* Crit Rev Anal Chem, 2016. **46**(3): p. 257-65.

47. Liu, F., et al., *Ethyl gallate as a novel ERK1/2 inhibitor suppresses patient-derived esophageal tumor growth.* Mol Carcinog, 2019. **58**(4): p. 533-543.

48. Cui, H., et al., *Ethyl gallate suppresses proliferation and invasion in human breast cancer cells via Akt-NF-kappaB signaling.* Oncol Rep, 2015. **33**(3): p. 1284-90.

49. Obiang-Obounou, B.W. and G.H. Ryu, *The effect of feed moisture and temperature on tannin content, antioxidant and antimicrobial activities of extruded chestnuts.* Food Chem, 2013. **141**(4): p. 4166-70.

50. Cai, Y., et al., *Recent Advances in Anticancer Activities and Drug Delivery Systems of Tannins.* Med Res Rev, 2017. **37**(4): p. 665-701.

51. Cofan, M. and E. Ros, *Use of Plant Sterol and Stanol Fortified Foods in Clinical Practice.* Curr Med Chem, 2019. **26**(37): p. 6691-6703.

52. Shin, E.J., et al., *Anti-tumour effects of beta-sitosterol are mediated by AMPK/PTEN/HSP90 axis in AGS human gastric adenocarcinoma cells and xenograft mouse models.* Biochem Pharmacol, 2018. **152**: p. 60-70.

53. Rhourri-Frih, B., et al., *Investigation of porous graphitic carbon for triterpenoids and natural resinous materials analysis by high performance liquid chromatography hyphenated to mass spectrometry.* J Chromatogr A, 2012. **1240**: p. 140-6.

54. Yadav, V.R., et al., *Targeting inflammatory pathways by triterpenoids for prevention and treatment of cancer.* Toxins (Basel), 2010. **2**(10): p. 2428-66.

55. Jiang, Y., et al., *Dereplication-guided isolation of novel hepatoprotective triterpenoid saponins from Celosiae Semen by high-performance liquid chromatography coupled with electrospray ionization tandem quadrupole-time-of-flight mass spectrometry.* J Pharm Biomed Anal, 2017. **132**: p. 148-155.

56. Pollier, J. and A. Goossens, *Oleanolic acid.* Phytochemistry, 2012. **77**: p. 10-5.

57. Mlala, S., et al., *Ursolic Acid and Its Derivatives as Bioactive Agents.* Molecules, 2019. **24**(15).

58. Belchamber, K.B.R. and L.E. Donnelly, *Targeting defective pulmonary innate immunity - A new therapeutic option?* Pharmacol Ther, 2020. **209**: p. 107500.

59. Katoh, M., *Multi-layered prevention and treatment of chronic inflammation, organ fibrosis and cancer associated with canonical WNT/beta-catenin signaling activation (Review).* Int J Mol Med, 2018. **42**(2): p. 713-725.

60. Kinoshita, T. and T. Goto, *Molecular Mechanisms of Pulmonary Fibrogenesis and Its Progression to Lung Cancer: A Review.* Int J Mol Sci, 2019. **20**(6).

61. Betensley, A., R. Sharif, and D. Karamichos, *A Systematic Review of the Role of Dysfunctional Wound Healing in the Pathogenesis and Treatment of Idiopathic Pulmonary Fibrosis.* J Clin Med, 2016. **6**(1).

62. Fang, J., T. Seki, and H. Maeda, *Therapeutic strategies by modulating oxygen stress in cancer and inflammation.* Adv Drug Deliv Rev, 2009. **61**(4): p. 290-302.

63. Glasauer, A. and N.S. Chandel, *Targeting antioxidants for cancer therapy.* Biochem Pharmacol, 2014. **92**(1): p. 90-101.

64. Prasad, S., et al., *Oxidative Stress and Cancer: Advances and Challenges.* Oxid Med Cell Longev, 2016. **2016**: p. 5010423.

65. Giannoni, E., M. Parri, and P. Chiarugi, *EMT and oxidative stress: a bidirectional interplay affecting tumor malignancy.* Antioxid Redox Signal, 2012. **16**(11): p. 1248-63.

66. Wang, Y., et al., *Fenofibrate Improved Interstitial Fibrosis of Renal Allograft through Inhibited Epithelial-Mesenchymal Transition Induced by Oxidative Stress.* Oxid Med Cell Longev, 2019. **2019**: p. 8936856.

67. Amanzadeh, E., et al., *Application of quercetin in neurological disorders: from nutrition to nanomedicine.* Rev Neurosci, 2019. **30**(5): p. 555-572.

68. Gormaz, J.G., S. Quintremil, and R. Rodrigo, *Cardiovascular Disease: A Target for the Pharmacological Effects of Quercetin.* Curr Top Med Chem, 2015. **15**(17): p. 1735-42.

69. Justice, J.N., et al., *Senolytics in idiopathic pulmonary fibrosis: Results from a first-in-human, open-label, pilot study.* EBioMedicine, 2019. **40**: p. 554-563.

70. Lu, H., et al., *Quercetin ameliorates kidney injury and fibrosis by modulating M1/M2 macrophage polarization.* Biochem Pharmacol, 2018. **154**: p. 203-212.

71. Nabavi, S.F., et al., *Luteolin as an anti-inflammatory and neuroprotective agent: A brief review.* Brain Res Bull, 2015. **119**(Pt A): p. 1-11.

72. Kalbolandi, S.M., et al., *Luteolin confers renoprotection against ischemia-reperfusion injury via involving Nrf2 pathway and regulating miR320.* Mol Biol Rep, 2019. **46**(4): p. 4039-4047.

73. Kwon, E.Y. and M.S. Choi, *Luteolin Targets the Toll-Like Receptor Signaling Pathway in Prevention of Hepatic and Adipocyte Fibrosis and Insulin Resistance in Diet-Induced Obese Mice.* Nutrients, 2018. **10**(10).

74. Zhou, Y., et al., *Natural Polyphenols for Prevention and Treatment of Cancer.* Nutrients, 2016. **8**(8).

75. Sak, K., *Cytotoxicity of dietary flavonoids on different human cancer types.* Pharmacogn Rev, 2014. **8**(16): p. 122-46.

76. Serafini, M., I. Peluso, and A. Raguzzini, *Flavonoids as anti-inflammatory agents.* Proc Nutr Soc, 2010. **69**(3): p. 273-8.

77. Lin, Y., et al., *Luteolin, a flavonoid with potential for cancer prevention and therapy.* Curr Cancer Drug Targets, 2008. **8**(7): p. 634-46.

78. Putteeraj, M., et al., *Flavonoids and its Neuroprotective Effects on Brain Ischemia and Neurodegenerative Diseases.* Curr Drug Targets, 2018. **19**(14): p. 1710-1720.

79. Schepetkin, I.A. and M.T. Quinn, *Botanical polysaccharides: macrophage immunomodulation and therapeutic potential.* Int Immunopharmacol, 2006. **6**(3): p. 317-33.

80. Ng, Y.P., T.C. Or, and N.Y. Ip, *Plant alkaloids as drug leads for Alzheimer's disease.* Neurochem Int, 2015. **89**: p. 260-70.

81. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15545-50.

82. Mootha, V.K., et al., *PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes.* Nat Genet, 2003. **34**(3): p. 267-73.

83. Hanzelmann, S., R. Castelo, and J. Guinney, *GSVA: gene set variation analysis for microarray and RNA-seq data.* BMC Bioinformatics, 2013. **14**: p. 7.

84. Buffa, F.M., et al., *Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene.* Br J Cancer, 2010. **102**(2): p. 428-35.

85. Ye, Y., et al., *Characterization of Hypoxia-associated Molecular Features to Aid Hypoxia-Targeted Therapy.* Nat Metab, 2019. **1**(4): p. 431-444.

86. Jiashuo, W.U., et al., *Integration strategy of network pharmacology in Traditional Chinese Medicine: a narrative review.* J Tradit Chin Med, 2022. **42**(3): p. 479-486.

87. Hopkins, A.L., *Network pharmacology.* Nat Biotechnol, 2007. **25**(10): p. 1110-1.

88. Roth, B.L., D.J. Sheffler, and W.K. Kroeze, *Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia.* Nat Rev Drug Discov, 2004. **3**(4): p. 353-9.

89. Hopkins, A.L., J.S. Mason, and J.P. Overington, *Can we rationally design promiscuous drugs?* Curr Opin Struct Biol, 2006. **16**(1): p. 127-36.

90. Morphy, R. and Z. Rankovic, *Fragments, network biology and designing multiple ligands.* Drug Discov Today, 2007. **12**(3-4): p. 156-60.

91. Paolini, G.V., et al., *Global mapping of pharmacological space.* Nat Biotechnol, 2006. **24**(7): p. 805-15.

92. Keiser, M.J., et al., *Relating protein pharmacology by ligand chemistry.* Nat Biotechnol, 2007. **25**(2): p. 197-206.

93. Zhao, J., et al., *Investigation of the therapy targets of Yi-Qi-Yang-Yin-Hua-Tan-Qu-Yu recipe on type 2 diabetes by serum proteome labeled with iTRAQ.* J Ethnopharmacol, 2018. **224**: p. 1-14.

94. Liu, C.M., et al., *The Chinese herbal formula Zhibai Dihuang Granule treat Yin-deficiency-heat syndrome rats by regulating the immune responses.* J Ethnopharmacol, 2018. **225**: p. 271-278.

95. Kola, I. and J. Landis, *Can the pharmaceutical industry reduce attrition rates?* Nat Rev Drug Discov, 2004. **3**(8): p. 711-5.

96. Guan, L., et al., *ADMET-score - a comprehensive scoring function for evaluation of chemical drug-likeness.* Medchemcomm, 2019. **10**(1): p. 148-157.

97. Wang, Y., et al., *In silico ADME/T modelling for rational drug design.* Q Rev Biophys, 2015. **48**(4): p. 488-515.

98. Fan, J. and I.A. de Lannoy, *Pharmacokinetics.* Biochem Pharmacol, 2014. **87**(1): p. 93-120.

99. Gaohua, L., X. Miao, and L. Dou, *Crosstalk of physiological pH and chemical pKa under the umbrella of physiologically based pharmacokinetic modeling of drug absorption, distribution, metabolism, excretion, and toxicity.* Expert Opin Drug Metab Toxicol, 2021. **17**(9): p. 1103-1124.

100. Hogben, C.A., et al., *Absorption of drugs from the stomach. II. The human.* J Pharmacol Exp Ther, 1957. **120**(4): p. 540-5.

101. Schanker, L.S., et al., *Absorption of drugs from the stomach. I. The rat.* J Pharmacol Exp Ther, 1957. **120**(4): p. 528-39.

102. van den Anker, J., et al., *Developmental Changes in Pharmacokinetics and Pharmacodynamics.* J Clin Pharmacol, 2018. **58 Suppl 10**: p. S10-S25.

103. Gardiner, S.J. and E.J. Begg, *Pharmacogenetics, drug-metabolizing enzymes, and clinical practice.* Pharmacol Rev, 2006. **58**(3): p. 521-90.

104. Hu, B., et al., *Structure-Property Relationships and Machine Learning Models for Addressing CYP3A4-Mediated Victim Drug-Drug Interaction Risk in Drug Discovery.* Mol Pharm, 2020. **17**(9): p. 3600-3608.

105. Daly, A.K., et al., *Pharmacogenomics of CYP2C9: Functional and Clinical Considerations.* J Pers Med, 2017. **8**(1).

106. Kusuhara, H. and Y. Sugiyama, *In vitro-in vivo extrapolation of transporter-mediated clearance in the liver and kidney.* Drug Metab Pharmacokinet, 2009. **24**(1): p. 37-52.

107. Feng, B., et al., *Renal clearance in drug discovery and development: molecular descriptors, drug transporters and disease state.* Expert Opin Drug Metab Toxicol, 2010. **6**(8): p. 939-52.

108. Hodgson, J., *ADMET--turning chemicals into drugs.* Nat Biotechnol, 2001. **19**(8): p. 722-6.

109. Lipinski, C.A., et al., *Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings.* Adv Drug Deliv Rev, 2001. **46**(1-3): p. 3-26.

110. Ghose, A.K., V.N. Viswanadhan, and J.J. Wendoloski, *A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases.* J Comb Chem, 1999. **1**(1): p. 55-68.

111. Oprea, T.I., *Property distribution of drug-related chemical databases.* J Comput Aided Mol Des, 2000. **14**(3): p. 251-64.

112. Muegge, I., S.L. Heald, and D. Brittelli, *Simple selection criteria for drug-like chemical matter.* J Med Chem, 2001. **44**(12): p. 1841-6.

113. Walters, W.P. and M. Namchuk, *Designing screens: how to make your hits a hit.* Nat Rev Drug Discov, 2003. **2**(4): p. 259-66.

114. Bhal, S.K., et al., *The Rule of Five revisited: applying log D in place of log P in drug-likeness filters.* Mol Pharm, 2007. **4**(4): p. 556-60.

115. Veber, D.F., et al., *Molecular properties that influence the oral bioavailability of drug candidates.* J Med Chem, 2002. **45**(12): p. 2615-23.

116. Kuntz, I.D., et al., *A geometric approach to macromolecule-ligand interactions.* J Mol Biol, 1982. **161**(2): p. 269-88.

117. Brendan J. McConkey, V.S., Marvin Edelman, *The performance of current methods in ligand–protein docking.* Current Science, 2002. **83**: p. 845-856.

118. Chen, Y.C., *Beware of docking!* Trends Pharmacol Sci, 2015. **36**(2): p. 78-95.

119. Garrett M. Morris, D.S.G., Robert S. Halliday, Ruth Huey, William E. Hart, Richard K. Belew, Arthur J. Olson, *Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function.* Journal of computational chemistry, 1998. **19**(14): p. 1639-1662.

120. Trott, O. and A.J. Olson, *AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading.* J Comput Chem, 2010. **31**(2): p. 455-61.

121. Jones, G., et al., *Development and validation of a genetic algorithm for flexible docking.* J Mol Biol, 1997. **267**(3): p. 727-48.

122. Verdonk, M.L., et al., *Improved protein-ligand docking using GOLD.* Proteins, 2003. **52**(4): p. 609-23.

123. Rarey, M., et al., *A fast flexible docking method using an incremental construction algorithm.* J Mol Biol, 1996. **261**(3): p. 470-89.

124. Friesner, R.A., et al., *Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy.* J Med Chem, 2004. **47**(7): p. 1739-49.

125. Willett, A.T.B.a.P., *Algorithms for the identification of three-dimensional maximal common substructures.* Chemical Structures, 1987. **27**(4): p. 152-158.

126. Fischer, D., et al., *Surface motifs by a computer vision technique: searches, detection, and implications for protein-ligand recognition.* Proteins, 1993. **16**(3): p. 278-92.

127. Norel, R., et al., *Molecular surface recognition by a computer vision-based technique.* Protein Eng, 1994. **7**(1): p. 39-46.

128. Oshiro, C.M., I.D. Kuntz, and J.S. Dixon, *Flexible ligand docking using a genetic algorithm.* J Comput Aided Mol Des, 1995. **9**(2): p. 113-30.

129. Ru, J., et al., *TCMSP: a database of systems pharmacology for drug discovery from herbal medicines.* J Cheminform, 2014. **6**: p. 13.

130. Xu, H.Y., et al., *ETCM: an encyclopaedia of traditional Chinese medicine.* Nucleic Acids Res, 2019. **47**(D1): p. D976-D982.

131. Missiuro, P.V., et al., *Information flow analysis of interactome networks.* PLoS Comput Biol, 2009. **5**(4): p. e1000350.

132. Barabasi, A.L. and Z.N. Oltvai, *Network biology: understanding the cell's functional organization.* Nat Rev Genet, 2004. **5**(2): p. 101-13.

133. Raman, K., N. Damaraju, and G.K. Joshi, *The organisational structure of protein networks: revisiting the centrality-lethality hypothesis.* Syst Synth Biol, 2014. **8**(1): p. 73-81.

134. Brandes, U., *A Faster Algorithm for Betweenness Centrality.* Journal of Mathematical Sociology, 2001. **25**: p. 163-177.

135. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nat Protoc, 2009. **4**(1): p. 44-57.

136. Newman, M.E.J., *A measure of betweenness centrality based on random walks.* Social Networks, 2003. **27**(1): p. 39-54.

137. Wan, Y., et al., *Utilising network pharmacology to explore the underlying mechanism of Wumei Pill in treating pancreatic neoplasms.* BMC Complement Altern Med, 2019. **19**(1): p. 158.

138. Ashburner, M., et al., *Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.* Nat Genet, 2000. **25**(1): p. 25-9.

139. Khatri, P. and S. Draghici, *Ontological analysis of gene expression data: current tools, limitations, and open problems.* Bioinformatics, 2005. **21**(18): p. 3587-95.

140. Huang, D.W., et al., *The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists.* Genome Biol, 2007. **8**(9): p. R183.

141. Pinero, J., et al., *The DisGeNET knowledge platform for disease genomics: 2019 update.* Nucleic Acids Res, 2020. **48**(D1): p. D845-D855.

142. Chin, C.H., et al., *cytoHubba: identifying hub objects and sub-networks from complex interactome.* BMC Syst Biol, 2014. **8 Suppl 4**(Suppl 4): p. S11.

143. Wang, S., et al., *UCSCXenaShiny: an R/CRAN package for interactive analysis of UCSC Xena data.* Bioinformatics, 2022. **38**(2): p. 527-529.

144. Lecomte, S., et al., *Deciphering the Molecular Mechanisms Sustaining the Estrogenic Activity of the Two Major Dietary Compounds Zearalenone and Apigenin in ER-Positive Breast Cancer Cell Lines.* Nutrients, 2019. **11**(2).

145. Bauer, D., E. Mazzio, and K.F.A. Soliman, *Whole Transcriptomic Analysis of Apigenin on TNFalpha Immuno-activated MDA-MB-231 Breast Cancer Cells.* Cancer Genomics Proteomics, 2019. **16**(6): p. 421-431.

146. Zhou, Y., et al., *Apigenin in cancer therapy: From mechanism of action to nano-therapeutic agent.* Food Chem Toxicol, 2022. **168**: p. 113385.

147. Motallebi, M., et al., *Naringenin: A potential flavonoid phytochemical for cancer therapy.* Life Sci, 2022. **305**: p. 120752.

148. Wang, X., et al., *The mechanism of anticancer action and potential clinical use of kaempferol in the treatment of breast cancer.* Biomed Pharmacother, 2019. **117**: p. 109086.

149. Tang, S.M., et al., *Pharmacological basis and new insights of quercetin action in respect to its anti-cancer effects.* Biomed Pharmacother, 2020. **121**: p. 109604.

150. Madunic, J., et al., *Apigenin: A dietary flavonoid with diverse anticancer properties.* Cancer Lett, 2018. **413**: p. 11-22.

151. Yang, C.S., et al., *Inhibition of carcinogenesis by dietary polyphenolic compounds.* Annu Rev Nutr, 2001. **21**: p. 381-406.

152. Wang, Y.C. and K.M. Huang, *In vitro anti-inflammatory effect of apigenin in the Helicobacter pylori-infected gastric adenocarcinoma cells.* Food Chem Toxicol, 2013. **53**: p. 376-83.

153. Ozcelik, B., M. Kartal, and I. Orhan, *Cytotoxicity, antiviral and antimicrobial activities of alkaloids, flavonoids, and phenolic acids.* Pharm Biol, 2011. **49**(4): p. 396-402.

154. Xu, M., et al., *Apigenin suppresses colorectal cancer cell proliferation, migration and invasion via inhibition of the Wnt/beta-catenin signaling pathway.* Oncol Lett, 2016. **11**(5): p. 3075-3080.

155. Huang, C., et al., *Chrysin, Abundant in Morinda citrifolia Fruit Water-EtOAc Extracts, Combined with Apigenin Synergistically Induced Apoptosis and Inhibited Migration in Human Breast and Liver Cancer Cells.* J Agric Food Chem, 2016. **64**(21): p. 4235-45.

156. Lee, Y.M., et al., *Inhibition of glutamine utilization sensitizes lung cancer cells to apigenin-induced apoptosis resulting from metabolic and oxidative stress.* Int J Oncol, 2016. **48**(1): p. 399-408.

157. Zhao, G., et al., *Apigenin inhibits proliferation and invasion, and induces apoptosis and cell cycle arrest in human melanoma cells.* Oncol Rep, 2017. **37**(4): p. 2277-2285.

158. Gupta, S., F. Afaq, and H. Mukhtar, *Involvement of nuclear factor-kappa B, Bax and Bcl-2 in induction of cell cycle arrest and apoptosis by apigenin in human prostate carcinoma cells.* Oncogene, 2002. **21**(23): p. 3727-38.

159. Angulo, P., et al., *Natural compounds targeting major cell signaling pathways: a novel paradigm for osteosarcoma therapy.* J Hematol Oncol, 2017. **10**(1): p. 10.

160. Yan, X., et al., *Apigenin in cancer therapy: anti-cancer effects and mechanisms of action.* Cell Biosci, 2017. **7**: p. 50.

161. Cardenas, H., et al., *Dietary Apigenin Exerts Immune-Regulatory Activity in Vivo by Reducing NF-kappaB Activity, Halting Leukocyte Infiltration and Restoring Normal Metabolic Function.* Int J Mol Sci, 2016. **17**(3): p. 323.

162. Erdogan, S., et al., *The flavonoid apigenin reduces prostate cancer CD44(+) stem cell survival and migration through PI3K/Akt/NF-kappaB signaling.* Life Sci, 2016. **162**: p. 77-86.

163. Cheong, J.W., et al., *Inhibition of CK2alpha and PI3K/Akt synergistically induces apoptosis of CD34+CD38- leukaemia cells while sparing haematopoietic stem cells.* Anticancer Res, 2010. **30**(11): p. 4625-34.

164. Bao, Y.Y., et al., *Inhibiting GLUT-1 expression and PI3K/Akt signaling using apigenin improves the radiosensitivity of laryngeal carcinoma in vivo.* Oncol Rep, 2015. **34**(4): p. 1805-14.

165. Shukla, S., et al., *Apigenin attenuates insulin-like growth factor-I signaling in an autochthonous mouse prostate cancer model.* Pharm Res, 2012. **29**(6): p. 1506-17.

166. Suh, Y.A., et al., *Inhibition of IL-6/STAT3 axis and targeting Axl and Tyro3 receptor tyrosine kinases by apigenin circumvent taxol resistance in ovarian cancer cells.* Int J Oncol, 2015. **46**(3): p. 1405-11.

167. Imran, M., et al., *Luteolin, a flavonoid, as an anticancer agent: A review.* Biomed Pharmacother, 2019. **112**: p. 108612.

168. Birt, D.F., S. Hendrich, and W. Wang, *Dietary agents in cancer prevention: flavonoids and isoflavonoids.* Pharmacol Ther, 2001. **90**(2-3): p. 157-77.

169. Martin, K.R., *Targeting apoptosis with dietary bioactive agents.* Exp Biol Med (Maywood), 2006. **231**(2): p. 117-29.

170. Gao, G., et al., *Luteolin exhibits anti-breast cancer property through up-regulating miR-203.* Artif Cells Nanomed Biotechnol, 2019. **47**(1): p. 3265-3271.

171. Talukdar, S., et al., *EGFR: An essential receptor tyrosine kinase-regulator of cancer stem cells.* Adv Cancer Res, 2020. **147**: p. 161-188.

172. Hynes, N.E. and H.A. Lane, *ERBB receptors and cancer: the complexity of targeted inhibitors.* Nat Rev Cancer, 2005. **5**(5): p. 341-54.

173. Scaltriti, M. and J. Baselga, *The epidermal growth factor receptor pathway: a model for targeted therapy.* Clin Cancer Res, 2006. **12**(18): p. 5268-72.

174. Yarden, Y. and M.X. Sliwkowski, *Untangling the ErbB signalling network.* Nat Rev Mol Cell Biol, 2001. **2**(2): p. 127-37.

175. Butti, R., et al., *Receptor tyrosine kinases (RTKs) in breast cancer: signaling, therapeutic implications and challenges.* Mol Cancer, 2018. **17**(1): p. 34.

176. Macdonald-Obermann, J.L. and L.J. Pike, *Different epidermal growth factor (EGF) receptor ligands show distinct kinetics and biased or partial agonism for homodimer and heterodimer formation.* J Biol Chem, 2014. **289**(38): p. 26178-26188.

177. Wilson, K.J., et al., *Functional selectivity of EGF family peptide growth factors: implications for cancer.* Pharmacol Ther, 2009. **122**(1): p. 1-8.

178. Evangelopoulos, M.E., J. Weis, and A. Kruttgen, *Signalling pathways leading to neuroblastoma differentiation after serum withdrawal: HDL blocks neuroblastoma differentiation by inhibition of EGFR.* Oncogene, 2005. **24**(20): p. 3309-18.

179. Lee, H.J., et al., *Prognostic and predictive values of EGFR overexpression and EGFR copy number alteration in HER2-positive breast cancer.* Br J Cancer, 2015. **112**(1): p. 103-11.

180. Liao, B.C., et al., *Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors for Non-Small-Cell Lung Cancer Patients with Leptomeningeal Carcinomatosis.* J Thorac Oncol, 2015. **10**(12): p. 1754-61.

181. Nicholson, R.I., J.M. Gee, and M.E. Harper, *EGFR and cancer prognosis.* Eur J Cancer, 2001. **37 Suppl 4**: p. S9-15.

182. Park, H.S., et al., *High EGFR gene copy number predicts poor outcome in triple-negative breast cancer.* Mod Pathol, 2014. **27**(9): p. 1212-22.

183. Zhu, H., et al., *Oncogenic EGFR signaling cooperates with loss of tumor suppressor gene functions in gliomagenesis.* Proc Natl Acad Sci U S A, 2009. **106**(8): p. 2712-6.

184. Brand, T.M., et al., *Nuclear EGFR as a molecular target in cancer.* Radiother Oncol, 2013. **108**(3): p. 370-7.

185. Morgillo, F., et al., *Mechanisms of resistance to EGFR-targeted drugs: lung cancer.* ESMO Open, 2016. **1**(3): p. e000060.

186. Elayoubi, J., et al., *A Review of Endocrine Therapy in Early-stage Breast Cancer: The Journey From Crudeness to Precision.* Am J Clin Oncol, 2023. **46**(5): p. 225-230.

187. Malagrino, M. and G. Zavatta, *Review of bone health in women with estrogen receptor-positive breast cancer receiving endocrine therapy.* Womens Health (Lond), 2023. **19**: p. 17455057221149493.

188. Riggs, B.L. and L.C. Hartmann, *Selective estrogen-receptor modulators -- mechanisms of action and application to clinical practice.* N Engl J Med, 2003. **348**(7): p. 618-29.

189. Cuzick, J., et al., *Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 10-year analysis of the ATAC trial.* Lancet Oncol, 2010. **11**(12): p. 1135-41.

190. Mouridsen, H., et al., *Phase III study of letrozole versus tamoxifen as first-line therapy of advanced breast cancer in postmenopausal women: analysis of survival and update of efficacy from the International Letrozole Breast Cancer Group.* J Clin Oncol, 2003. **21**(11): p. 2101-9.

191. Coombes, R.C., et al., *A randomized trial of exemestane after two to three years of tamoxifen therapy in postmenopausal women with primary breast cancer.* N Engl J Med, 2004. **350**(11): p. 1081-92.

192. Slamon, D.J., et al., *Phase III Randomized Study of Ribociclib and Fulvestrant in Hormone Receptor-Positive, Human Epidermal Growth Factor Receptor 2-Negative Advanced Breast Cancer: MONALEESA-3.* J Clin Oncol, 2018. **36**(24): p. 2465-2472.

193. Giuliano, M., et al., *Biological mechanisms and clinical implications of endocrine resistance in breast cancer.* Breast, 2011. **20 Suppl 3**: p. S42-9.

194. Schofield, C.J. and P.J. Ratcliffe, *Oxygen sensing by HIF hydroxylases.* Nat Rev Mol Cell Biol, 2004. **5**(5): p. 343-54.

195. Infantino, V., et al., *Cancer Cell Metabolism in Hypoxia: Role of HIF-1 as Key Regulator and Therapeutic Target.* Int J Mol Sci, 2021. **22**(11).

196. Palazon, A., et al., *An HIF-1alpha/VEGF-A Axis in Cytotoxic T Cells Regulates Tumor Progression.* Cancer Cell, 2017. **32**(5): p. 669-683 e5.

197. Pardo-Cabello, A.J., V. Manzano-Gamero, and E. Puche-Canas, *Vitamin B12: For more than just the treatment of megaloblastic anemia?* Rev Clin Esp (Barc), 2023. **223**(2): p. 114-119.

198. Green, R., *Vitamin B(12) deficiency from the perspective of a practicing hematologist.* Blood, 2017. **129**(19): p. 2603-2611.

199. Gonzalez-Montana, J.R., et al., *Relationship between Vitamin B12 and Cobalt Metabolism in Domestic Ruminant: An Update.* Animals (Basel), 2020. **10**(10).

200. Taddei, M.L., et al., *Role of tyrosine phosphorylation in modulating cancer cell metabolism.* Biochim Biophys Acta Rev Cancer, 2020. **1874**(2): p. 188442.

201. Hayashi, A., et al., *Role of protein tyrosine phosphorylation in rat corneal neovascularization.* Graefes Arch Clin Exp Ophthalmol, 1997. **235**(7): p. 460-7.

202. Anand, S., A. Burkenroad, and J. Glaspy, *Workup of anemia in cancer.* Clin Adv Hematol Oncol, 2020. **18**(10): p. 640-646.

203. Dicato, M., L. Plawny, and M. Diederich, *Anemia in cancer.* Ann Oncol, 2010. **21 Suppl 7**: p. vii167-72.

204. Kibble, M., et al., *Network pharmacology applications to map the unexplored target space and therapeutic potential of natural products.* Nat Prod Rep, 2015. **32**(8): p. 1249-66.

205. Noor, F., et al., *Network Pharmacology Approach for Medicinal Plants: Review and Assessment.* Pharmaceuticals (Basel), 2022. **15**(5).

206. Eden, E., et al., *GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists.* BMC Bioinformatics, 2009. **10**: p. 48.

207. Kanehisa, M. and S. Goto, *KEGG: kyoto encyclopedia of genes and genomes.* Nucleic Acids Res, 2000. **28**(1): p. 27-30.

208. Rosenfeld, R., S. Vajda, and C. DeLisi, *Flexible docking and design.* Annu Rev Biophys Biomol Struct, 1995. **24**: p. 677-700.

209. Raval, A., et al., *Refinement of protein structure homology models via long, all-atom molecular dynamics simulations.* Proteins, 2012. **80**(8): p. 2071-9.

210. Sharma, G.N., et al., *Various types and management of breast cancer: an overview.* J Adv Pharm Technol Res, 2010. **1**(2): p. 109-26.

211. Akram, M., et al., *Awareness and current knowledge of breast cancer.* Biol Res, 2017. **50**(1): p. 33.

212. Nolan, E., G.J. Lindeman, and J.E. Visvader, *Deciphering breast cancer: from biology to the clinic.* Cell, 2023. **186**(8): p. 1708-1728.

213. Tan, P.H., et al., *The 2019 World Health Organization classification of tumours of the breast.* Histopathology, 2020. **77**(2): p. 181-185.

214. Weigelt, B. and J.S. Reis-Filho, *Histological and molecular types of breast cancer: is there a unifying taxonomy?* Nat Rev Clin Oncol, 2009. **6**(12): p. 718-30.

215. Tarantino, P., et al., *HER2-Low Breast Cancer: Pathological and Clinical Landscape.* J Clin Oncol, 2020. **38**(17): p. 1951-1962.

216. Sasaki, Y. and H. Tsuda, *Clinicopathological characteristics of triple-negative breast cancers.* Breast Cancer, 2009. **16**(4): p. 254-9.

217. Dent, R., et al., *Pattern of metastatic spread in triple-negative breast cancer.* Breast Cancer Res Treat, 2009. **115**(2): p. 423-8.

218. Sher, G., et al., *Epigenetic and breast cancer therapy: Promising diagnostic and therapeutic applications.* Semin Cancer Biol, 2022. **83**: p. 152-165.

219. Higgins, M.J. and V. Stearns, *Understanding resistance to tamoxifen in hormone receptor-positive breast cancer.* Clin Chem, 2009. **55**(8): p. 1453-5.

220. Lafci, O., et al., *DCE-MRI Radiomics Analysis in Differentiating Luminal A and Luminal B Breast Cancer Molecular Subtypes.* Acad Radiol, 2023. **30**(1): p. 22-29.

221. Krishnamurti, U. and J.F. Silverman, *HER2 in breast cancer: a review and update.* Adv Anat Pathol, 2014. **21**(2): p. 100-7.

222. Figueroa-Magalhaes, M.C., et al., *Treatment of HER2-positive breast cancer.* Breast, 2014. **23**(2): p. 128-136.

223. Cejalvo, J.M., et al., *Intrinsic Subtypes and Gene Expression Profiles in Primary and Metastatic Breast Cancer.* Cancer Res, 2017. **77**(9): p. 2213-2221.

224. Kumar, P. and R. Aggarwal, *An overview of triple-negative breast cancer.* Arch Gynecol Obstet, 2016. **293**(2): p. 247-69.

225. Orrantia-Borunda, E., et al., *Subtypes of Breast Cancer*, in *Breast Cancer*, H.N. Mayrovitz, Editor. 2022: Brisbane (AU).

226. Sarvari, P., et al., *Advances of Epigenetic Biomarkers and Epigenome Editing for Early Diagnosis in Breast Cancer.* Int J Mol Sci, 2022. **23**(17).

227. Attalla, S., T. Taifour, and W. Muller, *Tailoring therapies to counter the divergent immune landscapes of breast cancer.* Front Cell Dev Biol, 2023. **11**: p. 1111796.

228. Nounou, M.I., et al., *Breast Cancer: Conventional Diagnosis and Treatment Modalities and Recent Patents and Technologies.* Breast Cancer (Auckl), 2015. **9**(Suppl 2): p. 17-34.

229. Boyages, J., *Radiation therapy and early breast cancer: current controversies.* Med J Aust, 2017. **207**(5): p. 216-222.

230. Joshi, S.C., et al., *Role of radiotherapy in early breast cancer: an overview.* Int J Health Sci (Qassim), 2007. **1**(2): p. 259-64.

231. Watkins, E.J., *Overview of breast cancer.* JAAPA, 2019. **32**(10): p. 13-17.

232. Puhalla, S., S. Bhattacharya, and N.E. Davidson, *Hormonal therapy in breast cancer: a model disease for the personalization of cancer care.* Mol Oncol, 2012. **6**(2): p. 222-36.

233. Awan, A. and K. Esfahani, *Endocrine therapy for breast cancer in the primary care setting.* Curr Oncol, 2018. **25**(4): p. 285-291.

234. Higgins, M.J. and J. Baselga, *Targeted therapies for breast cancer.* J Clin Invest, 2011. **121**(10): p. 3797-803.

235. Masoud, V. and G. Pages, *Targeted therapies in breast cancer: New challenges to fight against resistance.* World J Clin Oncol, 2017. **8**(2): p. 120-134.

236. Dhankhar, R., et al., *Advances in novel drug delivery strategies for breast cancer therapy.* Artif Cells Blood Substit Immobil Biotechnol, 2010. **38**(5): p. 230-49.

237. *Favourable and unfavourable effects on long-term survival of radiotherapy for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group.* Lancet, 2000. **355**(9217): p. 1757-70.

238. Akram, M. and S.A. Siddiqui, *Breast cancer management: past, present and evolving.* Indian J Cancer, 2012. **49**(3): p. 277-82.

239. Shao, N., et al., *Sequential versus concurrent anthracyclines and taxanes as adjuvant chemotherapy of early breast cancer: a meta-analysis of phase III randomized control trials.* Breast, 2012. **21**(3): p. 389-93.

240. den Hollander, P., M.I. Savage, and P.H. Brown, *Targeted therapy for breast cancer prevention.* Front Oncol, 2013. **3**: p. 250.

241. Kawalec, P., S. Lopuch, and A. Mikrut, *Effectiveness of targeted therapy in patients with previously untreated metastatic breast cancer: a systematic review and meta-analysis.* Clin Breast Cancer, 2015. **15**(2): p. 90-100 e1.

242. Miller, K., et al., *Paclitaxel plus bevacizumab versus paclitaxel alone for metastatic breast cancer.* N Engl J Med, 2007. **357**(26): p. 2666-76.

243. Huang, M., J.J. Lu, and J. Ding, *Natural Products in Cancer Therapy: Past, Present and Future.* Nat Prod Bioprospect, 2021. **11**(1): p. 5-13.

244. Yuan, L., et al., *Promoting Apoptosis, a Promising Way to Treat Breast Cancer With Natural Products: A Comprehensive Review.* Front Pharmacol, 2021. **12**: p. 801662.

245. Hematpoor, A., et al., *Phenylpropanoids isolated from Piper sarmentosum Roxb. induce apoptosis in breast cancer cells through reactive oxygen species and mitochondrial-dependent pathways.* Chem Biol Interact, 2018. **279**: p. 210-218.

246. Kushwaha, P.P., et al., *Induction of apoptosis in breast cancer cells by naphthylisoquinoline alkaloids.* Toxicol Appl Pharmacol, 2020. **409**: p. 115297.

247. Sindhu, R.K., et al., *Impacting the Remedial Potential of Nano Delivery-Based Flavonoids for Breast Cancer Treatment.* Molecules, 2021. **26**(17).

248. Quisbert-Valenzuela, E.O. and G.M. Calaf, *Apoptotic effect of noscapine in breast cancer cell lines.* Int J Oncol, 2016. **48**(6): p. 2666-74.

249. Athanasiou, E., et al., *The Association Between the Risk of Breast Cancer and Epigallocatechin- 3-Gallate Intake: A Literature Review of a Potential Chemopreventive Agent.* Curr Med Chem, 2022. **29**(40): p. 6169-6196.

250. Lamb, J., et al., *The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease.* Science, 2006. **313**(5795): p. 1929-35.

251. Brum, A.M., et al., *Connectivity Map-based discovery of parbendazole reveals targetable human osteogenic pathway.* Proc Natl Acad Sci U S A, 2015. **112**(41): p. 12711-6.

252. Liu, J., et al., *Treatment of obesity with celastrol.* Cell, 2015. **161**(5): p. 999-1011.

253. Zhang, M., et al., *Drug repositioning for diabetes based on 'omics' data mining.* PLoS One, 2015. **10**(5): p. e0126082.

254. Farooq, F., et al., *p38 Mitogen-activated protein kinase stabilizes SMN mRNA through RNA binding protein HuR.* Hum Mol Genet, 2009. **18**(21): p. 4035-45.

255. Dudley, J.T., et al., *Computational repositioning of the anticonvulsant topiramate for inflammatory bowel disease.* Sci Transl Med, 2011. **3**(96): p. 96ra76.

256. Churchman, M.L., et al., *Efficacy of Retinoids in IKZF1-Mutated BCR-ABL1 Acute Lymphoblastic Leukemia.* Cancer Cell, 2015. **28**(3): p. 343-56.

257. Li, B., D.H. Robinson, and D.F. Birt, *Evaluation of properties of apigenin and [G-3H]apigenin and analytic method development.* J Pharm Sci, 1997. **86**(6): p. 721-5.

258. Rajhard, S., et al., *Solubility of Luteolin and Other Polyphenolic Compounds in Water, Nonpolar, Polar Aprotic and Protic Solvents by Applying FTIR/HPLC.* Processes, 2021. **9**(11).

259. Chaicharoenaudomrung, N., P. Kunhorm, and P. Noisa, *Three-dimensional cell culture systems as an in vitro platform for cancer and stem cell modeling.* World J Stem Cells, 2019. **11**(12): p. 1065-1083.

260. Reimand, J., et al., *Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap.* Nat Protoc, 2019. **14**(2): p. 482-517.

261. Subramanian, A., et al., *A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles.* Cell, 2017. **171**(6): p. 1437-1452 e17.

262. Uva, P., et al., *Connectivity Map Analysis Indicates PI3K/Akt/mTOR Inhibitors as Potential Anti-Hypoxia Drugs in Neuroblastoma.* Cancers (Basel), 2021. **13**(11).

263. Gupta, V.K. and O. Chaudhuri, *Mechanical regulation of cell-cycle progression and division.* Trends Cell Biol, 2022. **32**(9): p. 773-785.

264. Jamasbi, E., et al., *The cell cycle, cancer development and therapy.* Mol Biol Rep, 2022. **49**(11): p. 10875-10883.

265. Bower, J.J., et al., *Patterns of cell cycle checkpoint deregulation associated with intrinsic molecular subtypes of human breast cancer cells.* NPJ Breast Cancer, 2017. **3**: p. 9.

266. Visconti, R., R. Della Monica, and D. Grieco, *Cell cycle checkpoint in cancer: a therapeutically targetable double-edged sword.* J Exp Clin Cancer Res, 2016. **35**(1): p. 153.

267. Druker, J., et al., *Role of Hypoxia in the Control of the Cell Cycle.* Int J Mol Sci, 2021. **22**(9).

268. Sun, Y., et al., *The Influence of Cell Cycle Regulation on Chemotherapy.* Int J Mol Sci, 2021. **22**(13).

269. Marcus, J.M., et al., *Loss of p53 expression in cancer cells alters cell cycle response after inhibition of exportin-1 but does not prevent cell death.* Cell Cycle, 2018. **17**(11): p. 1329-1344.

270. Karimian, A., Y. Ahmadi, and B. Yousefi, *Multiple functions of p21 in cell cycle, apoptosis and transcriptional regulation after DNA damage.* DNA Repair (Amst), 2016. **42**: p. 63-71.

271. Abbas, T. and A. Dutta, *p21 in cancer: intricate networks and multiple activities.* Nat Rev Cancer, 2009. **9**(6): p. 400-14.

272. Qian, Y. and X. Chen, *Tumor suppression by p53: making cells senescent.* Histol Histopathol, 2010. **25**(4): p. 515-26.

273. Ye, Y., et al., *Characterization of hypoxia-associated molecular features to aid hypoxia-targeted therapy.* Nature Metabolism, 2019. **1**(4): p. 431-444.

274. Masoud, G.N. and W. Li, *HIF-1alpha pathway: role, regulation and intervention for cancer therapy.* Acta Pharm Sin B, 2015. **5**(5): p. 378-89.

275. Kaelin, W.G., Jr. and P.J. Ratcliffe, *Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway.* Mol Cell, 2008. **30**(4): p. 393-402.

276. Chan, M.C., et al., *Tuning the Transcriptional Response to Hypoxia by Inhibiting Hypoxia-inducible Factor (HIF) Prolyl and Asparaginyl Hydroxylases.* J Biol Chem, 2016. **291**(39): p. 20661-73.

277. Wang, Y., et al., *Nuclear entry and export of FIH are mediated by HIF1alpha and exportin1, respectively.* J Cell Sci, 2018. **131**(22).

278. Myllyharju, J., *Prolyl 4-hydroxylases, master regulators of the hypoxia response.* Acta Physiol (Oxf), 2013. **208**(2): p. 148-65.

279. Mahon, P.C., K. Hirota, and G.L. Semenza, *FIH-1: a novel protein that interacts with HIF-1alpha and VHL to mediate repression of HIF-1 transcriptional activity.* Genes Dev, 2001. **15**(20): p. 2675-86.

280. Jin, I.J., et al., *Solubilization of oleanolic acid and ursolic acid by cosolvency.* Arch Pharm Res, 1997. **20**(3): p. 269-74.

281. Einbond, L.S., et al., *The growth inhibitory effect of actein on human breast cancer cells is associated with activation of stress response pathways.* Int J Cancer, 2007. **121**(9): p. 2073-83.

282. Lv, C., 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**(1): p. 352.

283. Wen, Z., et al., *Discovery of molecular mechanisms of traditional Chinese medicinal formula Si-Wu-Tang using gene expression microarray and connectivity map.* PLoS One, 2011. **6**(3): p. e18278.

284. Gaube, F., et al., *Gene expression profiling reveals effects of Cimicifuga racemosa (L.) NUTT. (black cohosh) on the estrogen receptor positive human breast cancer cell line MCF-7.* BMC Pharmacol, 2007. **7**: p. 11.

285. Bateman, H.R., et al., *Sparstolonin B inhibits pro-angiogenic functions and blocks cell cycle progression in endothelial cells.* PLoS One, 2013. **8**(8): p. e70500.

286. Ahmed, K., et al., *Chemical inducers of heat shock proteins derived from medicinal plants and cytoprotective genes response.* Int J Hyperthermia, 2012. **28**(1): p. 1-8.

287. Lee, C.C., et al., *The small molecule calactin induces DNA damage and apoptosis in human leukemia cells.* Eur J Cancer Prev, 2012. **21**(5): p. 467-73.

288. Nourmohammadi, S., et al., *Effect of Compound Kushen Injection, a Natural Compound Mixture, and Its Identified Chemical Components on Migration and Invasion of Colon, Brain, and Breast Cancer Cell Lines.* Front Oncol, 2019. **9**: p. 314.

289. Aung, T.N., et al., *Fractional Deletion of Compound Kushen Injection Indicates Cytokine Signaling Pathways are Critical for its Perturbation of the Cell Cycle.* Sci Rep, 2019. **9**(1): p. 14200.

290. Shen, H., et al., *Understanding the Mechanistic Contribution of Herbal Extracts in Compound Kushen Injection With Transcriptome Analysis.* Front Oncol, 2019. **9**: p. 632.

291. Shen, H., et al., *A New Strategy for Identifying Mechanisms of Drug-drug Interaction Using Transcriptome Analysis: Compound Kushen Injection as a Proof of Principle.* Sci Rep, 2019. **9**(1): p. 15889.

292. Qu, Z., et al., *Identification of candidate anti-cancer molecular mechanisms of Compound Kushen Injection using functional genomics.* Oncotarget, 2016. **7**(40): p. 66003-66019.

293. Kang, J.I., et al., *Pro-angiogenic Ginsenosides F1 and Rh1 Inhibit Vascular Leakage by Modulating NR4A1.* Sci Rep, 2019. **9**(1): p. 4502.

294. Cui, J., et al., *The effect of compound kushen injection on cancer cells: Integrated identification of candidate molecular mechanisms.* PLoS One, 2020. **15**(7): p. e0236395.

295. Makhafola, T.J., et al., *Apoptosis in Cancer Cells Is Induced by Alternative Splicing of hnRNPA2/B1 Through Splicing of Bcl-x, a Mechanism that Can Be Stimulated by an Extract of the South African Medicinal Plant, Cotyledon orbiculata.* Front Oncol, 2020. **10**: p. 547392.

296. Haider, M., et al., *Transcriptome analysis and connectivity mapping of Cissampelos pareira L. provides molecular links of ESR1 modulation to viral inhibition.* Sci Rep, 2021. **11**(1): p. 20095.

297. Bie, B., et al., *Baicalein, a Natural Anti-Cancer Compound, Alters MicroRNA Expression Profiles in Bel-7402 Human Hepatocellular Carcinoma Cells.* Cell Physiol Biochem, 2017. **41**(4): p. 1519-1531.

298. Dong, Y., et al., *Pao Pereira extract suppresses benign prostatic hyperplasia by inhibiting inflammation-associated NFkappaB signaling.* BMC Complement Med Ther, 2020. **20**(1): p. 150.

299. Zhan, Z., et al., *Oridonin alleviates hyperbilirubinemia through activating LXRalpha-UGT1A1 axis.* Pharmacol Res, 2022. **178**: p. 106188.

300. Mao, Z.J., et al., *Combined Use of Astragalus Polysaccharide and Berberine Attenuates Insulin Resistance in IR-HepG2 Cells via Regulation of the Gluconeogenesis Signaling Pathway.* Front Pharmacol, 2019. **10**: p. 1508.

301. Ng, Y.K. *The Effect of Cratoxylum Cochinchinense Lour (CCL) On Global Mrna Gene Expression In Hepg2 Liver Cancer Cells*. 2017.

302. Kuk, H., et al., *Glycyrrhetinic Acid Antagonizes Pressure-Induced Venous Remodeling in Mice.* Front Physiol, 2018. **9**: p. 320.

303. Liu, Y., et al., *Chinese herbal extract Su-duxing had potent inhibitory effects on both wild-type and entecavir-resistant hepatitis B virus (HBV) in vitro and effectively suppressed HBV replication in mouse model.* Antiviral Res, 2018. **155**: p. 39-47.

304. Shu, Y., et al., *DHOK Exerts Anti-Cancer Effect Through Autophagy Inhibition in Colorectal Cancer.* Front Cell Dev Biol, 2021. **9**: p. 760022.

305. Esser, D., et al., *Ayurvedic Herbal Preparation Supplementation Does Not Improve Metabolic Health in Impaired Glucose Tolerance Subjects; Observations from a Randomised Placebo Controlled Trial.* Nutrients, 2021. **13**(1).

306. Tsai, C.H., et al., *A standardized herbal extract mitigates tumor inflammation and augments chemotherapy effect of docetaxel in prostate cancer.* Sci Rep, 2017. **7**(1): p. 15624.

307. Thamsermsang, O., et al., *IL-1beta-induced modulation of gene expression profile in human dermal fibroblasts: the effects of Thai herbal Sahatsatara formula, piperine and gallic acid possessing antioxidant properties.* BMC Complement Altern Med, 2017. **17**(1): p. 32.

308. Chen, Y.C., et al., *Whole genome gene expression changes and hematological effects of rikkunshito in patients with advanced non-small cell lung cancer receiving first line chemotherapy.* Exp Ther Med, 2017. **14**(3): p. 2040-2052.

309. Marzotto, M., et al., *Extreme sensitivity of gene expression in human SH-SY5Y neurocytes to ultra-low doses of Gelsemium sempervirens.* BMC Complement Altern Med, 2014. **14**: p. 104.

310. Cheng, S., et al., *Pachymic acid inhibits growth and induces apoptosis of pancreatic cancer in vitro and in vivo by targeting ER stress.* PLoS One, 2015. **10**(4): p. e0122270.

311. Grinde, B., G. Hetland, and E. Johnson, *Effects on gene expression and viral load of a medicinal extract from Agaricus blazei in patients with chronic hepatitis C infection.* Int Immunopharmacol, 2006. **6**(8): p. 1311-4.

312. Li, C.Y., et al., *Gene expression profiling of dendritic cells in different physiological stages under Cordyceps sinensis treatment.* PLoS One, 2012. **7**(7): p. e40824.

313. Guo, S., et al., *(5R)-5-Hydroxytriptolide (LLDT-8) induces substantial epigenetic mediated immune response network changes in fibroblast-like synoviocytes from rheumatoid arthritis patients.* Sci Rep, 2019. **9**(1): p. 11155.

314. Savini, C., et al., *Folate Repletion after Deficiency Induces Irreversible Genomic and Transcriptional Changes in Human Papillomavirus Type 16 (HPV16)-Immortalized Human Keratinocytes.* Int J Mol Sci, 2019. **20**(5).

315. Tsukiyama, F., et al., *Gallate, the component of HIF-inducing catechins, inhibits HIF prolyl hydroxylase.* Biochem Biophys Res Commun, 2006. **351**(1): p. 234-9.

316. Iwai, K., et al., *Identification of the von Hippel-lindau tumor-suppressor protein as part of an active E3 ubiquitin ligase complex.* Proc Natl Acad Sci U S A, 1999. **96**(22): p. 12436-41.

317. Kallio, P.J., et al., *Regulation of the hypoxia-inducible transcription factor 1alpha by the ubiquitin-proteasome pathway.* J Biol Chem, 1999. **274**(10): p. 6519-25.

318. Maxwell, P.H., et al., *The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis.* Nature, 1999. **399**(6733): p. 271-5.

319. Wynn, T.A. and T.R. Ramalingam, *Mechanisms of fibrosis: therapeutic translation for fibrotic disease.* Nat Med, 2012. **18**(7): p. 1028-40.

320. Coussens, L.M. and Z. Werb, *Inflammation and cancer.* Nature, 2002. **420**(6917): p. 860-7.

321. Tang, X.H., et al., *Mitochondrial modulation is involved in the hepatoprotection of Limonium sinense extract against liver damage in mice.* J Ethnopharmacol, 2008. **120**(3): p. 427-31.

322. Das, S.K., *Cell cycle regulatory control for uterine stromal cell decidualization in implantation.* Reproduction, 2009. **137**(6): p. 889-99.

323. Malumbres, M. and M. Barbacid, *Cell cycle, CDKs and cancer: a changing paradigm.* Nat Rev Cancer, 2009. **9**(3): p. 153-66.

324. Reinhardt, H.C. and M.B. Yaffe, *Kinases that control the cell cycle in response to DNA damage: Chk1, Chk2, and MK2.* Curr Opin Cell Biol, 2009. **21**(2): p. 245-55.

325. Sclafani, R.A. and T.M. Holzen, *Cell cycle regulation of DNA replication.* Annu Rev Genet, 2007. **41**: p. 237-80.

326. Koshiji, M., et al., *HIF-1alpha induces cell cycle arrest by functionally counteracting Myc.* EMBO J, 2004. **23**(9): p. 1949-56.

327. Hackenbeck, T., et al., *HIF-1 or HIF-2 induction is sufficient to achieve cell cycle arrest in NIH3T3 mouse fibroblasts independent from hypoxia.* Cell Cycle, 2009. **8**(9): p. 1386-95.

328. Hubbi, M.E. and G.L. Semenza, *Regulation of cell proliferation by hypoxia-inducible factors.* Am J Physiol Cell Physiol, 2015. **309**(12): p. C775-82.

329. Janke, C. and M.M. Magiera, *The tubulin code and its role in controlling microtubule properties and functions.* Nat Rev Mol Cell Biol, 2020. **21**(6): p. 307-326.

330. Nogales, E. and H.W. Wang, *Structural mechanisms underlying nucleotide-dependent self-assembly of tubulin and its relatives.* Curr Opin Struct Biol, 2006. **16**(2): p. 221-9.

331. Akhmanova, A. and M.O. Steinmetz, *Tracking the ends: a dynamic protein network controls the fate of microtubule tips.* Nat Rev Mol Cell Biol, 2008. **9**(4): p. 309-22.

332. Nogales, E., S.G. Wolf, and K.H. Downing, *Structure of the alpha beta tubulin dimer by electron crystallography.* Nature, 1998. **391**(6663): p. 199-203.

333. Kingston, D.G., *Tubulin-interactive natural products as anticancer agents.* J Nat Prod, 2009. **72**(3): p. 507-15.

334. Dumontet, C. and M.A. Jordan, *Microtubule-binding agents: a dynamic field of cancer therapeutics.* Nat Rev Drug Discov, 2010. **9**(10): p. 790-803.

335. Wang, Y., et al., *Intrinsic disorder mediates the diverse regulatory functions of the Cdk inhibitor p21.* Nat Chem Biol, 2011. **7**(4): p. 214-21.

336. Gong, D. and J.E. Ferrell, Jr., *The roles of cyclin A2, B1, and B2 in early and late mitotic events.* Mol Biol Cell, 2010. **21**(18): p. 3149-61.

337. Lee, F.S. and M.J. Percy, *The HIF pathway and erythrocytosis.* Annu Rev Pathol, 2011. **6**: p. 165-92.

338. Schodel, J. and P.J. Ratcliffe, *Mechanisms of hypoxia signalling: new implications for nephrology.* Nat Rev Nephrol, 2019. **15**(10): p. 641-659.

339. De Bels, D., F. Corazza, and C. Balestra, *Oxygen sensing, homeostasis, and disease.* N Engl J Med, 2011. **365**(19): p. 1845; author reply 1846.

340. Semenza, G.L., *Hypoxia-inducible factors in physiology and medicine.* Cell, 2012. **148**(3): p. 399-408.

341. Fallah, J. and B.I. Rini, *HIF Inhibitors: Status of Current Clinical Development.* Curr Oncol Rep, 2019. **21**(1): p. 6.

342. Hong, M., et al., *Dual Effects of Chinese Herbal Medicines on Angiogenesis in Cancer and Ischemic Stroke Treatments: Role of HIF-1 Network.* Front Pharmacol, 2019. **10**: p. 696.

343. Darby, I.A. and T.D. Hewitson, *Hypoxia in tissue repair and fibrosis.* Cell Tissue Res, 2016. **365**(3): p. 553-62.

344. Kosuru, R.Y., et al., *Gallic Acid and Gallates in Human Health and Disease: Do Mitochondria Hold the Key to Success?* Mol Nutr Food Res, 2018. **62**(1).

345. Zhou, Y.D., et al., *Hypoxia-inducible factor-1 activation by (-)-epicatechin gallate: potential adverse effects of cancer chemoprevention with high-dose green tea extracts.* J Nat Prod, 2004. **67**(12): p. 2063-9.

346. *Green tea.* Altern Med Rev, 2000. **5**(4): p. 372-5.

347. Zhang, Y., et al., *The therapeutic role of Jingchuan tablet on ischaemic cerebral stroke via the HIF-1alpha/EPO/VEGFA signalling pathway.* Pharm Biol, 2022. **60**(1): p. 2110-2123.

348. Abdel-Hamed, A.R., et al., *Plicosepalus acacia Extract and Its Major Constituents, Methyl Gallate and Quercetin, Potentiate Therapeutic Angiogenesis in Diabetic Hind Limb Ischemia: HPTLC Quantification and LC-MS/MS Metabolic Profiling.* Antioxidants (Basel), 2021. **10**(11).

349. Kawano, Y., et al., *Effects of n-propyl gallate on neuronal survival after forebrain ischemia in rats.* Resuscitation, 2012. **83**(2): p. 249-52.

350. Guo, S., et al., *Glucose up-regulates HIF-1 alpha expression in primary cortical neurons in response to hypoxia through maintaining cellular redox status.* J Neurochem, 2008. **105**(5): p. 1849-60.

351. Ghosh, R., et al., *Targeting HIF-1alpha by Natural and Synthetic Compounds: A Promising Approach for Anti-Cancer Therapeutics Development.* Molecules, 2022. **27**(16).

352. Ragsdale, A., et al., *Paternal hypoxia exposure primes offspring for increased hypoxia resistance.* BMC Biol, 2022. **20**(1): p. 185.

353. Roesner, A., T. Hankeln, and T. Burmester, *Hypoxia induces a complex response of globin expression in zebrafish (Danio rerio).* J Exp Biol, 2006. **209**(Pt 11): p. 2129-37.

354. MacIntyre, N.R., *Tissue hypoxia: implications for the respiratory clinician.* Respir Care, 2014. **59**(10): p. 1590-6.

355. Lee, P., N.S. Chandel, and M.C. Simon, *Cellular adaptation to hypoxia through hypoxia inducible factors and beyond.* Nat Rev Mol Cell Biol, 2020. **21**(5): p. 268-283.

356. Leach, R.M. and D.F. Treacher, *Oxygen transport-2. Tissue hypoxia.* BMJ, 1998. **317**(7169): p. 1370-3.

357. Martin, D.S., et al., *Concepts in hypoxia reborn.* Crit Care, 2010. **14**(4): p. 315.

358. Zhao, Y., et al., *Nitric oxide in red blood cell adaptation to hypoxia.* Acta Biochim Biophys Sin (Shanghai), 2018. **50**(7): p. 621-634.

359. Ratcliffe, P.J., *Oxygen sensing and hypoxia signalling pathways in animals: the implications of physiology for cancer.* J Physiol, 2013. **591**(8): p. 2027-42.

360. Semenza, G.L., *Targeting HIF-1 for cancer therapy.* Nat Rev Cancer, 2003. **3**(10): p. 721-32.

361. Pritchard, K.A., Jr., et al., *Hypoxia-induced acute lung injury in murine models of sickle cell disease.* Am J Physiol Lung Cell Mol Physiol, 2004. **286**(4): p. L705-14.

362. Remensnyder, J.P. and G. Majno, *Oxygen gradients in healing wounds.* Am J Pathol, 1968. **52**(2): p. 301-23.

363. Rankin, E.B. and A.J. Giaccia, *The role of hypoxia-inducible factors in tumorigenesis.* Cell Death Differ, 2008. **15**(4): p. 678-85.

364. Dhani, N., et al., *The clinical significance of hypoxia in human cancers.* Semin Nucl Med, 2015. **45**(2): p. 110-21.

365. Bhandari, V., et al., *Divergent mutational processes distinguish hypoxic and normoxic tumours.* Nat Commun, 2020. **11**(1): p. 737.

366. Zundel, W., et al., *Loss of PTEN facilitates HIF-1-mediated gene expression.* Genes Dev, 2000. **14**(4): p. 391-6.

367. Bernardi, R., et al., *PML inhibits HIF-1alpha translation and neoangiogenesis through repression of mTOR.* Nature, 2006. **442**(7104): p. 779-85.

368. Luo, Z., et al., *Hypoxia signaling in human health and diseases: implications and prospects for therapeutics.* Signal Transduct Target Ther, 2022. **7**(1): p. 218.

369. Ancel, J., et al., *Hypoxia in Lung Cancer Management: A Translational Approach.* Cancers (Basel), 2021. **13**(14).

370. Wang, G.L., et al., *Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension.* Proc Natl Acad Sci U S A, 1995. **92**(12): p. 5510-4.

371. Jiang, B.H., et al., *Hypoxia-inducible factor 1 levels vary exponentially over a physiologically relevant range of O2 tension.* Am J Physiol, 1996. **271**(4 Pt 1): p. C1172-80.

372. Prabhakar, N.R. and G.L. Semenza, *Adaptive and maladaptive cardiorespiratory responses to continuous and intermittent hypoxia mediated by hypoxia-inducible factors 1 and 2.* Physiol Rev, 2012. **92**(3): p. 967-1003.

373. Dunwoodie, S.L., *The role of hypoxia in development of the Mammalian embryo.* Dev Cell, 2009. **17**(6): p. 755-73.

374. Huang, L.E., et al., *Regulation of hypoxia-inducible factor 1alpha is mediated by an O2-dependent degradation domain via the ubiquitin-proteasome pathway.* Proc Natl Acad Sci U S A, 1998. **95**(14): p. 7987-92.

375. Pugh, C.W., et al., *Activation of hypoxia-inducible factor-1; definition of regulatory domains within the alpha subunit.* J Biol Chem, 1997. **272**(17): p. 11205-14.

376. Wang, G.L. and G.L. Semenza, *Purification and characterization of hypoxia-inducible factor 1.* J Biol Chem, 1995. **270**(3): p. 1230-7.

377. Wenger, R.H., et al., *Nucleotide sequence, chromosomal assignment and mRNA expression of mouse hypoxia-inducible factor-1 alpha.* Biochem Biophys Res Commun, 1996. **223**(1): p. 54-9.

378. Wu, D., et al., *Structural integration in hypoxia-inducible factors.* Nature, 2015. **524**(7565): p. 303-8.

379. Iyer, N.V., S.W. Leung, and G.L. Semenza, *The human hypoxia-inducible factor 1alpha gene: HIF1A structure and evolutionary conservation.* Genomics, 1998. **52**(2): p. 159-65.

380. Tian, H., S.L. McKnight, and D.W. Russell, *Endothelial PAS domain protein 1 (EPAS1), a transcription factor selectively expressed in endothelial cells.* Genes Dev, 1997. **11**(1): p. 72-82.

381. Ema, M., et al., *A novel bHLH-PAS factor with close sequence similarity to hypoxia-inducible factor 1alpha regulates the VEGF expression and is potentially involved in lung and vascular development.* Proc Natl Acad Sci U S A, 1997. **94**(9): p. 4273-8.

382. Wiesener, M.S., et al., *Widespread hypoxia-inducible expression of HIF-2alpha in distinct cell populations of different organs.* FASEB J, 2003. **17**(2): p. 271-3.

383. Raval, R.R., et al., *Contrasting properties of hypoxia-inducible factor 1 (HIF-1) and HIF-2 in von Hippel-Lindau-associated renal cell carcinoma.* Mol Cell Biol, 2005. **25**(13): p. 5675-86.

384. Hu, C.J., et al., *Differential roles of hypoxia-inducible factor 1alpha (HIF-1alpha) and HIF-2alpha in hypoxic gene regulation.* Mol Cell Biol, 2003. **23**(24): p. 9361-74.

385. Albadari, N., S. Deng, and W. Li, *The transcriptional factors HIF-1 and HIF-2 and their novel inhibitors in cancer therapy.* Expert Opin Drug Discov, 2019. **14**(7): p. 667-682.

386. Pasanen, A., et al., *Hypoxia-inducible factor (HIF)-3alpha is subject to extensive alternative splicing in human tissues and cancer cells and is regulated by HIF-1 but not HIF-2.* Int J Biochem Cell Biol, 2010. **42**(7): p. 1189-200.

387. Maynard, M.A., et al., *Multiple splice variants of the human HIF-3 alpha locus are targets of the von Hippel-Lindau E3 ubiquitin ligase complex.* J Biol Chem, 2003. **278**(13): p. 11032-40.

388. Hara, S., et al., *Expression and characterization of hypoxia-inducible factor (HIF)-3alpha in human kidney: suppression of HIF-mediated gene expression by HIF-3alpha.* Biochem Biophys Res Commun, 2001. **287**(4): p. 808-13.

389. Diao, X., et al., *Identification of oleoylethanolamide as an endogenous ligand for HIF-3alpha.* Nat Commun, 2022. **13**(1): p. 2529.

390. Huang, L.E., et al., *Activation of hypoxia-inducible transcription factor depends primarily upon redox-sensitive stabilization of its alpha subunit.* J Biol Chem, 1996. **271**(50): p. 32253-9.

391. Maltepe, E., et al., *The role of ARNT2 in tumor angiogenesis and the neural response to hypoxia.* Biochem Biophys Res Commun, 2000. **273**(1): p. 231-8.

392. Wu, D., et al., *NPAS1-ARNT and NPAS3-ARNT crystal structures implicate the bHLH-PAS family as multi-ligand binding transcription factors.* Elife, 2016. **5**.

393. Sun, X., et al., *Structures of NPAS4-ARNT and NPAS4-ARNT2 heterodimers reveal new dimerization modalities in the bHLH-PAS transcription factor family.* Proc Natl Acad Sci U S A, 2022. **119**(46): p. e2208804119.

394. Pezzuto, A. and E. Carico, *Role of HIF-1 in Cancer Progression: Novel Insights. A Review.* Curr Mol Med, 2018. **18**(6): p. 343-351.

395. Yamashita, K., et al., *Molecular regulation of the endothelin-1 gene by hypoxia. Contributions of hypoxia-inducible factor-1, activator protein-1, GATA-2, AND p300/CBP.* J Biol Chem, 2001. **276**(16): p. 12645-53.

396. Masson, N., et al., *Independent function of two destruction domains in hypoxia-inducible factor-alpha chains activated by prolyl hydroxylation.* EMBO J, 2001. **20**(18): p. 5197-206.

397. Semenza, G.L., *HIF-1 and mechanisms of hypoxia sensing.* Curr Opin Cell Biol, 2001. **13**(2): p. 167-71.

398. Maxwell, P.H., et al., *Hypoxia-inducible factor-1 modulates gene expression in solid tumors and influences both angiogenesis and tumor growth.* Proc Natl Acad Sci U S A, 1997. **94**(15): p. 8104-9.

399. Goda, N., S.J. Dozier, and R.S. Johnson, *HIF-1 in cell cycle regulation, apoptosis, and tumor progression.* Antioxid Redox Signal, 2003. **5**(4): p. 467-73.

400. Lee, S.H., M. Golinska, and J.R. Griffiths, *HIF-1-Independent Mechanisms Regulating Metabolic Adaptation in Hypoxic Cancer Cells.* Cells, 2021. **10**(9).

401. Meijer, T.W., et al., *Targeting hypoxia, HIF-1, and tumor glucose metabolism to improve radiotherapy efficacy.* Clin Cancer Res, 2012. **18**(20): p. 5585-94.

402. Mouriaux, F., et al., *Increased HIF-1alpha expression correlates with cell proliferation and vascular markers CD31 and VEGF-A in uveal melanoma.* Invest Ophthalmol Vis Sci, 2014. **55**(3): p. 1277-83.

403. Kumar, H. and D.K. Choi, *Hypoxia Inducible Factor Pathway and Physiological Adaptation: A Cell Survival Pathway?* Mediators Inflamm, 2015. **2015**: p. 584758.

404. Semenza, G.L., *Pharmacologic Targeting of Hypoxia-Inducible Factors.* Annu Rev Pharmacol Toxicol, 2019. **59**: p. 379-403.

405. Talks, K.L., et al., *The expression and distribution of the hypoxia-inducible factors HIF-1alpha and HIF-2alpha in normal human tissues, cancers, and tumor-associated macrophages.* Am J Pathol, 2000. **157**(2): p. 411-21.

406. Jiang, B.H., et al., *Phosphatidylinositol 3-kinase signaling controls levels of hypoxia-inducible factor 1.* Cell Growth Differ, 2001. **12**(7): p. 363-9.

407. An, W.G., et al., *Stabilization of wild-type p53 by hypoxia-inducible factor 1alpha.* Nature, 1998. **392**(6674): p. 405-8.

408. Onnis, B., A. Rapisarda, and G. Melillo, *Development of HIF-1 inhibitors for cancer therapy.* J Cell Mol Med, 2009. **13**(9A): p. 2780-6.

409. Scheuermann, T.H., et al., *Allosteric inhibition of hypoxia inducible factor-2 with small molecules.* Nat Chem Biol, 2013. **9**(4): p. 271-6.

410. Jochmanova, I., et al., *Hypoxia-inducible factor signaling in pheochromocytoma: turning the rudder in the right direction.* J Natl Cancer Inst, 2013. **105**(17): p. 1270-83.

411. Zhang, Q., et al., *Role of hypoxia inducible factor-1 in cancer stem cells (Review).* Mol Med Rep, 2021. **23**(1).

412. Qiang, L., et al., *HIF-1alpha is critical for hypoxia-mediated maintenance of glioblastoma stem cells by activating Notch signaling pathway.* Cell Death Differ, 2012. **19**(2): p. 284-94.

413. Wang, R., S. Zhou, and S. Li, *Cancer therapeutic agents targeting hypoxia-inducible factor-1.* Curr Med Chem, 2011. **18**(21): p. 3168-89.

414. Davis, C.K., et al., *Hypoxia Mimetic Agents for Ischemic Stroke.* Front Cell Dev Biol, 2018. **6**: p. 175.

415. Yeh, T.L., et al., *Molecular and cellular mechanisms of HIF prolyl hydroxylase inhibitors in clinical trials.* Chem Sci, 2017. **8**(11): p. 7651-7668.

416. Chan, M.C., et al., *Pharmacological targeting of the HIF hydroxylases--A new field in medicine development.* Mol Aspects Med, 2016. **47-48**: p. 54-75.

417. Wilkins, S.E., et al., *Targeting Protein-Protein Interactions in the HIF System.* ChemMedChem, 2016. **11**(8): p. 773-86.

418. Fan, X.X., et al., *An integrated shotgun proteomics and bioinformatics approach for analysis of brain proteins from MCAO model using serial affinity chromatograph with four active ingredients from Shengmai preparations as ligands.* Neurochem Int, 2017. **103**: p. 45-56.

419. Gao, J., et al., *In vitro investigation of the mechanism underlying the effect of ginsenoside on the proliferation and differentiation of neural stem cells subjected to oxygen-glucose deprivation/reperfusion.* Int J Mol Med, 2018. **41**(1): p. 353-363.

420. He, Q., et al., *Total Flavonoids in Caragana (TFC) Promotes Angiogenesis and Enhances Cerebral Perfusion in a Rat Model of Ischemic Stroke.* Front Neurosci, 2018. **12**: p. 635.

421. Webster, A.C., et al., *Chronic Kidney Disease.* Lancet, 2017. **389**(10075): p. 1238-1252.

422. Lefebvre, P., et al., *Relationship between hemoglobin level and quality of life in anemic patients with chronic kidney disease receiving epoetin alfa.* Curr Med Res Opin, 2006. **22**(10): p. 1929-37.

423. Locatelli, F., et al., *Anaemia in haemodialysis patients of five European countries: association with morbidity and mortality in the Dialysis Outcomes and Practice Patterns Study (DOPPS).* Nephrol Dial Transplant, 2004. **19**(1): p. 121-32.

424. Kalantar-Zadeh, K., *History of Erythropoiesis-Stimulating Agents, the Development of Biosimilars, and the Future of Anemia Treatment in Nephrology.* Am J Nephrol, 2017. **45**(3): p. 235-247.

425. Hayat, A., D. Haria, and M.O. Salifu, *Erythropoietin stimulating agents in the management of anemia of chronic kidney disease.* Patient Prefer Adherence, 2008. **2**: p. 195-200.

426. Del Vecchio, L. and F. Locatelli, *Investigational hypoxia-inducible factor prolyl hydroxylase inhibitors (HIF-PHI) for the treatment of anemia associated with chronic kidney disease.* Expert Opin Investig Drugs, 2018. **27**(7): p. 613-621.

427. Palmer, S.C., et al., *Meta-analysis: erythropoiesis-stimulating agents in patients with chronic kidney disease.* Ann Intern Med, 2010. **153**(1): p. 23-33.

428. Koury, M.J. and V.H. Haase, *Anaemia in kidney disease: harnessing hypoxia responses for therapy.* Nat Rev Nephrol, 2015. **11**(7): p. 394-410.

429. Watts, E.R. and S.R. Walmsley, *Inflammation and Hypoxia: HIF and PHD Isoform Selectivity.* Trends Mol Med, 2019. **25**(1): p. 33-46.

430. Wong, B.W., et al., *Emerging novel functions of the oxygen-sensing prolyl hydroxylase domain enzymes.* Trends Biochem Sci, 2013. **38**(1): p. 3-11.

431. McDonough, M.A., et al., *Cellular oxygen sensing: Crystal structure of hypoxia-inducible factor prolyl hydroxylase (PHD2).* Proc Natl Acad Sci U S A, 2006. **103**(26): p. 9814-9.

432. Chowdhury, R., et al., *Structural basis for binding of hypoxia-inducible factor to the oxygen-sensing prolyl hydroxylases.* Structure, 2009. **17**(7): p. 981-9.

433. Khadem, S. and R.J. Marles, *Monocyclic phenolic acids; hydroxy- and polyhydroxybenzoic acids: occurrence and recent bioactivity studies.* Molecules, 2010. **15**(11): p. 7985-8005.

434. Juurlink, B.H., et al., *Hydroxybenzoic acid isomers and the cardiovascular system.* Nutr J, 2014. **13**: p. 63.

435. Majamaa, K., et al., *Partial identity of the 2-oxoglutarate and ascorbate binding sites of prolyl 4-hydroxylase.* J Biol Chem, 1986. **261**(17): p. 7819-23.

436. Siddiq, A., et al., *Hypoxia-inducible factor prolyl 4-hydroxylase inhibition. A target for neuroprotection in the central nervous system.* J Biol Chem, 2005. **280**(50): p. 41732-43.

437. Kasiganesan, H., V. Sridharan, and G. Wright, *Prolyl hydroxylase inhibitor treatment confers whole-animal hypoxia tolerance.* Acta Physiol (Oxf), 2007. **190**(2): p. 163-9.

438. Yu, Z., et al., *Discovery of prolyl hydroxylase 2 inhibitors with new chemical scaffolds as in vivo active erythropoietin inducers through a combined virtual screening strategy.* Chem Biol Drug Des, 2020. **95**(2): p. 270-278.

439. Viallard, C. and B. Larrivee, *Tumor angiogenesis and vascular normalization: alternative therapeutic targets.* Angiogenesis, 2017. **20**(4): p. 409-426.

440. Harris, A.L., *Hypoxia--a key regulatory factor in tumour growth.* Nat Rev Cancer, 2002. **2**(1): p. 38-47.

441. Semenza, G.L., *HIF-1 and human disease: one highly involved factor.* Genes Dev, 2000. **14**(16): p. 1983-91.

442. Nakhoul, G. and J.F. Simon, *Anemia of chronic kidney disease: Treat it, but not too aggressively.* Cleve Clin J Med, 2016. **83**(8): p. 613-24.

443. Li, Z., Q. You, and X. Zhang, *Small-Molecule Modulators of the Hypoxia-Inducible Factor Pathway: Development and Therapeutic Applications.* J Med Chem, 2019. **62**(12): p. 5725-5749.

444. Tang XH, X.L., Gao J, *Experimental Study on Antitumor Activity of Limonium sinense Extract.* Lishizhen Medicine and Materia Medica Research, 2010. **21**(04): p. 917-918.

445. Schafer, K.A., *The cell cycle: a review.* Vet Pathol, 1998. **35**(6): p. 461-78.

446. Molinari, M., *Cell cycle checkpoints and their inactivation in human cancer.* Cell Prolif, 2000. **33**(5): p. 261-74.

447. Holland, T.A., et al., *Subcellular localisation of cyclin D1 protein in colorectal tumours is associated with p21(WAF1/CIP1) expression and correlates with patient survival.* Int J Cancer, 2001. **95**(5): p. 302-6.

448. Yousefi, B., M. Rahmati, and Y. Ahmadi, *The roles of p53R2 in cancer progression based on the new function of mutant p53 and cytoplasmic p21.* Life Sci, 2014. **99**(1-2): p. 14-7.

449. Roninson, I.B., *Oncogenic functions of tumour suppressor p21(Waf1/Cip1/Sdi1): association with cell senescence and tumour-promoting activities of stromal fibroblasts.* Cancer Lett, 2002. **179**(1): p. 1-14.

450. Hsieh, Y.J., et al., *P21-driven multifusion gene system for evaluating the efficacy of histone deacetylase inhibitors by in vivo molecular imaging and for transcription targeting therapy of cancer mediated by histone deacetylase inhibitor.* J Nucl Med, 2014. **55**(4): p. 678-85.

451. Pavlides, S.C., et al., *TGF-beta activates APC through Cdh1 binding for Cks1 and Skp2 proteasomal destruction stabilizing p27kip1 for normal endometrial growth.* Cell Cycle, 2016. **15**(7): p. 931-47.

452. Lee, E.W., et al., *Differential regulation of p53 and p21 by MKRN1 E3 ligase controls cell cycle arrest and apoptosis.* EMBO J, 2009. **28**(14): p. 2100-13.

453. Malumbres, M. and M. Barbacid, *To cycle or not to cycle: a critical decision in cancer.* Nat Rev Cancer, 2001. **1**(3): p. 222-31.

454. Malumbres, M. and M. Barbacid, *Mammalian cyclin-dependent kinases.* Trends Biochem Sci, 2005. **30**(11): p. 630-41.

455. Shamloo, B. and S. Usluer, *p21 in Cancer Research.* Cancers (Basel), 2019. **11**(8).

456. el-Deiry, W.S., et al., *WAF1, a potential mediator of p53 tumor suppression.* Cell, 1993. **75**(4): p. 817-25.

457. Niculescu, A.B., 3rd, et al., *Effects of p21(Cip1/Waf1) at both the G1/S and the G2/M cell cycle transitions: pRb is a critical determinant in blocking DNA replication and in preventing endoreduplication.* Mol Cell Biol, 1998. **18**(1): p. 629-43.

458. Bertoli, C., J.M. Skotheim, and R.A. de Bruin, *Control of cell cycle transcription during G1 and S phases.* Nat Rev Mol Cell Biol, 2013. **14**(8): p. 518-28.

459. Georgakilas, A.G., O.A. Martin, and W.M. Bonner, *p21: A Two-Faced Genome Guardian.* Trends Mol Med, 2017. **23**(4): p. 310-319.

460. Gartel, A.L. and A.L. Tyner, *The role of the cyclin-dependent kinase inhibitor p21 in apoptosis.* Mol Cancer Ther, 2002. **1**(8): p. 639-49.

461. Warfel, N.A. and W.S. El-Deiry, *p21WAF1 and tumourigenesis: 20 years after.* Curr Opin Oncol, 2013. **25**(1): p. 52-8.

462. Mole, D.R., et al., *Genome-wide association of hypoxia-inducible factor (HIF)-1alpha and HIF-2alpha DNA binding with expression profiling of hypoxia-inducible transcripts.* J Biol Chem, 2009. **284**(25): p. 16767-16775.

463. Salceda, S. and J. Caro, *Hypoxia-inducible factor 1alpha (HIF-1alpha) protein is rapidly degraded by the ubiquitin-proteasome system under normoxic conditions. Its stabilization by hypoxia depends on redox-induced changes.* J Biol Chem, 1997. **272**(36): p. 22642-7.

464. Sutter, C.H., E. Laughner, and G.L. Semenza, *Hypoxia-inducible factor 1alpha protein expression is controlled by oxygen-regulated ubiquitination that is disrupted by deletions and missense mutations.* Proc Natl Acad Sci U S A, 2000. **97**(9): p. 4748-53.

465. Majmundar, A.J., W.J. Wong, and M.C. Simon, *Hypoxia-inducible factors and the response to hypoxic stress.* Mol Cell, 2010. **40**(2): p. 294-309.

466. Siddiq, A., L.R. Aminova, and R.R. Ratan, *Hypoxia inducible factor prolyl 4-hydroxylase enzymes: center stage in the battle against hypoxia, metabolic compromise and oxidative stress.* Neurochem Res, 2007. **32**(4-5): p. 931-46.

467. Lieb, M.E., et al., *Mammalian EGLN genes have distinct patterns of mRNA expression and regulation.* Biochem Cell Biol, 2002. **80**(4): p. 421-6.

468. Willam, C., et al., *HIF prolyl hydroxylases in the rat; organ distribution and changes in expression following hypoxia and coronary artery ligation.* J Mol Cell Cardiol, 2006. **41**(1): p. 68-77.

469. Negri, A.L., *Role of prolyl hydroxylase/HIF-1 signaling in vascular calcification.* Clin Kidney J, 2023. **16**(2): p. 205-209.

470. Gupta, N. and J.B. Wish, *Hypoxia-Inducible Factor Prolyl Hydroxylase Inhibitors: A Potential New Treatment for Anemia in Patients With CKD.* Am J Kidney Dis, 2017. **69**(6): p. 815-826.

471. Zheng, J., et al., *HIF-1alpha in myocardial ischemia-reperfusion injury (Review).* Mol Med Rep, 2021. **23**(5).

472. McMahon, S., et al., *Transforming growth factor beta1 induces hypoxia-inducible factor-1 stabilization through selective inhibition of PHD2 expression.* J Biol Chem, 2006. **281**(34): p. 24171-81.

473. Lun, J., et al., *Hypoxia inducible factor prolyl hydroxylases in inflammatory bowel disease.* Front Pharmacol, 2023. **14**: p. 1045997.

474. Zhang, H., R. Xu, and Z. Wang, *Contribution of Oxidative Stress to HIF-1-Mediated Profibrotic Changes during the Kidney Damage.* Oxid Med Cell Longev, 2021. **2021**: p. 6114132.

475. Requena-Ibanez, J.A., et al., *Prolyl Hydroxylase Inhibitors: a New Opportunity in Renal and Myocardial Protection.* Cardiovasc Drugs Ther, 2022. **36**(6): p. 1187-1196.

476. Ravi, R., et al., *Regulation of tumor angiogenesis by p53-induced degradation of hypoxia-inducible factor 1alpha.* Genes Dev, 2000. **14**(1): p. 34-44.

477. Gruber, M., et al., *Acute postnatal ablation of Hif-2alpha results in anemia.* Proc Natl Acad Sci U S A, 2007. **104**(7): p. 2301-6.

478. Rankin, E.B., et al., *Hypoxia-inducible factor-2 (HIF-2) regulates hepatic erythropoietin in vivo.* J Clin Invest, 2007. **117**(4): p. 1068-77.

479. Takeda, K., et al., *Regulation of adult erythropoiesis by prolyl hydroxylase domain proteins.* Blood, 2008. **111**(6): p. 3229-35.

480. Hickey, M.M., et al., *von Hippel-Lindau mutation in mice recapitulates Chuvash polycythemia via hypoxia-inducible factor-2alpha signaling and splenic erythropoiesis.* J Clin Invest, 2007. **117**(12): p. 3879-89.

481. Box, A.H. and D.J. Demetrick, *Cell cycle kinase inhibitor expression and hypoxia-induced cell cycle arrest in human cancer cell lines.* Carcinogenesis, 2004. **25**(12): p. 2325-35.

482. Gordan, J.D., et al., *HIF-2alpha promotes hypoxic cell proliferation by enhancing c-myc transcriptional activity.* Cancer Cell, 2007. **11**(4): p. 335-47.

483. Hubbi, M.E., et al., *MCM proteins are negative regulators of hypoxia-inducible factor 1.* Mol Cell, 2011. **42**(5): p. 700-12.

484. Baba, M., et al., *Loss of von Hippel-Lindau protein causes cell density dependent deregulation of CyclinD1 expression through hypoxia-inducible factor.* Oncogene, 2003. **22**(18): p. 2728-38.

485. Gardner, L.B., et al., *Hypoxia inhibits G1/S transition through regulation of p27 expression.* J Biol Chem, 2001. **276**(11): p. 7919-26.

486. Green, S.L., R.A. Freiberg, and A.J. Giaccia, *p21(Cip1) and p27(Kip1) regulate cell cycle reentry after hypoxic stress but are not necessary for hypoxia-induced arrest.* Mol Cell Biol, 2001. **21**(4): p. 1196-206.

487. Goda, N., et al., *Hypoxia-inducible factor 1alpha is essential for cell cycle arrest during hypoxia.* Mol Cell Biol, 2003. **23**(1): p. 359-69.

488. Semenza, G.L., *Defining the role of hypoxia-inducible factor 1 in cancer biology and therapeutics.* Oncogene, 2010. **29**(5): p. 625-34.

489. Balamurugan, K., *HIF-1 at the crossroads of hypoxia, inflammation, and cancer.* Int J Cancer, 2016. **138**(5): p. 1058-66.

490. Scholz, C.C. and C.T. Taylor, *Targeting the HIF pathway in inflammation and immunity.* Curr Opin Pharmacol, 2013. **13**(4): p. 646-53.

491. Cramer, T., et al., *HIF-1alpha is essential for myeloid cell-mediated inflammation.* Cell, 2003. **112**(5): p. 645-57.

492. Mantovani, A., et al., *Cancer-related inflammation.* Nature, 2008. **454**(7203): p. 436-44.

493. Mamlouk, S. and B. Wielockx, *Hypoxia-inducible factors as key regulators of tumor inflammation.* Int J Cancer, 2013. **132**(12): p. 2721-9.

494. Noman, M.Z., et al., *Microenvironmental hypoxia orchestrating the cell stroma cross talk, tumor progression and antitumor response.* Crit Rev Immunol, 2011. **31**(5): p. 357-77.

495. Reuter, S., et al., *Oxidative stress, inflammation, and cancer: how are they linked?* Free Radic Biol Med, 2010. **49**(11): p. 1603-16.

496. Ushio-Fukai, M. and Y. Nakamura, *Reactive oxygen species and angiogenesis: NADPH oxidase as target for cancer therapy.* Cancer Lett, 2008. **266**(1): p. 37-52.

497. Moniz, S., J. Biddlestone, and S. Rocha, *Grow(2): the HIF system, energy homeostasis and the cell cycle.* Histol Histopathol, 2014. **29**(5): p. 589-600.

498. Louis, K. and A. Hertig, *How tubular epithelial cells dictate the rate of renal fibrogenesis?* World J Nephrol, 2015. **4**(3): p. 367-73.

499. Sandau, K.B., H.G. Faus, and B. Brune, *Induction of hypoxia-inducible-factor 1 by nitric oxide is mediated via the PI 3K pathway.* Biochem Biophys Res Commun, 2000. **278**(1): p. 263-7.

500. Gomez, I.G., et al., *Anti-microRNA-21 oligonucleotides prevent Alport nephropathy progression by stimulating metabolic pathways.* J Clin Invest, 2015. **125**(1): p. 141-56.