

# **Bioinformatics in Traditional Chinese Medicine (TCM):**

Potential anti-cancer mechanisms of Compound Kushen  
Injection (CKI)

By  
Hanyuan Shen



THE UNIVERSITY  
*of* ADELAIDE

Department of Molecular and Biomedical Science  
School of Biological Sciences

A thesis presented for the degree of DOCTOR OF PHILOSOPHY

APRIL 2019

## **Abstract:**

With thousands of years of clinical practice, Traditional Chinese Medicine (TCM) is an enormous resource for both the pharmaceutical industry and daily health care. However, the wide popularization and application of TCM are hindered by the ambiguous explanation of mechanisms with ancient Chinese concepts. In addition, modern pharmacologic methods based on the interaction between single compound drug and target are inadequate to deal with the complex mixtures for TCM formulas which usually contain plant secondary metabolites from several or even dozens of herbs. New high-throughput technologies and bioinformatics methods can provide systematic and holistic ways to understand biological processes. Applying these methods to TCM research, we can clarify complex biological processes that result from hundreds or thousands of molecular interactions between components in TCM and targets in the organism. Therefore, the purpose of my project is to use models for the application of high-throughput sequencing technologies and bioinformatics methods in order to understand the molecular basis of TCM that involve drug-drug and compound-compound interactions.

My model TCM, Compound Kushen Injection (CKI) is an anticancer agent clinically used in China since 1995. It's commonly used as an adjuvant medicine in the treatment of carcinomas for pain relief, activation of innate immune response and reducing side effects of chemo or radiotherapy. Extracted from two herbs, CKI contains multiple alkaloids and flavonoids, which have been shown to be bioactive in previous studies. However, with the exception of several purified, well characterised compounds, the underlying mechanisms of action for CKI are still unclear.

In this thesis, I first applied transcriptome analysis and bioinformatics methods as part of a pipeline to investigate interactions between CKI and chemotherapy drugs. With this pipeline, the mechanisms for the opposing effects of CKI combined with doxorubicin compared to 5-fluorouracil (5-Fu) were determined, and potential interactions between CKI and chemotherapeutic anticancer agents were revealed. These results are closely related to the clinical usage of CKI and may help refine its clinical application. As my second approach, I applied transcriptome analysis to investigate the role of the two plant extracts that make up CKI

in order to determine which plant extract contains the primary bioactivity and to identify how the two plant extracts interact to generate the combined effects from CKI.

Altogether, this thesis presents approaches for the application of transcriptome analysis in order to identify the molecular mechanisms perturbed by CKI. I have successfully applied systems-biology based approaches to analyse herb-drug interactions and herbal compatibility and demonstrated these methods are valuable additions to TCM research. In addition, my results have indicated that high-throughput sequencing technologies and bioinformatics methods are powerful tools for linking TCM with modern pharmacologic methods.

## Dedication and Acknowledgements

As the first time going abroad, my Ph.D. study could not be accomplished without the support and guidance from many truly kind and exceptional people. I'd like to express my gratitude to all of them, particularly the following.

First of all, I would like to thank David Adelson and Dan Kortschak. I am so lucky to have you as my supervisors and you have given and taught me much more than just supervising during the last four years. Dave, from application to conducting experiments and until editing my manuscripts, you are always ready to help me no matter how full your schedule is. Your optimism to life and encouragement to my study makes me full of joy and motivation every day. Dan, your wisdom and wide learning sets a good model for my progress and I am so grateful for your suggestions and efforts to my research.

All the lab members are immensely helpful for all matters during these years, both personal and professional. Zhipeng and Yuka, your timely advice and patient help are the reasons I can conduct my research smoothly. Jian and Thazin, we studied and had fun together in Adelaide and I will always remember your help during four years. Also, I want to appreciate all the friends in the lab, you together makes the lab such an enjoyable place to stay every day.

I would also like to thank Anping Li, Minghua Li and Zhendong company. You introduced the opportunity and encouraged me to pursue a Ph.D. degree in Adelaide. My research project could not be finished without your support.

Special thanks to my family. For my parents, you sacrificed so much to raise me and words cannot express my gratitude to you. I will always remember the support from my parents in law and wish my father in law has peace in heaven. Shuping, my wife, I love you so much and admire your courage to give up your career in China to accompany me and bring our little angel to the world. Finally, to my daughter Wanyu, your lovely smile melted all the tiredness and upsets during my thesis writing.

## Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

SIGNED: .....

.....

DATE: .....

29/3/2019

# Table of Contents

<b>1 The Application of bioinformatics and -omics techniques in research for traditional Chinese medicine (TCM) and Compound Kushen Injection (CKI): a review</b> .....	<b>1</b>
<b>2 A New Strategy for Identifying Mechanisms of Drug-drug Interaction Using Transcriptome Analysis: Compound Kushen injection as a Proof of Principle</b> .....	<b>28</b>
<b>3 Comprehensive Understanding the Compatibility Mechanisms in Compound Kushen Injection with Transcriptome Analysis</b> .....	<b>62</b>
<b>4 The Effect of Compound Kushen Injection on Cancer Cells: Integrated Identification of Candidate Molecular Mechanisms</b> .....	<b>94</b>
<b>5 Cell Cycle, Glycolysis and DNA Repair Pathways in Cancer Cells are Suppressed by Compound Kushen Injection</b> .....	<b>124</b>
<b>6 Fractional Deletion of Compound Kushen Injection, a Natural Compound Mixture, Indicates Cytokine Signaling Pathways are Critical for Its Perturbation of The Cell Cycle</b> .....	<b>137</b>
<b>7 Conclusions and Future Directions</b> .....	<b>176</b>
<b>A Supplementary Tables and Data</b> .....	<b>179</b>
<b>B Supplementary Figures for Chapter 2</b> .....	<b>182</b>
<b>C Supplementary Figures for Chapter 3</b> .....	<b>188</b>
<b>D Supplementary Figures for Chapter 4</b> .....	<b>193</b>
<b>E Supplementary Figures for Chapter 5</b> .....	<b>199</b>
<b>F Supplementary Figures for Chapter 6</b> .....	<b>202</b>

# Chapter 1

## Introduction

In this chapter, I provide an overview of the application of bioinformatics to the study of TCM. First, online databases and network approaches for TCM research are introduced as the basis for -omics based research. Because -omics techniques can characterise biological processes at a whole system level, they are powerful tools for revealing the underlying mechanisms of TCM which is believed to act upon many targets in the system. I then review the current applications of large-scale methods in the study of TCM and introduce the current state of knowledge with respect to CKI and explain why it is a suitable model medicine in TCM research. Finally, my research hypothesis and aims are described.

# Statement of Authorship

Title of Paper.	The Application of bioinformatics and -omics techniques in research for traditional Chinese medicine (TCM) and Compound Kushen Injection (CKI): a review
Publication Status	<input type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input checked="" type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	

## Principal Author

Name of Principal Author (Candidate)	Hanyuan Shen		
Contribution to the Paper	Analysed literature, prepare figures and wrote manuscript		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	1/4/2019

## Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	David L. Adelson		
Contribution to the Paper	Supervised the development of work and assisted in writing the manuscript.		
Signature		Date	1/4/2019



# **The Application of bioinformatics and -omics techniques in research for traditional Chinese medicine (TCM) and Compound Kushen Injection (CKI): a review**

## **1. Literature review**

### **Abstract**

As a comprehensive medical practice system, traditional Chinese medicine (TCM) has been used in China for thousands of years. Even though widely applied in clinics, it is still a challenge to explain how TCM works. This difficulty impedes its acceptance and development. Due to the compositional complexity and relatively low concentration for majority of natural compounds in TCM formulas, it is hard to completely reveal the mechanism of TCM by following the drug-target-effect model in standard chemical pharmaceutical research. The emergence of systems biology has provided a potential solution to bridge TCM and modern science, and is drawing more and more attention in TCM research. This review summarizes the application of bioinformatics and integrative methods in the study of TCM mechanisms. It is likely that research combining TCM and bioinformatics can stimulate new approaches to complex problems in both areas as well as to modern medicine.

### **1.1 Introduction**

With the “back to nature” trend and popularity of complementary medicines, herbal remedies are widely used both in China and Western countries [1, 2]. Traditional Chinese medicine (TCM) is a valuable resource for the medical industry as well as human wellbeing given its use of

naturally derived products and rich accumulated clinical experience over thousands of years of practice [3]. Although TCM is consistently used by many people, the mechanism of most herbal medicines is still unable to be explained completely by modern scientific methods due to the complexity of their components [2].

Typically, TCM research borrows methods from chemistry and pharmacology [4-6]. Because of the complexity of the components in herbs and the constraints of analytical instruments, it is almost impossible to qualitatively and quantitatively analyze all compounds in a prescription, which may contain several or even dozens of herbs [7]. Therefore, separating single natural compounds by chemical purification followed by screens for bioactivity is a widely accepted research methodology to determine the molecular mode of action of TCM preparations. A well-known example is the discovery that artemisinin can cure malaria by Tu Youyou[8] who won the 2015 Nobel Prize. However, this classical research route is limited because the efficiency of this method is very low and single compounds usually do not represent the effects of the whole herbs or prescriptions [9, 10]. Hence, Chinese researchers also try to reveal the mechanism of TCM based on whole prescription and link TCM theory with modern medicine. However, except for combinatorial analysis of multiple compounds [11, 12], the pharmacological testing and animal model research on prescriptions and TCM theory have not gained acceptance in Western countries. Therefore, it is important to find a suitable way to link modern medical science and the ancient concept of holism to try to explain and develop TCM.

Prescription based TCM research should consider TCM theory, which regards the human body as a whole and treats disease by keeping the harmony between different parts [1]. Systems biology concepts of integration and system outcomes are a natural fit for the study of TCM [13,

14]. Specifically, this approach allows pharmaceutical researchers to overcome the restriction of specific single targets and single molecules and study the biological system with simultaneous perturbations of hundreds of linked targets [15]. Integrative techniques including genomics, transcriptomics, proteomics and metabolomics have recently begun to be applied to TCM research. At the same time, bioinformatics is a useful tool to deal with the massive data sets produced in these experiments [16]. With these new biological approaches and rapidly developing techniques, the molecular mechanisms of TCM may be explained in the near future.

Because of the complexity of TCM, it is important to choose a suitable formula for mechanistic research based on a whole prescription. I have chosen Compound Kushen Injection (CKI) as a candidate prescription for systems biology investigations, which is administered to more than 30,000 patients every day as one of the most widely used TCM injections for treatment of cancer in China [17,18]. There are three main reasons for choosing CKI for this. Firstly, compared with typical decoctions that may contain ingredients from dozens of sources, TCM injections are usually only produced from less than three ingredients and therefore are easier to analyze. Secondly, different modes of administration make TCM injections attractive as they can avoid the influence of intestinal flora and the different bioavailability of various natural compounds [19-21]. As a result, the compounds in the injection are those circulating *in vivo*, so cell-based experiments are a reasonable way of examining their activity. Finally, the manufacturing processes and raw materials of TCM injections are much stricter than traditional decoctions, improving the reproducibility of experimental results. In short, TCM injections are a good choice for beginning to understand the mechanisms of TCM. The methods used here may be subsequently applied to research on other TCM formulas.

## 1.2 Composition and current understanding of CKI

CKI (Compound Kushen Injection), a modern Chinese herb preparation based on TCM theory, is widely used in China as an adjuvant with chemotherapeutic or radiotherapeutic treatments for cancer [17, 18]. It has been reported that CKI can suppress the growth of tumors, relieve pain caused by advanced cancer and improve immunity. Zhao *et al.* showed that down-regulation of the phosphorylation of ERK and AKT kinases and blocking TRPV1 signaling by CKI was consistent with an anti-tumor effect as well as reducing inflammation and pain [22]. CKI can enhance the level of IgA, IgM, IgG, IL2, IL4, and IL10 and decrease IL6 and TNF- $\alpha$  levels *in vivo*, leading to improved immunity [23]. In combined use with chemotherapeutic agents, literature shows that CKI can significantly inhibit proliferation of MCF-7 side population (SP) cells through suppression of the Wnt/ $\beta$ -catenin pathway while cisplatin just suppressed non-SP cells, supporting the rationale of treating breast cancer by co-administration of CKI and cisplatin [24]. These different mechanisms indicate that as a TCM preparation, different bioactive compounds in CKI may regulate different targets to produce a systemic anti-tumor effect.

CKI is produced from extracts of two herbs: Kushen (*Radix Sophorae flavescens*) and Baituling (*Rhizoma Smilacis glabrae*), therefore CKI can be considered as a complex sample. Yue *et al.* applied high-performance liquid chromatography (HPLC) and LC-DAD-MS/MS to analyze 27 CKI samples from different batches and identified 21 components, 19 from Kushen and 2 from Baituling [25]. As one of the common herbs used in TCM, natural compounds in Kushen have previously been shown to exhibit various bioactivities. Alkaloids (mainly matrine and oxymatrine) have been shown to manifest anti-cancer [26, 27], antifungal [28], anti-inflammatory [29] and cell protective [30, 31] effects in different studies. More recently, flavonoids in Kushen have been reported to have better anti-tumor effects than Kushen alkaloids [20], as they can

suppress both the growth of cancer cells [32] and tumor angiogenesis [33]. Compared to Kushen, the literature on Baituling is relatively limited. From existing research, the active ingredients in Baituling are believed to be flavonoids such as taxifolin and astilbin[34]. As a major component in CKI and with no reported anti-tumor effects, Baituling may contribute to improved patients' quality of life and the effects of Kushen through improving immune system function [35-37] and anti-inflammatory activity [38, 39].

### **1.3 The application of bioinformatics to TCM research**

With the application of integrative bioinformatics techniques and new analytical methods, data sets in TCM research are increasing rapidly [16]. Bioinformatics, a fast developing interdisciplinary science of biology and computer science, is useful for storing, retrieving and analyzing these data and transforming data into knowledge [40, 41]. At the same time, as an obligatory tool in systems biology research, bioinformatics can help to identify the mechanisms of TCM in a network biology context.

#### **1.3.1 The databases of natural compounds and their targets**

Natural products are a very important source of new medicines, contributing to 60% of new drugs from 1981 to 2010 [45]. Although there are usually thousands of components in one herb or TCM formula, not all of them contribute to the effect of the formula, and it is likely that homologs may share similar bioactivity [42]. It is, therefore, a reasonable first step to identify the main compounds and their bioactivity to acquire basic knowledge. Furthermore, structural information of natural compounds is very important for molecular docking experiments [12, 43, 44]. As a result of decades of TCM research, a huge amount of data about natural compounds

in TCM has been accumulated and many databases (see Table 1.) are now searchable for components and bioactivity based on single compounds, herbs or even commonly used formulas.

Table 1 Databases about TCM and natural products

Name	Country	Contents	Websites
TCMD	China	Information about 9127 compounds with 3-D structure and 3922 TCM herbs	Commercial Database
CMEMCPD	China	About 50000 formulas, 1400 diseases with available medicine, 22000 TCM drugs and 19700 compounds	<a href="http://www.chemcpd.csdb.cn/scdb/main/tcm_introduce.asp">http://www.chemcpd.csdb.cn/scdb/main/tcm_introduce.asp</a>
CNPD	China	About 57000 natural products and their information	Commercial Database
NCMI	China	About 10000 compounds in 4500 TCM drugs	<a href="http://pharmdata.ncmi.cn/cnpc/index.asp">http://pharmdata.ncmi.cn/cnpc/index.asp</a>
TCM-ID [46]	Singapore	Contains 1197 formulas, 1098 herbs and 9852 ingredients	<a href="http://bidd.nus.edu.sg/group/TCMsite/">http://bidd.nus.edu.sg/group/TCMsite/</a>
Natural Products Alert[47]	USA	More than 20000 species of plants and related information	<a href="https://www.napralert.org">https://www.napralert.org</a>
KNApSack Family Databases [48]	Japan	Including 20741 species and 59048 metabolites	<a href="http://kanaya.naist.jp/KNApSack_Family/">http://kanaya.naist.jp/KNApSack_Family/</a>
CHEM-TCM	UK	Information about 9500 compounds from 350 TCM	<a href="http://www.chemtcm.com/">http://www.chemtcm.com/</a>

The effect of TCM relies on small molecules to bind to proteins or nucleic acid targets and then modify metabolic processes in human body. As a result, it is an important part of TCM research to identify the targets influenced by natural compounds. Bioinformatics provides two linked approaches to identify drug targets: computer prediction and database searching.

Computer prediction applies methods from computer-aided molecular drug design to predict potential targets based on the 3D structures of compounds and proteins. Commonly used computer prediction methods are inverse docking and chemical similarity searching. Inverse docking uses small molecules as probes to search databases for potential biological macromolecule binding partners [49]. Chemical similarity searching is based on the fact that small molecules with similar structures tend to bind to similar targets; searching for information on similar chemical compounds can help to predict targets of natural compounds [50]. These methods have been widely applied in TCM research from single natural compounds to formulas [51, 52].

Database searching is used to find target information for natural compounds based on experiments and computer predictions. Many such databases have been established in recent years. For TCM components, there are databases like HIT (Herbal Ingredients' Targets Database) and TCM-PTD (The potential target database of Traditional Chinese Medicine) that provide much useful target information. For example, HIT contains 586 natural compounds from 1300 herbs and their 1301 protein targets [53]. In addition to natural compounds, chemical medicine target databases can also provide information for reference in TCM research. The Drugbank database (<http://www.drugbank.ca>) collects 7795 FDA approved and experimental medicines and 4313 related proteins while the STITCH database (<http://stitch.embl.de>) contains interactions between 300,000 small molecules and 2.6 million proteins [54, 55]. The combined use of these databases can help to predict TCM mechanisms to some extent [56].

### 1.3.2 Network based mechanism of TCM activity.

Due to the complexity of TCM, it is almost impossible to explain its mechanism with one or two pathways. As a result, more and more TCM researchers are turning from traditional target-based mechanism studies to network pharmacology research. Network pharmacology tries to model the process of disease and the interaction between medicine and organism from the viewpoint of systems biology. It studies the mechanism and promotes innovation in medicine through high-throughput data in large-scale data capture research, molecular data on the Internet and computer analysis [57]. At present, network pharmacology research is constrained by low integrity and accuracy of data for computer modeling and the difficulty of testing predicted models based on networks. Therefore, the development of network pharmacology requires new theories, methods and techniques. At this stage, there are two main approaches in network pharmacology: one is building networks of drug effects and diseases based on public databases and literature, and then using its predictions to find new targets. The other approach is to apply large-scale methods to acquire data about the interaction between drugs and experimental models, in order to build a network model of medical targets and disease.

Intuitively, genes and proteins that participate in the disease process should be drug targets. However, by building and analyzing a network of drug targets and diseases, Yildirim *et al.* found that for some diseases related to endocrine function and blood, drug targets are more likely to be disease genes, while only a few targets are disease genes in cancer and gastrointestinal disease [58]. In the latter case, drugs usually target the proteins that interact with disease genes or block those interactions directly [59]. Therefore, recognizing disease genes and constructing disease networks is an important part of network pharmacology.



There are two main steps in constructing a disease network. First, gene information associated with diseases can be obtained from databases and the research literature. There are many databases that contain information on disease-related genes and are available to the public, such as OMIM (Online Mendelian Inheritance in Man) [60], GAD (Genetic Association Database)[61], DisGeNET [62] and GeneCards [63]. This databases allows one to extract information relating to the proteins encoded by these genes and construct networks. Sometimes, the genes directly related to known disease genes and their links can be extracted as a whole to build a complicated network, which can better express the progress of disease [64, 65]. Recently, researchers have also developed algorithms to predict disease-related genes based on known genes, such as direct network neighbors and neighborhood [66], k-step Markov method [67] and random walks [68].

Recognizing pathways or sub-networks modulated by drugs is the kernel of network pharmacology research. By identifying and analyzing the target network regulated by bioactive compounds in TCM, we can evaluate the effects on the disease network and then model the mechanism of TCM [69]. The target information acquired from TCM databases can be linked to pathway databases such as KEGG [70] and BioCarta [71] and the pathway enriched by targets can be considered as a candidate pathway regulated by the drug [72]. Once targets and their related pathways are recognized, they can be used to build the drug sub-network using similar methods for constructing disease networks [73, 74].

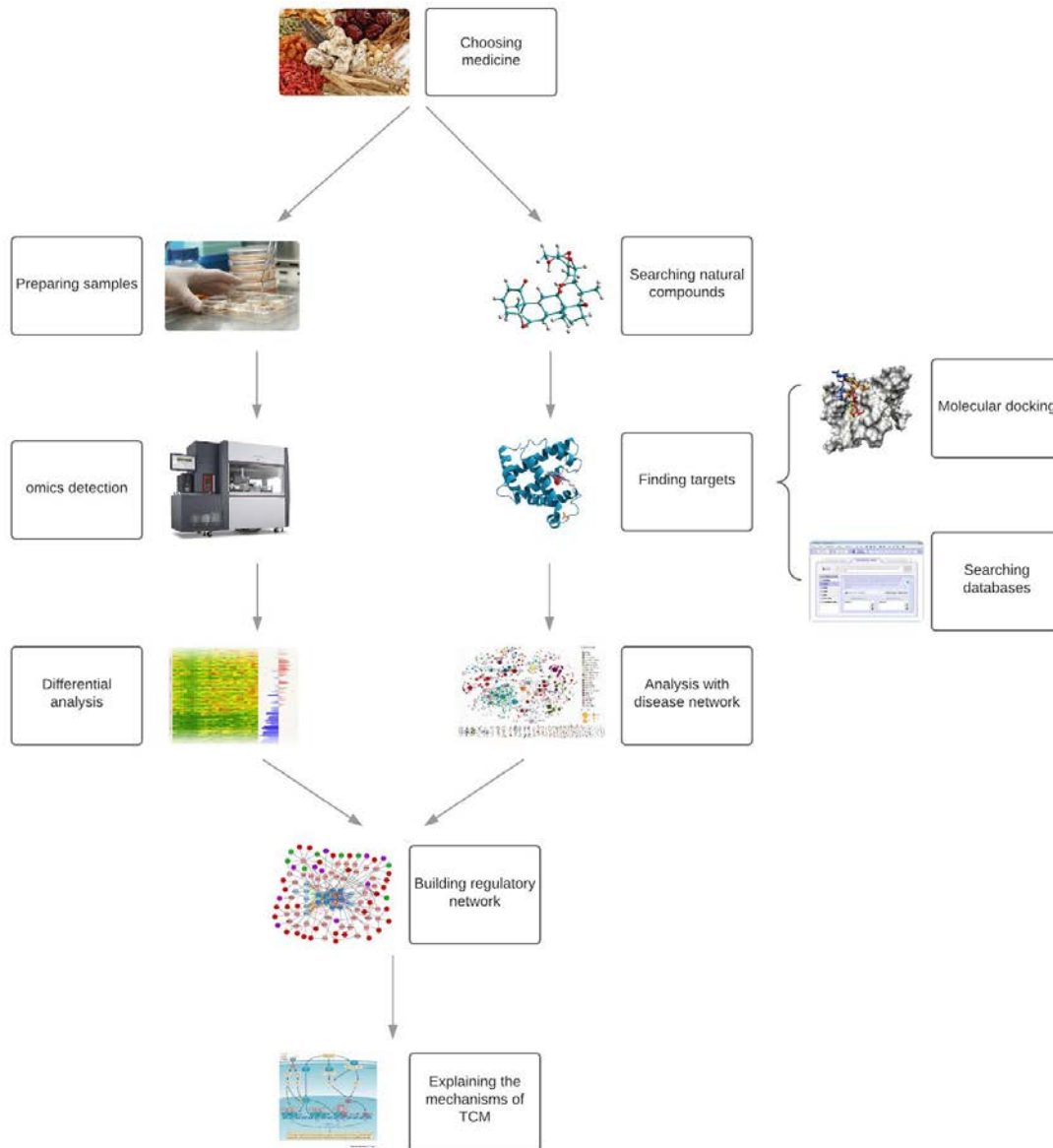


Figure 1: Two pipelines of explaining mechanisms of TCM based on networks.

#### 1.4 Large scale methods used in TCM research

In the past decade, systems biology based on large-scale data analysis methods have rapidly developed, especially with the publication of the human genome sequence and breakthroughs

in mass spectrometry (MS) technology. Genomics, proteomics and metabolomics have been widely used in pharmaceutical research in order to understand the responses to drugs at a whole system level [75]. Due to the component complexity of TCM, a system level approach should be a more suitable solution for TCM research than the traditional study of interactions between single genes or proteins and drugs. Furthermore, large-scale data may also provide solutions for personalized healthcare, something that is emphasized by TCM theory. As a result, there has been a significant amount of TCM research based on large-scale integrative methods.

#### 1.4.1 Epigenomics

Epigenetics refers to the inherited alteration in phenotype or gene expression without DNA sequence changes [76], and regulates the time, location and pattern of gene expression [77]. Through epigenetics, parental environmental factors can influence the phenotype of offspring by DNA methylation, histone modification, chromatin remodeling, gene silencing and RNA editing [78]. Disorders of epigenetics have been linked to many diseases such as cancer [79] and cardiovascular disease [80]. Because epigenetic changes are reversible, it is a new area for pharmaceutical design to treat diseases that cannot be cured by current drugs or gene therapy [81, 82].

Epigenomics refers to the analysis of epigenetic changes across the genome and at present is mainly focused on DNA methylation, histone modification and action of miRNAs. In 2003, the Human Epigenome Project (HEP) was started in order to identify, catalog and interpret the genome-wide DNA methylation pattern in all major human tissues [83]. Later, the emergence of next-generation sequencing (NGS) in 2005 accelerated epigenomics research, by allowing the generation of high resolution and throughput sequencing data using much less time and money

[84]. Combined with existing research methods, NGS can be used to detect epigenetically modified genome regions [85-87]. There are now a number of epigenomics research methods based on NGS such as whole genome bisulfite sequencing (WGBS), methylated DNA immunoprecipitation-sequencing (MeDIP-seq) for DNA methylation and chromatin immunoprecipitation-sequencing (ChIP-seq) for histone modification.

Epigenetic changes result from environmental effects and can cause many chronic diseases. For example, the initiation and progression of cancer are associated with multiple gene expression changes such as silencing of tumor suppressor genes or activation of oncogenes [88]. This phenomenon may be linked to the practice of disease prevention and food therapy in TCM theory. There is also accumulating evidence indicating that natural compounds and TCM formulas can influence gene expression [89, 90]. For example, triptolide from the thunder god vine can decrease histone methylation and increase acetylation leading to the induction of apoptosis in cancer cells [91] and genistein has anti-cancer activity by regulating miRNAs both *in vivo* and *in vitro* [92, 93]. Although at present there is limited epigenomics research on TCM formulations, these results indicate that epigenomics studies could be important in revealing the mechanisms of TCM and to help find new drug candidates that target the epigenome.

#### **1.4.2 Transcriptomics**

Transcriptomics is the study of the complete set of RNA transcripts, which is produced by the genome under specific circumstances or in a specific cell type [94]. RNA transcripts are collectively referred to as the transcriptome and include mRNA, and rRNA, tRNA and other non-coding RNAs [95]. The transcriptome can vary with physiological status or as a result of external factors. Transcriptomics studies can help to determine where and when genes are

turned on or off, and therefore, are very useful tools for characterizing the mechanisms of disease progression and therapeutic interventions.

The complexity and size of the transcriptome call for high-throughput data acquisition techniques, and there are two main approaches to acquire transcriptomics data. One method, microarray analysis, is based on hybridization of fluorescently labeled cDNA to an array of sequence probes [96]. However, more recently with the development of next-generation sequencing technique, RNA-seq has become the most commonly used method in transcriptomics [99].

RNA expression level is one of the most common experimental readouts in biomedical science. However, the number of genes for measuring expression is restricted for methods like quantitative real-time PCR and northern blotting. This is a particular limitation for TCM research, where hundreds of compounds have the potential to perturb tens or hundreds of targets. Transcriptomics overcomes this limitation and allows one to examine genome-wide gene expression changes, so it can be used in TCM research areas such as mechanism study [100, 101] and active compound detection [102]. For example, Li *et al.* explored the mechanism of the TCM formula Qi-Shen-Yi-Qi for the treatment of myocardial infarction. Using a microarray assay for gene expression, they found 55 potential targets of Qi-Shen-Yi-Qi within which 14 were confirmed in the literature [103]. Liu *et al.* also used microarrays to profile the transcriptome of breast cancer cells and found that Si-Wu-Tang has phytoestrogenic activity [104]. RNA-seq, with its increased sensitivity and ability to detect ncRNAs, is the current technology of choice for TCM-related transcriptomics.

### 1.4.3 Proteomics

The proteome is the entire set of proteins expressed by the whole genome at the cell, system or organism level, and proteomics refers to the study of the structures and functions of the proteome [105, 106]. Proteomics research mainly relies on two methodologies to separate and identify proteins. One is based on the separation of proteins using two-dimensional electrophoresis combined with fluorescence analysis for expression measurement and mass spectrometer based protein identification from the molecular weight of tryptic peptide [107]. The second method is based on peptides, and uses enzymes to break down proteins and then analyze with liquid chromatography-mass spectrometry (LC-MS) [108]. With the rapid development of mass spectrometer technology and a large demand for comparative analysis of different proteomes, quantitative proteomics has been applied widely in disease biomarker identification and pathway analysis as well as drug discovery [109, 110]. Stable isotope labeling and label-free methods are two common techniques in quantitative proteomics based on mass spectrometry [111]. At present, the label-free method is favored because of its simpler process and relatively low cost [112].

Proteomics has been used to study drug mechanisms, find therapeutic targets and identify new biomarkers of disease development and drug candidates. Combined with TCM research, proteomics has been used to find potential targets of TCM compounds or formulations by comparing the differences in the proteome for normal, pathological and post-treatment samples. For example, Liu *et al.* revealed the mechanism of tanshinone IIA activity against nephropathy induced by doxorubicin using proteomics [113]. In addition, proteomics data together with information from databases about protein interactions can be used to build pathways/networks

affected by TCM to model treatment mechanisms. These models can then be verified by validating the functions of key proteins in these networks [114, 115].

#### 1.4.4 Metabolomics

Metabolomics refers to the qualitative and quantitative analysis of all endogenous metabolites produced by biological systems [116]. By analyzing changes in metabolites and metabolic pathways, metabolomics can help explain the molecular phenotype and mechanism of biological events [117]. From nuclear magnetic resonance spectroscopy (NMR) to chromatography-mass spectrometry, multiple analytical instrument platforms have been used in metabolomics. NMR was the first high-throughput technology to be applied in metabolomics analysis, with the advantage that it can examine samples directly without destroying them [118]. However, higher sensitivity and throughput makes chromatography-mass spectrometry the most commonly used metabolomics platform including GC-MS and LC-MS for volatile and derivative metabolites [119, 120]. For maximum coverage of metabolites, the interactive use of different platforms is also widely used [121]. The complexity of the metabolome requires multiple bioinformatics methods such as pattern recognition techniques in data processing [122] and databases for pathway analysis [123].

Metabolomics may be the most commonly applied large-scale approach in TCM research. In addition to the system-wide view, the convenience of sample collection from urine or plasma makes it easier to observe dynamic changes *in vivo*, which relates better to TCM theory [124]. Bioactive compounds from herbs can be metabolized in the body or by intestinal flora, and sometimes these metabolic products may have better treatment activity or perhaps toxic effects

[125]. Therefore, metabolomics has been widely used to evaluate safety [126, 127], explain mechanisms [128, 129] and study formulations [130] in TCM research.

### **1.5 Integration of different large scale data approaches**

Disease progression and treatment effects can be expressed at different biological levels; genome, epigenome, transcriptome, proteome and metabolome. These different levels, while linked are not always well correlated or consistent. Previous work has shown that data from one large-scale approach cannot be used to predict effects seen with another, but those different technologies can provide complementary data [131-133]. Therefore, analysis using a single approach may only reflect changes in a limited range whereas multiple approaches can help attain more complete and systematic knowledge of disease processes, which can be widely applied in drug development, clinical diagnosis and personalized treatment [134].

Data processing is a major obstacle for the application of multi-approach analysis. To put it simply, the first step for integrating multi-approach analysis is to process and analyze the data at each level first and then connect the analyses across the different levels. Then, biomarkers or pathway networks could be screened or established based on the connections between different levels with the ultimate goal of building quantitative models for predicting the effects of various factors and treatments [135]. Integration of multiple levels is still in its infancy and the development of analysis methods in this area is extremely active.



## 1.6 Conclusion

While traditional methods based on pharmacological experiments and chemical analysis have achieved significant outcomes in TCM research, because multiple compounds affect multiple targets in TCM, there is still a huge gap in our understanding of its modes of action. Because of similar holistic views, systems biology methods may help fill this gap by providing a network context for TCM research. From virtual model building to different large-scale analyses, these new methods are slowly changing the direction of TCM research. Hopefully, TCM can be explained scientifically with the application of systems biology and contribute more widely to human health.

## 2. Hypothesis

The **central hypothesis** for this study is that CKI contains multiple anti-cancer components that can affect different pathways to generate anti-cancer effects and act synergistically with other cancer chemotherapeutic agents. These effects can be detected using next-generation transcriptome sequencing, and characterized by bioinformatic analysis.

## 3. Aims

**Aim 1: Determine whether CKI works synergistically with anti-cancer chemotherapeutic agents.**

The **working hypothesis** is that natural compounds in CKI can increase cancer cells' sensitivity to chemotherapeutic agents to induce better anti-tumor effects.

**Aim 2: Determine whether the two herbs used to make CKI have individual anti-cancer effects.**

The **working hypothesis** is that Kushen and Baitulin may individually have anti-cancer activity by causing apoptosis in cancer cells or stimulating immune cells.

**Aim 3: Determine whether the effects of CKI and single herb injections can be detected by transcriptomic analysis.**

The **working hypothesis** is that CKI and single herb injections can alter the cell transcriptome in order to help understand their effects.

## References

1. Chan, E., et al., *Interactions between traditional Chinese medicines and Western therapeutics*. *Curr Opin Drug Discov Devel*, 2010. **13**(1): p. 50-65.
2. Qiu, J., *Traditional medicine: a culture in the balance*. *Nature*, 2007. **448**(7150): p. 126-8.
3. Leung, A.Y., *Traditional toxicity documentation of Chinese Materia Medica--an overview*. *Toxicol Pathol*, 2006. **34**(4): p. 319-26.
4. Lu, S., et al., *The treatment of rheumatoid arthritis using Chinese Medicinal Plants: From pharmacology to potential molecular mechanisms*. *J Ethnopharmacol*, 2015.
5. Wu, L., H. Hao, and G. Wang, *LC/MS based tools and strategies on qualitative and quantitative analysis of herbal components in complex matrixes*. *Curr Drug Metab*, 2012. **13**(9): p. 1251-65.
6. Wang, C.Y., X.Y. Bai, and C.H. Wang, *Traditional Chinese medicine: a treasured natural resource of anti-cancer drug research and development*. *Am J Chin Med*, 2014. **42**(3): p. 543-59.
7. Song, Y., et al., *Large-scale qualitative and quantitative characterization of components in Shenfu injection by integrating hydrophilic interaction chromatography, reversed phase liquid chromatography, and tandem mass spectrometry*. *J Chromatogr A*, 2015. **1407**: p. 106-18.
8. Miller, L.H. and X. Su, *Artemisinin: discovery from the Chinese herbal garden*. *Cell*, 2011. **146**(6): p. 855-8.
9. Qiu, S., et al., *A green protocol for efficient discovery of novel natural compounds: Characterization of new ginsenosides from the stems and leaves of Panax ginseng as a case study*. *Anal Chim Acta*, 2015. **893**: p. 65-76.
10. Eberhardt, L., K. Kumar, and H. Waldmann, *Exploring and exploiting biologically relevant chemical space*. *Curr Drug Targets*, 2011. **12**(11): p. 1531-46.
11. Fu, H., et al., *Kinetic cellular phenotypic profiling: prediction, identification, and analysis of bioactive natural products*. *Anal Chem*, 2011. **83**(17): p. 6518-26.
12. Gong, P., et al., *Chemicalome and metabolome matching approach to elucidating biological metabolic networks of complex mixtures*. *Anal Chem*, 2012. **84**(6): p. 2995-3002.
13. Buriani, A., et al., *Omic techniques in systems biology approaches to traditional Chinese medicine research: present and future*. *J Ethnopharmacol*, 2012. **140**(3): p. 535-44.
14. Yan, S.K., et al., *"Omics" in pharmaceutical research: overview, applications, challenges, and future perspectives*. *Chin J Nat Med*, 2015. **13**(1): p. 3-21.
15. Auffray, C., Z. Chen, and L. Hood, *Systems medicine: the future of medical genomics and healthcare*. *Genome Med*, 2009. **1**(1): p. 2.
16. Gu, P. and H. Chen, *Modern bioinformatics meets traditional Chinese medicine*. *Brief Bioinform*, 2014. **15**(6): p. 984-1003.
17. Wu, L. et al. Synthesis and biological evaluation of matrine derivatives containing benzo- $\alpha$ -pyrone structure as potent anti-lung cancer agents. *Sci. Rep.* **6**, 35918 (2016).
18. Sun, Q., et al., *Meta-analysis: therapeutic effect of transcatheter arterial chemoembolization combined with compound kushen injection in hepatocellular carcinoma*. *Afr J Tradit Complement Altern Med*, 2012. **9**(2): p. 178-88.
19. Hao, D.C., et al., *Drug Metabolism and Pharmacokinetic Diversity of Ranunculaceae Medicinal Compounds*. *Curr Drug Metab*, 2015. **16**(4): p. 294-321.

20. Xu, J., et al., *Comparative metabolism of Radix scutellariae extract by intestinal bacteria from normal and type 2 diabetic mice in vitro*. J Ethnopharmacol, 2014. **153**(2): p. 368-74.
21. Zhang, L., et al., *Metabolic routes along digestive system of licorice: multicomponent sequential metabolism (MSM) method in rat*. Biomed Chromatogr, 2015.
22. Zhao, Z., et al., *Fufang Kushen injection inhibits sarcoma growth and tumor-induced hyperalgesia via TRPV1 signaling pathways*. Cancer Lett, 2014. **355**(2): p. 232-41.
23. Zhou, S.K., et al., *Antioxidant and immunity activities of Fufang Kushen Injection Liquid*. Molecules, 2012. **17**(6): p. 6481-90.
24. Xu, W., et al., *Compound Kushen Injection suppresses human breast cancer stem-like cells by down-regulating the canonical Wnt/beta-catenin pathway*. J Exp Clin Cancer Res, 2011. **30**: p. 103.
25. Yue Ma, H.G., Jing Liu, Liangmian Chen, Qiwei Zhang, Zhimin Wang, *Identification and determination of the chemical constituents in a herbal preparation, Compound Kushen Injection, by HPLC and LC-DAD-MS/MS*. Journal of Liquid Chromatography & Related Technologies, 2014. **37**: p. 207-220.
26. Liu, X.S. and J. Jiang, *Molecular mechanism of matrine-induced apoptosis in leukemia K562 cells*. Am J Chin Med, 2006. **34**(6): p. 1095-103.
27. Zhao, B., et al., *Effects of matrine on proliferation and apoptosis of cultured retinoblastoma cells*. Graefes Arch Clin Exp Ophthalmol, 2012. **250**(6): p. 897-905.
28. Liu, Q., et al., *Antifungal activity in plants from Chinese traditional and folk medicine*. J Ethnopharmacol, 2012. **143**(3): p. 772-8.
29. Zhang, B., et al., *Antiinflammatory effects of matrine in LPS-induced acute lung injury in mice*. Eur J Pharm Sci, 2011. **44**(5): p. 573-9.
30. Zhang, K., et al., *Neuroprotective effects of oxymatrine against excitotoxicity partially through down-regulation of NR2B-containing NMDA receptors*. Phytomedicine, 2013. **20**(3-4): p. 343-50.
31. Zhang, M., et al., *Oxymatrine protects against myocardial injury via inhibition of JAK2/STAT3 signaling in rat septic shock*. Mol Med Rep, 2013. **7**(4): p. 1293-9.
32. Sun, M.Y., et al., *[Anti-tumour activities of kushen flavonoids in vivo and in vitro]*. Zhong Xi Yi Jie He Xue Bao, 2008. **6**(1): p. 51-9.
33. Zhang, X.L., et al., *A novel flavonoid isolated from Sophora flavescens exhibited anti-angiogenesis activity, decreased VEGF expression and caused G0/G1 cell cycle arrest in vitro*. Pharmazie, 2013. **68**(5): p. 369-75.
34. Chen, L., et al., *Simultaneous quantification of five major bioactive flavonoids in Rhizoma smilacis glabrae by high-performance liquid chromatography*. J Pharm Biomed Anal, 2007. **43**(5): p. 1715-20.
35. Xu, Q., et al., *A new strategy for regulating the immunological liver injury--effectiveness of DTH-inhibiting agents on DTH-induced liver injury to picryl chloride*. Pharmacol Res, 1997. **36**(5): p. 401-9.
36. Cai, Y., T. Chen, and Q. Xu, *Astilbin suppresses collagen-induced arthritis via the dysfunction of lymphocytes*. Inflamm Res, 2003. **52**(8): p. 334-40.
37. Jiang, J. and Q. Xu, *Immunomodulatory activity of the aqueous extract from rhizome of Smilax glabra in the later phase of adjuvant-induced arthritis in rats*. J Ethnopharmacol, 2003. **85**(1): p. 53-9.
38. Jiang, J., et al., *Anti-inflammatory activity of the aqueous extract from Rhizoma smilacis glabrae*. Pharmacol Res, 1997. **36**(4): p. 309-14.
39. Hong, Q., et al., *Smilacis Glabrae Rhizoma reduces oxidative stress caused by hyperuricemia via upregulation of catalase*. Cell Physiol Biochem, 2014. **34**(5): p. 1675-85.

40. Schafer, J. and K. Strimmer, *An empirical Bayes approach to inferring large-scale gene association networks*. Bioinformatics, 2005. **21**(6): p. 754-64.
41. Quan, Y., et al., *Dissecting traditional Chinese medicines by omics and bioinformatics*. Nat Prod Commun, 2014. **9**(9): p. 1391-6.
42. Mayank and V. Jaitak, *Molecular docking study of natural alkaloids as multi-targeted hedgehog pathway inhibitors in cancer stem cell therapy*. Comput Biol Chem, 2015.
43. Hao, H., et al., *Global detection and identification of nontarget components from herbal preparations by liquid chromatography hybrid ion trap time-of-flight mass spectrometry and a strategy*. Anal Chem, 2008. **80**(21): p. 8187-94.
44. Jiang, L., et al., *Virtual Screening and Molecular Dynamics Study of Potential Negative Allosteric Modulators of mGluR1 from Chinese Herbs*. Molecules, 2015. **20**(7): p. 12769-86.
45. Newman, D.J. and G.M. Cragg, *Natural products as sources of new drugs over the 30 years from 1981 to 2010*. J Nat Prod, 2012. **75**(3): p. 311-35.
46. Chen, X., et al., *Database of traditional Chinese medicine and its application to studies of mechanism and to prescription validation*. Br J Pharmacol, 2006. **149**(8): p. 1092-103.
47. Ningthoujam, S.S., et al., *Challenges in developing medicinal plant databases for sharing ethnopharmacological knowledge*. J Ethnopharmacol, 2012. **141**(1): p. 9-32.
48. Afendi, F.M., et al., *KNAPSAcK family databases: integrated metabolite-plant species databases for multifaceted plant research*. Plant Cell Physiol, 2012. **53**(2): p. e1.
49. Lybrand, T.P., *Ligand-protein docking and rational drug design*. Curr Opin Struct Biol, 1995. **5**(2): p. 224-8.
50. Wassermann, A.M., E. Lounkine, and M. Glick, *Bioturbo similarity searching: combining chemical and biological similarity to discover structurally diverse bioactive molecules*. J Chem Inf Model, 2013. **53**(3): p. 692-703.
51. Zhao, J., et al., *Therapeutic effects of astragaloside IV on myocardial injuries: multi-target identification and network analysis*. PLoS One, 2012. **7**(9): p. e44938.
52. Zhao, S. and S. Li, *Network-based relating pharmacological and genomic spaces for drug target identification*. PLoS One, 2010. **5**(7): p. e11764.
53. Ye, H., et al., *HIT: linking herbal active ingredients to targets*. Nucleic Acids Res, 2011. **39**(Database issue): p. D1055-9.
54. Kuhn, M., et al., *STITCH: interaction networks of chemicals and proteins*. Nucleic Acids Res, 2008. **36**(Database issue): p. D684-8.
55. Wishart, D.S., et al., *DrugBank: a comprehensive resource for in silico drug discovery and exploration*. Nucleic Acids Res, 2006. **34**(Database issue): p. D668-72.
56. Fang, H., et al., *Bioinformatics analysis for the antirheumatic effects of huang-lian-jie-du-tang from a network perspective*. Evid Based Complement Alternat Med, 2013. **2013**: p. 245357.
57. Hopkins, A.L., *Network pharmacology*. Nat Biotechnol, 2007. **25**(10): p. 1110-1.
58. Yildirim, M.A., et al., *Drug-target network*. Nat Biotechnol, 2007. **25**(10): p. 1119-26.
59. Vassilev, L.T., et al., *In vivo activation of the p53 pathway by small-molecule antagonists of MDM2*. Science, 2004. **303**(5659): p. 844-8.
60. Hamosh, A., et al., *Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders*. Nucleic Acids Res, 2005. **33**(Database issue): p. D514-7.
61. Becker, K.G., et al., *The genetic association database*. Nat Genet, 2004. **36**(5): p. 431-2.
62. Bauer-Mehren, A., et al., *Gene-disease network analysis reveals functional modules in mendelian, complex and environmental diseases*. PLoS One, 2011. **6**(6): p. e20284.
63. Safran, M., et al., *GeneCards Version 3: the human gene integrator*. Database (Oxford), 2010. **2010**: p. baq020.

64. Li, S., et al., *Are Topological Properties of Drug Targets Based on Protein-Protein Interaction Network Ready to Predict Potential Drug Targets?* Comb Chem High Throughput Screen, 2015.
65. Hwang, S., et al., *A protein interaction network associated with asthma.* J Theor Biol, 2008. **252**(4): p. 722-31.
66. Oti, M., et al., *Predicting disease genes using protein-protein interactions.* J Med Genet, 2006. **43**(8): p. 691-8.
67. Chen, J., B.J. Aronow, and A.G. Jegga, *Disease candidate gene identification and prioritization using protein interaction networks.* BMC Bioinformatics, 2009. **10**: p. 73.
68. Kohler, S., et al., *Walking the interactome for prioritization of candidate disease genes.* Am J Hum Genet, 2008. **82**(4): p. 949-58.
69. Li, X., et al., *Valeriana jatamansi constituent IVHD-valtrate as a novel therapeutic agent to human ovarian cancer: in vitro and in vivo activities and mechanisms.* Curr Cancer Drug Targets, 2013. **13**(4): p. 472-83.
70. Du, J., et al., *KEGG-PATH: Kyoto encyclopedia of genes and genomes-based pathway analysis using a path analysis model.* Mol Biosyst, 2014. **10**(9): p. 2441-7.
71. Han, J., et al., *ESEA: Discovering the Dysregulated Pathways based on Edge Set Enrichment Analysis.* Sci Rep, 2015. **5**: p. 13044.
72. 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.
73. Li, S., B. Zhang, and N. Zhang, *Network target for screening synergistic drug combinations with application to traditional Chinese medicine.* BMC Syst Biol, 2011. **5 Suppl 1**: p. S10.
74. Huang, C., et al., *Artemisinin rewires the protein interaction network in cancer cells: network analysis, pathway identification, and target prediction.* Mol Biosyst, 2013. **9**(12): p. 3091-100.
75. Wang, L., *Pharmacogenomics: a systems approach.* Wiley Interdiscip Rev Syst Biol Med, 2010. **2**(1): p. 3-22.
76. Dupont, C., D.R. Armant, and C.A. Brenner, *Epigenetics: definition, mechanisms and clinical perspective.* Semin Reprod Med, 2009. **27**(5): p. 351-7.
77. McQuown, S.C. and M.A. Wood, *Epigenetic regulation in substance use disorders.* Curr Psychiatry Rep, 2010. **12**(2): p. 145-53.
78. Hamilton, J.P., *Epigenetics: principles and practice.* Dig Dis, 2011. **29**(2): p. 130-5.
79. Verma, M., *The Role of Epigenomics in the Study of Cancer Biomarkers and in the Development of Diagnostic Tools.* Adv Exp Med Biol, 2015. **867**: p. 59-80.
80. Baccarelli, A., M. Rienstra, and E.J. Benjamin, *Cardiovascular epigenetics: basic concepts and results from animal and human studies.* Circ Cardiovasc Genet, 2010. **3**(6): p. 567-73.
81. Al-Ali, H.K., et al., *Azacitidine in patients with acute myeloid leukemia medically unfit for or resistant to chemotherapy: a multicenter phase I/II study.* Leuk Lymphoma, 2012. **53**(1): p. 110-7.
82. Keating, G.M., *Azacitidine: a review of its use in higher-risk myelodysplastic syndromes/acute myeloid leukaemia.* Drugs, 2009. **69**(17): p. 2501-18.
83. Bradbury, J., *Human epigenome project--up and running.* PLoS Biol, 2003. **1**(3): p. E82.
84. Voelkerding, K.V., S.A. Dames, and J.D. Durtschi, *Next-generation sequencing: from basic research to diagnostics.* Clin Chem, 2009. **55**(4): p. 641-58.
85. Taylor, K.H., et al., *Ultradeep bisulfite sequencing analysis of DNA methylation patterns in multiple gene promoters by 454 sequencing.* Cancer Res, 2007. **67**(18): p. 8511-8.

86. Barski, A., et al., *High-resolution profiling of histone methylations in the human genome*. Cell, 2007. **129**(4): p. 823-37.
87. Mikkelsen, T.S., et al., *Genome-wide maps of chromatin state in pluripotent and lineage-committed cells*. Nature, 2007. **448**(7153): p. 553-60.
88. Baylin, S.B., et al., *Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer*. Hum Mol Genet, 2001. **10**(7): p. 687-92.
89. Supic, G., M. Jagodic, and Z. Magic, *Epigenetics: a new link between nutrition and cancer*. Nutr Cancer, 2013. **65**(6): p. 781-92.
90. Aggarwal, R., et al., *Natural Compounds: Role in Reversal of Epigenetic Changes*. Biochemistry (Mosc), 2015. **80**(8): p. 972-89.
91. Liu, Z., L. Ma, and G.B. Zhou, *The main anticancer bullets of the Chinese medicinal herb, thunder god vine*. Molecules, 2011. **16**(6): p. 5283-97.
92. Li, Y., et al., *Genistein depletes telomerase activity through cross-talk between genetic and epigenetic mechanisms*. Int J Cancer, 2009. **125**(2): p. 286-96.
93. Salerno, E., et al., *Correcting miR-15a/16 genetic defect in New Zealand Black mouse model of CLL enhances drug sensitivity*. Mol Cancer Ther, 2009. **8**(9): p. 2684-92.
94. Wang, Z., M. Gerstein, and M. Snyder, *RNA-Seq: a revolutionary tool for transcriptomics*. Nat Rev Genet, 2009. **10**(1): p. 57-63.
95. Lindberg, J. and J. Lundeberg, *The plasticity of the mammalian transcriptome*. Genomics, 2010. **95**(1): p. 1-6.
96. Qian, X., et al., *RNA-Seq technology and its application in fish transcriptomics*. OMICS, 2014. **18**(2): p. 98-110.
97. Matsumura, H., et al., *SuperSAGE*. Cell Microbiol, 2005. **7**(1): p. 11-8.
98. Brenner, S., et al., *Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays*. Nat Biotechnol, 2000. **18**(6): p. 630-4.
99. Chu, Y. and D.R. Corey, *RNA sequencing: platform selection, experimental design, and data interpretation*. Nucleic Acid Ther, 2012. **22**(4): p. 271-4.
100. Zhang, P., et al., *Genome wide expression analysis of the effect of the Chinese patent medicine Zilongjin tablet on four human lung carcinoma cell lines*. Phytother Res, 2011. **25**(10): p. 1472-9.
101. Wang, L., et al., *A network study of chinese medicine xuesaitong injection to elucidate a complex mode of action with multicomponent, multitarget, and multipathway*. Evid Based Complement Alternat Med, 2013. **2013**: p. 652373.
102. Wang, L., et al., *Dissecting active ingredients of Chinese medicine by content-weighted ingredient-target network*. Mol Biosyst, 2014. **10**(7): p. 1905-11.
103. Li, X., et al., *A network pharmacology study of Chinese medicine QiShenYiQi to reveal its underlying multi-compound, multi-target, multi-pathway mode of action*. PLoS One, 2014. **9**(5): p. e95004.
104. Liu, M., et al., *Transcriptional profiling of Chinese medicinal formula Si-Wu-Tang on breast cancer cells reveals phytoestrogenic activity*. BMC Complement Altern Med, 2013. **13**: p. 11.
105. Wilkins, M.R., et al., *From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis*. Biotechnology (N Y), 1996. **14**(1): p. 61-5.
106. Anderson, N.L. and N.G. Anderson, *Proteome and proteomics: new technologies, new concepts, and new words*. Electrophoresis, 1998. **19**(11): p. 1853-61.
107. Irar, S., et al., *Combination of 2DE and LC for plant proteomics analysis*. Methods Mol Biol, 2014. **1072**: p. 131-40.

108. Schubert, K.O., et al., *Proteome and pathway effects of chronic haloperidol treatment in mouse hippocampus*. *Proteomics*, 2015.
109. Zhao, Y., W.N. Lee, and G.G. Xiao, *Quantitative proteomics and biomarker discovery in human cancer*. *Expert Rev Proteomics*, 2009. **6**(2): p. 115-8.
110. Dong, M.Q., et al., *Quantitative mass spectrometry identifies insulin signaling targets in C. elegans*. *Science*, 2007. **317**(5838): p. 660-3.
111. Schulze, W.X. and B. Usadel, *Quantitation in mass-spectrometry-based proteomics*. *Annu Rev Plant Biol*, 2010. **61**: p. 491-516.
112. Zhu, W., J.W. Smith, and C.M. Huang, *Mass spectrometry-based label-free quantitative proteomics*. *J Biomed Biotechnol*, 2010. **2010**: p. 840518.
113. Liu, X., et al., *Proteomic assessment of tanshinone IIA sodium sulfonate on doxorubicin induced nephropathy*. *Am J Chin Med*, 2011. **39**(2): p. 395-409.
114. Feng, L.X., et al., *Clarifying the signal network of salvianolic acid B using proteomic assay and bioinformatic analysis*. *Proteomics*, 2011. **11**(8): p. 1473-85.
115. Zhang, A., et al., *Proteomics analysis of hepatoprotective effects for scoparone using MALDI-TOF/TOF mass spectrometry with bioinformatics*. *OMICS*, 2013. **17**(4): p. 224-9.
116. Nicholson, J.K., J.C. Lindon, and E. Holmes, 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica*, 1999. **29**(11): p. 1181-9.
117. Li, Y., et al., *An NMR metabolomics investigation of perturbations after treatment with Chinese herbal medicine formula in an experimental model of sepsis*. *OMICS*, 2013. **17**(5): p. 252-8.
118. Li, S. and B. Zhang, *Traditional Chinese medicine network pharmacology: theory, methodology and application*. *Chin J Nat Med*, 2013. **11**(2): p. 110-20.
119. Lindon, J.C. and J.K. Nicholson, *Spectroscopic and statistical techniques for information recovery in metabonomics and metabolomics*. *Annu Rev Anal Chem (Palo Alto Calif)*, 2008. **1**: p. 45-69.
120. Huan, T., et al., *DnsID in MyCompoundID for Rapid Identification of Dansylated Amine- and Phenol-Containing Metabolites in LC-MS-Based Metabolomics*. *Anal Chem*, 2015. **87**(19): p. 9838-45.
121. Crockford, D.J., et al., *Statistical heterospectroscopy, an approach to the integrated analysis of NMR and UPLC-MS data sets: application in metabonomic toxicology studies*. *Anal Chem*, 2006. **78**(2): p. 363-71.
122. Duran, A.L., et al., *Metabolomics spectral formatting, alignment and conversion tools (MSFACTs)*. *Bioinformatics*, 2003. **19**(17): p. 2283-93.
123. Wishart, D.S., et al., *HMDB 3.0--The Human Metabolome Database in 2013*. *Nucleic Acids Res*, 2013. **41**(Database issue): p. D801-7.
124. Lao, Y.M., J.G. Jiang, and L. Yan, *Application of metabonomic analytical techniques in the modernization and toxicology research of traditional Chinese medicine*. *Br J Pharmacol*, 2009. **157**(7): p. 1128-41.
125. Wan, J.Y., et al., *Biotransformation and metabolic profile of American ginseng saponins with human intestinal microflora by liquid chromatography quadrupole time-of-flight mass spectrometry*. *J Chromatogr A*, 2013. **1286**: p. 83-92.
126. Wei, L., et al., *Metabolic profiling studies on the toxicological effects of realgar in rats by (1)H NMR spectroscopy*. *Toxicol Appl Pharmacol*, 2009. **234**(3): p. 314-25.
127. Ma, C., et al., *Toxicology effects of morning glory seed in rat: a metabonomic method for profiling of urine metabolic changes*. *J Ethnopharmacol*, 2010. **130**(1): p. 134-42.



128. Wang, P.R., et al., *Neuroprotective effects of Huang-Lian-Jie-Du-Decoction on ischemic stroke rats revealed by (1)H NMR metabolomics approach*. J Pharm Biomed Anal, 2014. **88**: p. 106-16.
129. Gao, X.X., et al., *An investigation of the antidepressant action of xiaoyaosan in rats using ultra performance liquid chromatography-mass spectrometry combined with metabonomics*. Phytother Res, 2013. **27**(7): p. 1074-85.
130. Tan, G., et al., *Metabonomic profiles delineate the effect of traditional Chinese medicine sini decoction on myocardial infarction in rats*. PLoS One, 2012. **7**(4): p. e34157.
131. Zhang, B., et al., *Proteogenomic characterization of human colon and rectal cancer*. Nature, 2014. **513**(7518): p. 382-7.
132. de Sousa Abreu, R., et al., *Global signatures of protein and mRNA expression levels*. Mol Biosyst, 2009. **5**(12): p. 1512-26.
133. Balbin, O.A., et al., *Reconstructing targetable pathways in lung cancer by integrating diverse omics data*. Nat Commun, 2013. **4**: p. 2617.
134. Cisek, K., et al., *The application of multi-omics and systems biology to identify therapeutic targets in chronic kidney disease*. Nephrol Dial Transplant, 2015.
135. Yoon, S.H., et al., *Comparative multi-omics systems analysis of Escherichia coli strains B and K-12*. Genome Biol, 2012. **13**(5): p. R37.

## Chapter 2

### A New Strategy for Identifying Mechanisms of Drug-drug Interaction Using Transcriptome Analysis: Compound Kushen Injection as a Proof of Principle

CKI is typically used as an adjuvant drug for cancer treatment in clinical practice, so the research for its interactions with chemotherapy drugs is important in order to use CKI safely and effectively. In this chapter, we introduced a workflow to apply transcriptome analysis into research for drug-drug interactions (DDIs). With the workflow, we successfully explained the mechanisms of opposite combinational effects for CKI with doxorubicin and 5-fluorouracil (5-Fu). We found that pathways related to DNA synthesis and metabolism might be the main reason for the opposite effects and pathways related to organic biosynthetic and metabolic processes might be potential targets for CKI when interacting with doxorubicin and 5-Fu. In addition, we also found and verified *MYD-88* as an important gene in these interactions. As an important adaptor protein in immune response, *MYD-88* would be an interesting target to study the relationship between tumor environment and immune system. The results in this chapter not only demonstrated the interactions of CKI with chemotherapy drugs but also indicated our method is a powerful tool in DDIs research.

# Statement of Authorship

Title of Paper	A new strategy for identifying mechanisms of drug-drug interaction using transcriptome analysis: Compound Kushen Injection as a proof of principle
Publication Status	<input type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input checked="" type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	

## Principal Author

Name of Principal Author (Candidate)	Hanyuan Shen		
Contribution to the Paper	Experimental design, carried out experiments, analysed data, wrote paper		
Overall percentage (%)	60%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	13/3/2019

## Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Zhipeng Qu		
Contribution to the Paper	Experimental design, assisted with experiments, assisted with data analysis, wrote paper		
Signature		Date	19/03/19

Name of Co-Author	Yuka Harata-Lee <i>YH</i>		
Contribution to the Paper	Experimental design, assisted with experiments, assisted with data analysis, wrote paper		
Signature		Date	13/3/19

Name of Co-Author	Jian Cui		
Contribution to the Paper	Assisted with experiments		
Signature		Date	19/03/2019

Name of Co-Author	Thazin Nwe Aung		
Contribution to the Paper	Assisted with experiments		
Signature		Date	13/03/2019

Name of Co-Author	Wei Wang		
Contribution to the Paper	Assisted with experimental design, assisted with experiments		
Signature		Date	19/03/2019

Name of Co-Author	Robert Daniel Kortschak		
Contribution to the Paper	Experimental design, wrote paper		
Signature		Date	15/3/19

Name of Co-Author	David L. Adelson		
Contribution to the Paper	Supervised the research, acquired funding for the experiments, experimental design, wrote paper		
Signature		Date	1/4/2019

**A new strategy for identifying mechanisms of drug-drug interaction using transcriptome analysis: Compound Kushen injection as a proof of principle**

Hanyuan Shen<sup>1</sup>, Zhipeng Qu<sup>1</sup>, Yuka Harata-Lee<sup>1</sup>, Jian Cui<sup>1</sup>, Thazin Nwe Aung<sup>1</sup>, Wei Wang<sup>2</sup>, R. Daniel Kortschak<sup>1</sup>, and David L. Adelson\*<sup>1</sup>

<sup>1</sup> Zhendong Australia - China Centre for Molecular Chinese Medicine, School of Biological Sciences, University of Adelaide, Adelaide, South Australia, 5005.

<sup>2</sup> Zhendong Research Institute, Shanxi-Zhendong Pharmaceutical Co Ltd, Beijing, China.

\*Corresponding author: David L. Adelson

Department of Molecular and Biomedical Science, School of Biological Sciences, University of Adelaide, Adelaide, South Australia, 5005.

Telephone : +61 8 8303 7555

Email: [david.adelson@adelaide.edu.au](mailto:david.adelson@adelaide.edu.au)

Running title: Identification of drug-drug interaction mechanisms from transcriptome data

## **Abstract**

Drug-drug interactions (DDIs), especially with herbal medicines, are complex, making it difficult to identify potential molecular mechanisms and targets. We introduce a workflow to carry out DDI research using transcriptome analysis and interactions of a complex herbal mixture, Compound Kushen Injection (CKI), with cancer chemotherapy drugs, as a proof of principle. Using CKI combined with doxorubicin or 5-Fu on cancer cells as a model, we found that CKI enhanced the cytotoxic effects of doxorubicin on A431 cells while protecting MDA-MB-231 cells treated with 5-Fu. We generated and analysed transcriptome data from cells treated with single treatments or combined treatments and our analysis showed that opposite directions of regulation for pathways related to DNA synthesis and metabolism appeared to be the main reason for different effects of CKI when used in combination with chemotherapy drugs. We also found that pathways related to organic biosynthetic and metabolic processes might be potential targets for CKI when interacting with doxorubicin and 5-Fu. Through co-expression analysis correlated with phenotype results, we selected the MYD88 gene as a candidate major regulator for validation as a proof of concept for our approach. Inhibition of MYD88 reduced antagonistic cytotoxic effects between CKI and 5-Fu, indicating that MYD88 is an important gene in the DDI mechanism between CKI and chemotherapy drugs. These findings demonstrate that our pipeline is effective for the application of transcriptome analysis to the study of DDIs in order to identify candidate mechanisms and potential targets.

Key words: Drug-drug interactions, transcriptome, herb medicines, cancer treatments

## Introduction

Drug combinations or polypharmacy is a commonly used clinical strategy for elderly patients and chronic diseases like diabetes, cardiovascular disease and cancer, in order to overcome unwanted off-target effects and compensatory mechanisms for certain drugs <sup>1-3</sup>. However, the challenge for polypharmacy <sup>4</sup> is how to estimate the effects of drug combinations compared to single drugs, and avoid potentially serious adverse effects resulting from drug-drug interactions (DDIs). The most common strategy for identifying DDIs is through pharmacokinetic approaches. This is because, by affecting transporters and metabolizing enzymes, one drug's pharmacokinetic process (absorption, distribution, metabolism or excretion) can be changed by another drug. However, pharmacokinetic properties are not usually directly linked to pharmacodynamic effects and cannot show interactions with treatment targets or potential side effects. Furthermore, pharmacodynamic assays may not provide enough information for detecting potential interaction effects and interpreting their mechanisms <sup>5</sup>. This is a particular concern for drug interactions involving complementary and alternative medicines (CAM), where herbal extracts can contain over a hundred different, potentially bioactive, compounds.

Public acceptance of combining complementary and alternative medicine (CAM) with conventional medicines has increased significantly over the last few decades. In 2007, nearly 38% of American adults used CAM <sup>6</sup>, and in China, which has a long history of traditional herbal medicine, 93.4% of cancer patients use CAM <sup>7</sup>. These medicines, especially traditional Chinese medicines (TCM) which are usually made from several herbs, can also exert their effects on conventional medicines both through pharmacokinetic and pharmacodynamic effects. The complexity of components in these CAMs make it extremely difficult to predict potential

interactions with conventional medicines and explain these rationally. By providing opportunities to examine a broad range of biological information, omics-related techniques provide a more comprehensive way for the study of drug-drug or herb-drug interactions<sup>8</sup>. In this report, we apply these methods to the identification of interactions between Compound Kushen Injection (CKI), a complex herbal extract mixture, and chemotherapy drugs.

In this study CKI is used as a model complementary medicine. CKI was approved by the State Food and Drug Administration (SFDA) of China in 1995, CKI is used by more than 30,000 patients every day as part of their treatment for various types of cancers<sup>9</sup>. Previous reports have shown that CKI can sensitize cancer to chemotherapeutic drugs, and reduce side effects of chemotherapy and radiotherapy to improve treatment effects and quality of life for cancer patients<sup>10,11</sup>. CKI is extracted from two herbs, Kushen (*Radix Sophorae flavescens*) and Baituling (*Rhizoma Smilacis glabrae*), which contain many natural compounds including, but not limited to alkaloids and flavonoids. Matrine and oxymatrine have been implicated as the primary active components for cancer treatment<sup>12</sup>, but this is not supported by our previous research that showed that CKI, but not oxymatrine, can inhibit cancer cell proliferation and cause apoptosis by perturbing the cell cycle and other cancer related pathways<sup>13-15</sup>. However, to date, no reports have revealed how CKI or its active components interact with cancer chemotherapy drugs.

In order to better understand DDIs and deal with the difficulties caused by complex components in herb-drug interactions, we propose a pipeline to apply transcriptome analysis for the study of DDIs. CKI was used as a test drug in combination with different chemotherapy agents, and was found to have different effects on cancer cells when combined with doxorubicin or 5-Fu



(synergistic with doxorubicin and antagonistic with 5-Fu). Based on transcriptome data, we have identified hundreds of differentially expressed genes that are correlated with opposite effects of CKI and chemotherapy agents on cell viability or apoptosis. These genes indicate that several cancer related pathways, such as DNA replication and cell cycle, are perturbed differently by CKI under different medical circumstances. Compared to previous DDI studies focused on transporters, metabolizing enzymes and therapy targets, our methods can provide a comprehensive and deeper analysis of interactions, that may help to pinpoint potential therapeutic or side effects, and explain the mechanisms underlying DDIs.

## **Results**

### **Pipeline for the study of DDIs using transcriptome analysis**

Figure 1 shows the flowchart for transcriptome analysis of DDIs. First, assays for DDIs are selected that are suitable for RNA sequencing and phenotype readouts. Second, shared differentially expressed (DE) genes from treatment with the primary drug only and combined treatment of primary drug and interacting drug are identified and further classified based on their manner of regulation. Gene co-expression analysis can then be used to identify groups of genes whose regulation is correlated with phenotype. Finally, different annotation methods can be used to propose mechanisms for DDIs and predict potential interactions.

The differentially regulated genes for single drug treatment are calculated with respect to untreated samples, while combined treatments are compared to single treatment. In addition, to identify types of interactions, genes consistently up or down regulated in single treatment and in combined treatment are classified as positively interacting (in other words, the expression level of primary chemotherapy agent treatment is intermediate between untreated and combined

treatment). Negatively interacting genes have expression levels where the primary chemotherapy treatment causes either the highest or lowest expression compared to untreated cells, or combined treatment).

### **CKI enhances the effects of doxorubicin but protects cells when co-administered with 5-Fu**

CKI alone can inhibit proliferation, induce apoptosis and alter the cell cycle for various cancer cell lines<sup>13-15</sup>. In order to determine whether CKI can potentiate the anticancer effects of chemotherapy agents, we used the XTT assay as a preliminary screen for the interaction of CKI with different chemotherapy drugs (Supplementary Fig. 1). Results showed that CKI could have opposite effects in different chemical contexts. These effects were most obvious at relatively low doses of CKI and chemotherapy agents to treat MDA-MB-231(with 5-Fu) and A431 cells (with doxorubicin) for 48 hours. CKI increased the apoptotic effects of doxorubicin whereas it antagonized the cytotoxicity of 5-Fu (Fig. 2A). Flow cytometric analysis of propidium iodide (PI) stained cells was also used to assess alterations to the cell cycle and apoptosis for different treatment groups. In MDA-MB-231 cells, treatment with CKI caused the increased percentages of apoptosis from 5-Fu to be drawn back to the same level as untreated cells. However, the proportion of apoptotic cells increased significantly when CKI was combined with doxorubicin on A431 cells (Fig. 2B ). Also, compared to slight changes in the cell cycle caused by the combination of CKI and 5-Fu, CKI caused large decreases in G1 and S phases of the cell cycle compared to doxorubicin only treatment (Fig. 2C ). Altogether, these data suggest that CKI has opposite interactions with doxorubicin and 5-Fu *in vitro*.

### **Selecting DE genes involved in drug-drug interactions**

In order to understand the molecular mechanisms of the opposite interactions of CKI with doxorubicin and 5-Fu, we carried out transcriptome profiling from chemotherapy agents treatment, CKI treatment and combined CKI+chemotherapy using high-depth next generation sequencing. In order to correlate the gene expression results with phenotype results, we selected 48 hours as the treatment time with three biological replicates. After preliminary multidimensional scaling of all the samples, every treatment group clustered, and clusters were clearly separated, indicating that combined CKI treatment can change the transcriptome of cancer cells compared to chemotherapy or CKI alone (Supplementary Fig. 3 & 4).

Because we were primarily interested in determining the changes in gene expression between combined and single treatments, we identified DE genes by comparing the combined treatment to treatment with chemotherapy drug only. We also compared single treatments to untreated. This gave 4 sets of DE genes (A431 cell line: doxorubicin compared to untreated and doxorubicin + CKI compared to doxorubicin, MDA-MB-231 cell line: 5-Fu compared to untreated and 5-Fu + CKI compared to 5-Fu) with each set containing thousands of DE genes (Fig. 3A & Supplementary Table 2).

### **Identification of DE genes based on interaction and direction of regulation**

From the original four gene sets, we refined our results to get DE gene subsets related to drug interactions. Because of our primary focus on the mechanisms underlying opposite effects of CKI combined with doxorubicin or 5-Fu, we identified the set of common DE genes across the four sets of DE genes identified above. This subset of 2926 genes was selected for further analysis (Fig. 3A). Because differential expression can result from up or down regulation of expression, we included the direction of interaction as a means to separate DE genes involved

in DDIs. If one gene's change in expression level from untreated to single treatment was consistent with its direction of regulation for the combined treatment (either up regulated or down regulated) then we defined it as a positive interaction gene, either additive or synergistic. In contrast, if its direction of regulation was opposite in the single treatment compared to the combined treatment, it was defined as a negative interaction. When these criteria were applied to our subset of 2926 genes, while most of the DE genes underwent negative interaction across both cell lines/treatments, the proportion of positive interaction genes differed between treatment groups. In A431 cells treated with doxorubicin, CKI induced 30.9% positive interaction genes, whereas only 12.9% of the genes were positively interacted with CKI in MDA-MB-231 cells treated with 5-Fu (Fig. 3B).

In order to further characterise the genes with different directions of interaction, we performed functional enrichment analysis of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways for our set of shared 2926 DE genes and calculated the number of genes for negative and positive interaction with CKI in both treatment groups/cell lines. The results showed that in every pathway, the proportion of DE genes positively regulated by CKI treatment with doxorubicin (A431 cells) was larger than by CKI treatment with 5-Fu (MDA-MB-231 cells) (Fig. 3C). Strikingly, there were four pathways related to cell cycle that had over 50% the genes positively regulated in the A431 cells, including: "Base excision repair", "Cell cycle", "DNA replication" and "Homologous recombination". When we took into account the expectation that one third of the DE genes should fall into the positive interaction class (Fig. 1B), we used 33.33% as the cut off for distinguishing direction of interaction pathways. With this criterion, there were 13 pathways where CKI caused positive interactions with doxorubicin, but only 1 with 5-Fu (Fig. 3C). Furthermore, 9 pathways were consistently found to interact in a positive

manner, including immune pathways (“Bacterial invasion of epithelial cells”, “Human papillomavirus infection”, “Viral carcinogenesis”) and metabolic pathways (“Glyoxylate and dicarboxylate metabolism”, “Steroid biosynthesis”) and others. .

To have a comprehensive understanding of drug-drug interactions, samples treated with single chemotherapy agents were also annotated with KEGG pathways. With the common regulated subset, genes in 8 cell cycle related pathways were primarily regulated in the same directions (mainly down-regulated in 6 pathways and up-regulated in 2 pathways) both by doxorubicin and 5-Fu (Fig. 3D).

### **DE genes related to phenotype**

Based on the direction of regulation in each combined treatment group, we separated the 2926 shared DE genes into four groups (Fig. 4A & Supplementary Table 3). Group C in which genes were negatively interacted in both cell lines, contained the largest number of genes (1815, 62% for total gene number) followed by group A genes with 732 that are negatively interacted for 5-Fu and positively interacted for doxorubicin. The other two groups (B and D) only contained 208 and 171 genes respectively. Based on the phenotype results, genes in group A were more likely to relate to our study purpose, while groups C and D might reveal CKI’s overall effects on chemotherapy cancer drugs.

Functional enrichment analysis was also performed for each of the four groups (Fig. 4B, C & D & Supplementary Table 3). For group A, the Gene Ontology terms for 732 genes (Fig. 4C) were mainly related to “cell cycle” and “nucleobase-containing compound metabolic process”. In KEGG pathway analysis, except for pathways closely related to cell growth like “Cell Cycle” and

“DNA replication”, there were two cancer related pathways detected, “Bladder cancer” and “Chronic myeloid leukemia” and one immune related pathway “NF-kappa B signaling pathway”. Genes in these pathways, like cell cycle (Fig. 5A & Supplementary Fig. 2) are regulated in opposite direction by CKI compared to chemotherapy drugs. GO terms for 1815 genes in group C (Fig. 4D) were mainly clustered into “Organic substance biosynthetic process”, “regulation of cell cycle” and “organic substance catabolic process”, KEGG results also indicated that the majority of genes in this group belonged to pathways related to metabolism and biosynthesis. Gene numbers for groups B and D were much lower, with mainly immune related GO terms for group B and cell cycle related GO terms for group D . Only three KEGG pathways were significantly enriched for group D and none for group B.

In order to validate the gene expression changes with different directions of regulation in the doxorubicin and 5-Fu treatment groups, we estimated protein abundance using flow cytometry for 4 proteins in group A. Overall, the protein level changes were consistent with gene expression levels from transcriptome analysis (Fig. 5B).

### **Integrating information to select genes for validation**

In order to select genes for experimental validation with bench experiments, we constructed the co-expression networks for genes in group A. 732 genes in 8 treatment groups were separated into 14 co-expression modules. By including data from the XTT and apoptosis assays, we calculated the correlation coefficients for each gene module with the phenotype results. The red, black and purple modules were more highly correlated with phenotype results than other modules. Because it had the highest correlation coefficients with both traits, genes in the red module (45 genes) were picked for further investigation (Fig. 6A).

Protein interaction, GO and KEGG analyses were performed for genes in the red module. From the protein interaction network, MYD88 was the most connected/interacting protein (Fig. 6B & C, Fig. 7A). Furthermore, it was commonly shared across different functions or pathways in the GO and KEGG analyses. Considering that MYD88 is upstream of NF-kappa B which itself regulates the cell cycle and other cancer related pathways, we selected it as a candidate for validation by inhibition.

### **Inhibiting MYD88 partially affected the interaction between CKI and 5-Fu**

To validate our analysis results, we performed the cell viability and apoptosis assays again and included an MYD88 inhibitory peptide or a control peptide (Fig. 7B). Because MYD88 is one of the key regulators in NF-kappa B pathway and occupies a central position in the red module, we expected that inhibiting it would reduce cell proliferation and the opposite effects from CKI on chemotherapy drugs. Results showed that for the MDA-MB-231 cells, inhibiting MYD88 does not affect the overall cell viability or apoptosis rates. However, compared to the control peptide, the inhibitor significantly reduced cell viability for 5-Fu and CKI combined treatment, by weakening the antagonistic effect of CKI. In the apoptosis assay, the apoptosis rate for 5-Fu treatment was significantly lower when treated with the inhibitor, also suggesting a similar reduction in the antagonistic effect. For the doxorubicin group, no significant changed interaction was found. However, unlike the MDA-MB-231 cell line, A431 cells were sensitive to the MYD88 inhibitor as shown by the overall lower cell viability values compared to the control peptide group.

In summary, we were able to dissect and characterise a DDI with transcriptome analysis. With CKI as a model, we identified candidate mechanisms behind its opposite effects compared to different chemotherapy agents and revealed potential interactions with them. We also identified and verified MYD88 as a target/key regulator for DDIs between CKI and anticancer drugs. These results demonstrate the value of our pipeline for characterising and understanding the molecular basis of DDIs.

## Discussion

Drug-drug interactions are one of the main reasons for adverse events associated with medication. The traditional pharmacokinetic methods for studying DDIs are inadequate for discovering potential side effects or explaining complicated interaction mechanisms. Furthermore, the complex components for many complementary medicines and herbal medicines that are often used in conjunction with pharmaceutical drugs pose a significant challenge research on DDIs . Although high-throughput omics-related techniques have been widely used for identifying novel disease biomarkers or potential drug targets <sup>16,17</sup>, very limited research has applied them to the investigation of DDIs. Because transcriptome based approaches generate very large data sets, we adopted a hierarchical approach for our analysis of DDIs between CKI and chemotherapy drugs. First, instead of comparing every treatment sample to untreated control in order to identify DE genes, we decided to set the baseline for comparison of interactions as the main drug treatment. We then identified DE genes for the combined treatment based comparison to the main drug treatment. From the common set of DE genes found by comparing the main drug treatment to untreated control, and the combined treatment compared to the main drug treatment, we selected only genes that were differentially expressed and shared across the various comparisons. Second, we used “consistent directional



regulation” to separate the DE genes between multiple treatments into positive and negative interaction classes. These are more informative with respect to drug-drug interaction than simply up or down regulation. This also allowed genes to be separated based on consistent directional regulation to focus the scope of investigation. Finally, we applied gene co-expression network analysis to provide useful information for candidate gene selection. These methods combined with typical gene annotation analysis and protein interaction analysis, provided a rich profile for investigation of DDIs.

Diseases treated with drug combinations are usually complex and related to multiple genetic pathways. Furthermore, multiple active compounds in drug combinations, such as herbal medicines can affect a variety of targets <sup>1819</sup>. Therefore, by integrating the effects of interacting genes, analysis of pathways or networks may provide more useful evidence for characterising the mechanisms of drug interactions. For the mechanisms of opposite effects generated by CKI combination with drugs, the pathways related to DNA synthesis and metabolism, like “Base excision”, “DNA replication” “homologous recombination”, are oppositely interacted between the two treatment groups. Related to both chemotherapy drugs down regulate genes in these pathways, the opposite interacted effects from CKI can enhance effects from doxorubicin while reduce 5-Fu’s effects. Closely linked to these pathways, “cell cycle” and “apoptosis” also have large differences in their manner of interaction between the two groups. By correlating the results from cell viability and apoptosis assays, we can propose the opposite effects from CKI with doxorubicin or 5-Fu are primarily induced from pathways related to DNA synthesis and metabolism. As 5-fu and doxorubicin both target DNA replication and CKI’s cytotoxic effects have also been shown to increase DNA Double Strand Breaks, this indicates that therapeutic

results from drug combinations targeting the same or similar bioprocesses can be very different compared to what we might predict<sup>13,20,21</sup>.

Furthermore, by performing annotation analysis for functional or expression level clustering of gene groups, more information about interactions between CKI and chemotherapy drugs can be acquired. For group A, the annotation results are similar with opposite interacted results as discussed in the last paragraph, which is closely related to the phenotype results. In addition, the other three groups can help us to discover potential interactions which are not shown in our limited experiments. Group C displayed negative interaction of CKI to both doxorubicin and 5-fu and most annotation terms belonged to organic biosynthetic and metabolic processes. Many shared side effects from these two chemotherapy drugs are linked to disorders of metabolism, for example, cardiovascular and mucosal toxicity caused by cancer therapies are mainly caused by free radicals and oxidative stress<sup>22-24</sup>. Therefore, although more validation is needed, the results for group C might support the clinical reports that CKI can reduce the adverse effects of chemotherapy and radiotherapy in cancer treatment. In addition, we observed two pathways "steroid biosynthesis" and "Fluid shear stress and atherosclerosis" that suggest that doxorubicin and 5-Fu affect atherosclerosis in a manner opposite to that of CKI. We could find no existing literature that would corroborate this finding. Our results indicate that transcriptome analysis can not only reveal candidate molecular mechanisms altered by specific drugs, but can also provide clues about potential drug-drug interactions.

Transcriptome analysis can provide a far more comprehensive and complex candidate gene list than traditional approaches used in drug-drug interaction research. This makes it difficult to screen target genes for further study because of the gene specific assays required. In group A ,

we generated a list of 732 genes, including heme oxygenase 1 (HO-1) and E3 ubiquitin-protein ligase (CBL), which are involved in metabolism pathways, that were oppositely regulated when CKI was combined with doxorubicin or 5-Fu. In addition, genes like tumor necrosis factor alpha-induced protein 3 (TNFAIP3) and myeloid differentiation primary response protein (MYD88) from the NF-kappa B pathway and cyclin-dependent kinase inhibitor 1A (P21) in the cell cycle are regulated in the same manner. Although their functions are different, all these genes are important in carcinoma<sup>25-28</sup>. By using gene co-expression and protein interactions analysis, we chose MYD88 as a proof of concept for validation as it was highly correlated with phenotype results and interacted with more proteins in its WGCNA color module than others. Our prediction was that inhibiting MYD88 would decrease the antagonistic effect of CKI on 5-Fu and this prediction was confirmed. By using our approach, transcriptome analysis can not only be used for generating comprehensive gene lists for candidate mechanisms, it can also identify specific, potential targets for modulating drug-drug interactions.

In summary, we introduced a pipeline to integrate omics techniques into research for DDIs. By using transcriptome analysis to identify candidate mechanisms that might account for CKI's opposite effects on doxorubicin or 5-Fu in cancer cells, we have shown that our methods are effective and can be applied to complex situations, including drug interactions with complex mixtures or to compare different drug-drug interaction groups.

## **Methods**

### **Cell culture and drugs**

A431 and MDA-MB-231 cells were purchased from ATCC (VA, USA) and cultured in DMEM (Thermo Fisher Scientific, MA, USA) with 10 % fetal bovine serum (Thermo Fisher Scientific) at 37°C.

with 5 % CO<sub>2</sub>. CKI (total alkaloid concentration of 26.5 mg/ml) was provided by Zhendong pharmaceutical Co.Ltd (China) and used at a final concentration of 1 mg/ml. Fluorouracil (5-Fu) and doxorubicin were purchased from Sigma-Aldrich (MO, USA) and used at final concentrations of 10 ug/ml and 1 ng/ml, respectively. MYD88 inhibitor and control peptides were synthesised by GenScript (Hong Kong, China) with the following amino acid sequences with purity > 95%<sup>29,30</sup>; inhibitor: DRQIKIWFQNRRMKWKKRDVLPQT and control peptide: DRQIKIWFQNRRMKWKK .

For all *in vitro* assays 6-well or 96-well plates were used . The seeding density for both A431 and MDA-MB-231 cells was  $4 \times 10^5$  cells/well for 6-well plates . For 96-well plates, A431 cells were seeded at  $8 \times 10^4$  cells/well and MDA-MB-231 cells were  $1.6 \times 10^5$  cells/well. After seeding, cells were cultured overnight before being treated.

### **Cell viability assay**

Cells were seeded in 96-well plates with 50 µl of medium. For the MYD88 validation assay, the inhibitor or control peptide was added at the same time as cell seeding. After overnight culturing, 50 µl of CKI and/or chemotherapeutic agent at appropriate concentration were added and incubated for 48 hours. In order to measure the cell viability, 50 µl of XTT:PMS (at 1 mg/ml and 1.25mM, respectively, and combined at 50:1 ratio, Sigma-Aldrich) was added and incubated 4 hours before detecting absorbance of each well with a Biotrack II microplate reader

at 492 nm. Wells without cells were set up for each treatment for subtracting background absorbance.

### **Cell cycle assay**

Cells were cultured and treated in 6-well plates. After 48 hours of drug treatment, cells were harvested and stained with propidium iodide (PI) to examine cell cycle phases as previously described<sup>31</sup>. Stained cells were acquired on BD LSRFortessa-X20 (BD Biosciences, NJ, USA) and the data were analysed using FlowJo software (TreeStar Inc., OR, USA).

### **Flow cytometric quantification of protein expression**

Cells were cultured in 6-well plates and treated with drugs for 48 hours. The cells were subsequently harvested and stained with antibodies to detect intranuclear/intracellular protein levels. The antibodies were purchased from Abcam (UK) unless otherwise indicated: rabbit anti-CBL and rabbit IgG isotype control (Cell Signaling Technologies) detected with anti-rabbit IgG-PE (Cell Signaling Technologies); mouse anti-p21 and mouse IgG2b isotype control detected with anti-mouse IgG-Alexa Fluor 488 ; rabbit anti-TNFAIP3-Alexa Fluor 488 and rabbit IgG isotype control-Alexa Fluor 488 ; rabbit anti-HO-1-Alexa Fluor 568 and rabbit IgG isotype control-Alexa Fluor 568. Data was acquired with a BD Accuri (BD Biosciences) and analysed with FlowJo software.

### **RNA extraction and sequencing**

After being treated with drugs in 6-well plates for 48 hours, cells were harvested and snap-frozen with liquid nitrogen then stored at -80 °C. Total RNA was isolated with the RNA extraction kit (Thermo Fisher Scientific) and quantity and quality were measured with a

Bioanalyzer at the Cancer Genome Facility of the Australian Cancer Research Foundation (Australia) to ensure RINs > 7.0. Samples were sent to Novogene (China) and carried out on an Illumina HiSeq X platform with paired-end 150 bp reads.

### **Transcriptome data analysis**

Trim\_galore (v0.3.7, Babraham Bioinformatics) was used to trim adaptors and low-quality sequences in raw reads with parameters: `--stringency 5 --paired`. Then trimmed reads were aligned to reference genome (hg19, UCSC) using STAR (v2.5.3a) with parameters: `--outSAMstrandField intronMotif --outSAMattributes All --outFilterMismatchNmax 10 --seedSearchStartLmax 30`<sup>32</sup>. Differentially expressed genes between two groups were calculated with edgeR (v3.22.3) and selected with false discovery rate (FDR) < 0.05<sup>33</sup>.

ClueGO was used to perform the GO and KEGG over-representation analyses with following parameters: right-sided hypergeometric test for enrichment analysis; p values were corrected for multiple testing according to the Benjamini-Hochberg method and biological process at 3rd level for GO terms<sup>34</sup>. Then Cytoscape v3.6.0 were used to visualise selected terms or pathways<sup>35</sup>. Regulation profiles for specific pathways were visualised with the R Pathview package<sup>36</sup>.

Co-expression network analysis was performed with WGCNA with “16” as soft thresholding power and “5” as minimum gene size for module reconstruction<sup>37</sup>. String (V11.0) was used to show protein interactions with 0.4 for minimum interaction score<sup>38</sup>.

1. Lehár, J. *et al.* Erratum: Synergistic drug combinations tend to improve therapeutically relevant selectivity. *Nat. Biotechnol.* **27**, 864–864 (2009).
2. Palmer, A. C. & Sorger, P. K. Combination Cancer Therapy Can Confer Benefit via Patient-to-Patient Variability without Drug Additivity or Synergy. *Cell* **171**, 1678–1691.e13 (2017).
3. Guthrie, B., Makubate, B., Hernandez-Santiago, V. & Dreischulte, T. The rising tide of polypharmacy and drug-drug interactions: population database analysis 1995–2010. *BMC Med.* **13**, (2015).
4. van Leeuwen, R. W. F. *et al.* Drug-drug interactions in patients treated for cancer: a prospective study on clinical interventions. *Ann. Oncol.* **26**, 992–997 (2015).
5. van Leeuwen, R. W. F., van Gelder, T., Mathijssen, R. H. J. & Jansman, F. G. A. Drug–drug interactions with tyrosine-kinase inhibitors: a clinical perspective. *Lancet Oncol.* **15**, e315–e326 (2014).
6. Website. Available at: National Center for Complementary and Alternative Medicine, <http://nccam.nih.gov/about>. (Accessed: 9th January 2019)
7. Teng, L. *et al.* Use of complementary and alternative medicine by cancer patients at Zhejiang University Teaching Hospital Zhuji Hospital, China. *Afr. J. Tradit. Complement. Altern. Med.* **7**, 322–330 (2010).
8. Hopkins, A. L. Network pharmacology: the next paradigm in drug discovery. *Nat. Chem. Biol.* **4**, 682–690 (2008).
9. Wu, L. *et al.* Synthesis and biological evaluation of matrine derivatives containing benzo- $\alpha$ -pyrone structure as potent anti-lung cancer agents. *Sci. Rep.* **6**, 35918 (2016).
10. Wang, X.-Q., Liu, J., Lin, H.-S. & Hou, W. A multicenter randomized controlled open-label

trial to assess the efficacy of compound kushen injection in combination with single-agent chemotherapy in treatment of elderly patients with advanced non-small cell lung cancer: study protocol for a randomized controlled trial. *Trials* **17**, 124 (2016).

11. Wang, S., Lian, X., Sun, M., Luo, L. & Guo, L. Efficacy of compound Kushen injection plus radiotherapy on nonsmall-cell lungcancer: A systematic review and meta-analysis. *J. Cancer Res. Ther.* **12**, 1298–1306 (2016).
12. Wang, W. *et al.* Anti-tumor activities of active ingredients in Compound Kushen Injection. *Acta Pharmacol. Sin.* **36**, 676–679 (2015).
13. Qu, Z. *et al.* Identification of candidate anti-cancer molecular mechanisms of Compound Kushen Injection using functional genomics. *Oncotarget* **7**, 66003–66019 (2016).
14. Zhao, Z. *et al.* Fufang Kushen injection inhibits sarcoma growth and tumor-induced hyperalgesia via TRPV1 signaling pathways. *Cancer Lett.* **355**, 232–241 (2014).
15. Cui, J. *et al.* Cell Cycle, Energy Metabolism and DNA Repair Pathways in Cancer Cells are Suppressed by Compound Kushen Injection. (2018). doi:10.1101/348102
16. Barrett, C. L. *et al.* Systematic transcriptome analysis reveals tumor-specific isoforms for ovarian cancer diagnosis and therapy. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E3050–7 (2015).
17. Ciriello, G. *et al.* Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell* **163**, 506–519 (2015).
18. Leite, P. M., Martins, M. A. P. & Castilho, R. O. Review on mechanisms and interactions in concomitant use of herbs and warfarin therapy. *Biomed. Pharmacother.* **83**, 14–21 (2016).
19. Zhou, W. *et al.* Systems pharmacology exploration of botanic drug pairs reveals the mechanism for treating different diseases. *Sci. Rep.* **6**, 36985 (2016).
20. Momparler, R. L., Karon, M., Siegel, S. E. & Avila, F. Effect of adriamycin on DNA, RNA, and protein synthesis in cell-free systems and intact cells. *Cancer Res.* **36**, 2891–2895



- (1976).
21. Longley, D. B., Harkin, D. P. & Johnston, P. G. 5-fluorouracil: mechanisms of action and clinical strategies. *Nat. Rev. Cancer* **3**, 330–338 (2003).
  22. Mitry, M. A. & Edwards, J. G. Doxorubicin induced heart failure: Phenotype and molecular mechanisms. *Int J Cardiol Heart Vasc* **10**, 17–24 (2016).
  23. Focaccetti, C. *et al.* Effects of 5-fluorouracil on morphology, cell cycle, proliferation, apoptosis, autophagy and ROS production in endothelial cells and cardiomyocytes. *PLoS One* **10**, e0115686 (2015).
  24. Aghamohamamdi, A. & Hosseinimehr, S. J. Natural Products for Management of Oral Mucositis Induced by Radiotherapy and Chemotherapy. *Integr. Cancer Ther.* **15**, 60–68 (2016).
  25. Jozkowicz, A., Was, H. & Dulak, J. Heme oxygenase-1 in tumors: is it a false friend? *Antioxid. Redox Signal.* **9**, 2099–2117 (2007).
  26. Sanada, M. *et al.* Gain-of-function of mutated C-CBL tumour suppressor in myeloid neoplasms. *Nature* **460**, 904–908 (2009).
  27. Lerebours, F. *et al.* NF-kappa B genes have a major role in inflammatory breast cancer. *BMC Cancer* **8**, 41 (2008).
  28. Abbas, T. & Dutta, A. p21 in cancer: intricate networks and multiple activities. *Nat. Rev. Cancer* **9**, 400–414 (2009).
  29. Loiarro, M. *et al.* Peptide-mediated Interference of TIR Domain Dimerization in MyD88 Inhibits Interleukin-1-dependent Activation of NF-κB. *J. Biol. Chem.* **280**, 15809–15814 (2005).
  30. Derossi, D., Joliot, A. H., Chassaing, G. & Prochiantz, A. The third helix of the Antennapedia homeodomain translocates through biological membranes. *J. Biol. Chem.*

- 269**, 10444–10450 (1994).
31. Riccardi, C. & Nicoletti, I. Analysis of apoptosis by propidium iodide staining and flow cytometry. *Nat. Protoc.* **1**, 1458–1461 (2006).
  32. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
  33. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
  34. Bindea, G. *et al.* ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091–1093 (2009).
  35. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
  36. Luo, W. & Brouwer, C. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics* **29**, 1830–1831 (2013).
  37. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, (2008).
  38. Jensen, L. J. *et al.* STRING 8--a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res.* **37**, D412–6 (2009).

Figure 1: Experimental and data analysis workflow for applying omics to drug-drug interactions. A. The overall design of the study. B. Further details of 2 specific procedures indicated with broken-lined boxes in A. The black, blue and orange bars represent untreated, single treatment and combined treatment, respectively.

Figure 2: Opposite effects of CKI combined with doxorubicin or 5-Fu on cell viability and cell cycle. A. and B. The cell viability and percentage of apoptosis of the cancer cells treated with different drug combinations for 48 hours. C. Representative histograms of cell cycle phases for different treatments. Results are represented as means  $\pm$ SEM (n=9). Statistical analyses were performed by comparing treatments to untreated ( \*\*p< 0.01, \*\*\*p < 0.001, \*\*\*\* p < 0.0001) as well as 'CKI + Chemotherapy Agent' to 'Chemotherapy Agent only' (##p<0.01, ###p<0.001, ####p<0.0001) with one-way ANOVA .

Figure 3: Selection of differentially regulated shared genes and percentage of genes regulated in different fashions and their related pathways. A. Venn diagram showing the number of differentially regulated genes in cancer cells with different treatments. B. Percentage of genes that were regulated in 'synergistic' (yellow) and 'antagonistic' (blue) fashion in A431 (doxorubicin) and MDA-MB-231 (5-Fu) cells. C. Percentage of synergistically regulated genes in different KEGG pathways. D. Percentage of up-regulated and down-regulated genes for single chemotherapy drug treatment in different KEGG pathways.

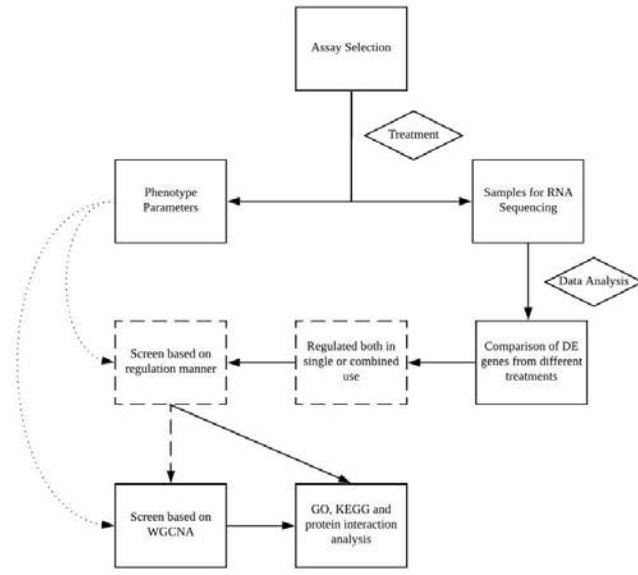
Figure 4: Grouping of genes based on type of regulation and annotation results for different gene groups. A. Criteria for separating 2926 genes into four groups. B. Table for over-represented KEGG terms and associated p-values for different groups of genes . C. and D. Over-represented GO terms (Biological Process at 3rd level) for genes in groups A and C.

Figure 5: Differentially regulated genes shown in pathway and validation of selected gene regulation. A. Comparison of types of regulation for CKI with doxorubicin and 5-Fu in the "Cell Cycle" pathway. Left half of the rectangle for each gene represents CKI with doxorubicin in A431 cells and the right half represents CKI with 5-Fu in MDA-MB-231 cells. Red and green colors mean synergistic and antagonistic regulation, respectively. B. Validation of gene regulation at protein level. Four genes (HO-1, TNFAIP3, P21 and CBL) with opposite types of regulation in A431 and MDA-MB-231 cells identified by transcriptome sequencing were selected and validated by flow cytometry. 'U', 'D' and 'U+D' represent untreated, single chemotherapy drug treatment and CKI plus chemotherapy drug treatment, respectively. Data are represented as means  $\pm$ SEM (n=9). Statistical analyses were performed between single drug treated or combined treated to untreated with one-way ANOVA ( \*p< 0.05, \*\*\*p < 0.001, \*\*\*\* p < 0.0001).

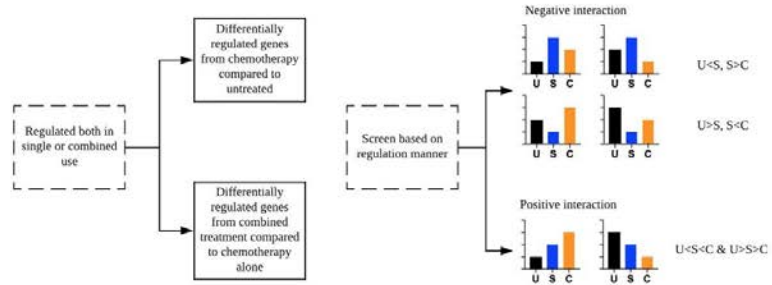
Figure 6: Co-expression analysis for genes in group A and related functional annotation. A. Clustering dendrogram for genes in group A and relationships with cell viability and apoptosis for each color module (red module is indicated by arrows). B. and C. Over-represented GO (Biological Process at 3rd level) and KEGG terms for genes in the red module.

Figure 7: Validation of MYD88 function. A. All proteins for genes in the red module in Figure 6 and their interactions as based on the STRING database. B. Cell viability and percentage of apoptosis as a result of different treatments combined with MYD88 inhibitor peptide or control peptide. Results are represented as means  $\pm$ SEM (n=9). Statistical analysis were performed with t-test (\*p< 0.05, \*\*p< 0.01).

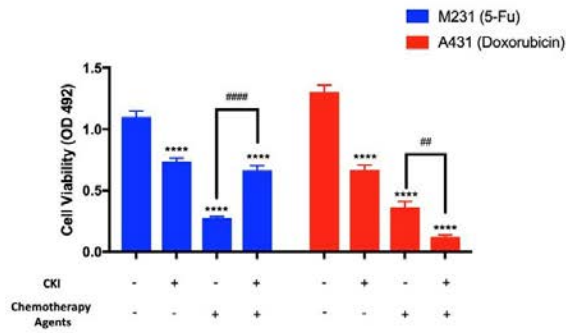
A



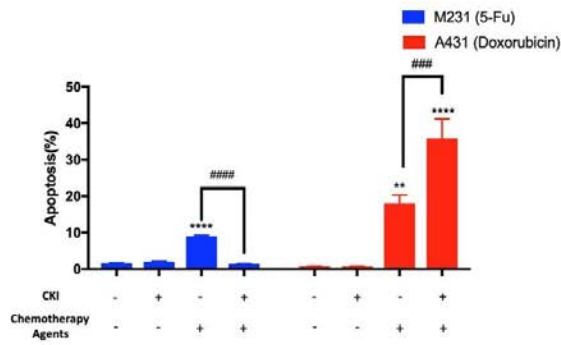
B



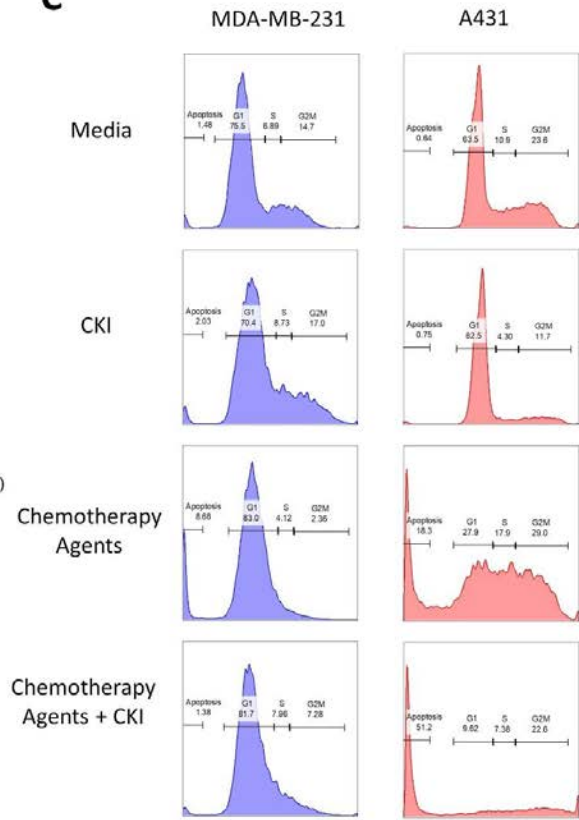
**A**

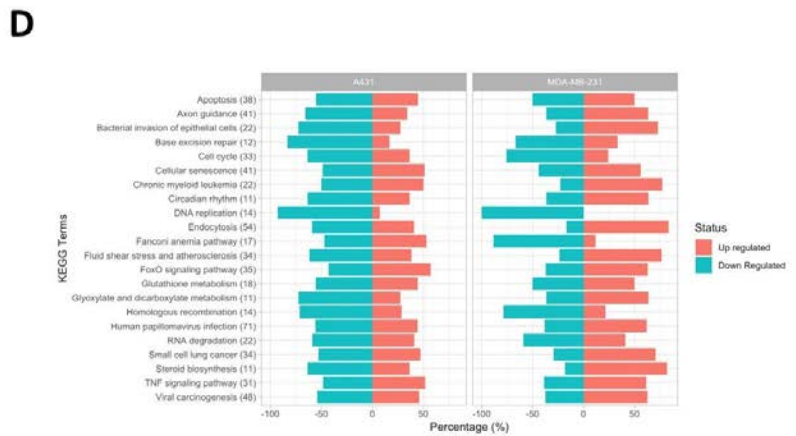
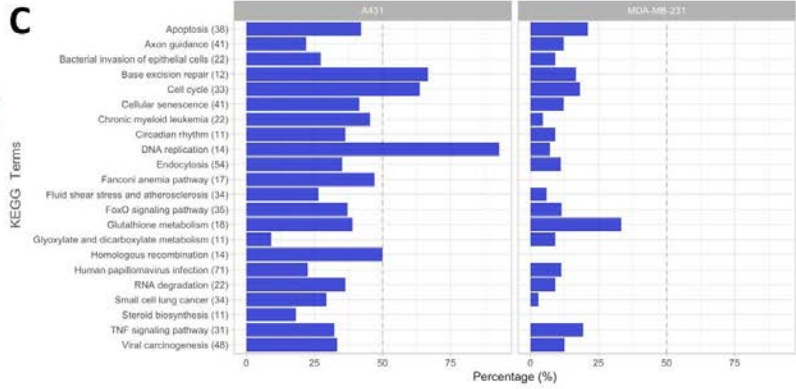
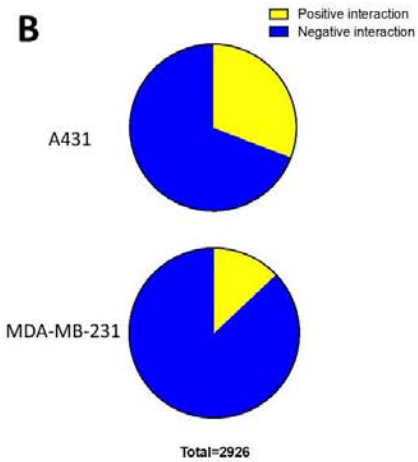
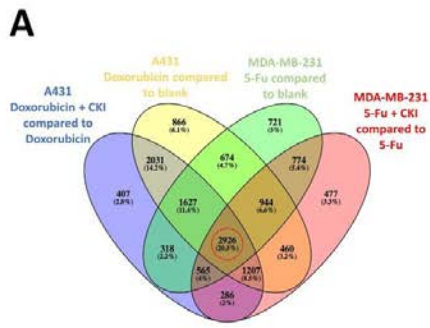


**B**

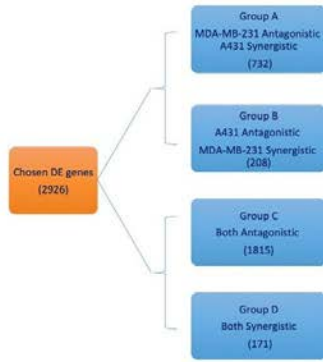


**C**





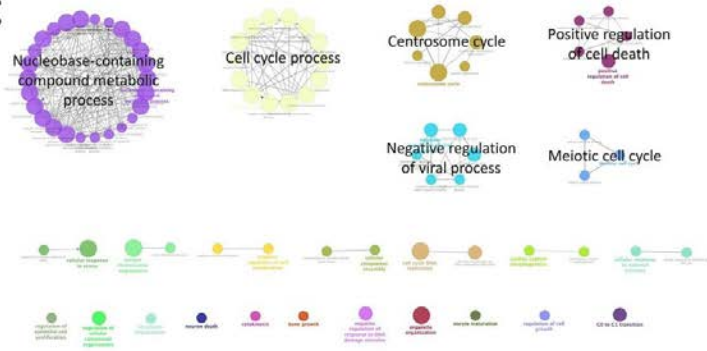
**A**



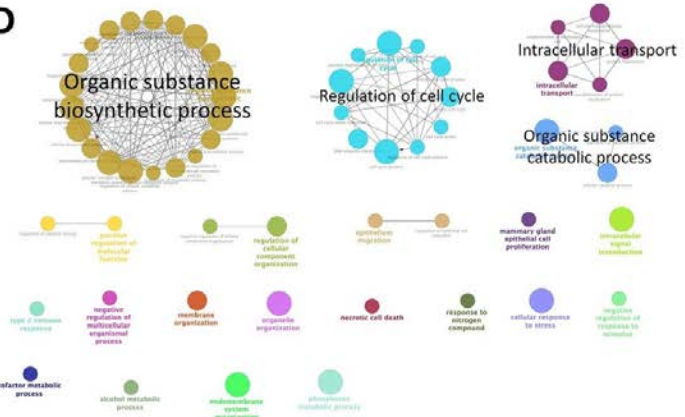
**B**

	Terms	P-value
Group A	DNA replication	3.96E-09
	Cell cycle	3.88E-06
	Base excision repair	1.93E-04
	Nucleotide excision repair	3.40E-04
	Homologous recombination	7.85E-04
	Fanconi anemia pathway	8.92E-04
	Chronic myeloid leukemia	0.00221
	Cellular senescence	0.00298
	NF-kappa B signaling pathway	0.00311
Group C	Bladder cancer	0.00418
	Colorectal cancer	0.00514
Group D	Small cell lung cancer	2.03E-06
	Steroid biosynthesis	1.73E-05
	Glyoxylate and dicarboxylate metabolism	2.16E-04
	Valine, leucine and isoleucine degradation	0.00101
	Fluid shear stress and atherosclerosis	0.00117

**C**

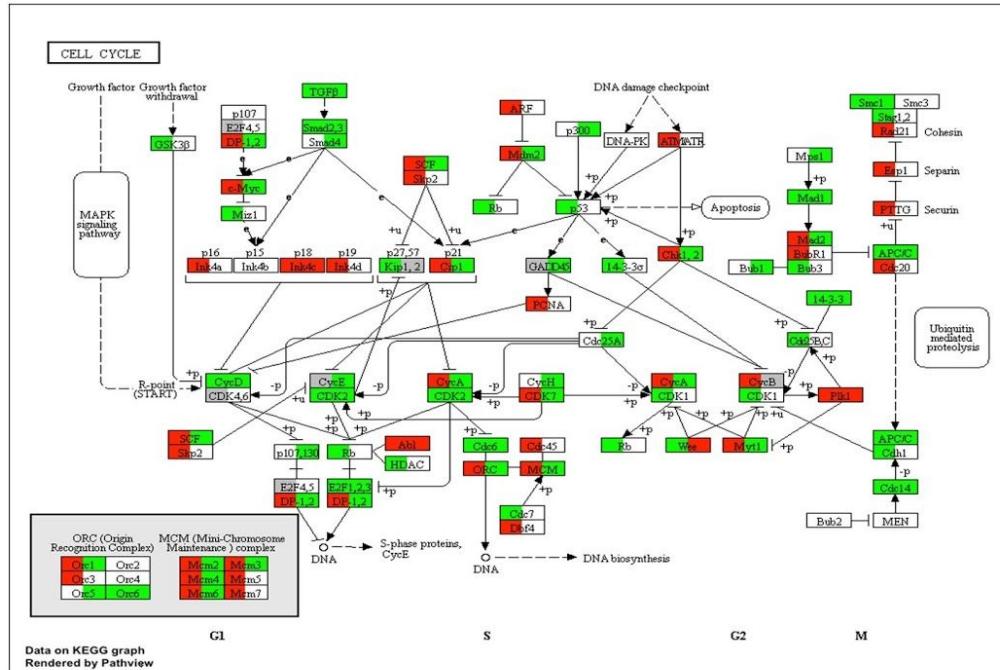


**D**

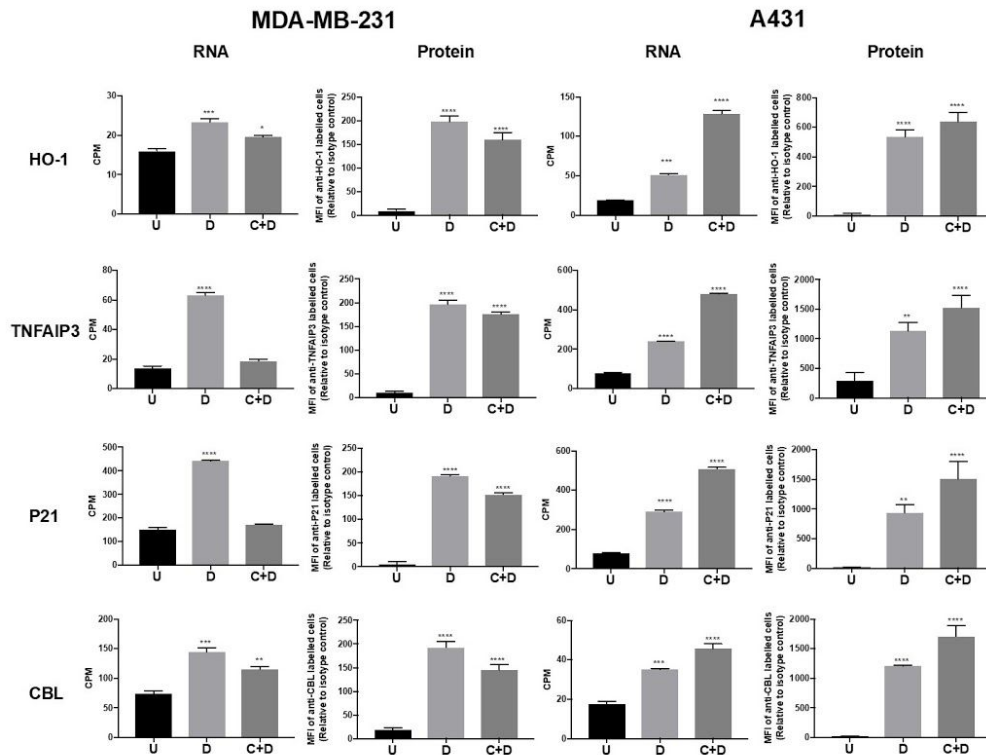


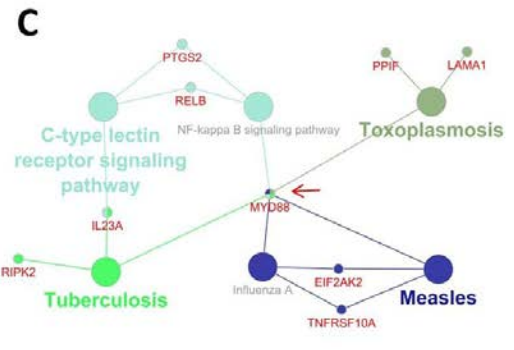
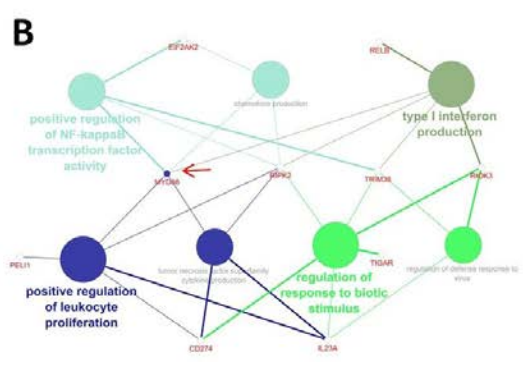
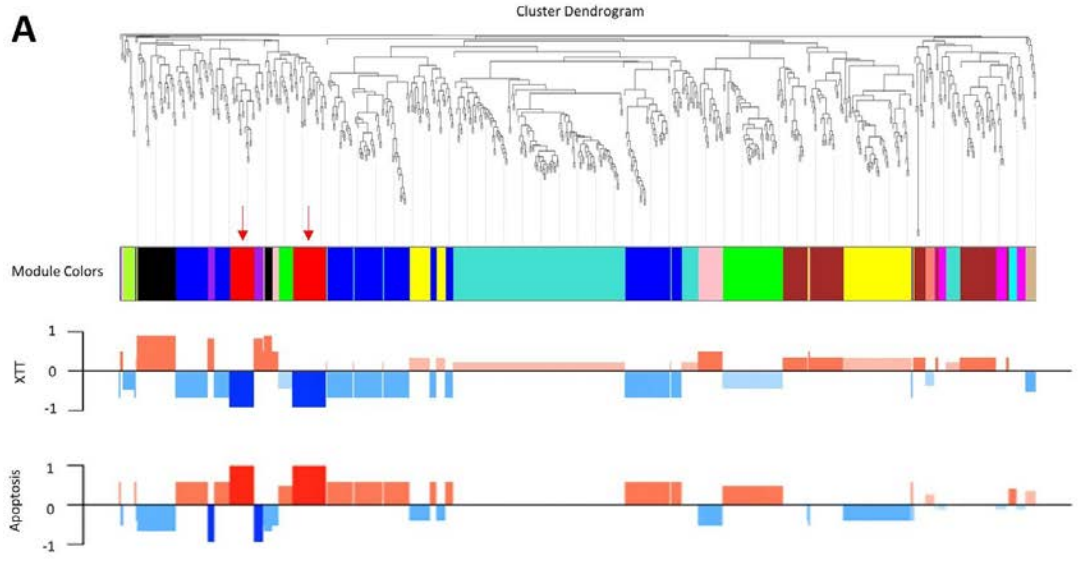


**A**

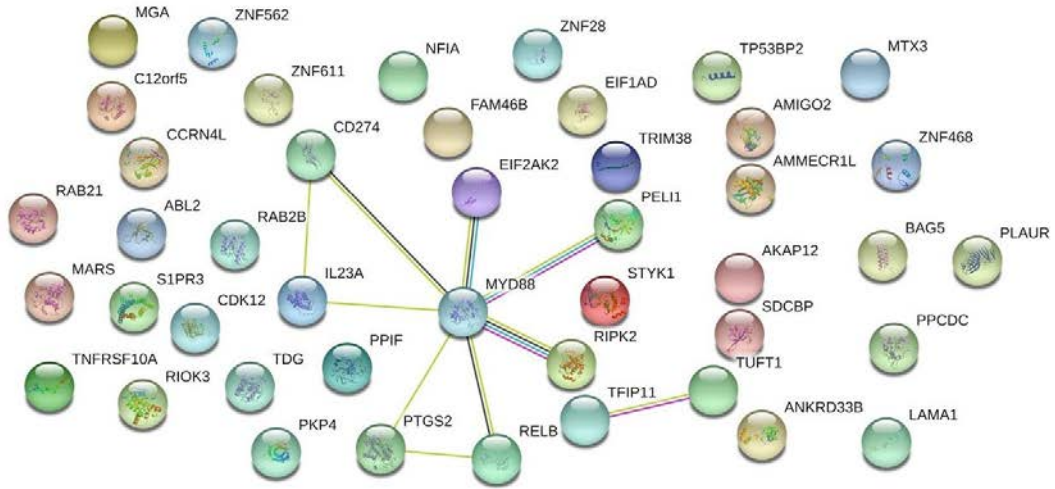


**B**

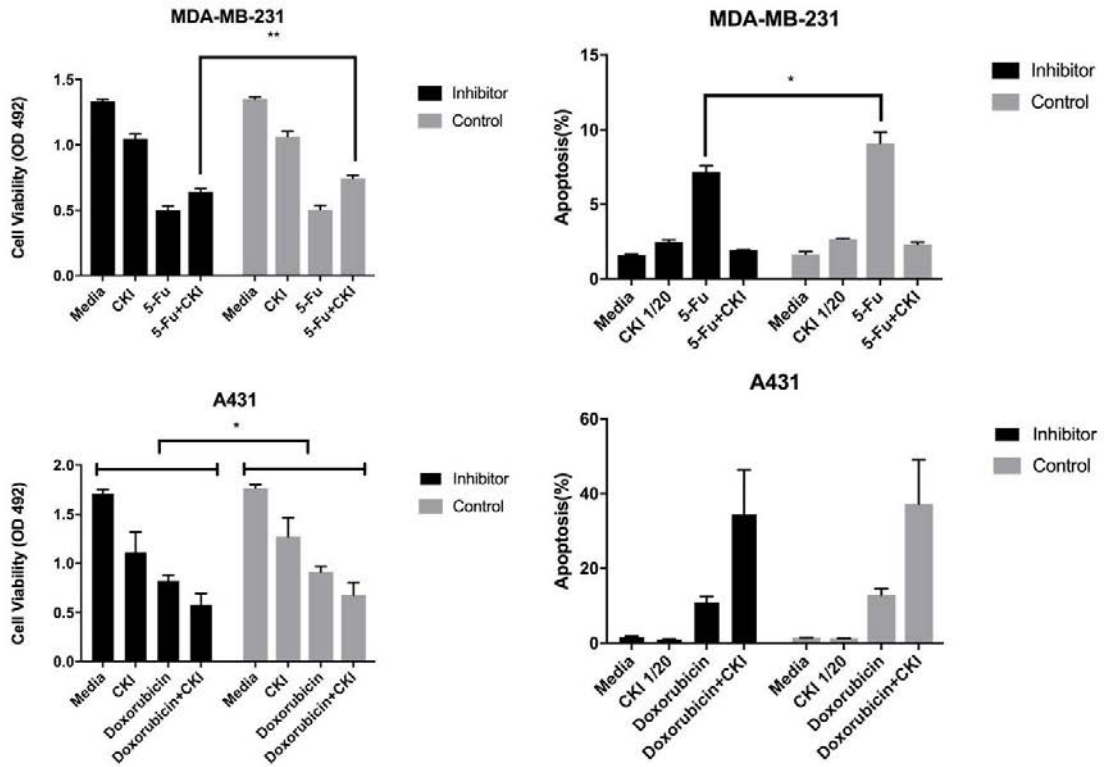




**A**



**B**



## Chapter 3

### Comprehensive Understanding the Compatibility Mechanisms in Compound Kushen Injection with Transcriptome Analysis

The study of herbal compatibility is an essential part of TCM research, as clinical formulas usually contain multiple herbal medicines. In this chapter, we introduced transcriptome analysis as a means to carry out herbal compatibility studies for CKI. Cellular and analytical chemistry results indicated that Kushen acts as the primary herb in the CKI formula, and contributes the major compounds and effects on cancer cells. Gene expression changes also confirmed that Kushen alone can perturb most pathways where CKI exerts its effects. However, we also demonstrated that Baituling can enhance the cytotoxic effects of Kushen and activate innate immune functions. These results are closely related to the description for these two herbs in TCM theory, which shows the network approach might be a good way to link modern science to ancient TCM theory.

# Statement of Authorship

Title of Paper	Understanding the mechanistic contribution of herbal extracts in Compound Kushen Injection with transcriptome analysis
Publication Status	<input type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input checked="" type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	

## Principal Author

Name of Principal Author (Candidate)	Hanyuan Shen		
Contribution to the Paper	Experimental design, carried out experiments, analysed data, wrote paper		
Overall percentage (%)	60%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	13/3/2019

## Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Zhipeng Qu		
Contribution to the Paper	Experimental design, assisted with experiments, assisted with data analysis, wrote paper		
Signature		Date	18/03/19

Name of Co-Author	Yuka Harata-Lee		
Contribution to the Paper	Experimental design, assisted with experiments, assisted with data analysis, wrote paper		
Signature		Date	13/3/19

Name of Co-Author	Thazin Nwe Aung		
Contribution to the Paper	Assisted with experiments		
Signature		Date	13/03/2019

Name of Co-Author	Jian Cui		
Contribution to the Paper	Assisted with experiments		
Signature		Date	19/03/2019

Name of Co-Author	Wei Wang		
Contribution to the Paper	Assisted with experimental design, assisted with experiments		
Signature		Date	19/03/2019

Name of Co-Author	Robert Daniel Kortschak		
Contribution to the Paper	Experimental design, wrote paper		
Signature		Date	15/3/19

Name of Co-Author	David L. Adelson		
Contribution to the Paper	Supervised the research, acquired funding for the experiments, experimental design, wrote paper		
Signature		Date	1/4/2019

# Understanding the Mechanistic Contribution of Herbal Extracts in Compound Kushen Injection with Transcriptome Analysis

Hanyuan Shen<sup>1</sup>, Zhipeng Qu<sup>1</sup>, Yuka Harata-Lee<sup>1</sup>, Thazin Nwe Aung<sup>1</sup>, Jian Cui<sup>1</sup>, Wei Wang<sup>2</sup>, R. Daniel Kortschak<sup>1</sup>, and David L. Adelson\*<sup>1</sup>

<sup>1</sup> Zhendong Australia - China Centre for Molecular Chinese Medicine, School of Biological Sciences, University of Adelaide, Adelaide, South Australia, 5005.

<sup>2</sup> Zhendong Research Institute, Shanxi-Zhendong Pharmaceutical Co Ltd, Beijing, China.

\*Corresponding author: David L. Adelson

Department of Molecular and Biomedical Science, School of Biological Sciences, University of Adelaide, Adelaide, South Australia, 5005.

Telephone : +61 8 8303 7555

Email: [david.adelson@adelaide.edu.au](mailto:david.adelson@adelaide.edu.au)

Running title: Identification of herbal compatibility mechanisms with transcriptome data

Key words: transcriptome, herb combination, traditional Chinese medicine, cancer treatment

## Abstract

Herbal compatibility is the knowledge of which herbs to combine in traditional Chinese medicine (TCM) formulations. The lack of understanding of herbal compatibility is one of the key problems for the application and popularization of TCM in western society. Because of the chemical complexity of herbal medicines, it is simpler to begin to conduct compatibility research based on herbs rather than component plant secondary metabolites. We have used transcriptome analysis to explore the effects and interactions of two plant extracts (Kushen and Baituling) combined in Compound Kushen Injection (CKI). Based on shared chemical compounds and *in vitro* cytotoxicity comparisons, we found that both the major compounds in CKI, and the cytotoxicity effects of CKI were mainly derived from the extract of Kushen (*Sophorae flavescens*). We generated and analyzed transcriptome data from MDA-MB-231 cells treated with single-herb extracts or CKI and results showed that Kushen contributed to the perturbation of the majority of cytotoxicity/cancer related pathways in CKI such as cell cycle and DNA replication. We also found that Baituling (*Heterosmilax yunnanensis Gagnep*) could not only enhance the cytotoxic effects of Kushen in CKI, but also activate immune-related pathways. Our analyses predicted that IL-1 $\beta$  gene expression was upregulated by Baituling in CKI and we confirmed that IL-1 $\beta$  protein expression was increased using an ELISA assay. Altogether, these findings help to explain the rationale for combining Kushen and Baituling in CKI, and show that transcriptome analysis using single herb extracts is an effective method for understanding herbal compatibility in TCM.



## Introduction

At present many complex and chronic diseases rely on therapies that combine modern pharmaceuticals. A similar multiple-herb strategy known as “Fufang” is an essential component in traditional Chinese medicine (TCM) theory and is used to achieve better therapeutic results, and reduce side effects and herbal toxicity<sup>1-2</sup>. As result of thousands of years’ of accumulated clinical practice, TCM has more than 100,000 formulae and abundant experience that has contributed to an understanding of which herbs should be combined in particular circumstances, herein referred to as herbal compatibility<sup>3,4</sup>. However, because herbal medicines are made up of complex mixtures of plant secondary metabolites, the mechanism of most TCM formulas has not been explored. This limitation of TCM has become one of the key problems for its modernization, and hinders the application and popularization of herbal medicines<sup>5</sup>.

Recent rapid developments in analytical chemistry and molecular biology have provided methods for researchers to tackle the complex mechanisms of herbal compatibility on a number of different levels. Usually, these methods focus on one or a few components within a complex mixture, in attempts to reveal how preparation/extraction for combined use can change their concentrations in products or pharmacokinetic processes *in vivo*<sup>6-9</sup>. However, as a complex mixture may contain thousands of compounds, it is unclear how changing one or several components in a TCM formula can explain and account for the principles and observations of herbal compatibility. Furthermore, pharmacological models that measure phenotypes associated with efficacy or proxies for efficacy are limited in their ability to explain potential therapeutic effects and mechanisms. New high-throughput technologies for measuring molecular phenotypes such as gene expression, and bioinformatic methods can provide systematic ways for refining and clarifying complex biological processes that result from

hundreds or thousands of molecular interactions. By applying these methods to the study of TCM, it is possible to transform the research paradigm from “main active compound that influences one target” to “multiple components that influence many network targets”<sup>10,11</sup>. Although it is now common to apply RNA-sequencing and systematic methods to study the effects of whole TCM formulae, no literature has applied these methods to study herbal compatibility. In this report, we apply transcriptome analysis to identify how the combination of Kushen (*Sophorae flavescens*) and Baituling (*Heterosmilax yunnanensis* Gagnep) extracts can account for the broader and increased effects observed in Compound Kushen Injection (CKI).

Our model system for dissecting herbal compatibility, CKI, is derived from an ancient Chinese formula and was approved by the State Food and Drug Administration (SFDA) of China in 1995. It is widely used as an adjuvant medicine in the treatment of carcinomas for pain relief, activation of innate immune response and reduced side effects in cancer therapy<sup>12,13</sup>. Our previous results have shown that CKI suppresses the growth of cancer cells by inhibiting cell cycle, energy metabolism, and DNA repair pathways<sup>11,14</sup>. Kushen is considered to be the principal herb and major contributor to the molecular effects of CKI. Many published studies have reported on the alkaloids and flavonoids contained in CKI, most of which are extracted from Kushen. These compounds have been reported to have a variety of bioactivities, including antitumor, antioxidant and anti-inflammatory activities<sup>15,16</sup>. However, there is no literature that mentions the role of Baituling in CKI. Therefore, current studies are not sufficient to provide a rational framework for CKI prescription or explain its molecular mechanisms.

In this report, we break down the formula into its individual components in order to study the herbal compatibility of CKI. By comparing transcriptome changes in MDA-MB-231 cells between

CKI and single herbal extract treatments, we found that Kushen extract alone, perturbed most of the pathways through which CKI exerts its effects on cancer cells. However, integrating Kushen with Baituling can enhance the effects of Kushen alone on cancer-related pathways and in addition can activate innate immune functions. These results support the TCM rationale behind the CKI formula and confirm RNA-sequencing as a useful tool for the identification of candidate mechanisms in TCM research.

## **Results**

### **HPLC comparison of the composition of CKI and single extract injections.**

In order to obtain information about plant specific components in CKI, we used high-performance liquid chromatography (HPLC) to compare the compounds in CKI and two single herb extracts/injections (Fig.1). From the chromatographic profile, it can be seen that Kushen injection contributes most of the major chemical components in CKI. In contrast, few compounds were detected in Baituling injection, which only contributes one major compound to CKI. Based on comparison with 9 reference standard compounds, 8 main compounds derived from Kushen (adenine, N-methylcytisine, sophorodine, matrine, sophocarpine, oxysophocarpine, oxymatrine and trifolirhizin) are shown to contribute to CKI. Macrozamin, which is used as a control marker for Baituling during manufacturing, only appeared in CKI and Baituling injection. These results indicate that CKI contains major compounds from both Kushen and Baituling, and Kushen contributes most of the major chemical components in CKI.

### **Comparison of the anticancer effects between CKI and single injections.**

Our previous results showed CKI suppressed the proliferation of and induced apoptosis in MCF-7 cells<sup>11</sup>. To determine whether single injections had similar phenotypic effects as CKI, we

conducted XTT assays to measure cell viability using three different cell lines; MDA-MB-231, A431 and HepG2. Results showed that Kushen had stronger cytotoxic effects than Baituling in all three cell lines. However, neither of the single injections had apoptotic effects comparable to CKI (Fig.2A) based on rates of apoptosis determined by flow cytometry with propidium iodide (PI) staining. Consistent with XTT cell viability results, more apoptotic cells were found in CKI than single injections treatments and Baituling had the smallest effect on apoptosis of the three injections (Fig.2B).

### **Comparison of MDA-MB-231 transcriptomes from CKI and single injection treatments.**

In order to elucidate the molecular mechanisms of herbal compatibility in CKI, we carried out transcriptome profiling from CKI and the single injection treated MDA-MB-231 cells. Triplicate samples for each treatment clustered well in multidimensional scaling plots and different treatments were clearly separated (Supplementary Fig 1). Although CKI and Kushen injection have similar chemical profiles, the inclusion of Baituling in CKI is sufficient to change the transcriptome of MDA-MB-231 cells compared to Kushen single injection treatment. We used edgeR<sup>17</sup> to identify differentially expressed (DE) genes for each injection treatment compared to untreated. In addition, we also identified DE genes for CKI treatment compared to Kushen treatment to identify the effects of Baituling in CKI (Supplementary Table 3).

To validate the results of transcriptome analysis, we performed quantitative PCR for several genes known to be important read-outs for the effects of CKI; TP53, CYD1A1 and CCND1. Their expression levels confirmed the interaction of Baituling with Kushen observed in the overall RNA sequencing results (Fig.2C).

### **Similar effects of CKI and Kushen single injection on MDA-MB-231 cells.**

Because Kushen is considered to be the primary active herb in CKI, we first examined the effects of Kushen and CKI to see if they had similar effects on genes. By comparing the DE gene set between Kushen and CKI, the shared DE gene group accounted for 63% and 81% of DE genes from CKI and Kushen respectively (Fig.3A), indicating that Kushen contributed to the majority of effects from CKI.

In order to better understand the functions of shared DE genes between Kushen and CKI, we performed over-representation analysis using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways for these 2039 genes (Fig.3B&C, Supplementary Table 5). The results showed that cell cycle- and DNA replication-related pathways and terms were largely down-regulated by both Kushen and CKI. Based on our previous publications, perturbed regulation of genes in these pathways and annotated by these terms was associated with the observed cell viability and apoptotic effects CKI on cancer cells<sup>11,14</sup>. Other terms showing perturbation of metabolic processes and cell migration identified in this study, such as 'pyrimidine metabolism', 'steroid biosynthesis' and 'positive regulation of locomotion', also showed up in our previous research on the effects of CKI. Altogether, these results indicated that Kushen was very important to the major molecular consequences of CKI treatment, and perturbed most of the functions perturbed by CKI, including reduced viability and apoptosis in cancer cells.

### **Different effects of CKI and single injections on MDA-MB-231 cells**

Although the above results for enrichment analysis showed that Kushen and CKI mainly regulate the same pathways, they did not specifically show the magnitude and direction of these perturbations. We, therefore, performed Signalling Pathway Impact Analysis (SPIA) to compare the significantly perturbed functional pathways across the different treatments. Ninety two pathways were found to be significantly perturbed by CKI while only 30 of them were shown to be activated. For Kushen and Baituling, the ratios of activated/inhibited were 29/100 and 24/58 respectively (Supplementary Table 4). Clearly, Baituling perturbed fewer pathways, but the ratio of activated/inhibited was higher than for Kushen or CKI. Interestingly, most pathways were perturbed in the same way (inhibited or activated) by CKI and Kushen as shown in Fig.4A, with very few pathways showing different types of perturbation.

This similarity of effects between Kushen and CKI could also be seen in the high level of correlation for pathway perturbation between Kushen and CKI (0.83 correlation coefficient) compared to (0.33 correlation coefficient) for Baituling and CKI. CKI also had stronger perturbation effects on most pathways (Fig.4C). In pathways contributing to cytotoxic effects in cancer cells, such as cell cycle, p53 signaling pathway, proteoglycans in cancer and pathways in cancer, Baituling perturbed the pathways in the same direction as Kushen and seemed to reinforce those effects in CKI. However, the cytokine-cytokine receptor interaction pathway was very interesting as it was perturbed in an opposite fashion in Baituling compared to Kushen and did not show up as significantly perturbed by CKI treatment (Supplementary Fig. 2). By comparing DE genes for Baituling and Kushen in the cytokine-cytokine receptor interaction pathway (Fig.5), we observed that many genes were oppositely regulated by the two single injections, such as genes in the CXC subfamily and IL6/12-like cytokine and receptor genes, which supported the SPIA results.

## **The functions of Baituling in CKI.**

In order to investigate the function of Baituling in CKI, we identified the DE genes of CKI treatment compare to Kushen treatment. Only 308 DE genes were found (Fig.6A). KEGG analysis of these genes showed that with the exception of steroid hormone biosynthesis and transcriptional misregulation in cancer, all other pathways were related to immune function, and most genes in this set were up-regulated (Fig.6B, Supplementary Table 5). This result was consistent with findings in the SPIA analysis that Baituling tended to activate immune-related pathways. The over-represented GO terms for this gene set also included interferon-gamma production, organ or tissue specific immune response and interleukin-2 production, all aspects of immune function and which only contained up-regulated DE genes (Fig.6C).

The genes in the common set between Kushen and CKI compared to Kushen (Fig. 6A) were originally changed with Kushen treatment and then further significantly regulated in combination with Baituling (107 genes). Only three pathways (IL-17 signaling, salmonella infection and steroid hormone biosynthesis) were over-expressed by genes in this set, indicating Baituling could also modify functions related to immune system and hormone function upregulated by Kushen in CKI (Supplementary Fig 3.). Furthermore, 24 genes appeared in both Baituling and CKI compared to Kushen set (Fig. 6A), which can be regarded as the direct contribution from Baituling to the effects of CKI. Analysis of known protein-protein interactions in this gene set identified the IL-1 family and interacting proteins known to modulate immune function (Fig.7A). To verify that IL-1 protein levels were also changing as predicted, we carried out an ELISA assay to measure the IL-1 beta levels in different treatments. We were able to demonstrate that the increased IL-1 beta level in CKI treatment compared to untreated was mainly related to the

effects of Baituling (Fig.7B). Altogether, these results showed that Baituling contributed to the effects of CKI primarily by altering functions related to the immune system.

In summary, we characterized the herbal compatibility of Kushen and Baitulin in CKI by comparing the individual effects of the herbal extracts to the combined extract using transcriptome analysis. In this fashion, we were able to explain the origin of CKI's different effects in MDA-MB-231 cells. In addition, we also showed that Baituling could enhance the reduction of cell viability and increased apoptosis effects from Kushen in CKI. These results not only explained the specific molecular basis of the TCM rationale of combining Kushen with Baituling but also illustrate a general method to apply transcriptome analysis to study herbal compatibility in TCM.

## **Discussion**

It is undeniable that chemical composition is the basis of therapeutic effects from herbal TCM. However, because the identification and quantification of all compounds for even a single herb are still extremely difficult if not impossible, we need alternative means to conduct TCM research, particularly with respect to the study of herbal compatibility<sup>18,19</sup>. Herbal compatibility has a basis in TCM theory, but TCM theory is not generally accepted in Western medicine and it is difficult to map concepts from TCM theory to Western medicine. Methods that can identify the molecular consequences of TCM formulations and individual herbs can begin to provide such a map. One view of TCM is that it perturbs multiple targets or pathways with multiple low activity components to generate relatively strong effects. This is in contrast to the standard approach for pharmaceutical drug development which seeks to identify single compounds that inhibit a single pathway or target. However, this method of using one or several active compounds as



representative of single herbs is problematic<sup>10</sup> for understanding the roles of individual plant extracts in TCM formulations. This is illustrated by our results; although containing similar amounts of the main chemical compounds, CKI has much stronger effects than Kushen extract alone. Therefore, a formula disassembly approach that uses single herbs is a practical way to study herbal compatibility. Combined with omics techniques and network analysis, we can represent the mechanisms of TCM herbal compatibility as interactions between target networks familiar to Western medicine. In this report, we took CKI, a prescription containing only two herbs, as a proof of principle. However, related methods can easily be applied to more complex formulae.

As a frequently used herb, and because it is considered to contain the main bioactive components in CKI, Kushen or its main alkaloids and flavonoids are commonly used to represent CKI in studies<sup>20,21</sup>. Furthermore, TCM theory also regards Kushen as the primary herb in CKI. Our results support these hypotheses and TCM theory at different levels. First, at the chemical level, HPLC profiles showed that the source of most major components in CKI is Kushen. Second, in terms of overall efficacy, Kushen has much stronger cytotoxic effects than Baituling on various cell lines. Third, at the gene level, DE genes shared by Kushen and CKI account for 81% and 63% of total DE genes in Kushen and CKI treatments, and most of them are consistently up- or down-regulated. Furthermore, important genes are also regulated both by Kushen and CKI. Cytochrome P450 family 1 subfamily A member 1 (CYP1A1) gene, a steroid metabolizing enzyme which is important for steroid hormone responsive cancers and shown as the most over-expressed gene with CKI in our previous results, is also highly overexpressed by Kushen but not Baituling<sup>22,23</sup>. Also, the down-regulation of the *TP53* gene primarily results from Kushen. Our previous results based on CKI treated cells underwent

apoptosis while expression of *TP53* decreased, which indicates the apoptosis induced by Kushen and CKI is not *TP53*-dependent<sup>11</sup>. Finally, GO and KEGG over-representation analysis also indicated that the genes in major cancer related pathways and terms perturbed by CKI, including cell cycle, DNA replication, and cell migration, are induced by the shared DE gene set between Kushen and CKI. Taken together, our findings agree with TCM theory which considers Kushen as the principle herb in CKI, and they map this specific part of TCM theory to effects on specific genetic networks and pathways.

On the other hand, we could find no Baituling literature and studies on macrozamin to support its application in CKI<sup>24</sup>. From our HPLC and cytotoxic assay results, Baituling injection does not contain many components and has no significant effects on cancer cells. In addition, RNAseq analysis of Baituling-treated samples are close to the untreated samples in the multidimensional scaling plot, and only 253 DE genes compared to untreated were detected. However, after comparing the differences between CKI and Kushen, we found Baituling has a strong reinforcing effect on Kushen. The SPIA results showed Baituling can enhance many pathways' perturbation strength compared to Kushen treatment, including cell cycle, pathways in cancer and proteoglycans in cancer. In addition, many DE genes in immune-related pathways and GO terms are over-represented in CKI compared to Kushen. Together with the opposing direction of perturbation for the cytokine-cytokine receptor interaction pathway caused by Kushen and Baituling, we can conclude Baituling may also contribute to the immune regulatory effects of CKI. This was confirmed by our measurements of IL-1 $\beta$ , which was significantly up-regulated by Baituling and CKI. In summary, our results indicate that Baituling, an adjuvant herb in CKI according to TCM theory, may enhance the anticancer effects of Kushen and contribute to immune regulation.

In conclusion, we explained the TCM herbal compatibility of CKI in the context of pathway perturbations using transcriptome analysis. Kushen primarily contributes to CKI effects on cancer cells by perturbing cell cycle regulation and other functions<sup>11,14</sup>, meanwhile, Baituling enhanced potential anticancer effects for Kushen and activated the immune system. Therefore, the two herbs in CKI collaborate with each other in effects, which is very similar to formulation theory in TCM. Compared to previous studies on herbal compatibility, our method can explain the beneficial interaction pattern of herbs in TCM formulae in a more systematic and comprehensive fashion.

## Methods

### Cell culture and drugs

MDA-MB-231, HepG2 and A431 cells were purchased from ATCC (VA, USA). All cell lines were cultured at 37°C with 5% CO<sub>2</sub> in DMEM (Thermo Fisher Scientific, MA, USA) with 10% fetal bovine serum (Thermo Fisher Scientific). CKI, Baituling, and Kushen injections were provided by Zhendong pharmaceutical Co.Ltd (China). Baituling and Kushen injections were manufactured using the same processes as CKI and diluted to the equivalent concentration of CKI (total alkaloid concentration at 2 mg/ml).

All *in vitro* assays were conducted in 6-well or 96-well plates. The seeding density was  $4 \times 10^5$  for 6-well plates across all three cell lines. For 96-well plates, MDA-MB-231 cells, A431 cells and HepG2 cells were seeded at  $1.6 \times 10^5$  cells/well,  $8 \times 10^4$  cells/well and  $4 \times 10^3$  cells/well

respectively. Cells were cultured overnight before drug treatment and the treatment time was 48 hours for all assays.

### **Components comparison with HPLC**

CKI, Kushen or Baituling injection was diluted 1:20 with MilliQ water and then analyzed on a photodiode-array UV-Vis detector equipped Shimadzu HPLC (Japan) with a preparative C<sub>18</sub> column (5 µm, 250 x 10 mm, Phenomenex, CA, USA). The recording range is from 200 nm to 280 nm, with monitoring at 215 nm. 0.01M ammonium acetate (adjusted to pH 8.0, solvent A) and acetonitrile + 0.09 % trifluoroacetic acid (solvent B) were used as mobile phase and flow rate is 2 ml/min with linear gradient elution (0 min, 100 % A; 60 min, 65 % A, 70 min, 100 % A). Nine Standard compounds, including Oxymatrine, Oxysophocarpine, N-methylcytisine, Matrine, Sophocarpine, Trifolirhizin, Adenine, Sophoridine (Beina Biotechnology Institute Co., Ltd, China), and macrozamin (Zhendong Pharmaceutical Co.Ltd, China), were used to characterize peaks in the HPLC profile.

### **Cell viability assay**

Cells were cultured and treated in 96-well plates. After 48 hours drug treatment, 50 µl of XTT:PMS (at 1 mg/ml and 1.25 mM, respectively, and combined at 50:1 ratio, Sigma-Aldrich, MO, USA) was added into each well and incubated 4 hours for the measurement of cell viability. A Biotrack II microplate reader was used to detect the absorbance at 492 nm.

### **Apoptosis rate with cell cycle assay**

After treatment, cells were harvested from 6-well plates and stained with propidium iodide (PI; Sigma-Aldrich) as previously described<sup>25</sup>. The stained cells were quantified on a BD LSR

Fortessa-X20 (BD Biosciences, NJ, USA) and the data were analyzed with FlowJo software (TreeStar Inc., OR, USA).

### **qPCR for transcriptome validation**

The assay was performed as previously described<sup>11</sup>. The sequences of all primers are shown in the Supplementary Table.

### **RNA extraction and sequencing**

After treatment with injections, MDA-MB-231 cells were harvested from 6-well plates and snap-frozen with liquid nitrogen. Total RNA was isolated using the PureLink RNA mini kit (Thermo Fisher Scientific). Quality and quantity of RNAs were measured with a Bioanalyzer at the Cancer Genome Facility of the Australian Cancer Research Foundation (Australia) to ensure RIN>7.0 and sent to Novogene (China) for sequencing with paired-end 150 bp reads on an Illumina HiSeq X platform.

### **Transcriptome data analysis**

The adaptors and low-quality sequences in raw reads were trimmed with Trim\_galore (v0.3.7, Babraham Bioinformatics) using parameters: --stringency 5 --paired. STAR (v2.5.3a) was used to align reads to reference genome (hg19, UCSC) with parameters: --outSAMstrandField intronMotif --outSAMattributes All --outFilterMismatchNmax 10 --seedSearchStartLmax 30<sup>26</sup>. Differentially expressed genes were calculated with edgeR (v3.22.3) and selected with false discovery rate (FDR) < 0.05 and Log fold change >1 or <-1<sup>17</sup>.

The GO and KEGG over-representation analyses were performed with ClueGO and visualized with Cytoscape v3.6.0 with following parameters: right-sided hypergeometric test for enrichment analysis; p values were corrected for multiple testing according to the Benjamini-Hochberg method and biological process at 3rd level for GO terms<sup>27,28</sup>. The Signalling Pathway Impact Analysis (SPIA) package in R was used to conduct the pathway perturbation analysis using all DE genes(FDR<0.05)<sup>29</sup>. R Pathview package was used to visualize specific KEGG pathways<sup>30</sup>. String (V11.0) was used to identify protein-protein interactions with a threshold of 0.4 for minimum interaction score<sup>31</sup>.

### **ELISA for IL-1 $\beta$ level**

A431 cells were treated with different injections for 48 hours in 96-well plates and the cell culture supernatant was collected and tested for the level of IL-1 $\beta$  by ELISA using human interleukin-1 beta ELISA kit (Biosensis, CA, USA) according to the kit protocol. The absorbance at 450nm was detected with Multiskan Ascent Plate Reader.

1. Huang, J., Tang, X., Ye, F., He, J. & Kong, X. Clinical Therapeutic Effects of Aspirin in Combination with Fufang Danshen Diwan, a Traditional Chinese Medicine Formula, on Coronary Heart Disease: A Systematic Review and Meta-Analysis. *Cell. Physiol. Biochem.* **39**, 1955–1963 (2016).
2. Hu, X.-Q., Sun, Y., Lau, E., Zhao, M. & Su, S.-B. Advances in Synergistic Combinations of Chinese Herbal Medicine for the Treatment of Cancer. *Curr. Cancer Drug Targets* **16**, 346–356 (2016).
3. Wang, S. *et al.* Compatibility art of traditional Chinese medicine: from the perspective of herb pairs. *J. Ethnopharmacol.* **143**, 412–423 (2012).
4. Zhou, M., Hong, Y., Lin, X., Shen, L. & Feng, Y. Recent pharmaceutical evidence on the compatibility rationality of traditional Chinese medicine. *J. Ethnopharmacol.* **206**, 363–375 (2017).
5. Cai, F.-F., Zhou, W.-J., Wu, R. & Su, S.-B. Systems biology approaches in the study of Chinese herbal formulae. *Chin. Med.* **13**, (2018).
6. Zhou, J. *et al.* Identification and Analysis of Compound Profiles of Sinisan Based on ‘Individual Herb, Herb-Pair, Herbal Formula’ before and after Processing Using UHPLC-Q-TOF/MS Coupled with Multiple Statistical Strategy. *Molecules* **23**, (2018).
7. Liu, P., Shang, E.-X., Zhu, Y., Qian, D.-W. & Duan, J.-A. Volatile component interaction effects on compatibility of Cyperi Rhizoma and Angelicae Sinensis Radix or Chuanxiong Rhizoma by UPLC-MS/MS and response surface analysis. *J. Pharm. Biomed. Anal.* **160**, 135–143 (2018).
8. Jin, Y. *et al.* Pharmacokinetic Comparison of Seven Major Bio-Active Components in Normal and Blood Stasis Rats after Oral Administration of Herb Pair Danggui-Honghua by UPLC-TQ/MS. *Molecules* **22**, (2017).

9. Zhang, Y., Zhang, Z. & Song, R. The Influence of Compatibility of Rhubarb and Radix Scutellariae on the Pharmacokinetics of Anthraquinones and Flavonoids in Rat Plasma. *Eur. J. Drug Metab. Pharmacokinet.* **43**, 291–300 (2018).
10. 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.* **7**, 352 (2017).
11. Qu, Z. *et al.* Identification of candidate anti-cancer molecular mechanisms of Compound Kushen Injection using functional genomics. *Oncotarget* **7**, 66003–66019 (2016).
12. Ao, M., Xiao, X. & Li, Q. Efficacy and safety of compound Kushen injection combined with chemotherapy on postoperative Patients with breast cancer: A meta-analysis of randomized controlled trials. *Medicine* **98**, e14024 (2019).
13. Ma, X. *et al.* The Therapeutic Efficacy and Safety of Compound Kushen Injection Combined with Transarterial Chemoembolization in Unresectable Hepatocellular Carcinoma: An Update Systematic Review and Meta-Analysis. *Front. Pharmacol.* **7**, 70 (2016).
14. Cui, J. *et al.* Cell cycle, energy metabolism and DNA repair pathways in cancer cells are suppressed by Compound Kushen Injection. *BMC Cancer* **19**, 103 (2019).
15. Wang, W. *et al.* Anti-tumor activities of active ingredients in Compound Kushen Injection. *Acta Pharmacol. Sin.* **36**, 676–679 (2015).
16. Hu, H. *et al.* Synthesis and in vitro inhibitory activity of matrine derivatives towards pro-inflammatory cytokines. *Bioorg. Med. Chem. Lett.* **20**, 7537–7539 (2010).
17. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
18. Huang, T. *et al.* Approaches in studying the pharmacology of Chinese Medicine formulas: bottom-up, top-down-and meeting in the middle. *Chin. Med.* **13**, 15 (2018).



19. He, X.-R. *et al.* High-performance liquid chromatography coupled with tandem mass spectrometry technology in the analysis of Chinese Medicine Formulas: A bibliometric analysis (1997-2015). *J. Sep. Sci.* **40**, 81–92 (2017).
20. Gao, L. *et al.* Uncovering the anticancer mechanism of Compound Kushen Injection against HCC by integrating quantitative analysis, network analysis and experimental validation. *Sci. Rep.* **8**, 624 (2018).
21. Jin, Y. *et al.* Compound kushen injection suppresses human acute myeloid leukaemia by regulating the Prdxs/ROS/Trx1 signalling pathway. *J. Exp. Clin. Cancer Res.* **37**, 277 (2018).
22. Kim, J. *et al.* Abstract 2232: Dietary flavonoids, CYP1A1 genetic variants, and the risk of colorectal cancer. *Cancer Res.* **77**, 2232–2232 (2017).
23. Nguyen, C. H. *et al.* AHR/CYP1A1 interplay triggers lymphatic barrier breaching in breast cancer spheroids by inducing 12(S)-HETE synthesis. *Hum. Mol. Genet.* **25**, 5006–5016 (2016).
24. Lythgoe, B. & Riggs, N. V. Macrozamin, a toxic nitrogen-containing primeveroside. *Experientia* **5**, 471 (1949).
25. Riccardi, C. & Nicoletti, I. Analysis of apoptosis by propidium iodide staining and flow cytometry. *Nat. Protoc.* **1**, 1458–1461 (2006).
26. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
27. Bindea, G. *et al.* ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091–1093 (2009).
28. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
29. Tarca, A. L. *et al.* A novel signaling pathway impact analysis. *Bioinformatics* **25**, 75–82

(2008).

30. Luo, W. & Brouwer, C. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics* **29**, 1830–1831 (2013).
31. Jensen, L. J. *et al.* STRING 8--a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res.* **37**, D412–6 (2009).

Figure 1: HPLC profiles of CKI, Kushen and Baituling injections. Nine component compounds characterized using standard compounds are marked with red arrows.

Figure 2: Comparison of the effect of CKI and single herb injections on cancer cell lines. A. and B. Viability and the percentage of apoptotic cells treated with different injections for 48 hours. C. q-PCR validation of RNA sequencing results. Results are represented as means  $\pm$ SEM (n=9). Statistical analyses were performed using the t-test compared to untreated "Media" (\*p<0.05, \*\*p<0.01, \*\*\*p<0.001, \*\*\*\*p<0.0001)

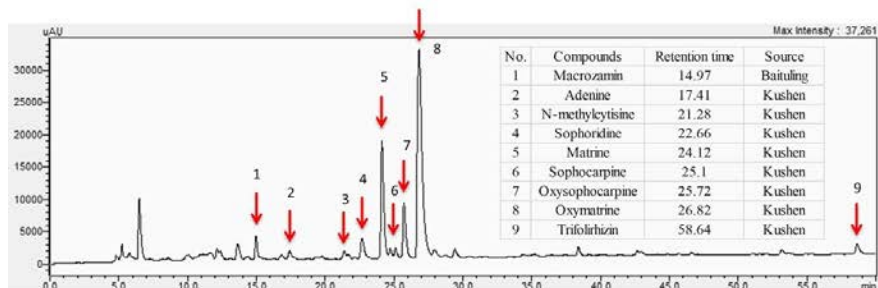
Figure 3: Significantly differentially expressed genes shared by Kushen and CKI treated cells and their functional enrichment analysis. A. Venn diagram showing the number of differentially regulated genes in MDA-MB-231 cells treated with Kushen (KS; blue) or CKI (yellow) compared to untreated cells. B. and C. Over-represented KEGG pathways and GO terms (Biological Process at 3rd level) for genes similarly regulated by Kushen and CKI. Node size is proportional to the statistical significance of over-representation and colors represent the proportion of up or down-regulated genes (yellow=up-regulated and blue=down-regulated). Similar GO terms are clustered (with representative terms shown in bold) and connected with edges.

Figure 4: Pathway perturbation analysis for CKI and single injections on MDA-MB-231 cells. A. and B. Perturbation accumulation and significance of perturbation for different KEGG pathways treated with different injections. Positive perturbation accumulation values mean the pathway is activated and *vice versa*. Dot colors indicate whether the pathway was significantly perturbed by CKI and/or Kushen. C. Heatmap showing the perturbation value of shared significantly perturbed pathways for the three injections. Table on the right-bottom corner shows the correlation coefficients for the perturbation value for the three injections. "\*" and "\*\*\*" represent p<0.05 and p< 0.001 respectively (Pearson's correlation test).

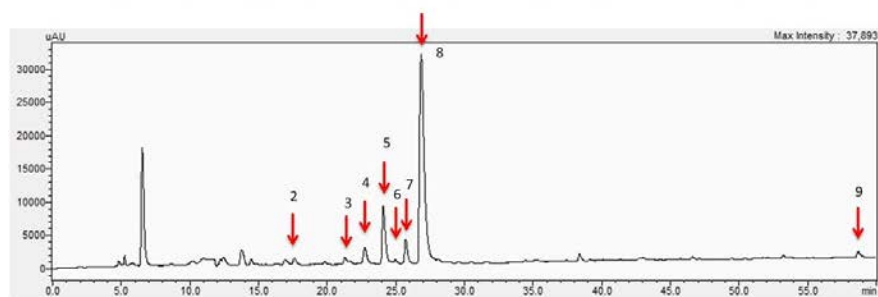
Figure 5: Comparison of gene expression changes caused by Baituling and Kushen treatments in the cytokine-cytokine receptor interaction pathway. Left half of each box represents the gene expression change with Baituling treatment and the right half represents the effect of Kushen treatment. White or grey colors indicate no significant change in gene expression as a function of treatment.

Figure 6: DE genes regulated by Baituling in MDA-MB-231 cells line and their functional enrichment analysis. A. Venn diagram showing the number of differentially regulated genes with CKI compared to Kushen (CKI-KS - Blue) and single herb injections (KS - Green or BTL - Yellow) compared to untreated. C. and D. Over-represented KEGG pathways and GO terms (Biological Process at 3rd level) for DE genes calculated by CKI compared to Kushen treated. Node size is proportional to the statistical significance of over-representation and colors represent the proportion of up or down-regulated genes (yellow=up-regulated and blue=down-regulated). Terms are clustered based on similar GO group (shown in bold) and related ones are connected with edges.

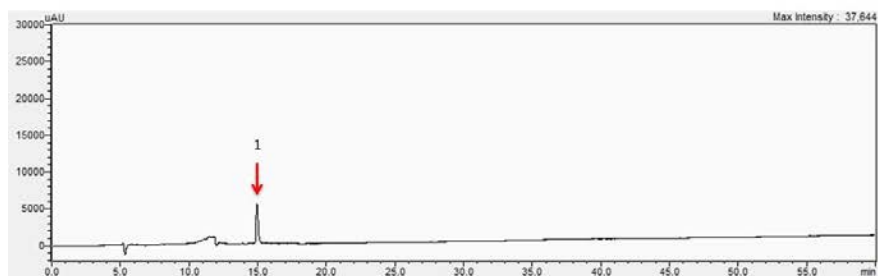
Figure 7: Validation of IL-1 $\beta$  expression changes regulated by Baituling. A. Diagram showing protein-protein interactions for common DE genes between CKI compared to Kushen, and Baituling compared to untreated. B. Comparison of expression levels and protein concentration for IL-1 $\beta$  with different treatments. Top panel; IL-1 $\beta$  gene expression levels determined by RNA sequencing, bottom panel; levels of IL-1 $\beta$  in culture supernatant measured by ELISA.



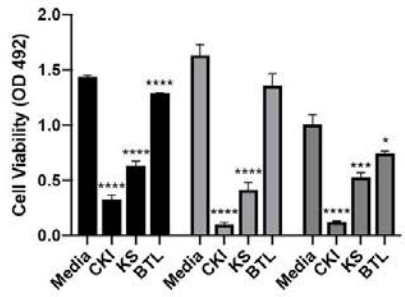
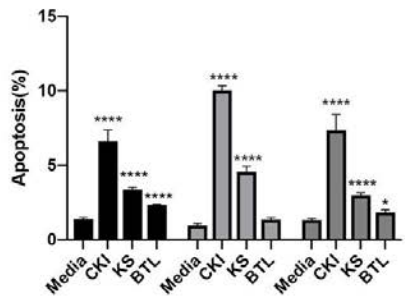
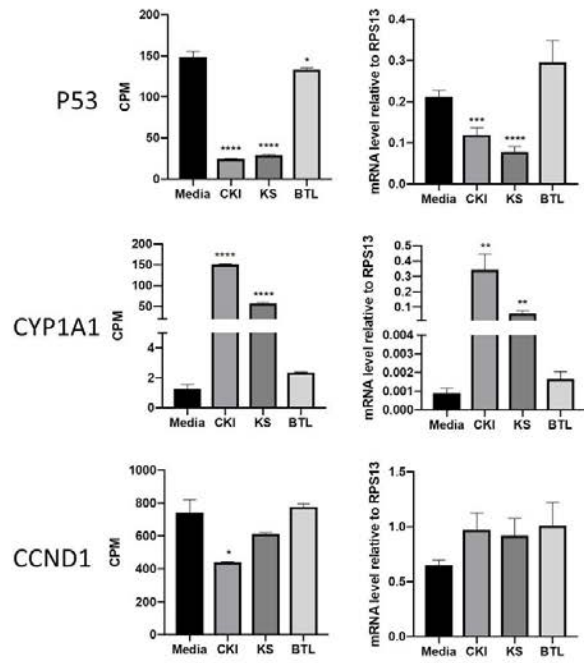
CKI

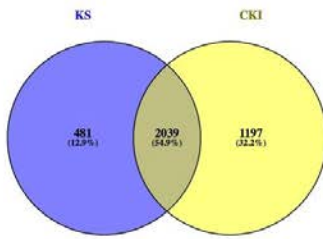
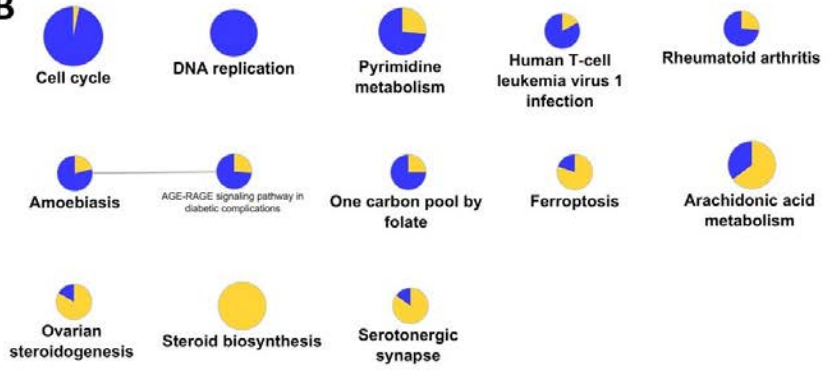
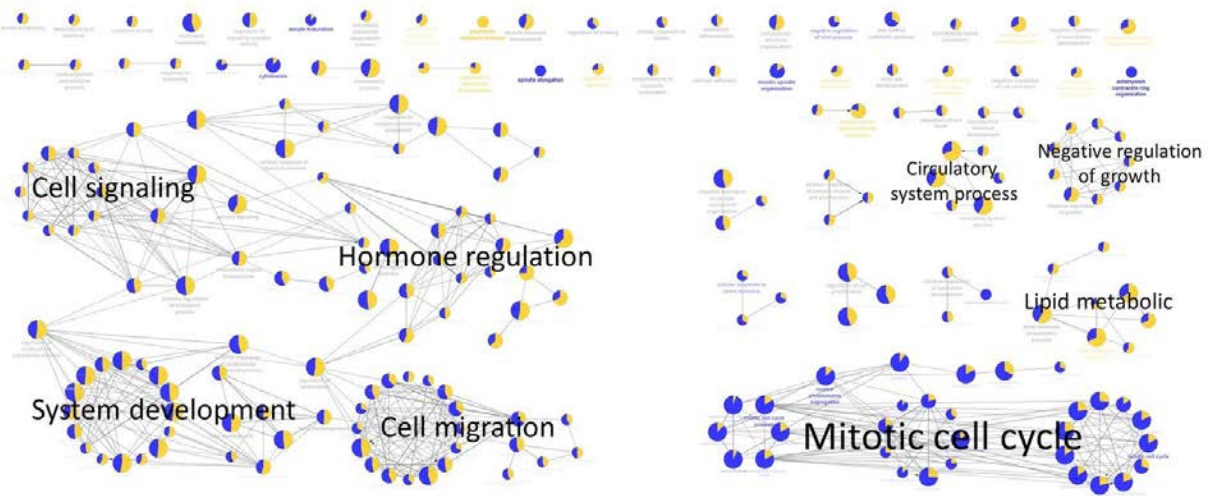


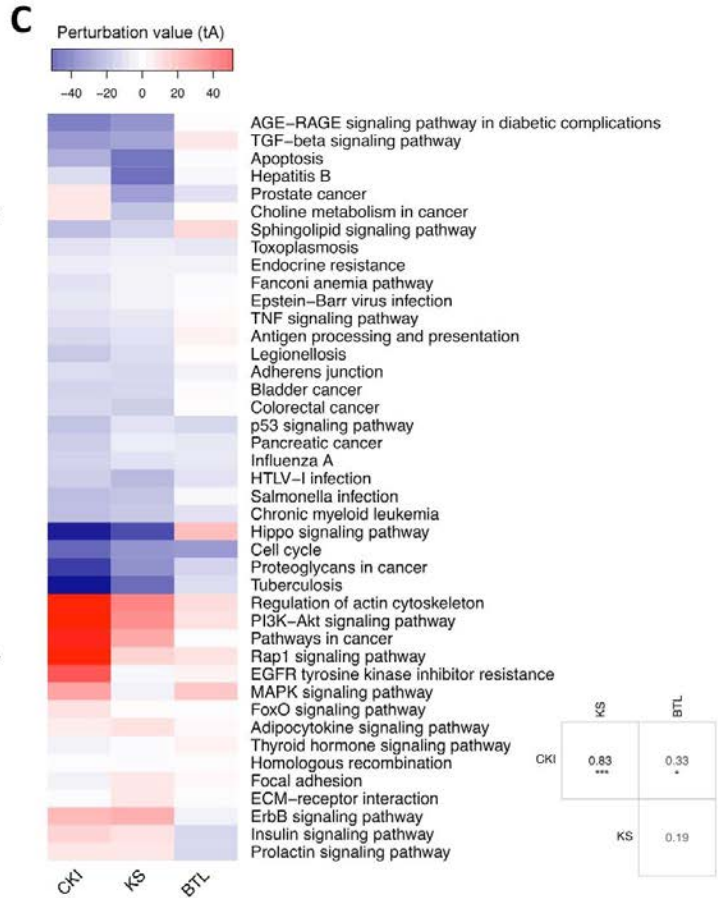
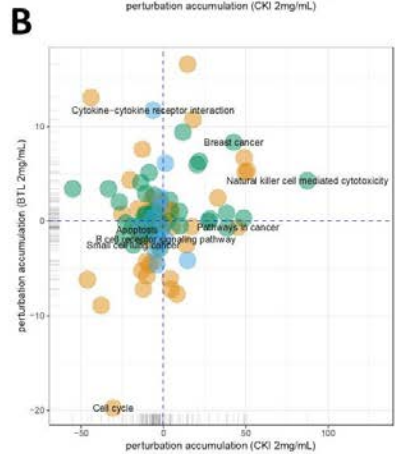
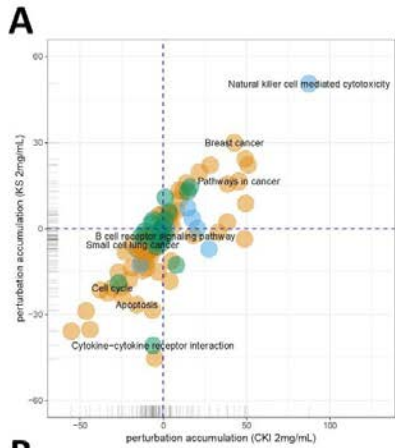
Kushen



Baituling

**A****B****C**

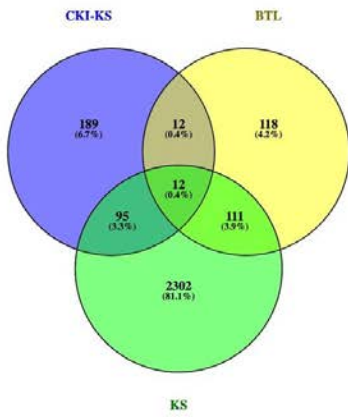
**A****B****C**



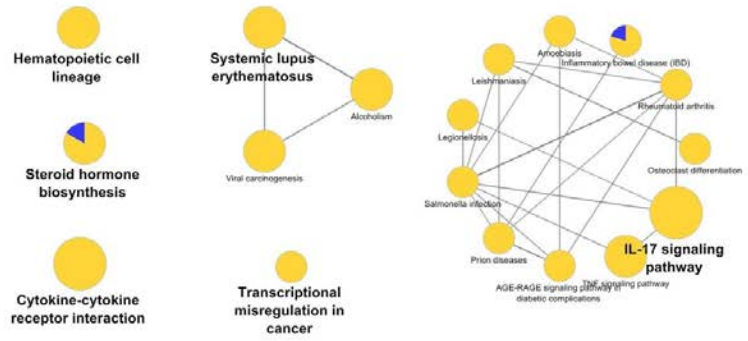




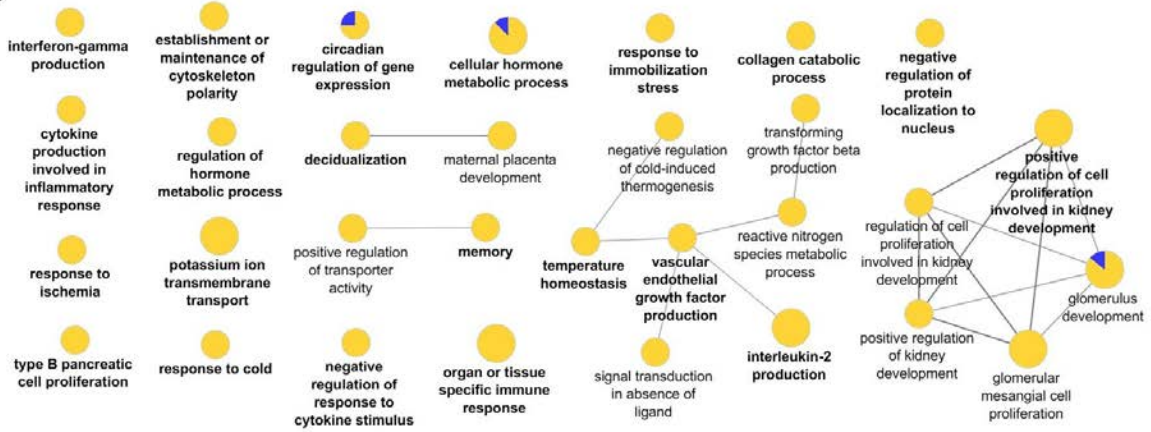
**A**

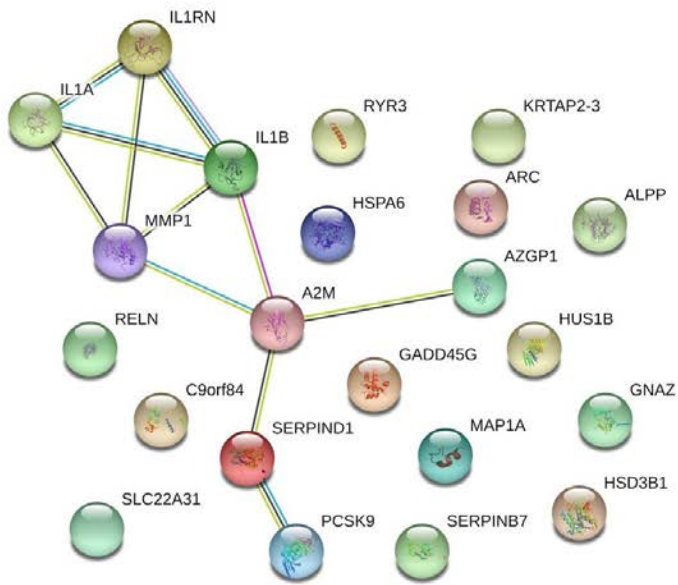
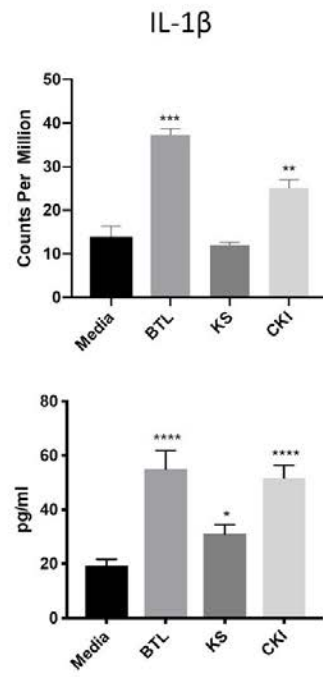


**B**



**C**



**A****B**

## Chapter 4

### **The Effect of Compound Kushen Injection on Cancer Cells: Integrated Identification of Candidate Molecular Mechanisms.**

In this chapter, I contributed to the exploration of transcriptome changes from CKI treatment on MDA-MB-231 and HEPG2 cells. Together with previously published data set from MCF-7 cells, the common specific genes/ pathways which CKI exerts its effects on cancer cells were identified. 363 common genes were found significantly regulated among three cell lines and mainly associated with cell cycle, apoptosis, DNA replication/ repair, and various cancer pathways. Many of the common DE genes were related to natural compounds in CKI based on an online database. This chapter is in the format of a manuscript that has been submitted to *BMC Complementary and Alternative Medicine*.

# Statement of Authorship

Title of Paper	The Effect of Compound Kushen Injection on Cancer Cells: Integrated Identification of Candidate Molecular Mechanisms.
Publication Status	<input type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input checked="" type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	

## Principal Author

Name of Principal Author (Candidate)	Hanyuan Shen		
Contribution to the Paper	Assisted with experiments		
Overall percentage (%)	20%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the fourth author of this paper.		
Signature		Date	13/3/2019

## Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Jian Cui		
Contribution to the Paper	Experimental design, carried out experiments, analysed data, wrote paper		
Signature		Date	19/03/2019

Name of Co-Author	Zhipeng Qu		
Contribution to the Paper	Experimental design, assisted with experiments, assisted with data analysis, wrote paper		
Signature		Date	18/03/2019

Name of Co-Author	Yuka Harata-Lee		
Contribution to the Paper	Experimental design, assisted with experiments, assisted with data analysis, wrote paper		
Signature		Date	13/3/19

Name of Co-Author	Thazin Nwe Aung		
Contribution to the Paper	Assisted with experiments		
Signature		Date	13/03/2019

Name of Co-Author	Wei Wang		
Contribution to the Paper	Assisted with experimental design, assisted with experiments		
Signature		Date	19/03/2019

Name of Co-Author	Robert Daniel Kortschak		
Contribution to the Paper	Experimental design		
Signature		Date	15/3/19

Name of Co-Author	David L. Adelson		
Contribution to the Paper	Supervised the research, acquired funding for the experiments, experimental design, wrote paper		
Signature		Date	1/4/2019

# The Effect of Compound Kushen Injection on Cancer Cells: Integrated Identification of Candidate Molecular Mechanisms.

Jian Cui<sup>1</sup>, Zhipeng Qu<sup>1</sup>, Yuka Harata-Lee<sup>1</sup>, Hanyuan Shen<sup>1</sup>, Thazin Nwe Aung<sup>1</sup>, Wei Wang<sup>1,2</sup>, R. Daniel Kortschak<sup>1</sup>, David L Adelson<sup>1,\*</sup>

*1 Department of Molecular and Biomedical Science, The University of Adelaide, North Terrace, Adelaide, South Australia, Australia, 5005*

*2 Zhendong Research Institute, Shanxi-Zhendong Pharmaceutical Co Ltd, Beijing, China*

*\* Corresponding author:*

*David L. Adelson,*

*Room 261, The Braggs Bldg, School of Biological Sciences, the University of Adelaide, South Australia, 5005, AUSTRALIA*

*Ph: +61 8 8303 7555*

*Fax: +61 8 8303 5338*

*E-mail: [david.adelson@adelaide.edu.au](mailto:david.adelson@adelaide.edu.au)*

---

## Abstract

**Background:** Because Traditional Chinese Medicine (TCM) preparations are often combinations of multiple herbs containing hundreds of compounds, they have been difficult to study. Compound Kushen Injection (CKI) is a complex mixture cancer treatment used in Chinese hospitals for over twenty years.

**Purpose:** To demonstrate that a systematic analysis of molecular changes resulting from complex mixtures of bioactives from TCM can identify a core set of differentially expressed (DE) genes and a reproducible set of candidate pathways.

**Study Design:** We used a cancer cell culture model to measure the effect of CKI on cell cycle phases, apoptosis and correlate those phenotypes with CKI induced changes in gene expression.

**Methods:** We treated cancer cells with CKI in order to generate and analyse high-throughput transcriptome data from two cancer cell lines. We integrated these differential gene expression results with previously reported results.

**Results:** CKI induced cell-cycle arrest and apoptosis and altered the expression of 363 core candidate genes associated with cell cycle, apoptosis, DNA replication/repair and various cancer

pathways. Of these, 7 are clinically relevant to cancer diagnosis or therapy and 14 are cell cycle regulators, and most of these 21 candidates are downregulated by CKI. Comparison of our core candidate genes to a database of plant medicinal compounds and their effects on gene expression identified one-to-one, one-to-many and many-to-many regulatory relationships between compounds in CKI and DE genes.

**Conclusions:** By identifying promising candidate pathways and genes associated with CKI based on our transcriptome-based analysis, we have shown this approach is useful for the systematic analysis of molecular changes resulting from complex mixtures of bioactives.

*Keywords:* Compound Kushen Injection, cancer cell, transcriptome, multiple targets, cell cycle, apoptosis

---

## Abbreviations:

DE, differentially expressed; TCM, traditional Chinese medicine; CKI, compound Kushen injection; GO, Gene Ontology; DO, Disease Ontology; KEGG, Kyoto Encyclopedia of Gene and Genomes; PI, propidium iodide.

## 1 Introduction

2 The treatments of choice for cancer are often radiotherapy and/or chemotherapy, and while  
3 these can be effective, they can cause quite serious side-effects, including death. These side-effects  
4 have driven the search for adjuvant therapies to both mitigate side-effects and/or potentiate the  
5 effectiveness of existing therapies. Traditional Chinese Medicine (TCM) is one of the options for  
6 adjuvant therapies, particularly in China, but increasingly so in the West. While clinical trial data  
7 on the effectiveness of TCM is currently limited, it remains an attractive option because of its long  
8 history and because its potential effectiveness is believed to result from the cumulative effects of  
9 multiple compounds on multiple targets [13]. Because TCM often has not been subjected to  
10 rigorous evidence-based assessment and because it is based on an alternative theoretical system  
11 compared to Western medicine, adoption of its plant derived therapeutics has been slow.

12 In this report, we continue to characterize the molecular effects of Compound Kushen Injection



13 (CKI) on cancer cells. CKI has been approved by the State Food and Drug Administration  
14 (SFDA) of China for clinical use since 1995 [31] (State medical license no. Z14021231). CKI is an  
15 herbal extract from two TCM plants, Kushen (*Sophora flavescens*) and Baituling (*Smilax Glabra*)  
16 and contains more than 200 different chemical compounds. These compounds include alkaloids  
17 and flavonoids such as matrine, oxymatrine and kurarinol that have been reported to have  
18 anti-cancer activities [31, 47, 35, 48]. Some of these activities have been shown to influence the  
19 expression of TP53, BAX, BCL2 and other key genes known to be important in cancer cell growth  
20 and survival [45, 38, 25, 17].

21 We have previously characterized the effect of CKI on the transcriptome of MCF-7 breast  
22 carcinoma cells and in this report, we extend our previous results to two additional human cancer  
23 cell lines (MDA-MB-231, breast carcinoma and Hep G2, hepatocellular carcinoma). Both cell lines  
24 have also been shown to undergo apoptosis in response to the ingredients of CKI [35, 48, 38, 44].  
25 Hep G2 is one of the most sensitive cancer cell lines with respect to exposure to CKI [39] and CKI  
26 is often used in conjunction with Western chemotherapy drugs for the treatment of liver cancer  
27 patients in China. While the specific mechanism of action of CKI is unknown, several recent  
28 studies have reported that CKI or its primary compounds affect the regulation/expression of  
29 oncogene products including  $\beta$ -catenin, TP53, STAT3 and AKT [31, 35, 22, 18, 41].

30 However, these and other reports did not evaluate the entire range of molecular changes from  
31 treatment with a multi-component mixture such as CKI [10, 8]. Whilst several research databases  
32 and tools for TCM research have been developed [32, 34, 5], they are limited by the fact that most  
33 of the studies that contribute to the corpus of these databases are from different experimental  
34 systems, use single compounds or measure effects based on one or a handful of genes/gene  
35 products.

36 In contrast to previous studies, our strategy was to carry out comprehensive transcriptome  
37 profiling and network reconstruction from cancer cells treated with CKI. Instead of focusing on  
38 specific genes or pathways in order to design experiments, we have linked phenotypic assessment  
39 and RNA-seq analysis to CKI treatment. This allows us to present an unbiased, comprehensive  
40 analysis of CKI specific responses of biological networks associated with cancer. Our results  
41 indicate that different cancer cell lines that undergo apoptosis in response to CKI treatment can

42 exhibit different CKI induced gene expression profiles that nonetheless implicate similar core genes  
43 and pathways in multiple cell lines.

44 The current study presents the effects of CKI on gene expression in cancer cells with an aim to  
45 identify candidate pathways and regulatory networks that may be perturbed by CKI *in vivo*. To  
46 this end we primarily use concentrations of CKI higher than used *in vivo* in order to be able to  
47 detect effects in the short time frames available to tissue culture experiments. We also combine  
48 our current analysis with previously published data to focus on a shared, much smaller set of  
49 candidate genes and pathways.

## 50 **Material and Methods**

### 51 *Cell culture and reagents*

52 CKI (total alkaloids concentration of 25 mg/ml) in 5 ml ampoules was provided by Zhendong  
53 Pharmaceutical Co. Ltd. (Beijing, China). Chemotherapeutic agent, Fluorouracil (5-FU) was  
54 purchased from Sigma-Aldrich (MO, USA). A human breast adenocarcinoma cell line,  
55 MDA-MB-231 and a hepatocellular carcinoma cell line Hep G2 were purchased from American  
56 Type Culture Collection (ATCC, VA, USA). The cells were cultured in Dulbecco's Modified Eagle  
57 Medium (Thermo Fisher Scientific, MA, USA) supplemented with 10% fetal bovine serum  
58 (Thermo Fisher Scientific). Both cell lines were cultured at 37°C with 5% CO<sub>2</sub>.

59 For all *in vitro* assays,  $4 \times 10^5$  cells were seeded in 6-well trays and cultured overnight before  
60 being treated with either CKI (at 1 mg/ml and 2 mg/ml of total alkaloids) or 5-FU (150  $\mu$ g/ml for  
61 Hep G2 and 20  $\mu$ g/ml for MDA-MB-231). As a negative control, cells were treated with medium  
62 only and labelled as "untreated". After 24 and 48 hours of treatment, cells were harvested and  
63 subjected to the downstream experiments.

### 64 *Cell cycle and apoptosis assay*

65 The assay was performed as previously described [28]. For each cell line, three operators  
66 replicated the assay twice in order to ensure reproducibility of the observations. The results were  
67 obtained by flow cytometry using either FACScanto or LSRII (BD Biosciences, NJ, US).

## 68 *RNA isolation and sequencing*

69 The treated cells were harvested, and the cell pellets were snap frozen with liquid nitrogen and  
70 stored at -80°C. Total RNA was isolated with PureLink™ RNA Mini Kit (Thermo Fisher  
71 Scientific) according to the manufacturer’s protocol. After quantified using a NanoDrop  
72 Spectrophotometer ND-1000 (Thermo Fisher Scientific), the quality of the total RNA was verified  
73 on a Bioanalyzer by Cancer Genome Facility (SA, Australia) ensuring all samples had RINs>7.0.

74 For both cell lines, the sequencing was performed in Ramaciotti Centre for Genomics (NSW,  
75 Australia). The sample preparation for each cell line was TruSeq Stranded mRNA-seq with dual  
76 indexed, on the NextSeq500 v2 platform. The parameter was 75bp paired-end High Output. The  
77 fastq files were generated and trimmed through Basespace with application FASTQ Generation  
78 *v1.0.0*.

## 79 *Bioinformatics analysis of RNA sequencing*

80 The clean Hep G2 reads were aligned to reference genome (hg38) using STAR v2.5.1 with  
81 following parameters: `-outFilterMultimapNmax 20 -outFilterMismatchNmax 10 -outSAMtype`  
82 `BAM SortedByCoordinate -outSAMstrandField intronMotif` [6]. The clean MDA-MB-231 reads  
83 were aligned to reference genome (hg19) using TopHat2 v2.1.1 with following parameters:  
84 `-read-gap-length 2 -read-edit-dist 2` [15]. Differential expression analysis for reference genes was  
85 performed with edgeR and differentially expressed (DE) genes were selected with a False Discovery  
86 Rate<0.05 [29].

87 The DE genes in common for both Hep G2 and MDA-MB-231 cell lines at 24 hours and 48  
88 hours after CKI treatment were selected as “shared” genes. These shared genes were utilized to  
89 describe the major anti-cancer functions and principal mechanisms of CKI.

90 Gene Ontology (GO), and Kyoto Encyclopedia of Gene and Genomes (KEGG)  
91 over-representation analyses of both cell lines were carried out using the online database system  
92 ConsensusPathDB [14] with the following settings: “Biological process” at third level (for GO); q  
93 values (<0.01) were corrected for multiple testing with the system default settings. Disease  
94 ontology (DO) over-representation analyses of both cell lines were performed by using the  
95 Bioconductor R package clusterProfiler v3.5.1 [40]. For the function analyses of shared/core genes,

96 the method was as similar as our previous study [28] using ClueGO app 2.2.5 in Cytoscape v3.6.0.  
97 We enriched our GO terms in the biological process category level 3 and KEGG pathways,  
98 showing only terms/pathways with  $p$  values less than 0.01. Specific Over-represented  
99 terms/pathways and gene expression status mapping in KEGG pathways were visualised with the  
100 R package “Pathview” [21].

### 101 *Gene expression-based investigation of bioactive components in CKI*

102 To integrate with previous data from MCF-7 cells [28], all the shared DE genes regulated by  
103 CKI identified in all three cell lines using edgeR were mapped to the BATMAN-TCM database  
104 [19]. The pharmacophore modelling method [16] was used to generate the interaction network  
105 between the key genes and TCM components using R package igraph [4].

### 106 *Reverse transcription quantitative polymerase chain reaction (RT-qPCR)*

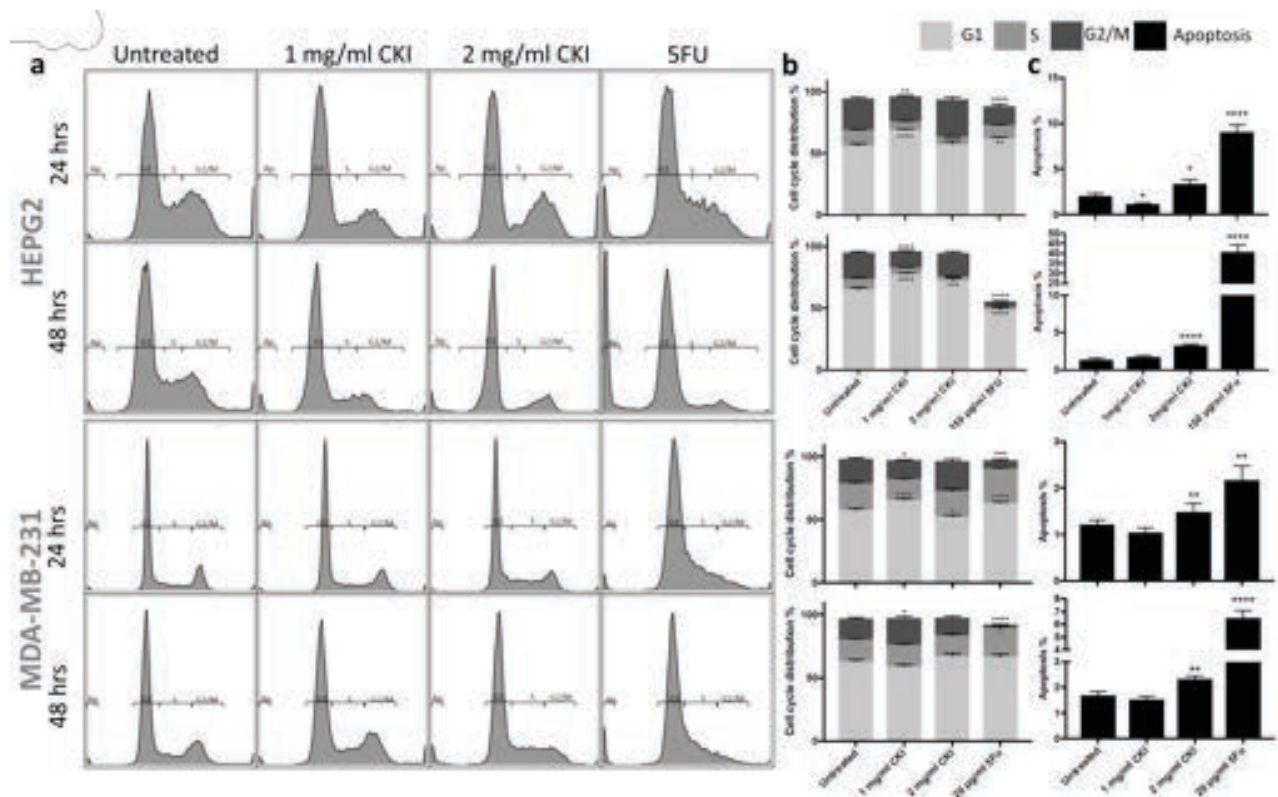
107 RT-qPCR was performed as previously described [28]. The list of target genes selected for this  
108 study and the sequences of all primers are shown in Additional file 1: Table S1.

## 109 **Results**

### 110 *Effect of CKI on the cell cycle and apoptosis*

111 In our previous study, CKI significantly perturbed/suppressed cancer cell target genes/networks.  
112 In the current study we present results that confirm and generalize our previous work. We  
113 observed in the MCF-7 study, low concentrations of CKI in our short-term cell assay showed  
114 no/little phenotypic effect within 48 hours, and very high doses resulted in excessive cell death at  
115 48 hours precluding the isolation of sufficient RNA for transcriptome analysis [28]. Therefore, in  
116 our current study with the two additional cell lines, to ensure consistency, we also selected 1  
117 mg/ml and 2 mg/ml total alkaloid concentrations of CKI for our assays because they generated  
118 reproducible and significant phenotypic effects in our cell culture assay.

119 We used flow cytometric analysis of propidium iodide stained cells to assess both CKI induced  
120 alterations to the cell cycle and apoptosis. In Hep G2 cells, CKI treatment resulted in an overall



**Figure 1. Effects of different treatment on cell cycle and apoptosis of Hep G2 and MDA-MB-231 cells.** A) The apoptosis and cell cycle distribution of each cell line after 24- and 48-hour treatments with CKI or 5-Fu assessed PI staining. B) Percentages of cells in different phases of cell cycle resulting from treatment. C) Percentage of apoptotic cells after treatment. Results shown are mean  $\pm$ SEM (n=6). Statistically significant differences from untreated control were identified using two-way ANOVA (\*p<0.05, \*\*p<0.01, \*\*\*\*p<0.0001).

121 increase in the proportion of cells in G1 phase and decrease in S phase (Fig. 1a and b). Similarly,  
 122 in MDA-MB-231 cells, although a consistent increase in G1 phase was not observed, CKI caused a  
 123 decrease in S phase particularly at the 24-hour time point (Fig. 1a and b) indicating possible  
 124 incidence of cell cycle arrest at G1 phase. Furthermore, at 2 mg/ml of total alkaloids, CKI  
 125 consistently induced significantly higher level of apoptosis in both cell lines at both time points  
 126 compared to untreated controls (Fig. 1c). These data together suggest that CKI has effects on the  
 127 cell cycle by interfering with the transition between G1 to S phase as well as by acting on the  
 128 apoptosis pathway and promoting cell death.

### 129 *CKI perturbation of gene expression*

130 In order to elucidate the molecular mechanisms of action of CKI on these cancer cells,  
 131 transcriptome analysis of CKI treated cells was performed. As mentioned above, RNA samples  
 132 from two cell lines were sequenced with 2 $\times$ 75 bp paired-end reads. We had previously sequenced

133 transcriptomes from CKI treated MCF-7 cells [28] and have included those results for comparison  
134 below. The samples from each cell line contained 7 groups at 3 time points (Fig. 2a), in triplicate  
135 for every group. In the multidimensional scaling (MDS) analysis, each cell line clustered  
136 independently and generally, within the cell line clusters, untreated cells clustered apart from  
137 treated cells (Additional file 2: Fig. S1).

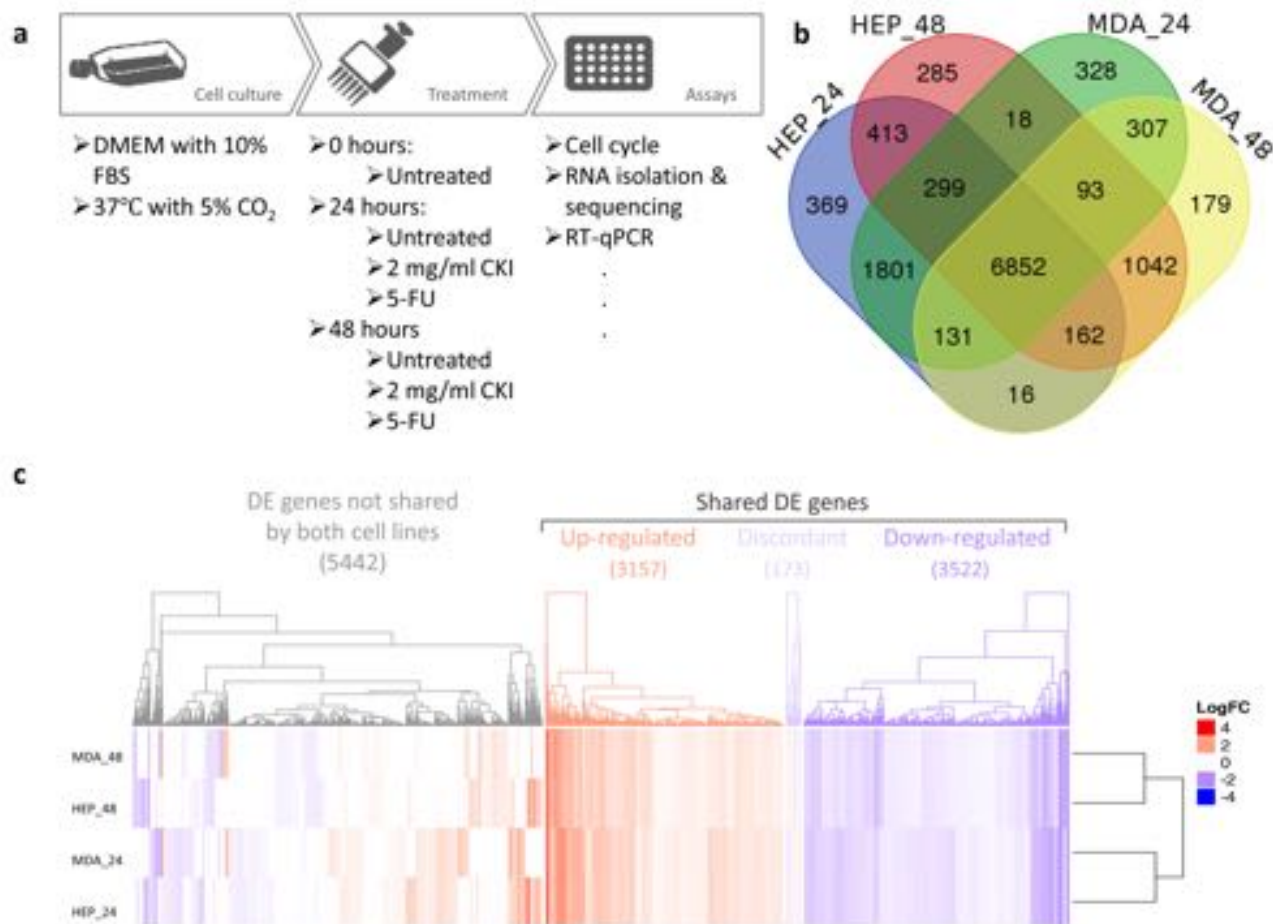
138 With the mapping rate were around 90% (Additional file 3: Table S2), a *p*-value based ranked  
139 list of DE genes (compared to untreated from each time point) was generated for both cell lines  
140 (Additional file 4: Table S3, sheet 1-4). This list was used to select the shared DE genes. This  
141 analysis generated thousands of DE genes (Additional file 4: Table S3, sheet 5) across Hep G2 and  
142 MDA-MB-231 cell lines.

143 Because for each cell line the respective treatment groups clustered together on the MDS plot,  
144 there were large numbers of shared genes between them. As a result, we identified a set of 6852  
145 shared DE genes by identifying common DE genes from Hep G2 and MDA-MB-231 cell lines, at 24  
146 hours and 48 hours (Fig. 2b). These shared genes might predict a common molecular signature for  
147 CKI's activity. However, there were still a large number of DE genes that were not shared by both  
148 cell lines, as seen in the heatmap in Fig. 2c. The expression of the shared gene set in both Hep G2  
149 and MDA-MB-231 is highly consistent. Interestingly, this consistency is with respect to treatment  
150 time, rather than with respect to cell line.

#### 151 *RT-qPCR validation and dose response of gene expression to CKI*

152 Based on our previous results [28], and analysis below, we selected the 4 top ranked DE genes  
153 expressed in G1-S phase of the cell cycle (TP53 and CCND1 for expression level validation and  
154 E2F2 and PCNA for low dose response), as well as the proliferation and differentiation relevant ras  
155 subfamily encoding gene (RAP1GAP1) for low dose response. We also selected a prominently  
156 expressed gene (CYP1A1) for validation because of its sensitivity to CKI treatment. CYP1A1,  
157 TP53 and CCND1 expression changes were validated with RT-qPCR with all three genes showing  
158 similar patterns of expression in the transcriptome data and RT-qPCR (Fig. 3a).

159 Because low dose treatment with CKI did not cause significant gross phenotypic effects in either  
160 cell line, we decided to use gene expression as a more sensitive measure of phenotype to look at



**Figure 2. DE genes shared in both cell lines at both time points.** A) Work flow diagram showing experimental design and sample collection. B) Venn diagram showing the number of shared DE genes between Hep G2 and MDA-MB-231. C) Heatmap presenting the overall gene expression pattern in both cell lines treated with CKI. Heatmap is split into four parts based on gene content and expression pattern: 5442 differentially regulated genes with expression not shared between the two cell lines, 3157 upregulated genes shared between both cell lines, 3522 down-regulated genes shared between both cell lines, and 173 discordantly regulated genes with differential expression shared between both cell lines.

161 the effect of lower doses of CKI. We used 0.125 mg/ml, 0.25 mg/ml, 0.5 mg/ml and 1 mg/ml  
162 concentrations to look for dose dependency of gene expression. Our results showed an obvious  
163 dose-dependent expression trend (Fig. 3b) in both cell lines. Because the 0.125 mg/ml  
164 concentration of CKI is equivalent to what cancer patients are treated with, our results are  
165 potentially clinically relevant.

### 166 *Function enrichment analysis*

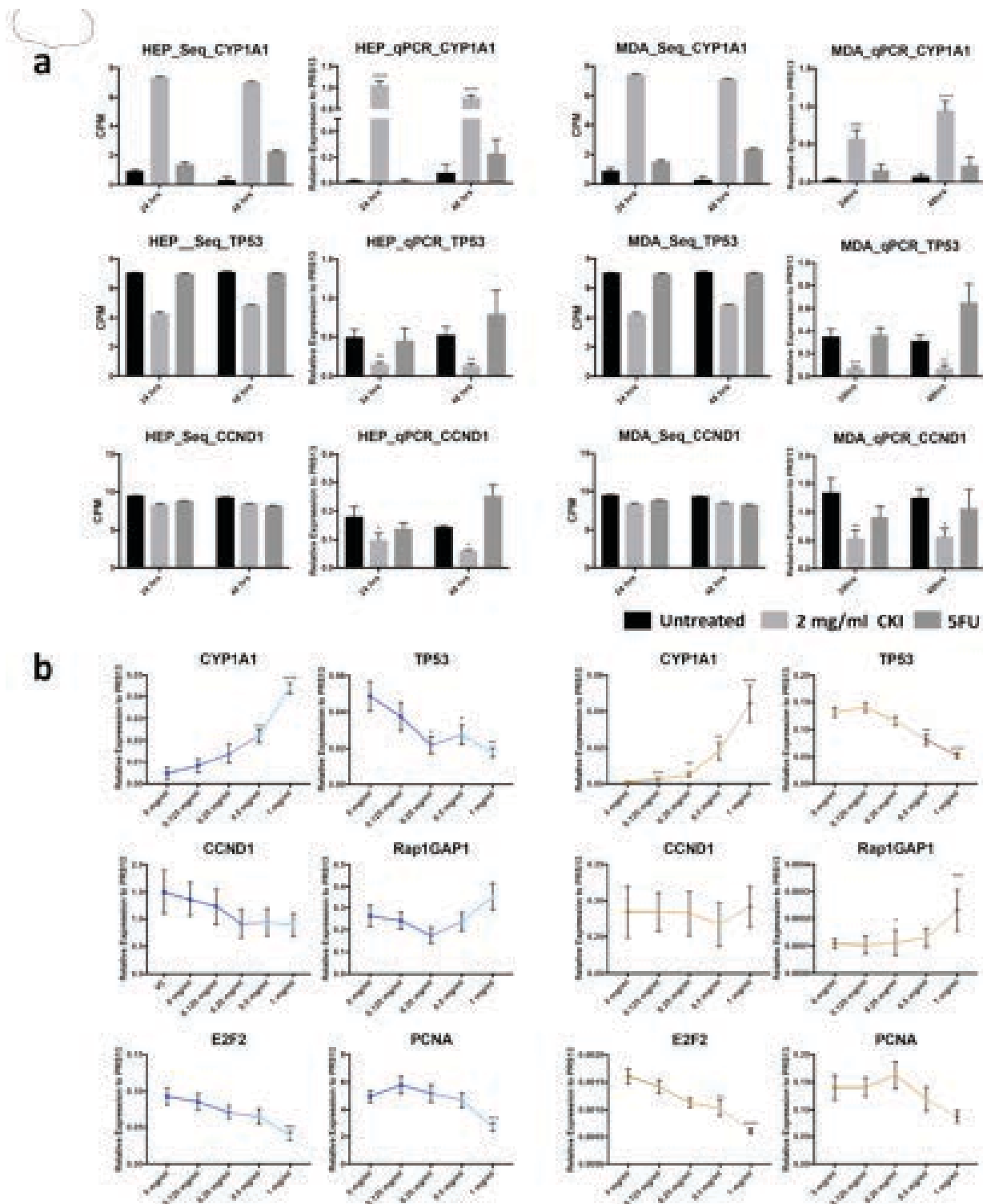
167 To identify candidate mechanisms of action of CKI, we carried out functional enrichment  
168 analysis. We used ConsensusPathDB [14] and Clusterprofiler [16] along with GO and KEGG  
169 pathways for over-representation analysis, along with disease ontology (DO) [30] enrichment.

170 GO over-representation test was determined based on Biological Process level 3 and q value  
171 <0.01. The results for both cell lines at both time points were summarized and visualized based  
172 on semantic analysis of terms in Fig. 4a. From this result, it was obvious that there were a large  
173 proportion of enriched GO terms relating to cell cycle, such as “cell cycle checkpoint”,  
174 “negative/positive regulation of cell cycle process” and so on prominently featured for all data sets  
175 (Additional file 5: Fig. S2, Additional file 6: Table S4, sheet 1-4).

176 We then used KEGG pathways to determine the specific pathways altered by CKI in cancer.  
177 The most regulated over-representative KEGG pathways are summarized according to KEGG  
178 Orthology (KO) (Fig. 4b). Cell cycle related pathways such as “cell cycle”, “DNA replication”,  
179 and “apoptosis” were also consistently seen in the KEGG enrichment results (Additional file 6:  
180 Table S4, sheet5-8) at both 24 and 48 hours. Moreover, in addition to the cell cycle relevant  
181 pathways, some cancer related pathways were also observed, such as “prostate cancer” and  
182 “chronic myeloid leukaemia”, and a large number of DE genes (283) from the two cell lines were  
183 relevant in “pathways in cancer”.

184 Because the KEGG enrichment revealed many pathways relating to diseases, most of which  
185 were cancers, we decided to explore the enrichment of DE genes with respect to DO terms (Fig.  
186 4c). In the DO list (Additional file 6: Table S4, sheet 9-12), all top ranked terms listed are cancers.  
187 Interestingly, most cancer types listed are from the lower abdomen, for example “ovarian cancer”,  
188 “urinary bladder cancer” and “prostate cancer” etc. occurring in genitourinary organs (Additional





**Figure 3. Validation of gene expression and effects of low dose CKI using RT-qPCR.** A) Comparison of DE genes between RNA-seq results (left) and RT-qPCR validation (right) for each cell line at 2 time-points. Three DE genes (CYP1A1, TP53 and CCND1) were chosen for validation. Gene expression was generally consistent between transcriptome data and qPCR data. B) Dose response of CKI using a subset of genes with conserved expression in Hep G2 (left), and MDA-MB-231 (right) from 0 mg/ml to 1 mg/ml of total alkaloids. Six genes (CYP1A1, TP53, CCND1, Rap2GAP1, E2F2 and PCNA) were selected based on their relevance to important pathways perturbed by CKI. RT-qPCR results are presented as expression relative to RPS13. Data are represented as mean  $\pm$  SEM ( $n > 3$ ). A t-test was used to compare CKI doses with “untreated” (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ ).

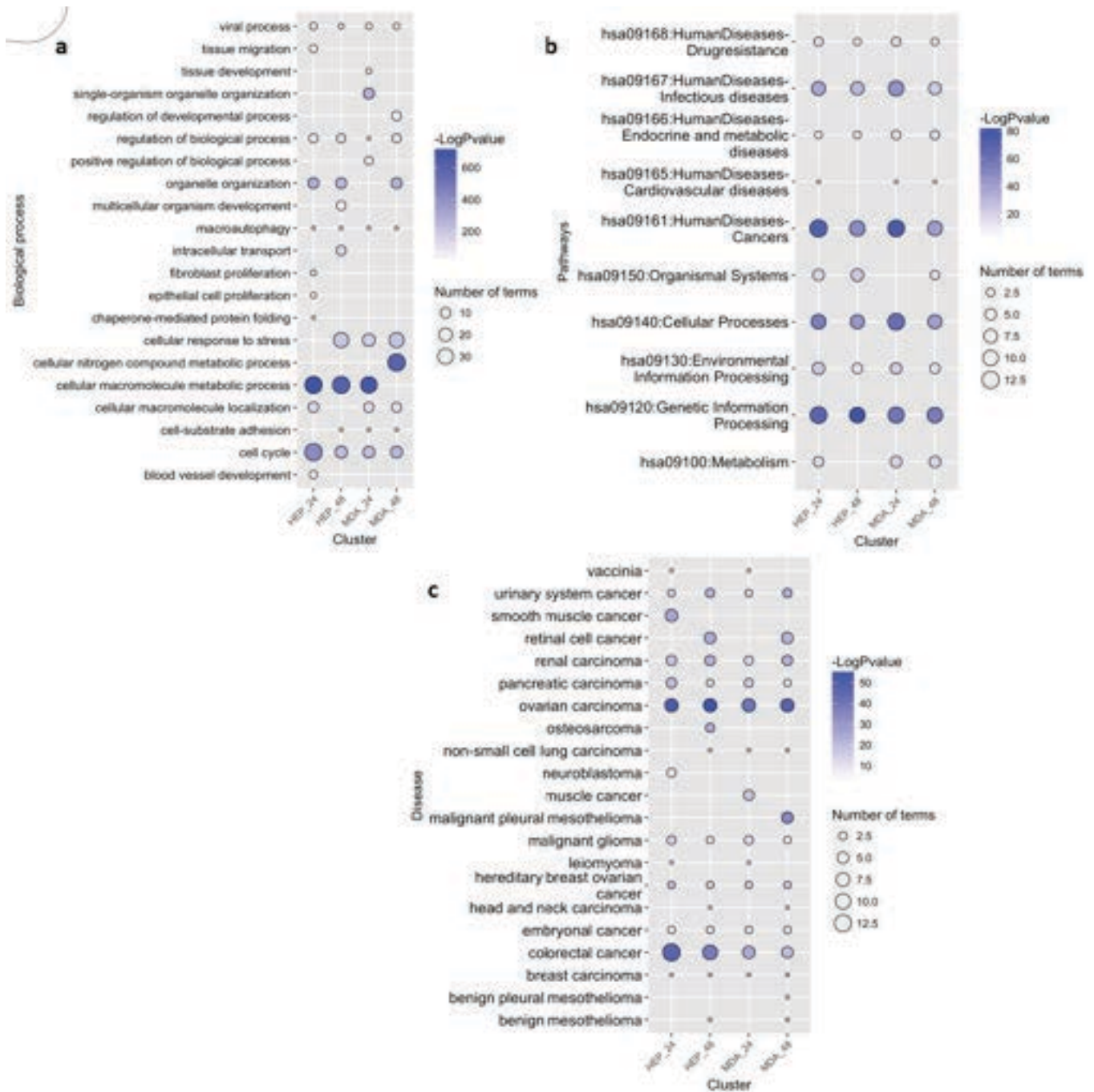
189 file 6: Table S4, sheet 9-12). For both KEGG pathway and DO enrichment, the effects of CKI on  
190 both cell lines were similar.

191 In addition to cell line specific functional enrichment of DE genes, we also analyzed the  
192 over-represented GO terms for shared DE genes (Fig. 5a). The most significant clusters were  
193 highly relevant to metabolic process, such as “cellular macromolecule metabolic process”, as well as  
194 the corresponding positive/negative regulatory biological process (Additional file 6: Table S4, sheet  
195 13). Moreover, various signaling pathways, though not forming a large cluster, were also significant,  
196 for example, “regulation of signal transduction” and “intracellular receptor signaling pathway”.  
197 Finally, some “cell cycle” related terms constituted relatively large sub-clusters, including “cell  
198 division” and “mitotic cell cycle process”. The enriched GO analysis was consistent with the cell  
199 line specific enriched results, and with our previous analysis of MCF-7 cells [28]. It is worth noting  
200 that for “cell cycle” related terms, most of the participating genes were down-regulated by CKI.

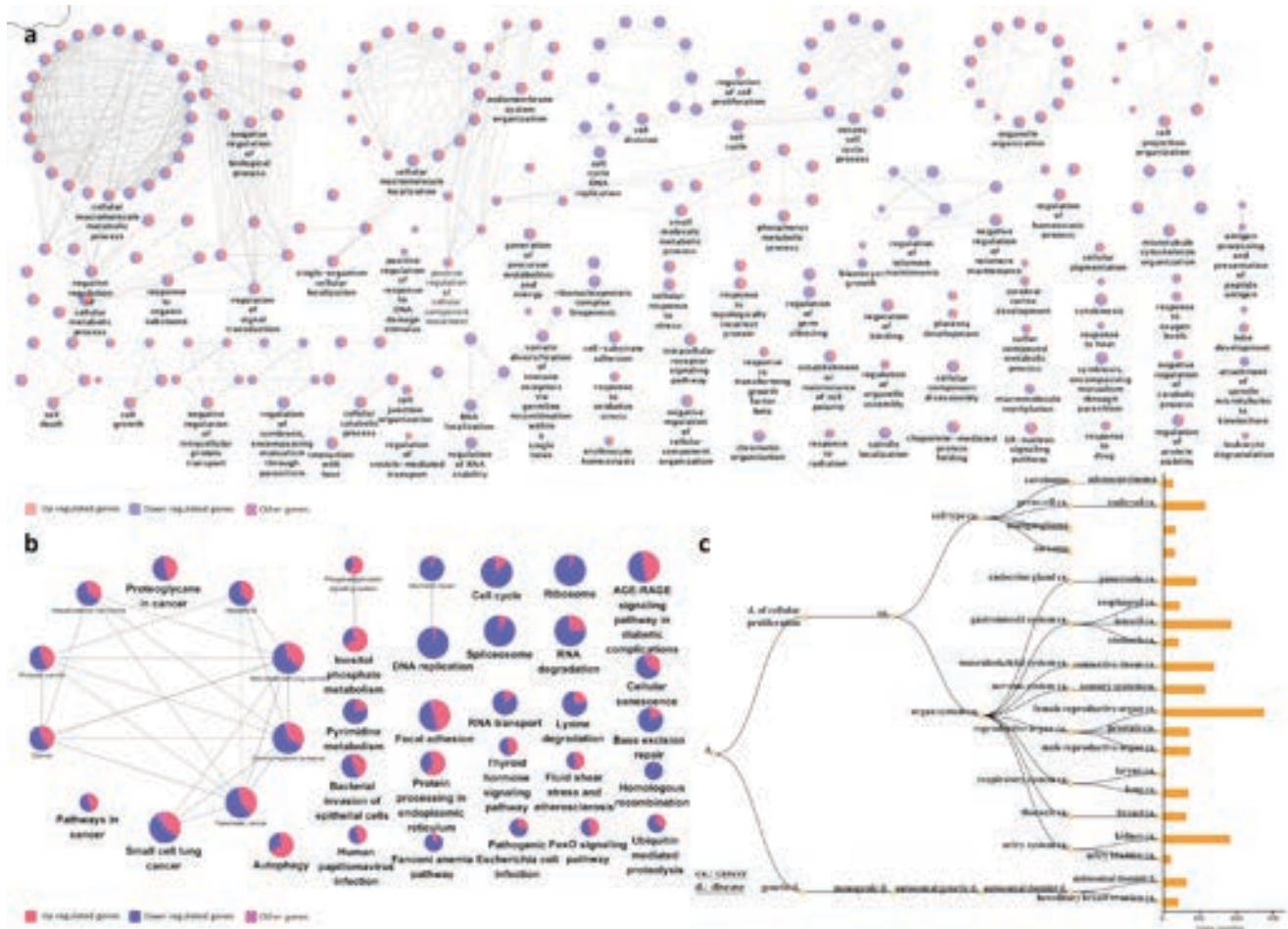
201 Similar results were observed from KEGG analysis (Fig. 5b, and Additional file 6: Table S4,  
202 sheet 14) of shared genes. Various pathways related to cancer, formed a large cluster. Pathways  
203 such as “DNA replication”, “Ribosome” and “cell cycle” were mostly down-regulated, while  
204 up-regulated pathways included “inositol phosphate metabolism” and “protein processing in  
205 endoplasmic reticulum”.

206 We also carried out over-representation analysis of DO terms (Fig. 5c) for all shared DE genes.  
207 The analysis results were consistent with the single cell line DO term analysis with mostly cancer  
208 related terms; in particular genitourinary or breast cancer terms. While this was also partially  
209 similar to the KEGG results for shared DE genes, there were some differences in the KEGG  
210 results for disease pathways compared to the DO results, such as “bacterial invasion of epithelial  
211 cells”, “Fanconi anemia pathway” and “AGE-RAGE pathway in diabetic complications”.

212 Specific to the therapeutic potential of CKI for cancer treatment, we applied our data set  
213 mapping to KEGG cancer pathways: pathways in cancer - homo sapiens (Additional file 7: Fig.  
214 S3). The R package Pathview [21] was used to integrate log fold change values of all the gene  
215 expression patterns into these target pathways. Within the 21 pathways in cancer, the “cell cycle”  
216 still featured prominently (Fig. 6a). The expression of almost every gene in the cell cycle pathway  
217 was affected by CKI, with most of them suppressed. We did not observe this kind of overall



**Figure 4. Functional annotation of DE genes for each cell line as a result of CKI treatment.** Summary of over-represented A) GO terms for Biological Process, B) KEGG pathways and C) DO terms for DE genes as a result of CKI treatment in each cell line at two time points. For GO semantic and enrichment analysis, Lin’s algorithm was applied to cluster and summarize similar functions based on GO terms found in every treatment. Similarly, by back-tracing the upstream categories in the KEGG Ontology, we were able to obtain a more generalized summary of KEGG pathways for each treatment. The size of each bubble represents the number of GO terms/pathways, and the colour shows the statistical significance of the relevant function or pathways. The DO summary for each treatment was determined by back-tracing to parent terms.



**Figure 5. Functional annotation of DE genes with shared expression in both cell lines as a result of CKI treatment.** Over-representation analysis was performed to determine A) GO terms for Biological Process, B) KEGG pathways, and C) DO terms for DE genes shared in both cell lines. In nodes for both GO terms and KEGG pathways, node size is proportional to the statistical significance of over-representation. For DO terms, all the enriched terms are statistically significant ( $p < 1 \times 10^{-5}$ ) in each category, and the bar length represents the number of expressed genes that map to the term.

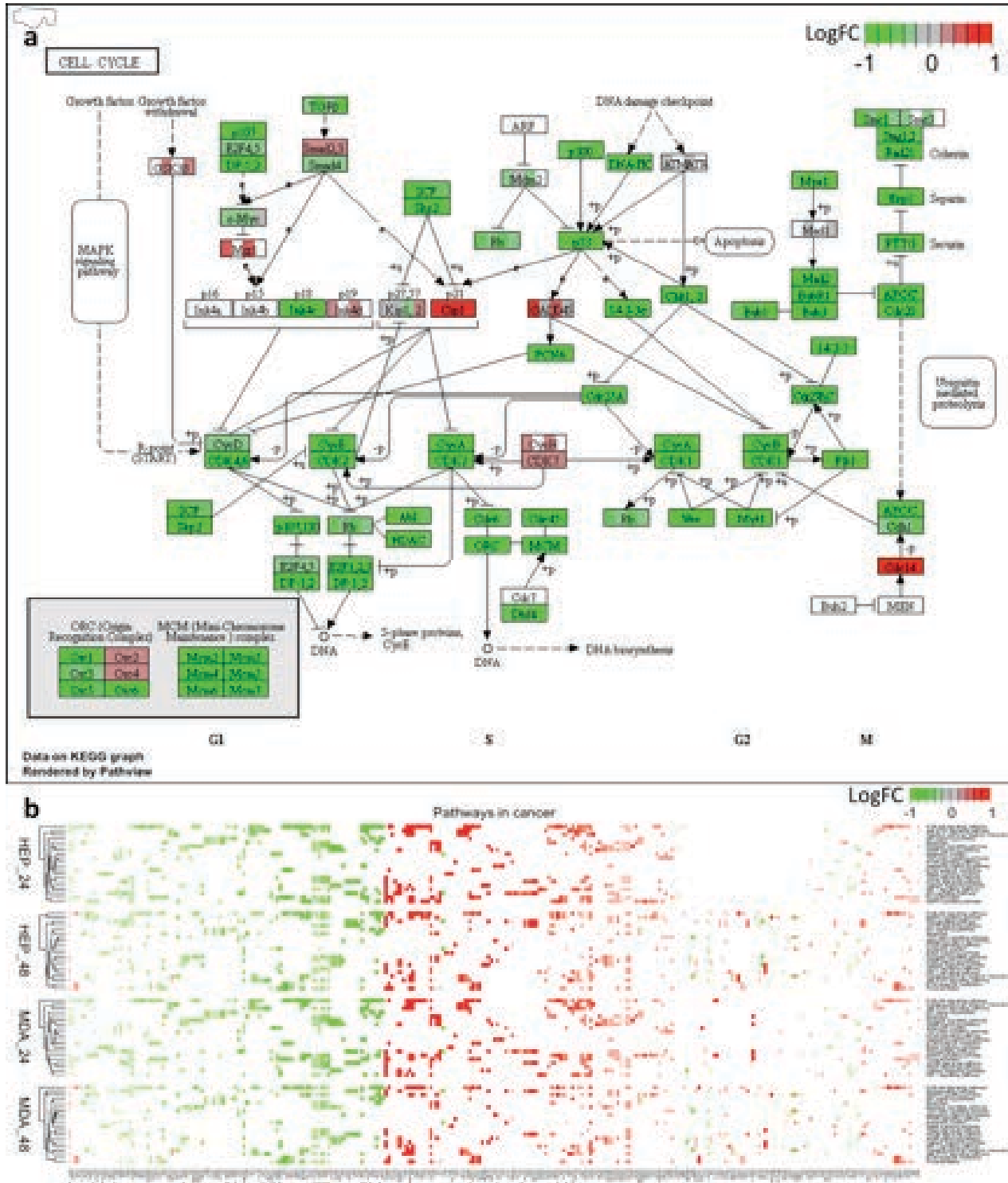
218 pathway suppression in any of the other pathways. We have displayed the summaries for the  
219 remaining 20 pathways in the heatmap in Fig. 6b. Although all the pathways were all perturbed  
220 by CKI, they include both over and under expressed genes in roughly equal proportions.

221 Collectively, these results suggest a direct anti-cancer effect of CKI, and implicate specific  
222 candidate mechanisms of action based on the perturbed molecular networks. The most obvious  
223 example is the cell cycle, where G1-S phase is significantly altered, resulting in the induction of  
224 apoptosis. The downstream process triggered by CKI is the suppression of gene expression of cell  
225 cycle regulators, including TP53 and CCND1. The other perturbed cancer pathways provide  
226 additional candidate mechanisms of action for CKI. In the following section we integrate these  
227 results with previous results reported in the literature to refine the core set of genes and pathways  
228 perturbed by CKI.

## 229 Discussion

230 Although Hep G2 (liver cancer – mesodermal tissue origin) and MDA-MB-231 (mammary  
231 epithelial adenocarcinoma – ectodermal tissue origin) are different cancer types, they shared a  
232 large number of CKI DE genes with similar expression profiles, presumably these shared genes  
233 include CKI response genes that are essential to the apoptotic response triggered by CKI.  
234 However, the number of shared CKI DE genes is too high to allow straight forward identification  
235 of genes critical to the CKI response. We therefore decided to combine these data with previously  
236 reported CKI DE genes from MCF-7 cells [28] in order to reduce the number of core CKI response  
237 genes. The intersection of MCF-7 CKI DE genes with the shared CKI DE genes yielded 363 core  
238 CKI DE genes (Additional file 8: Fig. S4).

239 Among the 363 core CKI DE genes, cytochrome P450 family 1 subfamily A member 1  
240 (CYP1A1) gene is the most over-expressed. This gene is consistently up-regulated by CKI in all  
241 three cell lines, and showed significant dose response. In liver cancer cells, over-expression of  
242 CYP1A1 induced by plant natural products has been associated with Aryl-hydrocarbon Receptor  
243 transformation [2, 49]. Furthermore, as a steroid-metabolizing enzyme, CYP1A1 is part of cancer  
244 metabolic processes relevant to steroid hormone responsive tumors, such as breast cancer, ovarian  
245 cancer and prostate cancer [24, 26, 23, 27]. Therefore, CYP1A1 may be of particular interest for



**Figure 6. Comparison of shared genes expression in specific pathways across two cell lines.** A) Cell cycle pathways, where each coloured box is separated into 4 parts, from left to right representing 24 hour CKI treated Hep G2, 48 hour CKI treated Hep G2, 24 hour CKI treated MDA-MB-231 as well as 48 hour CKI treated MDA-MB-231. B) Heatmap of pathways in cancer. The top two heatmaps summarise the effects of CKI on Hep G2 cells for two time-points, and the bottom two heatmaps show the effects of CKI on MDA-MB-231 cells. In addition to the cell cycle pathway, there were 21 associated pathways in cancer that were perturbed by CKI. The effects of CKI on both cell lines were similar, with changes in TARGET database genes indicated by arrows. Compared to other pathways in cancer, the effects of CKI on the cell cycle pathway showed overall down-regulation.

246 understanding the mechanism of action of CKI on cancer cells.

247 Comparison of the 363 core genes to the 135 Tumor Alterations Relevant for Genomics-driven  
248 Therapy (TARGET) genes (version 3) from The Broad Institute (<https://www.broadinstitute.org/cancer/cga/target>) identified 7 DE genes that were shared across the three cell lines and two time  
249 points (Fig. 7a). Of these seven genes, six (TP53, CCND1, MYD88 (Myeloid differentiation  
250 primary response gene 88) , EWSR1, TMPRSS2 and IDH1 (isocitrate dehydrogenase 1) were  
251 similarly regulated (either always over-expressed or under-expressed), while CCND3 was  
252 over-expressed in all three cell lines at both time points except at 48 hours in MCF-7 cells, where  
253 it was under-expressed.

255 The TP53 gene encodes a tumor suppressor protein, that can induce apoptosis [11]. However,  
256 in all cell lines TP53 was down-regulated, and all cell lines showed increased apoptosis. This  
257 suggests that CKI induced apoptosis was not TP53-dependent. Support for this comes from the  
258 fact that transcripts for PCNA (proliferating cell nuclear antigen), and a group of transcription  
259 factors: MCM (mini-chromosome maintenance) complex and the E2F family are down-regulated.  
260 The E2F transcription factors regulate the cell cycle and TP53-dependent and -independent  
261 apoptosis [37, 12, 42, 33]. In addition, other core genes present in the TARGET database have  
262 also been shown to induce apoptosis. For example, inhibition of MYD88 induces apoptosis in both  
263 triple negative breast cancer and bladder cancer [3, 43]. The increased expression of IDH1 may be  
264 important, as IDH1 is frequently mutated in cancers [16] and when mutated, it causes loss of  
265  $\alpha$ -ketoglutarate production and may be important for the Warburg effect. TMPRSS2  
266 (transmembrane protease, serine 2) has also been shown to regulate apoptosis in cancer [1].  
267 Therefore, CKI may induce apoptosis through a variety of means.

268 In the GO (Fig. 7b) and KEGG (Fig. 7c) over-representation analysis of the 363 core genes  
269 yielded enrichment for cell cycle and cancer pathways. In the GO enriched genes, cell cycle and  
270 related pathways accounted for the majority of functional sub-clusters. In the KEGG enriched  
271 pathways, cell cycle and cancer pathways predominated in a single cluster. Most of the core genes  
272 in GO and KEGG clusters were down-regulated by CKI. In addition to the cell cycle, CKI  
273 treatment also caused enrichment for terms or pathways related to cancer progression, such as  
274 “focal adhesion” and “blood vessel development”. (Additional file 6: Table S4, sheet 5-8). These

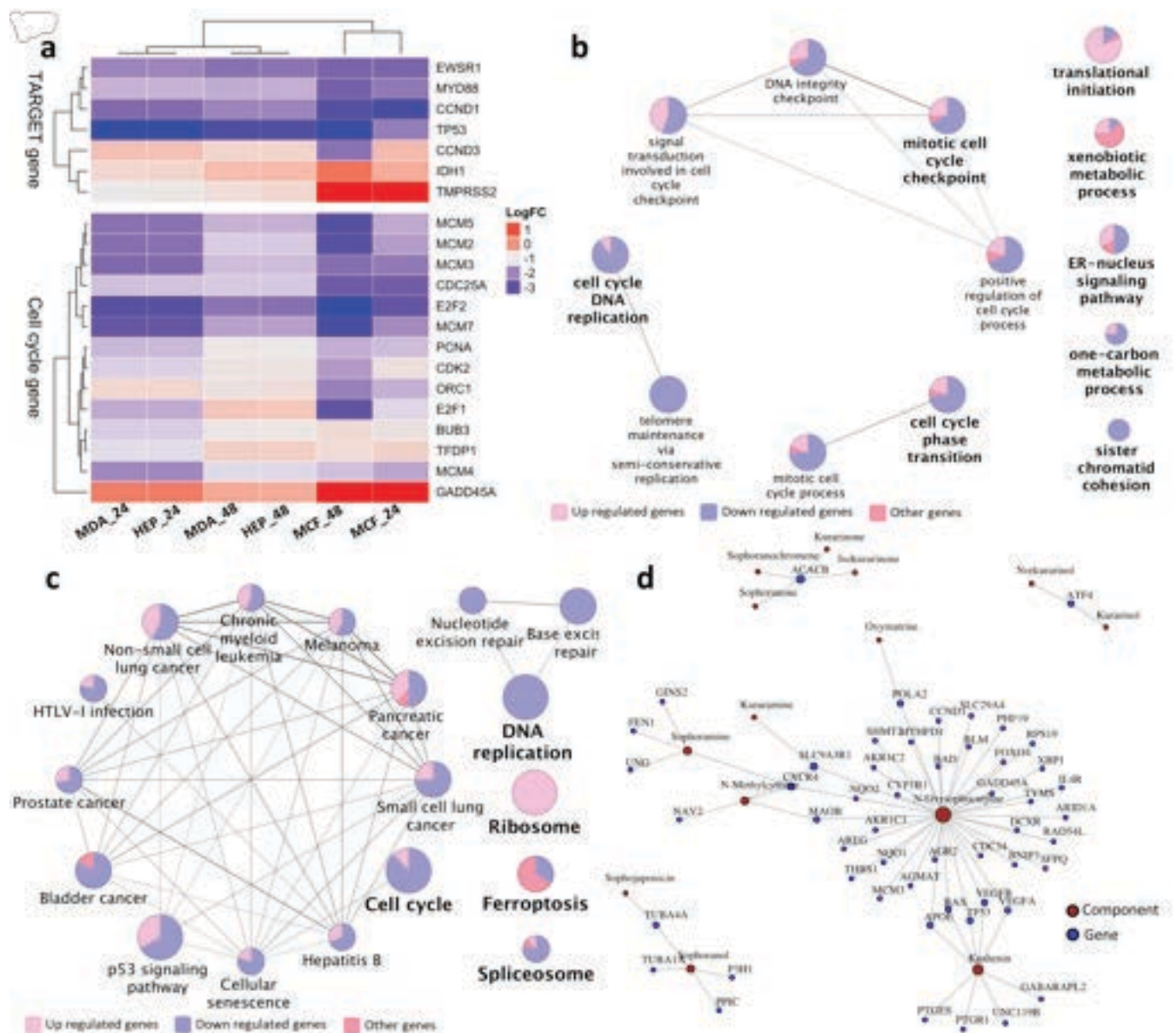
275 developmental processes contribute to tumorigenesis and metastasis [20, 36]. It is tempting to  
276 speculate that CKI may alter these functions *in vivo*, possibly altering angiogenesis which is  
277 critical for tumors [9]. In addition, there were metabolic pathways and terms that were also  
278 identified as perturbed by CKI. Effects on many targets/pathways is one of the expected features  
279 of TCM drugs which likely hit multiple targets [7].

280 We have examined the effect of a complex mixture of plant natural products (CKI) on different  
281 cancer cell lines and have identified specific, consistent effects on gene expression resulting from  
282 this mixture. However, the complexity of CKI makes it difficult to determine the mechanism of  
283 action of individual components, and often testing of individual components has resulted in either  
284 no effect or contradictory results in the research literature. In spite of this complexity, it is possible  
285 to map our results on to a pre-existing corpus of work that links individual natural compounds to  
286 changes in gene expression. We have used BATMAN [19], an online TCM database of curated  
287 links between compounds and gene expression. Based on this resource, we have identified 14  
288 components of CKI that have been linked to the regulation of 52 of our core genes (Fig. 7d). We  
289 can see from the network diagram in Fig. 7d that one to one, one to many and many to many  
290 relationships exist between CKI components and genes which is consistent with previous studies  
291 [38, 18, 46]. As more information becomes available for individual components, we will be able to  
292 construct a more comprehensive model of CKI mechanism based on network analysis.

## 293 Conclusion

294 Our systematic analysis of gene expression changes in cancer cells caused by a complex herbal  
295 extract used in TCM has proven to be effective at identifying candidate molecular pathways. CKI  
296 has consistent and specific effects on gene expression across multiple cancer cell lines and it also  
297 consistently induces apoptosis *in vitro*. These effects show that CKI can suppress the expression of  
298 cell cycle regulatory genes and other well characterized cancer related genes and pathways.  
299 Validation of a subset of DE genes at lower doses of CKI has shown a dose-response relationship  
300 that suggests that CKI may have similar effects *in vivo* at clinically relevant concentrations. Our  
301 results provide a molecular basis for further investigation of the mechanism of action of CKI.





**Figure 7. Analysis of CKI regulated core genes from this report combined with previous available data.** A) Fold changes of TARGET and cell cycle regulatory gene expression in MDA-MB-231, Hep G2 and MCF-7 [28] cell lines 24 and 48 hours after CKI treatment. Only seven TARGET genes are affected by CKI in all three cell lines. Most of the 14 cell cycle regulatory genes differentially expressed in all three cell lines are down-regulated. B) GO term enrichment analysis of 363 core genes from MDA-MB-231, Hep G2 and MCF-7 cell lines. C) KEGG pathway enrichment of 363 core genes from MDA-MB-231, Hep G2 and MCF-7 cell lines. D) Some individual compounds present in CKI linked to genes they regulate that are also found in this report and our previous study [28]. Node size is proportional to the number of related components/genes.

302 **Authors' e-mail addresses**

Jian Cui [jian.cui@adelaide.edu.au](mailto:jian.cui@adelaide.edu.au)  
Zhipeng Qu [zhipeng.qu@adelaide.edu.au](mailto:zhipeng.qu@adelaide.edu.au)  
Yuka Harata-Lee [yuka.harata-lee@adelaide.edu.au](mailto:yuka.harata-lee@adelaide.edu.au)  
Hanyuan Shen [hanyuan.shen@adelaide.edu.au](mailto:hanyuan.shen@adelaide.edu.au)  
303 Thazin Nwe Aung [thazin.nweaung@adelaide.edu.au](mailto:thazin.nweaung@adelaide.edu.au)  
Wei Wang [wangw@zhendongpharm.com](mailto:wangw@zhendongpharm.com)  
R.Daniel Kortschak [dan.kortschak@adelaide.edu.au](mailto:dan.kortschak@adelaide.edu.au)  
David L Adelson [david.adelson@adelaide.edu.au](mailto:david.adelson@adelaide.edu.au)

304 **Conflict of Interest**

305 We wish to draw the attention of the Editor to the following facts which may be considered as  
306 potential conflicts of interest and to significant financial contributions to this work. While a  
307 generous donation was used to set up the Zhendong Centre by Shanxi Zhendong Pharmaceutical  
308 Co Ltd, they did not determine the research direction for this work or influence the analysis of the  
309 data. JC: no competing interests, ZQ: no competing interests, YHL: no competing interests, HS:  
310 no competing interests, TNA: no competing interests, WW: is an employee of Zhendong Pharma  
311 seconded to Zhendong Centre to learn bioinformatics methods, RDK: no competing interests, DLA:  
312 Director of the Zhendong Centre which was set up with a generous donation from the Zhendong  
313 Pharmaceutical Co Ltd. Zhendong Pharmaceutical has had no control over these experiments,  
314 their design or analysis and have not exercised any editorial control over the manuscript.

315 **Author contributions statement**

316 JC experimental design, carried out experiments, analyzed data, wrote paper, ZQ experimental  
317 design, assisted with experiments, assisted with data analysis, wrote paper, YHL experimental  
318 design, assisted with experiments, assisted with data analysis, wrote paper, HS assisted with  
319 experiments, TNA assisted with experiments, WW assisted with experimental design, assisted  
320 with experiments, RDK experimental design, and DLA supervised the research, acquired funding  
321 for the experiments, experimental design, wrote paper.

322 **Data Availability**

323 All RNA-seq data raw and processed data were deposited at the Gene Expression Omnibus  
324 (GEO) data repository (XXXXXXX).

325 **Role of the funding source**

326 Funding for this study was provided by the Zhendong Centre for Molecular Chinese Medicine.

327 **References**

328 **References**

- 329 [1] Afar, D. E., Vivanco, I., Hubert, R. S., Kuo, J., Chen, E., Saffran, D. C., Raitano, A. B.,  
330 Jakobovits, A., 2001. Catalytic cleavage of the androgen-regulated tmprss2 protease results in  
331 its secretion by prostate and prostate cancer epithelia. *Cancer research* 61 (4), 1686–1692.
- 332 [2] Anwar-Mohamed, A., El-Kadi, A. O., 2009. Sulforaphane induces cyp1a1 mrna, protein, and  
333 catalytic activity levels via an ahr-dependent pathway in murine hepatoma hepa 1c1c7 and  
334 human hepg2 cells. *Cancer letters* 275 (1), 93–101.
- 335 [3] Christensen, A. G., Ehmsen, S., Terp, M. G., Batra, R., Alcaraz, N., Baumbach, J., Noer,  
336 J. B., Moreira, J., Leth-Larsen, R., Larsen, M. R., 2017. Elucidation of altered pathways in  
337 tumor-initiating cells of triple-negative breast cancer: A useful cell model system for drug  
338 screening. *STEM CELLS*.
- 339 [4] Csardi, G., Nepusz, T., 2006. The igraph software package for complex network research.  
340 *InterJournal, Complex Systems* 1695 (5), 1–9.
- 341 [5] Cui, M., Li, H., Hu, X., 2014. Similarities between “big data” and traditional chinese  
342 medicine information. *Journal of Traditional Chinese Medicine* 34 (4), 518–522.
- 343 [6] Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson,  
344 M., Gingeras, T. R., 2013. Star: ultrafast universal rna-seq aligner. *Bioinformatics* 29 (1),  
345 15–21.
- 346 [7] Efferth, T., Li, P. C., Konkimalla, V. S. B., Kaina, B., 2007. From traditional chinese  
347 medicine to rational cancer therapy. *Trends in molecular medicine* 13 (8), 353–361.
- 348 [8] Gao, L., Wang, K.-x., Zhou, Y.-z., Fang, J.-s., Qin, X.-m., Du, G.-h., 2018. Uncovering the  
349 anticancer mechanism of compound kushen injection against hcc by integrating quantitative  
350 analysis, network analysis and experimental validation. *Scientific reports* 8 (1), 624.
- 351 [9] Hanahan, D., Folkman, J., 1996. Patterns and emerging mechanisms of the angiogenic switch  
352 during tumorigenesis. *cell* 86 (3), 353–364.

- 353 [10] Hao, D. C., Xiao, P. G., 2014. Network pharmacology: a rosetta stone for traditional chinese  
354 medicine. *Drug development research* 75 (5), 299–312.
- 355 [11] Harris, S. L., Levine, A. J., 2005. The p53 pathway: positive and negative feedback loops.  
356 *Oncogene* 24 (17), 2899.
- 357 [12] Hollern, D. P., Honeysett, J., Cardiff, R. D., Andrechek, E. R., 2014. The e2f transcription  
358 factors regulate tumor development and metastasis in a mouse model of metastatic breast  
359 cancer. *Molecular and cellular biology* 34 (17), 3229–3243.
- 360 [13] Jiang, W.-Y., 2005. Therapeutic wisdom in traditional chinese medicine: a perspective from  
361 modern science. *Trends in pharmacological sciences* 26 (11), 558–563.
- 362 [14] Kamburov, A., Wierling, C., Lehrach, H., Herwig, R., 2008. Consensuspathdb—a database for  
363 integrating human functional interaction networks. *Nucleic acids research* 37 (suppl\_1),  
364 D623–D628.
- 365 [15] Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., Salzberg, S. L., 2013. Tophat2:  
366 accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions.  
367 *Genome Biol* 14 (4), R36.
- 368 [16] Li, X., Xu, X., Wang, J., Yu, H., Wang, X., Yang, H., Xu, H., Tang, S., Li, Y., Yang, L., 2012.  
369 A system-level investigation into the mechanisms of chinese traditional medicine: Compound  
370 danshen formula for cardiovascular disease treatment. *PloS one* 7 (9), e43918.
- 371 [17] Liao, C.-Y., Lee, C.-C., Tsai, C.-c., Hsueh, C.-W., Wang, C.-C., Chen, I., Tsai, M.-K., Liu,  
372 M.-Y., Hsieh, A.-T., Su, K.-J., 2015. Novel investigations of flavonoids as chemopreventive  
373 agents for hepatocellular carcinoma. *BioMed research international* 2015.
- 374 [18] Liu, Y., Xu, Y., Ji, W., Li, X., Sun, B., Gao, Q., Su, C., 2014. Anti-tumor activities of  
375 matrine and oxymatrine: literature review. *Tumor Biology* 35 (6), 5111–5119.
- 376 [19] Liu, Z., Guo, F., Wang, Y., Li, C., Zhang, X., Li, H., Diao, L., Gu, J., Wang, W., Li, D., 2016.  
377 Batman-tcm: a bioinformatics analysis tool for molecular mechanism of traditional chinese  
378 medicine. *Scientific reports* 6.

- 379 [20] Luo, M., Guan, J.-L., 2010. Focal adhesion kinase: a prominent determinant in breast cancer  
380 initiation, progression and metastasis. *Cancer letters* 289 (2), 127–139.
- 381 [21] Luo, W., Brouwer, C., 2013. Pathview: an r/bioconductor package for pathway-based data  
382 integration and visualization. *Bioinformatics* 29 (14), 1830–1831.
- 383 [22] Ma, X., Li, R.-S., Wang, J., Huang, Y.-Q., Li, P.-Y., Wang, J., Su, H.-B., Wang, R.-L.,  
384 Zhang, Y.-M., Liu, H.-H., 2016. The therapeutic efficacy and safety of compound kushen  
385 injection combined with transarterial chemoembolization in unresectable hepatocellular  
386 carcinoma: an update systematic review and meta-analysis. *Frontiers in pharmacology* 7.
- 387 [23] Mitsui, Y., Chang, I., Kato, T., Hashimoto, Y., Yamamura, S., Fukuhara, S., Wong, D. K.,  
388 Shiina, M., Imai-Sumida, M., Majid, S., 2016. Functional role and tobacco smoking effects on  
389 methylation of *cyp1a1* gene in prostate cancer. *Oncotarget* 7 (31), 49107.
- 390 [24] Nandekar, P. P., Khomane, K., Chaudhary, V., Rathod, V. P., Borkar, R. M., Bhandi, M. M.,  
391 Srinivas, R., Sangamwar, A. T., Guchhait, S. K., Bansal, A. K., 2016. Identification of leads  
392 for antiproliferative activity on mda-mb-435 human breast cancer cells through  
393 pharmacophore and *cyp1a1*-mediated metabolism. *European journal of medicinal chemistry*  
394 115, 82–93.
- 395 [25] Ninomiya, M., Koketsu, M., 2013. *Minor flavonoids (chalcones, flavanones, dihydrochalcones,  
396 and aurones)*. Springer, pp. 1867–1900.
- 397 [26] Ou, C., Zhao, Y., Liu, J.-H., Zhu, B., Li, P.-Z., Zhao, H.-L., 2016. Relationship between  
398 aldosterone synthase *cyp1a1* mspi gene polymorphism and prostate cancer risk. *Technology in  
399 cancer research & treatment*, 1533034615625519.
- 400 [27] Piotrowska-Kempisty, H., Klupczyńska, A., Trzybulska, D., Kulcenty, K., Sulej-Suchomska,  
401 A. M., Kucińska, M., Mikstacka, R., Wierzchowski, M., Murias, M., Baer-Dubowska, W.,  
402 2017. Role of *cyp1a1* in the biological activity of methylated resveratrol analogue, 3, 4, 5,  
403 4'-tetramethoxystilbene (dmu-212) in ovarian cancer a-2780 and non-cancerous hose cells.  
404 *Toxicology letters* 267, 59–66.

- 405 [28] Qu, Z., Cui, J., Harata-Lee, Y., Aung, T. N., Feng, Q., Raison, J. M., Kortschak, R. D.,  
406 Adelson, D. L., 2016. Identification of candidate anti-cancer molecular mechanisms of  
407 compound kushen injection using functional genomics. *Oncotarget* 7 (40), 66003–66019.
- 408 [29] Robinson, M. D., McCarthy, D. J., Smyth, G. K., 2010. edgeR: a bioconductor package for  
409 differential expression analysis of digital gene expression data. *Bioinformatics* 26 (1), 139–140.
- 410 [30] Schriml, L. M., Arze, C., Nadendla, S., Chang, Y.-W. W., Mazaitis, M., Felix, V., Feng, G.,  
411 Kibbe, W. A., 2011. Disease ontology: a backbone for disease semantic integration. *Nucleic  
412 acids research* 40 (D1), D940–D946.
- 413 [31] Shu, G., Yang, J., Zhao, W., Xu, C., Hong, Z., Mei, Z., Yang, X., 2014. Kurarinol induces  
414 hepatocellular carcinoma cell apoptosis through suppressing cellular signal transducer and  
415 activator of transcription 3 signaling. *Toxicology and applied pharmacology* 281 (2), 157–165.
- 416 [32] Song, Y.-N., Zhang, G.-B., Zhang, Y.-Y., Su, S.-B., 2013. Clinical applications of omics  
417 technologies on zheng differentiation research in traditional chinese medicine. *Evidence-Based  
418 Complementary and Alternative Medicine* 2013.
- 419 [33] Sun, H., Xu, Y., Yang, X., Wang, W., Bai, H., Shi, R., Nayar, S. K., Devbhandari, R. P., He,  
420 Y., Zhu, Q., 2013. Hypoxia inducible factor 2 alpha inhibits hepatocellular carcinoma growth  
421 through the transcription factor dimerization partner 3/e2f transcription factor 1-dependent  
422 apoptotic pathway. *Hepatology* 57 (3), 1088–1097.
- 423 [34] Wang, P., Chen, Z., 2013. Traditional chinese medicine zheng and omics convergence: A  
424 systems approach to post-genomics medicine in a global world. *Omics: a journal of integrative  
425 biology* 17 (9), 451–459.
- 426 [35] Wang, W., You, R.-l., Qin, W.-j., Hai, L.-n., Fang, M.-j., Huang, G.-h., Kang, R.-x., Li, M.-h.,  
427 Qiao, Y.-f., Li, J.-w., 2015. Anti-tumor activities of active ingredients in compound kushen  
428 injection. *Acta Pharmacologica Sinica* 36 (6), 676–679.
- 429 [36] Warren, R. S., Yuan, H., Matli, M. R., Gillett, N. A., Ferrara, N., 1995. Regulation by  
430 vascular endothelial growth factor of human colon cancer tumorigenesis in a mouse model of  
431 experimental liver metastasis. *The Journal of clinical investigation* 95 (4), 1789–1797.

- 432 [37] Woods, K., Thomson, J. M., Hammond, S. M., 2007. Direct regulation of an oncogenic  
433 micro-rna cluster by e2f transcription factors. *Journal of Biological Chemistry* 282 (4),  
434 2130–2134.
- 435 [38] Wu, C., Huang, W., Guo, Y., Xia, P., Sun, X., Pan, X., Hu, W., 2015. Oxymatrine inhibits  
436 the proliferation of prostate cancer cells in vitro and in vivo. *Molecular medicine reports*  
437 11 (6), 4129–4134.
- 438 [39] Yang, X., Huang, M., Cai, J., Lv, D., Lv, J., Zheng, S., Ma, X., Zhao, P., Wang, Q., 2017.  
439 Chemical profiling of anti-hepatocellular carcinoma constituents from *caragana tangutica*  
440 *maxim.* by off-line semi-preparative hplc-nmr. *Natural product research* 31 (10), 1150–1155.
- 441 [40] Yu, G., Wang, L.-G., Han, Y., He, Q.-Y., 2012. clusterprofiler: an r package for comparing  
442 biological themes among gene clusters. *OMICS : a Journal of Integrative Biology* 16 (5),  
443 284–287.  
444 URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3339379/>
- 445 [41] Yu, P., Liu, Q., Liu, K., Yagasaki, K., Wu, E., Zhang, G., 2009. Matrine suppresses breast  
446 cancer cell proliferation and invasion via vegf-akt-nf- $\kappa$ b signaling. *Cytotechnology* 59 (3),  
447 219–229.
- 448 [42] Zaldua, N., Llaverro, F., Artaso, A., Gálvez, P., Lacerda, H. M., Parada, L. A., Zugaza, J. L.,  
449 2016. Rac1/p21-activated kinase pathway controls retinoblastoma protein phosphorylation  
450 and e2f transcription factor activation in b lymphocytes. *FEBS journal* 283 (4), 647–661.
- 451 [43] Zhang, H., Ye, Y., Li, M., Ye, S., Huang, W., Cai, T., He, J., Peng, J., Duan, T., Cui, J.,  
452 2016. Cxcl2/mif-cxcr2 signaling promotes the recruitment of myeloid-derived suppressor cells  
453 and is correlated with prognosis in bladder cancer. *Oncogene*.
- 454 [44] Zhang, J.-Q., Li, Y.-M., Liu, T., He, W.-T., Chen, Y.-T., Chen, X.-H., Li, X., Zhou, W.-C.,  
455 Yi, J.-F., Ren, Z.-J., 2010. Antitumor effect of matrine in human hepatoma g2 cells by  
456 inducing apoptosis and autophagy. *World journal of gastroenterology: WJG* 16 (34), 4281.
- 457 [45] Zhang, L. P., Jiang, J. K., Tam, J. W. O., Zhang, Y., Liu, X. S., Xu, X. R., Liu, B. Z., He,  
458 Y. J., 2001. Effects of matrine on proliferation and differentiation in k-562 cells. *Leukemia*



- 459 Research 25 (9), 793–800.
- 460 URL <http://www.sciencedirect.com/science/article/pii/S0145212600001454>
- 461 [46] Zhang, X., Yu, H., 2016. Matrine inhibits diethylnitrosamine-induced hcc proliferation in rats  
462 through inducing apoptosis via p53, bax-dependent caspase-3 activation pathway and  
463 down-regulating mlck overexpression. Iranian Journal of Pharmaceutical Research: IJPR  
464 15 (2), 491.
- 465 [47] Zhang, X.-L., Cao, M.-A., Pu, L.-P., Huang, S.-S., Gao, Q.-X., Yuan, C.-S., Wang, C.-M.,  
466 2013. A novel flavonoid isolated from sophora flavescens exhibited anti-angiogenesis activity,  
467 decreased vegf expression and caused g0/g1 cell cycle arrest in vitro. Die Pharmazie-An  
468 International Journal of Pharmaceutical Sciences 68 (5), 369–375.
- 469 [48] Zhao, Z., Fan, H., Higgins, T., Qi, J., Haines, D., Trivett, A., Oppenheim, J. J., Wei, H., Li,  
470 J., Lin, H., 2014. Fufang kushen injection inhibits sarcoma growth and tumor-induced  
471 hyperalgesia via trpv1 signaling pathways. Cancer letters 355 (2), 232–241.
- 472 [49] Zhou, Y., Li, Y., Zhou, T., Zheng, J., Li, S., Li, H.-B., 2016. Dietary natural products for  
473 prevention and treatment of liver cancer. Nutrients 8 (3), 156.

## Chapter 5

### Cell Cycle, Glycolysis and DNA Repair Pathways in Cancer Cells are Suppressed by Compound Kushen Injection

In the previous chapter, we have identified the pathways that CKI shows its anticancer effects. Here, cell cycle, energy metabolism and DNA repair pathways were selected for verification on MDA-MB-231 and HEPG2 cells. We have confirmed the protein level with transcriptome data for these pathways and found CKI can reduce energy metabolism and increase DNA double-strand breaks (DSB). Besides, we have confirmed that oxymatrine, usually considered as the main active component in CKI, has different effects to CKI in these pathways. This chapter is in the format of the manuscript that was published in *BMC Cancer* (<https://doi.org/10.1186/s12885-018-5230-8>).

# Statement of Authorship

Title of Paper	Cell cycle, energy metabolism and DNA repair pathways in cancer cells are suppressed by Compound Kushen Injection
Publication Status	<input checked="" type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Jian Cui, Zhipeng Qu, Yuka Harata-Lee, Thazin Nwe Aung, Hanyuan Shen, Wei Wang, David L. Adelson. "Cell cycle, energy metabolism and DNA repair pathways in cancer cells are suppressed by Compound Kushen Injection". BMC Cancer. 2019; 19:103. <a href="https://doi.org/10.1186/s12885-018-5230-8">https://doi.org/10.1186/s12885-018-5230-8</a>

## Principal Author

Name of Principal Author (Candidate)	Hanyuan Shen		
Contribution to the Paper	Assited with experiments		
Overall percentage (%)	20%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the fifth author of this paper.		
Signature		Date	13/3/2019

## Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Jian Cui		
Contribution to the Paper	Experimental design, carried out experiments, analysed data, wrote paper		
Signature		Date	19/03/2019

Name of Co-Author	Zhipeng Qu		
Contribution to the Paper	Experimental design, assisted with experiments, assisted with data analysis, wrote paper		
Signature		Date	18/03/2019

Name of Co-Author	Yuka Harata-Lee		
Contribution to the Paper	Experimental design, assisted with experiments, assisted with data analysis, wrote paper		
Signature		Date	13/3/19

Name of Co-Author	Thazin Nwe Aung		
Contribution to the Paper	Assisted with experiments		
Signature		Date	13/03/2019

Name of Co-Author	Wei Wang		
Contribution to the Paper	Assisted with experimental design, assisted with experiments		
Signature		Date	19/03/2019


Name of Co-Author	David L. Adelson		
Contribution to the Paper	Supervised the research, acquired funding for the experiments, experimental design, wrote paper		
Signature		Date	1/4/2019

RESEARCH ARTICLE

Open Access



# Cell cycle, energy metabolism and DNA repair pathways in cancer cells are suppressed by Compound Kushen Injection

Jian Cui<sup>1,2</sup>, Zhipeng Qu<sup>1,2</sup>, Yuka Harata-Lee<sup>1,2</sup>, Thazin Nwe Aung<sup>1,2</sup>, Hanyuan Shen<sup>1,2</sup>, Wei Wang<sup>2,3</sup> and David L. Adelson<sup>1,2\*</sup> 

## Abstract

**Background:** In this report we examine candidate pathways perturbed by Compound Kushen Injection (CKI), a Traditional Chinese Medicine (TCM) that we have previously shown to alter the gene expression patterns of multiple pathways and induce apoptosis in cancer cells.

**Methods:** We have measured protein levels in Hep G2 and MDA-MB-231 cells for genes in the cell cycle pathway, DNA repair pathway and DNA double strand breaks (DSBs) previously shown to have altered expression by CKI. We have also examined energy metabolism by measuring [ADP]/[ATP] ratio (cell energy charge), lactate production and glucose consumption. Our results demonstrate that CKI can suppress protein levels for cell cycle regulatory proteins and DNA repair while increasing the level of DSBs. We also show that energy metabolism is reduced based on reduced glucose consumption and reduced cellular energy charge.

**Results:** Our results validate these pathways as important targets for CKI. We also examined the effect of the major alkaloid component of CKI, oxymatrine and determined that it had no effect on DSBs, a small effect on the cell cycle and increased the cell energy charge.

**Conclusions:** Our results indicate that CKI likely acts through the effect of multiple compounds on multiple targets where the observed phenotype is the integration of these effects and synergistic interactions.

**Keywords:** Alkaloid, Matrine, Cyclin, Ku70, Ku80, Cell-cycle

## Background

Compound Kushen Injection (CKI) is a complex mixture of plant bioactives extracted from Kushen (*Sophora flavescens*) and Baituling (*Smilax Glabra*) that has been approved for use in China since 1995 by the State Food and Drug Administration of China (State medical license no. Z14021231).

Kushen has a long history in Chinese Medicine and is a very commonly used plant in the Chinese Materia Medica (CMM). This leguminous plant is widely distributed in Russia, Japan, India, North Korea, and some provinces

and regions in mainland China [1]. The medicinal part of Kushen is its dry root that is used mainly to treat inflammation, eczema, parasites and similar afflictions [2, 3]. Modern pharmacological research suggests that the various alkaloids and flavonoids contained in Kushen have anticancer activity, especially with respect to arresting tumor growth and relieving cancer pain [4]. Compared to Kushen, Baituling is distributed in some regions of southern China, and was only used clinically in limited applications [5]. Because of this limited clinical usage, there are only a small number of research reports focused on Baituling.

The combination of above two herbs' extracts, CKI is widely used in China as an adjunct for both radiotherapy and chemotherapy in cancer. While most of the data supporting its use have been anecdotal and there is little clinical trial data demonstrating its efficacy, it has been

\*Correspondence: david.adelson@adelaide.edu.au

<sup>1</sup>Department of Molecular and Biomedical Science, The University of Adelaide, 5005 North Terrace, Adelaide, South Australia, Australia

<sup>2</sup>Zhendong Australia - China Centre for Molecular Chinese Medicine, The University of Adelaide, 5005 North Terrace, Adelaide, South Australia, Australia  
Full list of author information is available at the end of the article



shown to be effective at reducing sarcoma growth and cancer pain in an animal model [6] and cancer pain in patients [4].

CKI contains over 200 chemical compounds including alkaloids and flavonoids such as matrine, oxymatrine and kurarinol, and has previously been shown to affect the cell cycle and induce apoptosis in cancer cells [2, 4, 6–10]. Furthermore, functional genomic characterisation of the effect of CKI on cancer cells using transcriptome data indicated that multiple pathways were most likely affected by CKI [8]. These observations support a model wherein many/all of the individual compounds present in CKI can act on many single targets or on multiple targets to induce apoptosis.

Based on previously reported work [8] and our currently unpublished work (Cui et al) [11], specific pathways were selected for follow up experiments to validate their response to CKI in order to formulate more specific hypotheses regarding the mechanism of action of CKI on cancer cells. We had previously shown that CKI altered the cell cycle and induced apoptosis while altering the expression of many cell cycle genes in three cancer cell lines [8, 11]. We had also shown that DNA repair pathway genes were significantly down-regulated by CKI and that energy production related to NAD(P)H synthesis from glycolysis and oxidative phosphorylation was reduced by CKI. As a result we focused on the following candidate pathways: cell cycle, DNA repair and glucose metabolism to validate their alteration by CKI. We used two cell lines for these validation experiments, one relatively insensitive to CKI (MDA-MB-231) and one sensitive to CKI (Hep G2). Furthermore, while the literature shows varying effects for major compounds present in CKI on cancer cells [12, 13], we also tested oxymatrine, the major alkaloid found in CKI and widely believed to be very important for the effects of CKI, on our selected pathways.

## Methods

### Cell culture and chemicals

CKI with a total alkaloid concentration of 26.5 mg/ml in 5 ml ampoules was provided by Zhendong Pharmaceutical Co. Ltd. (Beijing, China). Cell culture methods have been previously described [8]. The concentration of total alkaloids in CKI was determined using HPLC.

A human breast adenocarcinoma cell line, MDA-MB-231 and a hepatocellular carcinoma cell line Hep G2 were purchased from American Type Culture Collection (VA, USA). The cells were cultured in Dulbecco's Modified Eagle Medium (Thermo Fisher Scientific, MA, USA) supplemented with 10% fetal bovine serum (Thermo Fisher Scientific). Both cell lines were cultured at 37 °C with 5% CO<sub>2</sub>. Cells were split twice per week with trypsinization, defined as two passages. Both cell lines were discarded when passage number was more than 15.

For all in vitro assays, cells were cultured overnight before being treated with CKI (either at 1 mg/ml and 2 mg/ml of total alkaloids). As a negative control, cells were treated with medium only and labelled as "untreated". After 24 and 48 h of treatment, cells were harvested and subjected to the downstream experiments.

All the in vitro assays employed either 6-well plates or 96-well plates. The seeding density for 6-well plates for both cell lines was  $4 \times 10^5$  cells and treatment methods were as previously described [8]. The seeding density of Hep G2 cells for 96 well plates was  $4 \times 10^3$  cells per well and for MDA-MB-231 cells was  $8 \times 10^4$  cells per well, and used the same treatment method as above: after seeding and culturing overnight, cells were treated with 2 mg/ml CKI diluted with complete medium for the specified time.

### Glucose consumption assay

Glucose consumption was assessed in both cell lines in 6-well plates. Glucose consumption was determined by using a glucose oxidase test kit (GAGO-20, Sigma Aldrich, MO, USA). After culturing for different durations (3, 6, 12, 24 and 48 h), 50  $\mu$ l of culture medium was collected from untreated groups and treated groups. The cells were trypsinized for cell number determination using trypan blue exclusion assay and the number of bright, viable cells were counted using a hemocytometer. Collected suspension, blank medium and 2 mg/ml CKI, were all filtered and diluted 100 fold with MilliQ water. The absorbance at 560 nm was converted to glucose concentration using a 5  $\mu$ g/ml glucose standard from the kit as a single standard. Glucose consumption was calculated by subtracting the blank medium value from treated/collected medium values. Glucose consumption per cell was calculated from the number of cells determined above.

### Measurement of [ADP]/[ATP] ratio

Cells were cultured in white 96-well plates with clear bottoms. The [ADP]/[ATP] ratio of both cell lines was determined immediately after the incubation period (24 and 48 h) using an assay kit (MAK135; Sigma Aldrich) according to the manufacturer's instructions. Levels of luminescence from the luciferase-mediated reaction was measured using a plate luminometer (PerkinElmer 2030 multilabel reader, MA, USA for CKI experiments or Promega, WI, USA for oxymatrine experiments). The [ADP]/[ATP] ratio was calculated from the luminescence values using a formula provided by the kit manufacturer.

### Lactate content assay

The concentration of lactate, the end product of glycolysis, was determined using a lactate colorimetric assay kit (Abcam, MA, USA). Cells were cultured in 6-well plates, and then harvested and deproteinized according to the manufacturer's protocol. The optical density was

measured at 450 nm and a standard curve plot (nmol/well vs. OD 450 nm) was generated using serial dilutions of lactate. Lactate concentrations were calculated with formula provided by the kit manufacturer.

### Cell cycle assay

Cells were cultured in 6-well plates and treated with 2 mg/ml CKI or 0.5 mg/ml oxymatrine. After culturing for 3, 6, 12, 24 and 48 h, cells were harvested and subjected to cell cycle analysis by propidium iodide staining as previously reported [8]. Data were obtained by flow cytometry using Accru (BD Biosciences, NJ, USA) and analysed using FlowJo software (Tree Star Inc, OR, USA).

### Microscopy

After treating for 48 h on 22×22 Deckglaser cover glasses placed in 6-well plates, control and treated cells were fixed in 1% paraformaldehyde for 10 min at room temperature, and permeabilized with 0.5% Triton X-100 for 10 min. After fixation and permeabilization, cells were blocked with 5% fetal bovine serum for 30 min. Permeabilized cells were stained with 10 μg/ml of Alexa Fluor®594 conjugated anti-γ-H2AX Phospho (red) (BioLegend, CA, USA) in 5% fetal bovine serum overnight at 4°C and mounted with 4,6-diamidino-2-phenylindole (DAPI).

Stained cells were visualized with an Olympus FV3000 (Olympus Corporation, Tokyo, Japan) confocal microscope using a 20× objective. Fluorescence intensity was automated pictured collection with ArrayScan XTI Live High Content Platform (Thermo Fisher Scientific) and software based nuclear analysis (HCS studio 3.0 Cell Analysis Software; Thermo Fisher Scientific) was implemented as for data acquirement of CircSpots under 20× focusing. The acquired number of each cell line in one replicate was, approximately 5000 cells for Hep G2 and approximately 2000 cells for MDA-MB-231.

### Detection of intranuclear/intracellular proteins by flow cytometry

Cells were cultured in 6-well plates, treated with or without CKI, and harvested on different time points to detect intranuclear/intracellular levels of proteins involved in cell cycle, DNA repair and DNA DSBs pathways using the following antibodies. To measure levels of proteins involved in cell cycle, rabbit anti-CDK1 (Abcam), rabbit anti-p53, rabbit anti-cyclin D1, rabbit anti-CDK2 along with rabbit IgG isotype control (Cell Signalling Technologies, MA, USA) were used and these were detected with anti-rabbit IgG-PE (Cell Signalling Technologies). In addition, β-catenin levels were detected with rabbit anti-β-catenin-Alexa Fluor 647 with rabbit IgG-Alexa Fluor 647 isotype control (Abcam). The levels of proteins involved in DNA repair pathway were measured with rabbit anti-Ku70-Alexa Fluor 647 or with rabbit anti-Ku80-Alexa Fluor 647

(Abcam) along with rabbit IgG-Alexa Fluor 647 isotype control. For the detection of γ-H2AX involved in DSBs pathway, mouse anti-γ-H2AX-PE and mouse IgG1-PE isotype control (BioLegend) were used.

The cells prepared for the detection of cell cycle and DNA repair pathways were fixed and permeabilised using Nuclear Factor Fixation and Permeabilization Buffer Set (BioLegend) according to the manufacturer's instructions. The cells prepared for the detection of DNA DSBs were fixed and permeabilised using chilled 70% ethanol. For samples under single time point and treatment, 2×10<sup>5</sup> cells were labelled either with target antibody or corresponding isotype. The data was acquired on a FACS Canto (BD Biosciences, NJ, USA) or Accru, and analysed using FlowJo software.

### Cell cycle functional enrichment re-analysis

In order to identify the phases of the cell cycle affected by CKI, differentially expressed gene data from Qu et al. and Cui et al. [8, 11] was submitted to the Reactome database [14], and used to identify functionally enriched genes.

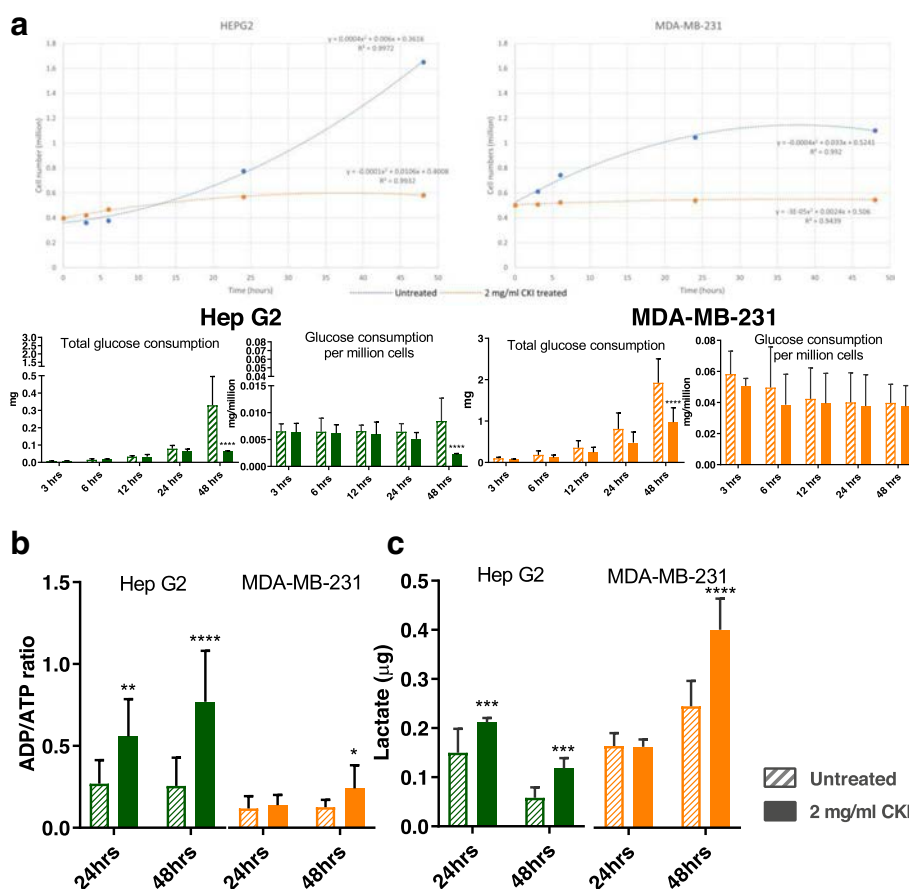
### Statistical analysis

All assays above were performed in triplicate and repeated at least three times. Statistical significance was defined as *p*-value less than 0.05, and determined by *t*-test for microscopy and two-way ANOVA test for rest of the assays with GraphPad Prism (v7.03, Graphpad Software Inc., CA, USA); error bars represent standard deviation.

## Results

### Pathway validation

Based on our previous results indicating that CKI could suppress NAD(P)H synthesis [8] and (Additional file 1: Figure S1), we examined the effect of CKI on energy metabolism by measuring glucose uptake, [ADP]/[ATP] ratio and lactate production. We measured glucose uptake in both CKI treated and untreated cells from 0 to 48 h after treatment and observed a reduction in glucose uptake (Fig. 1a). The growth curves for both cell lines were relatively flat after CKI treatment, in contrast to untreated cells. MDA-MB-231 cells, which are less sensitive to CKI in terms of apoptosis, had a higher level of glucose uptake than Hep G2 cells, which are more sensitive to CKI. Because the overall glucose uptake was consistent with the cell growth curves, the glucose consumption per million cells for each cell line under treatment was different. Untreated Hep G2 cells maintained a relatively flat rate of glucose consumption per million cells, while for CKI treated Hep G2 cells the rate of glucose consumption per million cells decreased with time, becoming significantly less towards 48 h. The glucose consumption variance for both untreated and treated MDA-MB-231 cells was



**Fig. 1** The energy metabolism determination assays in the two cell lines. **a** Growth curve (top panels) and comparison of glucose consumption analysis (lower panels) between the two cell lines at 3, 6, 12, 24 and 48 h. Overall glucose consumption is divided by cell number to calculate the consumption of glucose per million cells. **b** [ADP]/[ATP] ratio assay result for the two cell lines at 24 and 48 h. **c** Lactate content detection for the two cell lines at 24 and 48 h. Statistical analyses were performed using two-way ANOVA comparing treated with untreated (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ ); bars show one standard deviation from the mean

high, but both overall glucose consumption and glucose consumption per million cells appeared to decrease over time.

Because changes in glucose consumption are mirrored by other aspects of energy metabolism, we assessed the energy charge of both CKI treated and untreated cells by measuring the [ADP]/[ATP] ratio at 24 and 48 h after treatment (Fig. 1b). Hep G2 cells had a lower energy charge (higher [ADP]/[ATP] ratio) compared to MDA-MB-231 cells and after CKI treatment both cell lines showed a decrease in energy charge, consistent with our previous measurements using a 2,3-Bis(2-methoxy-4-nitro-5-sulfonyl)-2H-tetrazolium-5-carboxanilide inner salt (XTT) assay (Additional file 1: Figure S1). However the decrease in energy charge was earlier and much more pronounced for Hep G2 cells compared to MDA-MB-231 cells.

The flip side of glucose consumption is the production of lactate via glycolysis, which is the initial pathway

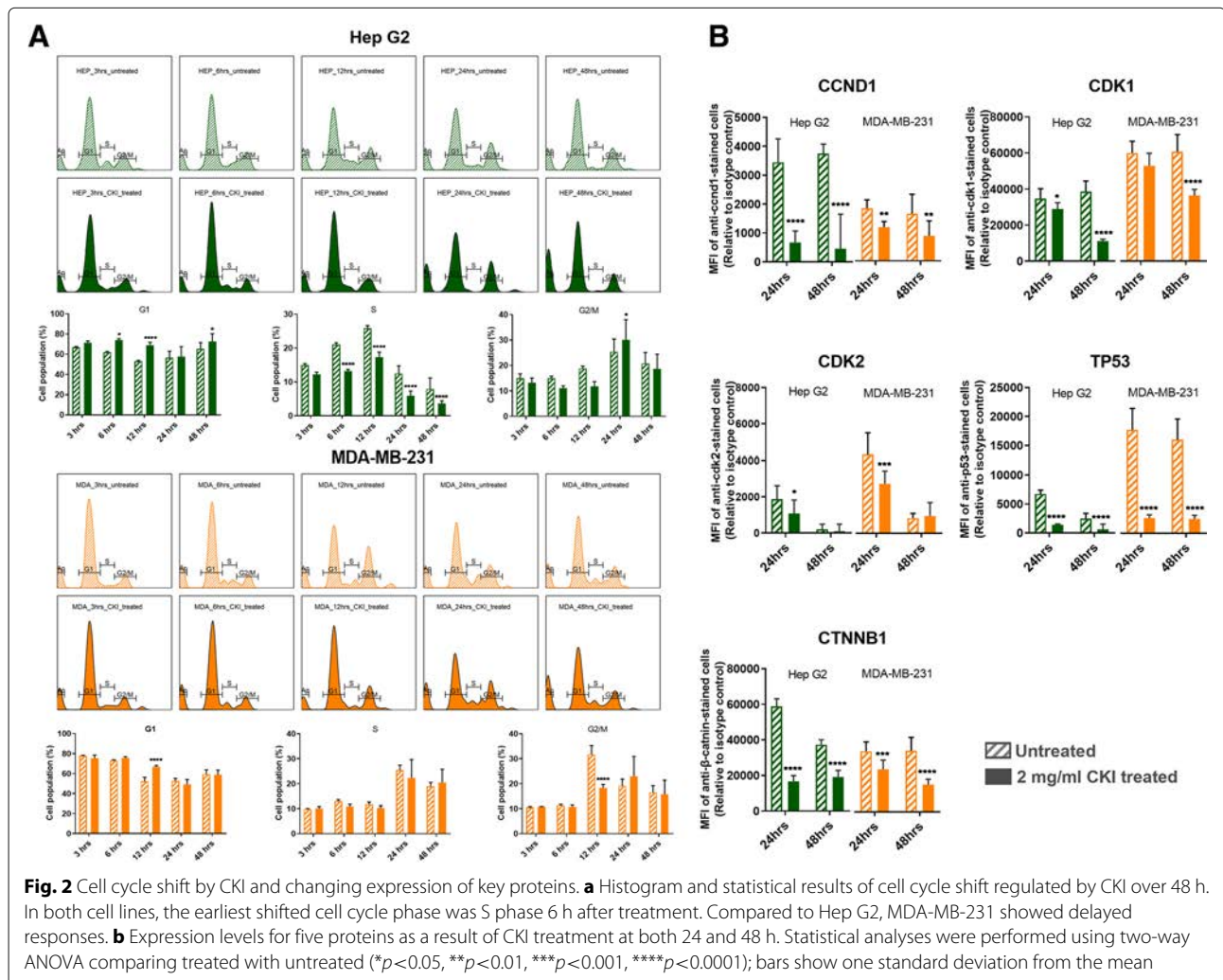
for glucose metabolism. We therefore measured lactate production in order to determine if the observed decreases in energy charge and glucose consumption were directly attributable to reduced glycolytic activity. We measured intracellular lactate concentration in both CKI treated and untreated cells at 24 and 48 h after treatment (Fig. 1c) and found that lactate concentrations increased as a function of CKI treatment in both cell lines. This result is consistent with a build up of lactate due to an inhibition of the Tricarboxylic Acid (TCA) cycle leading to decreased oxidative phosphorylation and lower cellular energy charge. CKI must therefore inhibit cellular energy metabolism downstream of glycolysis, most likely at the level of the TCA cycle. Decreased energy charge can have widespread effects on a number of energy hungry cellular processes involved in the cell cycle, such as DNA replication.

Having validated the effect of CKI on cellular energy metabolism, we proceeded to examine the perturbation of



cell cycle and expression of cell cycle proteins, as these are energy intensive processes. We had previously identified the cell cycle as a target for CKI based on transcriptome data from CKI treated cells [8, 11]. We carried out cell cycle profiling on CKI treated and untreated cells using propidium iodide staining and flow cytometry (Fig. 2a) as described in Materials and Methods. The two cell lines displayed slightly different profiles to each other, but their response to CKI was similar in terms of an increase in the proportion of cells in G1-phase. For Hep G2 cells, CKI caused consistent reductions in the proportion of cells in S-phase accompanied by corresponding increases in the proportion of cells in G1-phase. This is indicative of a block in S-phase leading to accumulation of cells in G1-phase. For MDA-MB-231 cells, CKI did not promote a significant decrease in the proportion of cells in S-phase, but did cause an increase in the percentage of cells in G1 phase at 24 h and a pronounced decrease in cells in G2/M phase at 12 h.

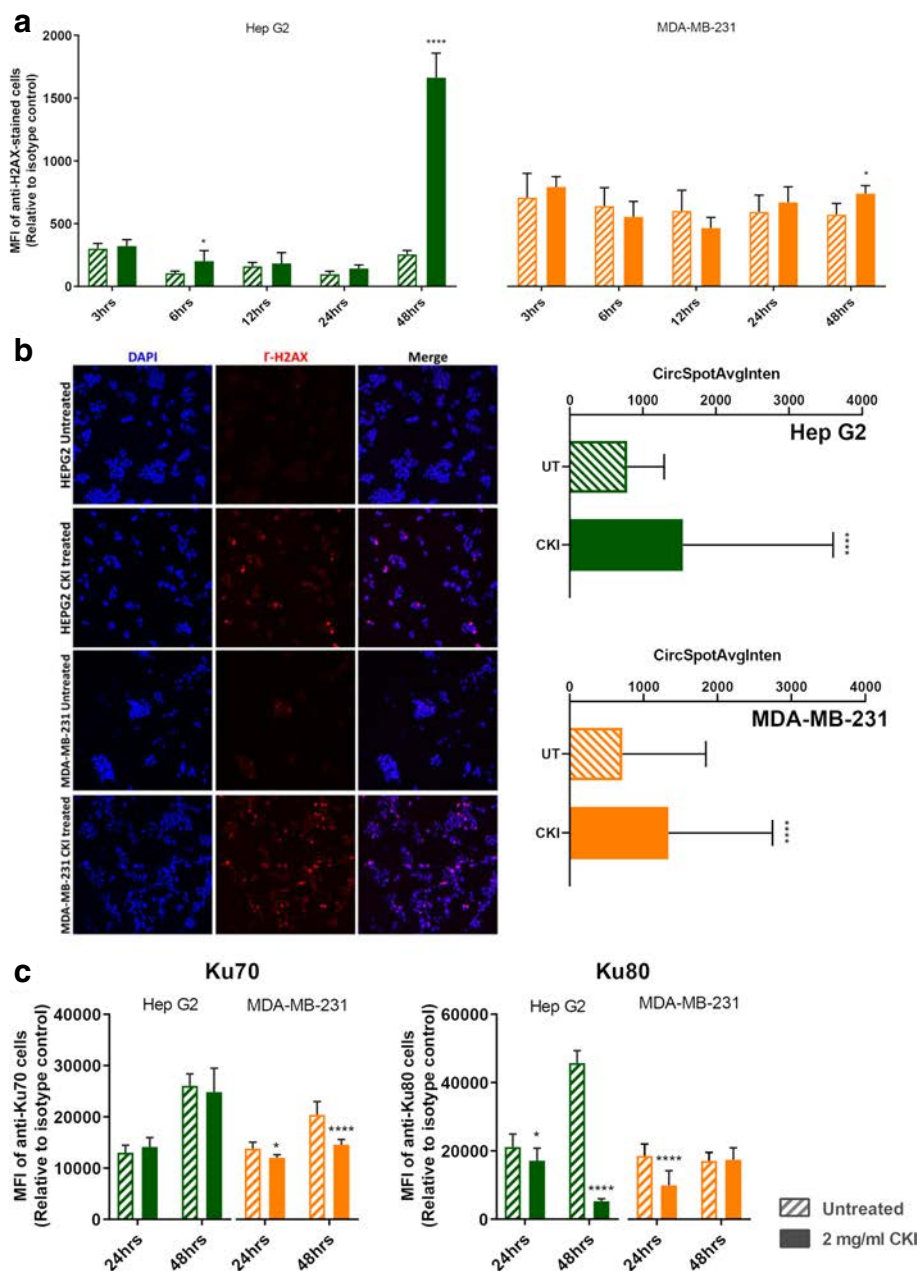
We also examined the levels of key proteins involved in the cell cycle pathway (Cyclin D1:CCND1, Cyclin Dependent Kinase 1:CDK1, Cyclin Dependent Kinase 2:CDK2, Tumor Protein p53:TP53 and Catenin Beta 1:CTNNB1) at 24 and 48 h after CKI treatment previously shown to have altered transcript expression by CKI (Fig. 2b). Both cell lines showed similar results for all five proteins, with decreased levels caused by CKI, and validated previous RNAseq data [8, 11]. *CCND1* regulates the cell-cycle during G1/S transition. *CDK-1* promotes G2-M transition, and regulates G1 progress and G1-S transition. *CDK-2* acts at the G1-S transition to promote the E2F transcriptional program and the initiation of DNA synthesis, and modulates G2 progression. *TP53* acts to negatively regulate cell division. *CTNNB1* acts as a negative regulator of centrosome cohesion. Down-regulation of these proteins is therefore consistent with cell cycle arrest/dysregulation and the cell cycle result in Fig. 2a. These results indicate that CKI alters cell cycle regulation consistent with cell



cycle arrest. Cell cycle arrest is also an outcome that can result from DNA damage such as DSBs [15].

We had previously observed that DNA repair genes had lower transcript levels in CKI treated cells [8, 11], so hypothesised that this might result in increased numbers of DSBs. We measured the expression of  $\gamma$ -H2AX

in both cell lines (Fig. 3a) and found that it was only over-expressed at 48 h in CKI treated Hep G2 cells. We also carried out localization of  $\gamma$ -H2AX using quantitative immunofluorescence microscopy and determined that the level of  $\gamma$ -H2AX increased in nuclei of CKI treated cells in both cell lines (Fig. 3b). These results



**Fig. 3** DSBs were increased by CKI treatment. **a**  $\gamma$ -H2AX expression from 3 to 48 h after treatment with 2 mg/ml CKI in two cell lines. **b** Localization of  $\gamma$ -H2AX in two cell lines after CKI treatment for 48 h. Blue is DAPI staining of nuclei, pink/red is staining of DSBs with antibody to  $\gamma$ -H2AX. The bar graph shows a quantification of the average number of  $\gamma$ -H2AX foci per cell detected in immunofluorescence images of 2 mg/ml CKI treated and untreated groups of 3 independent replicate experiments. **c** Expression of DSBs repair proteins, Ku70 and Ku80, as a result of treatment with 2 mg/ml CKI in two cell lines. Statistical analyses were performed using two-way ANOVA or *t*-test (for microscopy) comparing treated with untreated (\* $p$ <0.05, \*\* $p$ <0.01, \*\*\* $p$ <0.001, \*\*\*\* $p$ <0.0001); bars show one standard deviation from the mean

indicated an increase in DSBs as a result of CKI treatment. In order to confirm if reduced expression of DNA repair proteins was correlated with the increase in DSBs we measured levels of Ku70/Ku80 proteins in CKI treated cells (Fig. 3c). In Hep G2 cells, Ku80, a critical component of the Non-Homologous End Joining (NHEJ) DNA repair pathway was significantly down-regulated at both 24 and 48 h after CKI treatment. In MDA-MB-231 cells, Ku70 expression was down-regulated at both 24 and 48 h after CKI treatment, and Ku80 was down-regulated at 24 h. Because Ku70/Ku80 are subunits of a required DNA repair complex, reduced expression of either subunit will result in decreased DNA repair. Our results therefore support a suppressive effect of CKI on DNA repair, likely resulting in an increased level of DSBs.

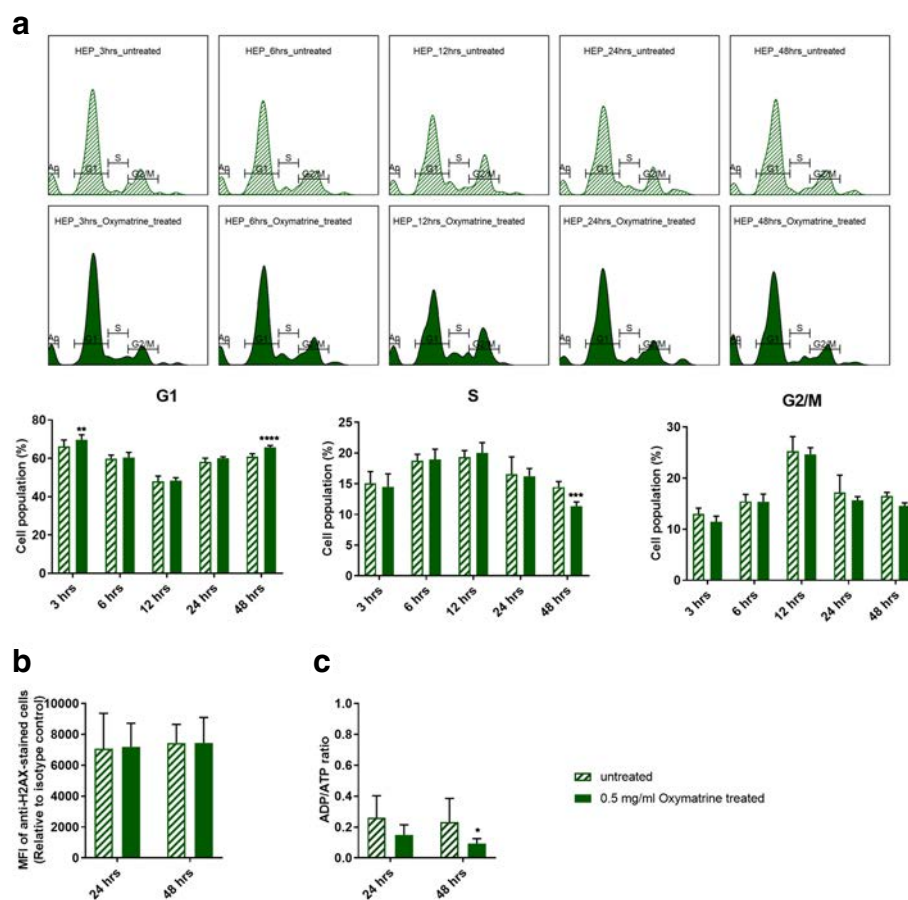
#### Effect of oxymatrine, the principal alkaloid in CKI

Because CKI is a complex mixture of many plant secondary metabolites that may have many targets and there

is little known about its molecular mode of action, we examined the effects of the most abundant single compound found in CKI, oxymatrine, on the most sensitive cell line, Hep G2. Oxymatrine is an alkaloid that has previously been reported to have effects similar to CKI, so we expected it might have an effect on one or more of our three validated pathways.

Oxymatrine, at 0.5 mg/ml which is equivalent to the concentration of oxymatrine in 2 mg/ml CKI, did not have an equivalent effect on the cell cycle compared to CKI (Fig. 4a vs Fig. 2a). Oxymatrine caused only minor changes to the cell cycle with small but significant increases in the proportion of cells in G1-phase at 3 and 48 h and a small but significant decrease in the proportion of cells in S1-phase at 48 h. Oxymatrine also caused a significant increase in the proportion of cells undergoing apoptosis in Hep G2 cells, albeit at a lower level than CKI (Additional file 1: Figure S2).

Oxymatrine had no effect on  $\gamma$ -H2AX levels in Hep G2 cells (Fig. 4b). This was in stark contrast to the effect of



**Fig. 4** Effect of oxymatrine alone on validated pathways. Oxymatrine was tested at 0.5 mg/mL which is equivalent to its concentration in CKI. **a** Histogram and statistical results of cell cycle affected by oxymatrine over 48 h. **b** Effect of oxymatrine on  $\gamma$ -H2AX levels after 24 and 48 h. **c** Effect of oxymatrine on [ADP]/[ATP] ratio after 24 and 48 h. Statistical analyses were performed using two-way ANOVA comparing treated with untreated ( $*p < 0.05$ ,  $**p < 0.01$ ,  $***p < 0.001$ ,  $****p < 0.0001$ ); bars show one standard deviation from the mean

CKI (Fig. 3a) at 48 h and indicated that oxymatrine alone had no effect on the level of DSBs.

Surprisingly, oxymatrine had the opposite effect on energy metabolism compared to CKI, causing a decrease in [ADP]/[ATP] ratio indicating a large increase in the energy charge of the cells (Fig. 4c).

### Integration of results

The effect of CKI on cancer cells was validated in all three of our candidate pathways: cell cycle, energy metabolism and DNA repair. Because these pathways are not isolated, but instead are integrated aspects of cell physiology CKI may act through targets in some or all of these three pathways or may act through other targets that either directly or indirectly suppress these pathways. CKI may also act through the synergistic effects of multiple compounds on multiple targets in our candidate pathways. This possibility is consistent with the partial and minor effects of oxymatrine on our candidate pathways.

### Discussion

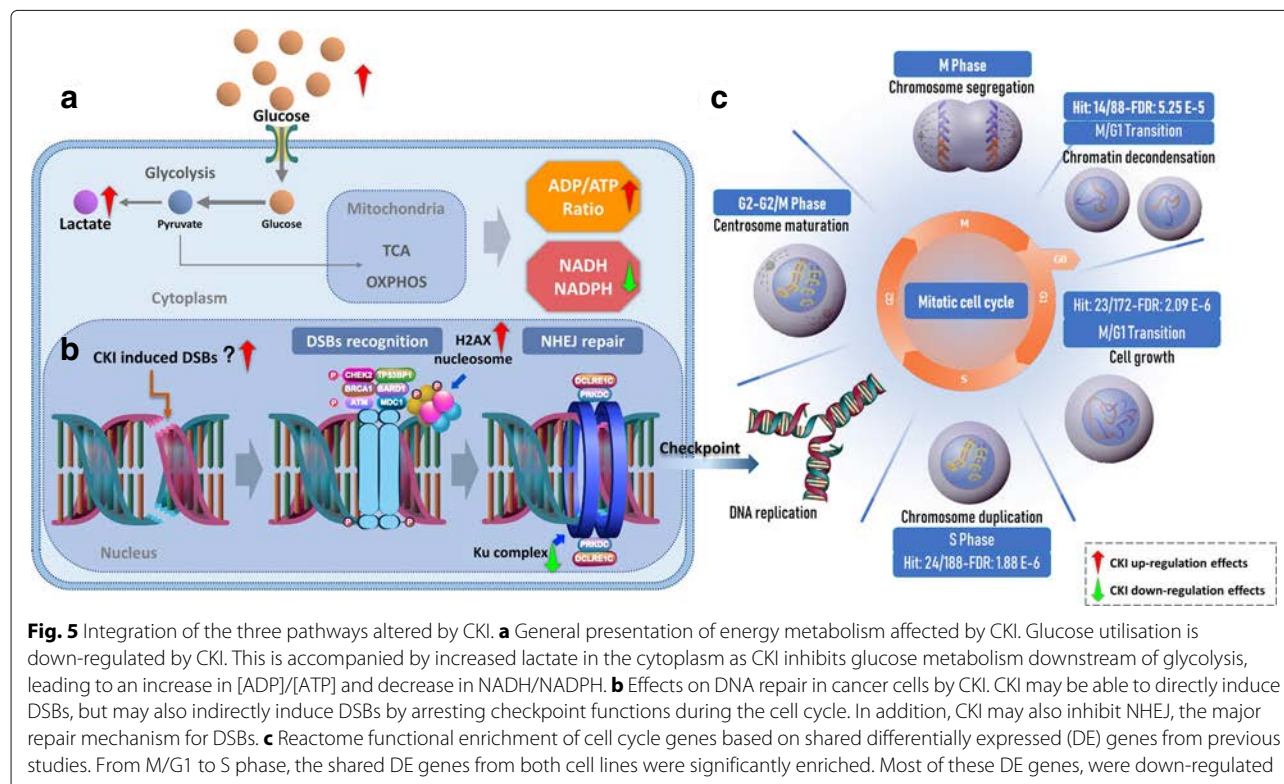
We have validated three pathways (cell cycle, energy metabolism and DNA repair) that are perturbed by CKI and that can be used as the focus for further investigations to identify specific molecular targets that mediate the perturbations.

### Cell cycle perturbation by CKI

Our results show that CKI can perturb the cell cycle by altering the proportions of cells in G1-phase, S-phase and G2/M-phase. This result is similar to what we have observed before [8, 11], but has not been widely reported in the literature. The alkaloid oxymatrine, the most abundant compound present in CKI, has also been shown previously to perturb a number of signaling pathways [16] and alter/arrest the cell cycle in a variety of cancer cells [17–21] and we have confirmed this observation (Fig. 4a) in Hep G2 cells. Our results permit direct comparison with CKI because our experiments have been done using equivalent concentrations of oxymatrine alone or in CKI. While oxymatrine has an effect on the cell cycle, it is not as effective at perturbing the cell cycle as is CKI. This indicates that oxymatrine must interact with other compounds in CKI to have a stronger effect on the cell cycle.

### Energy metabolism suppression by CKI

We have shown for the first time that CKI can inhibit energy metabolism as demonstrated by lower levels of NADH/NADPH and a higher [ADP]/[ATP] ratio. These results, combined with lower glucose utilisation and higher lactate levels indicate that this suppression was likely due to inhibition of the TCA cycle or oxidative phosphorylation. Previously, Gao et al. [7] have reported that



CKI significantly increased the concentration of pyruvate in the medium and this observation in combination with our results supports a decrease in metabolic flux through the TCA cycle as the likely cause of the reported suppression of energy metabolism. Interestingly, oxymatrine on its own had the opposite effect on [ADP]/[ATP] ratio compared to CKI, indicating that it can enhance energy metabolism and increase the energy charge of the cell.

### DNA repair suppression by CKI

There is only one report in the literature of oxymatrine inducing DSBs [22] and no reports with respect to CKI. Our results show for the first time that not only does CKI induce DSBs, but that is also likely inhibits DNA repair by decreasing the expression of the Ku70/Ku80 complex required for NHEJ mediated DNA repair. It is worth noting however, that the reported effect of oxymatrine on DSBs [22] uses significantly higher (4–8 fold) concentrations of oxymatrine compared to our experiments. In our hands oxymatrine alone at 0.5 mg/ml showed no effect on DSBs as judged by the level of  $\gamma$ -H2AX after 24 or 48 h.

### Conclusions

CKI causes suppression of energy metabolism and DNA repair along with altered cell cycle (summarized in Fig. 5). CKI has also previously been reported to induce apoptosis in cancer cells [8]. The overarching question is if CKI has independent effects on these three pathways or if the primary effect of CKI is through a single pathway that propagates effects to other, physiologically linked pathways. It may be that CKI suppresses energy metabolism thus disrupting downstream, energy hungry processes such as DNA replication and DNA repair. Alternatively, there could be independent effects on DNA repair leading to checkpoint induced cell cycle perturbation/arrest. Our results based on oxymatrine treatment of Hep G2 cells indicate that the cell-cycle is likely directly affected by oxymatrine and thus CKI. However oxymatrine alone had no effect on DNA repair and boosted, rather than reduced the energy charge of the cell. Taken together, these results support a model of many compounds/many targets [23] for the mode of action of CKI, where multiple compounds affect multiple targets and the synergistic, observed effect is significantly different to that seen with individual components.

### Additional file

**Additional file 1:** Contains supplementary figures as referred to in the main body of the paper. (PDF 721 kb)

### Abbreviations

CKI: Compound Kushen injection; CMM: Chinese materia medica; DE: Differentially expressed; DSBs: Double strand breaks; Ku70/Ku80: The Ku heterodimer proteins; NHEJ: Non-homologous end joining; TCM: Traditional Chinese medicine

### Acknowledgements

We would like to thank Prof. Frank Grutzner for providing DAPI, Adelaide Microscopy for training and equipment use and Jue Zhang, Bo Han and Dan Kortschak for helpful discussions.

### Funding

This project is supported by The Special International Cooperation Project of Traditional Chinese Medicine (GZYYGJ2017035) and The University of Adelaide - Zhendong Australia - China Centre for Molecular Chinese Medicine. Funding bodies played no role in study design, data collection, analysis or interpretation of the data.

### Availability of data and materials

All data analyzed in this study are available from the public source NCBI (<https://www.ncbi.nlm.nih.gov/>) and details can be found in the supplementary file.

### Authors' contributions

JC, ZQ, YH-L and DLA designed research; JC, YH-L, ZQ, TNA, HS and WW performed research; and JC, ZQ, YH-L and DLA wrote the paper. All authors have read and approved the manuscript.

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests. While a generous donation was used to set up the Zhendong Centre by Shanxi Zhendong Pharmaceutical Co Ltd, they did not determine the research direction for this work or influence the analysis of the data. JC: no competing interests, ZQ: no competing interests, YHL: no competing interests, HS: no competing interests, TNA: no competing interests, WW: is an employee of Zhendong Pharma seconded to Zhendong Centre to learn bioinformatics methods and is a PI on The Special International Cooperation Project of Traditional Chinese Medicine (GZYYGJ2017035), DLA: Director of the Zhendong Centre which was set up with a generous donation from the Zhendong Pharmaceutical Co Ltd. Zhendong Pharmaceutical has had no control over these experiments, their design or analysis and have not exercised any editorial control over the manuscript.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Author details

<sup>1</sup>Department of Molecular and Biomedical Science, The University of Adelaide, 5005 North Terrace, Adelaide, South Australia, Australia. <sup>2</sup>Zhendong Australia - China Centre for Molecular Chinese Medicine, The University of Adelaide, 5005 North Terrace, Adelaide, South Australia, Australia. <sup>3</sup>Zhendong Research Institute, Shanxi-Zhendong Pharmaceutical Co Ltd, Beijing, China.

Received: 11 July 2018 Accepted: 17 December 2018

Published online: 24 January 2019

### References

- Zhang C, Wang Y-M, Zhavo F-C, Chen L-M, Zhang Q-W, Gao H-M, Wang Z-M. Phenolic metabolites from the stems and leaves of *sophora flavescens*. *Helvetica Chimica Acta*. 2014;97(11):1516–25.
- Sun M, Cao H, Sun L, Dong S, Bian Y, Han J, Zhang L, Ren S, Hu Y, Liu C, Xu L, Liu P. Antitumor activities of kushen: literature review. *Evid Based Complement Alternat Med*. 2012;2012:373219. <https://doi.org/doi:10.1155/2012/373219>.

3. Zhang L, Liu W, Zhang R, Wang Z, Shen Z, Chen X, Bi K. Pharmacokinetic study of matrine, oxymatrine and oxysophocarpine in rat plasma after oral administration of sophora flavescens ait. extract by liquid chromatography tandem mass spectrometry. *J Pharm Biomed Anal.* 2008;47(4-5):892–8.
4. Wang W, You R-I, Qin W-j, Hai L-n, Fang M-j, Huang G-h, Kang R-x, Li M-h, Qiao Y-f, Li J-w, et al. Anti-tumor activities of active ingredients in compound kushen injection. *Acta Pharmacologica Sinica.* 2015;36(6):676.
5. Tang Y, He X, Quanlan C, Lanlan F, Jianye Z, Zhongzhen Z, Dong L, Zhitao L, Yi T, Chen H. A mixed microscopic method for differentiating seven species of "bixie"-related chinese materia medica. *Microsc Res Tech.* 2014;77(1):57–70.
6. Zhao Z, Fan H, Higgins T, Qi J, Haines D, Trivett A, Oppenheim JJ, Wei H, Li J, Lin H, et al. Fufang kushen injection inhibits sarcoma growth and tumor-induced hyperalgesia via trpv1 signaling pathways. *Cancer Lett.* 2014;355(2):232–41. <https://doi.org/10.1016/j.canlet.2014.08.037>.
7. Gao L, Wang K-X, Zhou Y-Z, Fang J-S, Qin X-M, Du G-H. Uncovering the anticancer mechanism of compound kushen injection against hcc by integrating quantitative analysis, network analysis and experimental validation. *Sci Rep.* 2018;8(1):624. <https://doi.org/10.1038/s41598-017-18325-7>.
8. Qu Z, Cui J, Harata-Lee Y, Aung TN, Feng Q, Raison JM, Kortschak RD, Adelson DL. Identification of candidate anti-cancer molecular mechanisms of compound kushen injection using functional genomics. *Oncotarget.* 2016;7(40):66003–19. <https://doi.org/10.18632/oncotarget.11788>.
9. Zhou S-K, Zhang R-L, Xu Y-F, Bi T-N. Antioxidant and immunity activities of fufang kushen injection liquid. *Molecules.* 2012;17(6):6481–90. <https://doi.org/10.3390/molecules17066481>.
10. Xu W, Lin H, Zhang Y, Chen X, Hua B, Hou W, Qi X, Pei Y, Zhu X, Zhao Z, Yang L. Compound kushen injection suppresses human breast cancer stem-like cells by down-regulating the canonical wnt/ $\beta$ -catenin pathway. *J Exp Clin Cancer Res.* 2011;30:103. <https://doi.org/10.1186/1756-9966-30-103>.
11. Cui J, Qu Z, Harata-Lee Y, Shen H, Aung TN, Wang W, Kortschak RD, Adelson DL. The Effect of Compound Kushen Injection on Cancer Cells: Integrated Identification of Candidate Molecular Mechanisms. *bioRxiv* 503318. <https://doi.org/10.1101/503318>.
12. Ma X, Li R-S, Wang J, Huang Y-Q, Li P-Y, Wang J, Su H-B, Wang R-L, Zhang Y-M, Liu H-H, et al. The therapeutic efficacy and safety of compound kushen injection combined with transarterial chemoembolization in unresectable hepatocellular carcinoma: An update systematic review and meta-analysis. *Front Pharmacol.* 2016;7 <https://doi.org/10.3389/fphar.2016.00070>.
13. Zhongquan Z, Hehe L, Ying J. Effect of compound kushen injection on t-cell subgroups and natural killer cells in patients with locally advanced non-small-cell lung cancer treated with concomitant radiochemotherapy. *J Tradit Chin Med.* 2016;36(1):14–8. [https://doi.org/doi:10.1016/S0254-6272\(16\)30002-4](https://doi.org/doi:10.1016/S0254-6272(16)30002-4).
14. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, Haw R, Jassal B, Korninger F, May B, et al. The reactome pathway knowledgebase. *Nucleic Acids Res.* 2018;46(D 1):649–55.
15. Shaltiel IA, Krenning L, Bruinsma W, Medema RH. The same, only different - dna damage checkpoints and their reversal throughout the cell cycle. *J Cell Sci.* 2015;128(4):607–20. <https://doi.org/10.1242/jcs.163766>.
16. Lu M-L, Xiang X-H, Xia S-H. Potential signaling pathways involved in the clinical application of oxymatrine. *Phytother Res.* 2016;30(7):1104–12. <https://doi.org/10.1002/ptr.5632>.
17. Li W, Yu X, Tan S, Liu W, Zhou L, Liu H. Oxymatrine inhibits non-small cell lung cancer via suppression of egfr signaling pathway. *Cancer Med.* 2018;7(1):208–18. <https://doi.org/10.1002/cam4.1269>.
18. He M, Jiang L, Li B, Wang G, Wang J, Fu Y. Oxymatrine suppresses the growth and invasion of mg63 cells by up-regulating pten and promoting its nuclear translocation. *Oncotarget.* 2017;8(39):65100–10. <https://doi.org/10.18632/oncotarget.17783>.
19. Li S, Zhang Y, Liu Q, Zhao Q, Xu L, Huang S, Huang S, Wei X. Oxymatrine inhibits proliferation of human bladder cancer t24 cells by inducing apoptosis and cell cycle arrest. *Oncol Lett.* 2017;13(6):4453–8. <https://doi.org/10.3892/ol.2017.6013>.
20. Wu J, Cai Y, Li M, Zhang Y, Li H, Tan Z. Oxymatrine promotes s-phase arrest and inhibits cell proliferation of human breast cancer cells in vitro through mitochondria-mediated apoptosis. *Biol Pharm Bull.* 2017;40(8):1232–9. <https://doi.org/10.1248/bpb.b17-00010>.
21. Ying X-J, Jin B, Chen X-W, Xie J, Xu H-M, Dong P. Oxymatrine downregulates hpv16e7 expression and inhibits cell proliferation in laryngeal squamous cell carcinoma hep-2 cells in vitro. *Biomed Res Int.* 2015;2015:150390. <https://doi.org/doi:10.1155/2015/150390>.
22. Wang Z, Xu W, Lin Z, Li C, Wang Y, Yang L, Liu G. Reduced apurinic/aprimidinic endonuclease activity enhances the antitumor activity of oxymatrine in lung cancer cells. *Int J Oncol.* 2016;49(6):2331–40.
23. Li F-S, Weng J-K. Demystifying traditional herbal medicine with modern approach. *Nat Plants.* 2017;3:17109. <https://doi.org/10.1038/nplants.2017.109>.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)



## Chapter 6

### Fractional Deletion of Compound Kushen Injection, a Natural Compound Mixture, Indicates Cytokine Signaling Pathways are Critical for Its Perturbation of The Cell Cycle

Plant secondary metabolites are the building blocks of TCM, and because of the extreme complexity of these compounds, it is usually not clear which components are responsible for the effects of TCM and how they interact with other compounds. In this chapter, we introduce an approach based on applied analytical chemistry with a cell based assay and transcriptome analysis to explain the chemical basis of TCM. By knocking out different numbers of primary components and testing the activity of residual mixtures, we concluded that many compounds are necessary for CKI's anticancer effects and no single compound can account for the entirety of its effects. Furthermore, our methods are more comprehensive and reliable than previous TCM research that relied on one or two main components to represent the whole prescription. This chapter is in the format of a manuscript that has been submitted to *Scientific Reports*.

# Statement of Authorship

Title of Paper	Fractional deletion of Compound Kushen Injection, a natural compound mixture, indicates cytokine signalling pathways are critical for its perturbation of the cell cycle
Publication Status	<input type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input checked="" type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	

## Principal Author

Name of Principal Author (Candidate)	Hanyuan Shen		
Contribution to the Paper	Assisted with experiments		
Overall percentage (%)	20%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the sixth author of this paper.		
Signature		Date	13/3/2019

## Co-Author Contributions

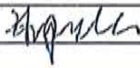
By signing the Statement of Authorship, each author certifies that:


- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.


Name of Co-Author	Thazin Nwe Aung		
Contribution to the Paper	Experimental design, carried out experiments, analysed data, wrote paper		
Signature		Date	13/03/2019


Name of Co-Author	Saeed Nourmohammadi		
Contribution to the Paper	Experimental design, carried out experiments, analysed data, wrote paper		
Signature		Date	14/03/2019





Name of Co-Author	Zhipeng Qu		
Contribution to the Paper	Experimental design, assisted with experiments, assisted with data analysis, wrote paper		
Signature		Date	18/03/2019

Name of Co-Author	Yuka Harata-Lee		
Contribution to the Paper	Experimental design, assisted with experiments, assisted with data analysis, wrote paper		
Signature		Date	13/3/19

Name of Co-Author	Jian Cui		
Contribution to the Paper	Assisted with experiments		
Signature		Date	19/03/2019

Name of Co-Author	Andrea J Yool		
Contribution to the Paper	Assisted with experiments, wrote paper		
Signature		Date	19/03/2019

Name of Co-Author	Tara Pukala		
Contribution to the Paper	Assisted with experiments, wrote paper		
Signature		Date	20/3/19

Name of Co-Author	Hong Du		
Contribution to the Paper	Experimental design		
Signature		Date	19/3/2019

Name of Co-Author	Robert Daniel Kortschak		
Contribution to the Paper	Experimental design, wrote paper		
Signature		Date	15/3/19

Name of Co-Author	David L. Adelson		
Contribution to the Paper	Supervised the research, acquired funding for the experiments, experimental design, wrote paper		
Signature		Date	1/4/2019

## **Fractional Deletion Of Compound Kushen Injection, A Natural Compound Mixture, Indicates Cytokine Signaling Pathways Are Critical For Its Perturbation Of The Cell Cycle.**

**Aung TN<sup>1</sup>†, Nourmohammadi S<sup>2</sup>†, Qu Z<sup>1</sup>, Harata-Lee Y<sup>1</sup>, Cui J<sup>1</sup>, Shen HY<sup>1</sup>, Yool AJ<sup>2</sup>, Pukala T<sup>3</sup>, Hong Du<sup>4</sup>, Kortschak RD<sup>1</sup>, and Adelson DL<sup>\*1</sup>.**

<sup>(1)</sup> Department of Molecular and Biomedical Science, School of Biological Sciences, University of Adelaide, Adelaide, South Australia, 5005.

<sup>(2)</sup> Adelaide Medical School, University of Adelaide, Adelaide, South Australia, 5005.

<sup>(3)</sup> School of Physical Sciences, University of Adelaide, Adelaide, South Australia, 5005.

<sup>(4)</sup> School of Chinese Materia Medica, Beijing University of Chinese Medicine, Beijing, 100029, P.R.China.

**†Co-1st authors: Aung TN and Nourmohammadi S**

**\*Corresponding author: Adelson DL**

Keywords: *Sophora flavescens*, *Smilax glabra*, alkaloid, matrine

Running title: Fractional deletion of Compound Kushen Injection

## ABSTRACT

We have used computational and experimental biology approaches to identify candidate mechanisms of action of a traditional Chinese medicine. Compound Kushen Injection (CKI), in a breast cancer cell line in which CKI causes apoptosis. Because CKI is a complex mixture of plant secondary metabolites, we used a high-performance liquid chromatography (HPLC) fractionation and reconstitution approach to define chemical fractions required for CKI to induce apoptosis in MDA-MB-231 cells. Our initial fractionation separated major from minor compounds, and showed that the major compounds accounted for little of the activity of CKI. By systematically perturbing the major compounds in CKI we found that removal of no single major compound could alter the effect of CKI on cell viability and apoptosis. However, simultaneous removal of two major compounds identified oxymatrine and oxysophocarpine as critical compounds with respect to CKI activity. We then used RNA sequencing and transcriptome analysis to correlate compound removal with gene expression and phenotype data. We determined that many compounds in CKI are required for its effectiveness in triggering apoptosis but that significant modulation of its activity is conferred by a small number of compounds. In conclusion, CKI may be typical of many plant based extracts that contain many compounds in that no single compound is responsible for all of the bioactivity of the mixture and that many compounds interact in a complex fashion to influence a network containing many targets.

## INTRODUCTION

Natural compounds are chemically diverse and have long served as resources for the identification of drugs ([Harvey et al., 2015](#)). However, the standard approach of fractionating natural product extracts to identify a single compound's biological activity can fail because the original activity of the mixture is not present in single compounds after fractionation. This failure to identify single compounds implies that some natural product mixtures derive their activity from the interaction of several bioactive compounds within the mixture. Characterising the mode of action of natural product mixtures has remained a difficult task as the combinatorial complexity of such mixtures makes it unfeasible to screen all combinations of the compounds in the mixture.

We introduce here a “subtractive fractionation approach” using high performance liquid chromatography (HPLC) that can pinpoint significant interacting compounds within a mixture when coupled with a suitable bioassay. We combined this approach with RNAseq (RNA sequencing) characterisation of our bioassay, correlating the removal of interacting compounds

with concomitant alterations in gene expression. This allows us to identify specific combinations of compounds associated with specific pathways and regulatory interactions. In this report, we have applied this approach for the first time to a particular Traditional Chinese Medicine (TCM) formulation: Compound Kushen Injection (CKI), which is used to treat approximately 30,000 cancer patients/day in China in conjunction with Western chemotherapy. CKI is composed primarily of alkaloids and flavonoids extracted from two herbal medicinal plants: Kushen (*Radix Sophorae Flavescentis*) and Baituling (*Rhizoma Smilacis Glabrae*). Twenty-one chromatographic peaks have been identified from CKI with eight compounds being recognized as major components on the basis of their abundance ([Ma et al., 2014](#)).

The extract containing the most abundant compounds in CKI is derived from Kushen herb which has a long history in the treatment of patients suffering immune function disorders ([Xu et al., 2005](#); [Cheng et al., 2006](#)). The main component of CKI, macrozamin, is a derivative of baituling which has been a suggested therapeutic agent for the treatment of inflammatory disease ([Jiang et al., 1997](#)). Gao and colleagues showed that treatment with each of four of the main compounds of CKI (oxymatrine, matrine, sophoridine and N-methycytisine) at 4 mg/ml significantly decreased cell viability ([Gao et al., 2018](#)). However, these concentrations are relatively high when compared to the contributing concentration of these four main compounds in CKI ([Ma et al., 2014](#)). The two main components of CKI, matrine and oxymatrine, may have significant anticancer activities in various types of solid tumors including breast cancer non small lung cancer, cervical cancer, prostate cancer, synovial sarcoma, and hepatocellular carcinoma ([Yu et al., 2009](#); [Li et al., 2015](#); [Wu et al., 2015](#); [Cai et al., 2016](#); [Wu et al., 2016](#); [Aung et al., 2017](#); [Gao et al., 2018](#); [Zhou et al., 2018](#)). In contrast, toxicity of medicinal herbs containing matrine and oxymatrine as main components has been reported ([Wang and Yang, 2003](#)). Administration of matrine 150 mg/kg and oxymatrine 360 mg/kg significantly increased cytochrome P450 family protein CYPB1/2 in rats demonstrating a potential therapeutic drawback of these two compounds ([Yuan et al., 2010](#)). Overall, understanding the effects of CKI based on the effects of single compounds present in CKI has been at best, partially successful.

Alternatively, by removing one, two or three compounds, we have been able to map the effects of these compounds and their interactions to effects on specific pathways based on altered gene expression profiles in a cell-based assay. This has illuminated the roles of several major compounds of CKI, which on their own have little or no activity in our bioassay. This approach can be used to dissect the roles and interactions of individual compounds from complex natural compound mixtures whose biological activity cannot be attributed to single purified compounds.

## RESULTS

### *Subtractive fractionation overview*

Well resolved chromatographic separation of CKI was used to collect all of the major components of CKI as individual fractions (Figure 1A). We then reconstituted all of the separated fractions except for those we wished to subtract. We tested the reconstituted combination of compounds/peaks to see if removal of a single (N-1) or multiple compounds, (N-2 or N-3, where N represents the number of compounds in CKI), or removal of all major peaks (minor, MN) or depletion of all minor peaks (major, MJ) significantly altered the effect of CKI in our cell based assays. Our cell based assays ([Qu et al., 2016](#)) measured MDA-MB-231 (human breast adenocarcinoma) cell viability, cell cycle phase and cell apoptosis. A summary of the subtractive fractions used in the cell based assays is shown in Table 1. We then carried out RNA isolation of cells treated with CKI, individual compounds or CKI deletions for RNAseq. Differentially expressed (DE) genes in these samples allowed the association of specific compounds with cell phenotype and underlying alterations in gene regulation. By comparing DE genes across treatment combinations we identified specific candidate pathways that were altered by removal of single or multiple compounds, as detailed below.

### *HPLC fractions and content identification using LC-MS/MS*

HPLC fractionation and reconstitution was used to generate a number of N-1, N-2, N-3, MJ and MN mixtures, (Figure 1A, B, C and Supplementary Figure 1) with specific combinations and their components shown in Table 1. The concentrations of known compounds in CKI and reconstituted subtractive fractions were determined from standard curves (Supplementary Data1) for nine available reference compounds, using cytosine as an internal standard (Table 2). The combined concentration of 9 reference compounds from CKI was approximately 10.461 mg/ml, whereas subtractive fractions N-OmtOspc and N-MacOmtOspc had concentrations of reference compounds of 3.045 mg/ml and 2.335 mg/ml which were equivalent to the concentrations of these compounds in unfractionated CKI. The depleted OmtOspc and MacOmtOspc were not observed in the N-OmtOspc and N-MacOmtOspc respectively. These collectively suggested any effects observed after the treatments of N-OmtOspc and N-MacOmtOspc were not influenced by the concentrations. A total of 9 (N-1), 4 (N-3) and 9 (N-2) combinations, along with MJ and MN deletions were tested in our cell based assays (Table 1).

### *Phenotypic changes associated with compound deletion*

Fractionation and full reconstitution caused no changes in cell viability compared to original CKI (see methods) at either 24- or 48-hours in MDA-MB-231 cells. Both reconstituted CKI and CKI caused significantly reduced viability compared to untreated (UT) cells (Supplementary Figure 2A). The MJ subtractive fraction contained a total of 9 compounds including eight previously identified MJ peaks ([Ma et al., 2014](#)) and adenine (unpublished data from Ma Yue) (Figure 1A) and the MN fraction contained the remaining peaks (Figure 1C). MJ had no effect on cell viability, while MN reduced cell viability to the same extent as CKI (Figure 2A). The 9 major compounds were individually depleted from CKI and tested as 9 (N-1) subtractive fractions, with no significant alterations in cell viability compared to CKI (Figure 2B). We then assessed the interaction effects of single MJ compounds by adding them back to the MN subtractive fraction. No change in cell viability compared to MN was observed (Supplementary Figure 2B). Sets of 3 compounds from the 9 major/standard compounds of CKI were depleted to generate 3 (N-3) subtractive fractions. The nine reference compounds were allocated into three groups, one of which contained structurally similar compounds (Omt, Ospc, Spc) and two other groups ([Mac, Ade, Tri] and [Nme, Mt, Spr]) that contained structurally different compounds. Of these three fractions, N-OmtOspcSpc decreased cell viability significantly ( $P < 0.05$ ) more than CKI after 48 hours (Figure 2C) while none of the sets of three compounds on their own had any effect on cell viability (Supplementary Figure 2). We then generated 9 (N-2) subtractive fractions based on the N-3 subtractive fractions (Table 1). Out of 9 (N-2) subtractive fractions (Supplementary Figure 2), only N-OmtOspc significantly decreased proliferation compared to CKI ( $P < 0.05$ ) (Figure 2C). We then depleted macrozamin, the only major compound derived from Baituling, together with OmtOspc as N-3 (N-MacOmtOspc) in order to determine if there was an additional effect when compared to CKI. N-OmtOspc and N-MacOmtOspc both decreased cell proliferation to the same extent (Figures 2C and 2D).

While no change in cell viability was found across all N-1 treatments, cell cycle analysis was performed to identify more subtle differences. There was no statistically significant difference in phases of the cell cycle of MDA-MB-231 cells for many of the N-1 treatments compared to CKI except for a statistically significant change in G1 phase by N-Omt after 48 hours (Figure 3A). On the other hand, N-OmtOspc treatment significantly altered the cell cycle for MDA-MB-231 cells and induced significant higher apoptosis from 0.25 mg/ml through 2 mg/ml treatments as compared to CKI at both time points (Figure 3B and Supplementary Figure 3). N-MacOmtOspc treatment also significantly altered the cell cycle at both time-points with generally similar effects to N-OmtOspc (Figure 3C).

Annexin V/PI apoptosis assays were performed using subtractive fractions on MDA-MB-231, HEK-293 (human embryonic kidney cells) and HFF (primary human foreskin fibroblasts) cell lines. While CKI at 2 mg/ml caused increased apoptosis in MDA-MB-231 cells at both 24- and 48-hour after treatment, N-OmtOspc and N-MacOmtOspc subtractive fractions at concentrations equivalent to CKI 2 mg/ml significantly increased the percentage of apoptotic cells at 24-hour with increasing apoptosis at the 48-hour timepoint, indicating that N-OmtOspc and N-MacOmtOspc significantly enhanced apoptosis compared to CKI (Figures 4A, 4E and Supplementary Figure 3A). Although CKI did not generally cause apoptosis in HEK-293 or HFF cells, N-OmtOspc and N-MacOmtOspc subtractive fractions significantly induced apoptosis ( $P^{***} < 0.001$ ) at 24-hour and 48-hour ( $P^{****} < 0.0001$ ) in both HEK-293 and HFF cells. CKI only induced apoptosis of HEK-293 ( $P^* < 0.05$ ) at 48-hour and showed no significant apoptotic induction in HFF (Figures 4B, 4C and Supplementary Figures 3B, 3C). These results indicated that the N-OmtOspc and N-MacOmtOspc subtractive fractions induced apoptosis not only in cancerous cells but also in non-cancerous cell lines. In contrast to this, no significant apoptosis was triggered by CKI on HFF cells. A small but significant apoptotic induction was observed for HEK-293.

Because of the significant decreased viability accompanied by increased apoptosis triggered by subtractive fractions, cytotoxicity tests were carried out for all three cell lines using CKI (2 mg/ml) and N-OmtOspc and N-MacOmtOspc subtractive fractions at concentrations equivalent to CKI 2 mg/ml. N-OmtOspc and N-MacOmtOspc at equivalent concentration to CKI 2 mg/ml were significantly cytotoxic to both non-cancerous cell lines (Figure 4D).

Overall, these results indicated that removal of combinations of specific compounds from CKI had unpredictable effects on the ability of CKI to kill cells. While removal of all major compounds from CKI caused no loss of activity and removal of all minor compounds caused total loss of activity, removal of selected major compounds (N-OmtOspc) paradoxically caused major, significant increases in the ability of CKI to reduce viability and killed cells.

### *Differential gene expression*

In order to understand the interactions of the components in CKI as a result of depletion, we carried out RNAseq of MDA-MB-231 cells treated with CKI and subtractive fractions. Out of nine (N-1) subtractive fractions, four, namely N-Omt, N-Mac, N-Tri and N-Nme, were selected due to their structural differences to determine their effects on transcript levels. N-OmtOspc and N-MacOmtOspc, OmtOspc, MacOmtOspc and CKI treated cells were sequenced at 24 and 48-hour time points.



After normalization, clear clustering of the replicates was observed (Figure 5A and Supplementary Figures 4, 5, 6 and 7), indicating that all 4 (N-1) treatments show comparable downstream gene expression patterns. Likewise, OmtOspc and MacOmtOspc groups and N-OmtOspc and N-MacOmtOspc groups showed similar changes in gene expression, except for one replicate (N-MacOmtOspc, 24-hour) that clustered with UT, OmtOspc and MacOmtOspc. The number of differentially expressed (DE) genes associated with each treatment was calculated using pair-wise comparative analysis. CKI treatment was used as a baseline to compare all other treatments in order to emphasize the effect of depleted compounds and CKI treatment was compared to UT.

There were thousands of upregulated and downregulated genes at 24 and 48 hours in most pairwise comparisons (Figure 5B). However DE genes between OmtOspc and MacOmtOspc treatments were not observed and there were almost no DE genes between N-Mac, N-Nme and N-Tri treatments (Figure 5B) indicating that these three subtractive fractions had very similar effects on gene expression.

When we compared the DE genes found between treatments, there were a large number of DE genes (~71.3%) shared between all four (N-1) treatments (Supplementary Figures 8 and 9 and Supplementary Table 3). A similar number of shared DE genes (~24.6%) between four (N-1), OmtOspc and MacOmtOspc and between four (N-1), N-OmtOspc and N-MacOmtOspc as compared to CKI at 48-hours indicated that gene expression patterns from N-1 treatments were mostly different from N-OmtOspc, N-MacOmtOspc, OmtOspc and MacOmtOspc treated cells. 55% of the DE genes between UT, OmtOspc and MacOmtOspc were shared. When the four (N-1) treatments were compared to CKI treatment, 42.8% of DE genes were shared, and when N-OmtOspc and N-MacOmtOspc treatments were compared to CKI, 50.1% DE genes were shared, indicating that N-OmtOspc and N-MacOmtOspc treatments appeared to be more similar to CKI than N-1 treatments.

The overall levels of similarity in DE genes were as follows: 1) All N-1 treatments had approximately 70 % similar gene expression patterns, 2) OmtOspc and MacOmtOspc treatments were approximately 50% similar to UT and 33% similar to N-1 treatments, 3) downstream gene expression patterns between N-1, N-OmtOspc and N-MacOmtOspc were approximately 37% similar.

#### *Gene ontology and pathway annotation of DE genes*

DE genes were analysed for over-representation in our data sets with respect to biological function using Gene Ontology (GO) annotation. We looked for shared DE genes between treatments and identified over-represented genes in these shared genes. The only common

function enriched across all comparisons was for "cell cycle checkpoint" (Figure 6A). This confirmed earlier results ([Qu et al., 2016](#)) and was consistent with the phenotype data for CKI.

#### *Subtracted fractionation altered pathways*

We also performed pathway based analysis to look for pathway level perturbation by comparing DE genes within Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways between treatments. We used Signaling Pathway Impact Analysis (SPIA) to identify pathways with statistically significant perturbation values expected to alter pathway flux. We identified 86 pathways (Supplementary Figure 11) with statistically significant ( $P < 0.05$ ) perturbations of gene expression and of these, 15 pathways were most obviously linked to our phenotypes of cell viability, cell cycle and apoptosis (Figure 6B). By comparing the pathway gene expression perturbation scores (pG) between treatments three specific observations could be made: 1) N-1 fractional deletions vs CKI had significant effects on flux in some pathways without phenotypic effects, 2) N-OmtOspc vs CKI which had a pronounced phenotypic effect at both 24- and 48-hours, had a significant effect on reducing estimated pathway flux for Cytokine-Cytokine Receptor, Cell Cycle and TGF-Beta signaling pathways, 3) comparison of N-1 fractional deletions vs fractional deletions of N-OmtOspc/N-MacOmtOspc showed consistent pathway perturbations for Cytokine-Cytokine Receptor and p53 signaling pathways. On this basis, we inferred that different major compounds could be deleted with very similar effects, indicating that they may have similar targets. In contrast, deleting Omt and Ospc simultaneously caused a significant shift in phenotype and was accompanied by specific perturbations in pathways that regulate inflammation, cell cycle and apoptosis. The combined deletion of Omt and Ospc had a synergistic effect on viability, cell-cycle and apoptosis and a synergistic effect on gene expression, consistent with the observed changes in pathway specific perturbation of gene expression. Because this double compound deletion potentiated the cell killing effect of CKI we hypothesised that the compounds in CKI have multiple targets leading to a phenotypic effect that reflects the integration of stimulation and inhibition across all those targets. Removal of Omt and Ospc alter the balance of stimulation and inhibition leading to an integrated effect for the remaining compounds in the mixture that caused more cell death than CKI.

More detailed examination of some of these interactions within significantly perturbed pathways highlighted the gene-specific changes in expression for some key regulators of inflammation and the cell-cycle. Most effects on gene expression from deletion of single vs two compounds were similar, suggesting that the enhanced cell killing by N-OmtOspc was due

to additive effects of the compound deletions. However, by comparing differences in pairwise comparisons between treatments at the gene level within the Cytokine-Cytokine Receptor Interaction and Cell Cycle pathways we identified a subset of genes that had opposite changes in gene expression when comparing single compound deletions to N-OmtOspc deletion. In the Cytokine-Cytokine Receptor Interaction pathway (Figure 7) these genes are IL1-R1 (Interleukin-1 Receptor), IL-27RA (Interleukin-27 Receptor alphasubunit), TNFRSF1B (Tumor Necrosis Factor Receptor Superfamily Member 1B), TNFRSF14 (Tumor Necrosis Factor Receptor Superfamily, Member 14) and OSMR (Oncostatin M Receptor/IL-31 Receptor Subunit Beta) and they all transduce inflammatory ligand signals to the NFkB pathway and/or the apoptosis pathway. In the Cell-Cycle pathway (Figure 8) these genes are CDKN1C (Cyclin-Dependent Kinase Inhibitor 1C (P57, Kip2), CDC25B (Cell Division Cycle 25B), ATR (ATR Serine/Threonine Kinase), CDKN1B (Cyclin-Dependent Kinase Inhibitor 1B (P27, Kip1)), CDKN2D (Cyclin-Dependent Kinase Inhibitor 2D (P19, Inhibits CDK4)), TGFBI (Transforming Growth Factor Beta 1), FZR1 (Fizzy And Cell Division Cycle 20 Related 1), CDC20 (Cell Division Cycle 20), CDC27 (Cell Division Cycle 27), ORC2 (Origin Recognition Complex Subunit 2), ANAPC4 (Anaphase Promoting Complex Subunit 4), ZBTB17 (Zinc Finger And BTB Domain Containing 17) and ABL1 (ABL Proto-Oncogene 1, Non-Receptor Tyrosine Kinase). The opposite changes in gene expression stimulated by N-OmtOspc compared to N-1 subfractions provides support for the idea that multiple major compounds can have similar effects on specific genes but that the combination of Omt and Ospc can have synergistic and opposite effects on those same genes. This means that multiple compounds with overlapping targets (based on their structural similarities) can either reinforce a single outcome or exhibit unpredictable and opposite effects when combined.

Overall our results support the concept of multi-compound/multi-target interactions for plant extract based drugs that contain many plant secondary metabolites. Biological effects of complex plant extracts may result from interactions of multiple compounds, with negligible effects from single compounds alone. This has implications for how we assess the functional evidence for such extracts.

## DISCUSSION

Previous studies have demonstrated that CKI can alter the cell cycle, induce apoptosis and reduce proliferation in various cancer cell lines ([Xu et al., 2011](#); [Qu et al., 2016](#); [Gao et al., 2018](#)). CKI also killed leukaemia cells via the Prdxs/ROS/Trx1 signalling pathway in an acute

myeloid leukaemia patient-derived xenograft model and caused cell cycle arrest in U937 leukaemia derived cells ([Jin et al., 2018](#)). Cell cycle arrest by CKI at checkpoints is correlated with the induction of double strand breaks by CKI treatment ([Cui et al., 2018](#)). In contrast to our experiments reported above, oxymatine was previously shown to arrest the cell cycle and induce apoptosis in human glioblastoma cells through EGFR/PI3K/Akt/mTOR signaling pathway ([Dai et al., 2018](#)) and inhibit the proliferation of laryngeal squamous cell carcinoma Hep-2 cells ([Ying et al., 2015](#)). As shown in this report, oxymatine or oxysophocarpine or combined OmtOspc treatment caused no significant change in cell viability, the cell cycle or apoptosis, in agreement with prior work that showed oxymatine and oxysophocarpine exerting no significant effect on apoptosis, cell cycle or cell proliferation in HCT116 human colon cancer cells ([Zhang et al., 2014](#)).

The paradoxical result that removal of OmtOspc caused a striking increase in apoptosis is most simply explained by a model based on integrating effects of multiple compounds on many targets. The interactions between compounds in the mixture can be synergistic and antagonistic such that if two compounds are removed that have a synergistic effect that is antagonistic to the remainder of the mixture, the resulting depleted mixture will be dis-inhibited compared to CKI. This is illustrated by our studies and others that show single compounds alone had no or little effect compared to CKI. For instance, while CKI treatment resulted in increased DNA double strand breaks and affected the cell cycle resulting in decreased cancer cell proliferation, oxymatine alone exhibited only a small effect in the same assay ([Cui et al., 2018](#)). Gao and colleagues also reported that oxysophocarpine at 4 mg/ml had no effect, oxymatine at 4 mg/ml (\*P<0.05) and CKI at 2 mg/ml (\*\*P<0.001) significantly reduced the proliferation of hepatocellular carcinoma SMMC-7721 cells *in vitro* ([Gao et al., 2018](#)). Although significant inhibition of proliferation by oxymatine occurred, the concentration used in this experiment was ~ 8x times higher than that of oxymatine in 2 mg/ml of CKI. These studies agree with our experimental outcomes that oxymatine and oxysophocarpine individually had no or little effect compared to CKI treatment.

At the level of gene expression in our study, gene ontology analysis indicated that genes for “cell cycle checkpoint” were significantly enriched in cells treated with all fractionated mixtures or mixtures of Omt and Ospc. Consistent with other studies, our results also demonstrated that these compounds had little or no phenotypic effect on their own, but that when both were deleted, the remaining compounds unexpectedly had significantly greater effects on phenotype and gene expression. When examined in the context of specific pathways,

treatment with OmtOspc or N-OmtOspc which had strikingly different effects on phenotype, had similar effects on the perturbation of the “Cytokine-Cytokine Receptor Interaction” pathway, the most commonly perturbed pathway seen in our analysis, that interestingly did not show up when comparing CKI to UT. This is consistent with previous information showing that CKI induced cytokines IL4 and IL10 in cancer patients with acute leukaemia ([Tu et al., 2016](#)). In contrast to this observation, IL4 and IL10 levels were significantly decreased in transgenic mice treated with oxymatrine at a dose of 200 mg/kg ([Dong et al., 2002](#)). In our experiment, we also observed that while CKI and many of the depleted fractions had significant effect on the genes in the “Cytokine-Cytokine Receptor Interaction” pathway, OmtOspc and MacOmtOspc had little effect on the genes in that pathway. The observation that many genes in the “Cytokine-Cytokine Receptor Interaction” pathway were not affected by OmtOspc and MacOmtOspc compared to deletion fractions confirmed that removal of compounds rather than treatment with single or a few compounds can be more informative of the role and significance of individual compounds as part of mixtures/extracts.

In summary, Our approach allowed the identification of both synergistic and antagonistic interactions within the drug mixture. Viewed as a network where the compounds and the targets are nodes and the interactions between compounds and targets, and between targets are edges, it is clear that the edges (interactions) determine the overall effect of the compound mixture. By removing one or two compounds from a mixture, we can potentially perturb the target network(s) to either reduce the effect of the mixture for some outcome or potentiate it for another. We believe this approach may be of general use for the study of herbal medicines/extracts, avoiding failures that stem from exclusive reliance on the identification of a single compound that accounts for most of the biological activity in mixtures.

## **MATERIALS AND METHODS**

### **Cell lines**

MDA-MB-231 cells were purchased from the American Type Culture Collection (ATCC, Manassas, VA). HEK-293 and HFF were kindly provided by Prof. Andrea Yool (Medical School, University of Adelaide). Cells were cultured in DMEM (Dulbecco's Modified Eagle's Medium, Invitrogen) with 10% FBS (Fetal bovine serum, Thermo Fisher Scientific) at 37°C with 5% CO<sub>2</sub>.

### **Compound fractionation by HPLC**

CKI (Batch No: 20170322, total alkaloid concentration of 26.5 mg/ml) was provided by Zhendong Pharmaceutical Co.Ltd (Shanxi, China). CKI (N) was processed to deplete Single (N-1), double (N-2) and triple (N-3) compounds using HPLC by standardizing using nine compounds, namely Oxymatrine (Omt), Oxysophocarpine (Ospc), N-Methylcytisine (Nme), Matrine (Mt), Sophocarpine (Spc), Trifolirhizin (Tri), Adenine (Ade), Sophoridine (Spr) (Beina Biotechnology Institute Co., Ltd, China), and macrozamin (Zhendong Pharmaceutical Co.Ltd) which were previously reported to be found in published and unpublished data ([Ma et al., 2014](#)). HPLC fractionation separated Minor (MN) and Major (MJ) peaks to determine the principle and secondary components. The MJ mixture contained the nine standard compounds mentioned above and MN contained the remaining CKI components. In addition, nine N-1 fractional deletions, nine N-2 fractional deletions and three N-3 fractional deletions were produced.

HPLC separation was achieved using an HPLC instrument (Shimadzu, Japan) equipped with a photodiode-array UV-Vis detector with preparative C<sub>18</sub> column (5 µm, 250 x 10 mm) (Phenomenex, Australia). The following mobile phase was used to fractionate the CKI mixture: 0.01 M ammonium acetate (adjusted to pH 8.0, solvent A) and acetonitrile + 0.09 % trifluoroacetic acid (solvent B). The flow rate was 2 ml/min and linear gradient was adopted as follows; 0 min, 100% A; 60 min, 65% A, 70 min, 100%A. The chromatogram was recorded from 200 nm to 280 nm, with monitoring at 215 nm. Samples were frozen and lyophilised using a Christ Alpha 1-2 LD lyophilizer (Martin Christ Gefriertrocknungsanlagen GmbH, Germany). Several cycles of lyophilisation and resuspension were used to remove all remaining HPLC solvents and final reconstitution was carried out using MilliQ water buffered with 10 mM Hepes (Gibco, Life technologies, USA) and adjusted to pH 6.8-7.0. Lyophilised samples were resuspended to create an equivalent dilution for compounds in the sample compared to CKI.

### **Identification of reconstituted mixtures by LC-MS/MS**

Agilent 6230 TOF mass spectrometer was used to determine the concentration of the known compounds from the CKI and reconstituted N-OmtOspc and N-MacOmtOspc mixtures. 10µL sample was injected with the flow rate of 0.8 ml/min, a gradient program of 0 min, 100 % A; 40 % B; 25 min, 60 % B, 35 min, and solvents H<sub>2</sub>O + 0.1 % formic acid (solvent A) and acetonitrile + 0.1% formic acid (solvent B). The column used was C18 (5µ, 150 x 4.6 mm, Diamonsil, Dkimatech). The recovered contents of the samples was measured by spike-in compound cytosine. Gas phase ions were generated with an electrospray source, with with key instrument parameters: gas temperature, 325; sheath gas temperature, 350; vCap, 3500;

fragmentor, 175; acquisition range (m/z) 60-17000. Calibration curves for 9 standard compounds containing various concentrations were shown in Supplementary Data 1.

### **Cell viability Assay**

XTT [2,3-bis-(2-methoxy-4-nitro-5-sulfophenyl)-2*H*-tetrazolium-5-carboxanilide] and PMS (N-methyl dibenzopyrazine methyl sulfate) (50:1, Sigma-Aldrich) assay was used to assess cell viability following. Briefly, 8,000 cells in 50  $\mu$ l of medium were plated in 96 wells trays overnight prior to drug treatments in triplicate. Cells were subsequently treated with 50  $\mu$ l of drug mixtures to provide final concentrations of 0.25, 0.5, 2 and 2 mg/ml in 100  $\mu$ l. Cell viability was then measured at 24- and 48-hours after drug treatment by the addition of 50  $\mu$ l of XTT:PMS (50:1 ratio). An equal volume of medium and treating agents plus XTT:PMS was used to subtract the background optical density (OD). The absorbance of each well was recorded using a Biotrack II microplate reader at 492 nm.

### **Annexin V/PI apoptosis assay**

Apoptosis, or programmed cell death, resulting from treatment was determined using an Annexin V-FITC apoptosis detection kit (eBioscience™ Annexin V-FITC Apoptosis Detection Kit, Thermofisher Scientific) according to the manufacturer's protocol. Briefly,  $4 \times 10^5$  cells were seeded in 6-well plates in triplicate overnight prior to treatment. On the following day, cells were treated with the agents as described for 24- and 48-hours. Data were acquired with a BD LSR-FORTESSA (NJ, USA) flow cytometer, and FlowJo software (TreeStar Inc., OR, USA) was used to analyse the acquired data and produce percent apoptosis values.

### **Cell cycle assay**

Cell culture and drug treatments were performed as described above for cell cycle analysis. A Propidium Iodide (PI) staining protocol ([Riccardi and Nicoletti, 2006](#)) was used to detect the changes in cell cycle as a result of treatment after 24- and 48-hours. The characteristics of stained cells were measured using a BD LSR-FORTESSA (NJ, USA) flow cytometer, and acquired data were analysed using FlowJo software.

### **Cytotoxicity assay**

MDA-MB-231, HEK-239 and HFF cells were seeded in 96-well plates at a density of  $2.5 \times 10^3$  cells per well in triplicate. CKI and fractionated mixtures to produce a final concentration of 1 mg/ml and 2 mg/ml were added to each well and after 24-hours of incubation and viable cells

were measured using the Alamar Blue assay (Molecular Probes, Eugene, OR). Mercuric chloride (Sigma-Aldrich, St. Louis, MO) (5  $\mu$ M) was used as a positive control and wells without cells were set as a negative control in the same plate.

### **Sample preparation and RNA sequencing**

Cells were plated in 6 well plates with a density of  $2 \times 10^5$  cells/ml overnight prior to drug treatments. On the following day, CKI (to give a final concentration of 2 mg/ml) and fractionated mixtures (equivalent dilutions of CKI) were added. Total RNA was isolated by using an RNA extraction kit (Thermo Fisher Scientific) according to the manufacturer's instructions and RNA samples were quantified and quality determined using a Bioanalyzer at the Cancer Genome Facility of the Australian Cancer Research Foundation (SA, Australia). RNA samples with RNA integrity number (RINs)  $> 7.0$  were sent to be sequenced at Novogene (China). Briefly, after QC procedures were performed, mRNA was isolated using oligo(dT) beads and randomly fragmented by adding fragmentation buffer, followed by cDNA synthesis primed with random hexamers. Next, a custom second-strand synthesis buffer (Illumina), dNTPs, RNase H and DNA polymerase I were added for second-strand synthesis. After end repair, barcode ligation and sequencing adaptor ligation, the double-stranded cDNA library was size selected and PCR amplified. Sequencing was carried out on an Illumina HiSeq X platform with paired-end 150 bp reads.

### **Transcriptome data processing**

FastQC (v0.11.4, Babraham Bioinformatics) was used to check the quality of raw reads before proceeding with downstream analysis. Trim\_galore (v0.3.7, Babraham Bioinformatics) with the parameters: `--stringency 5 --paired --fastqc_args` was used to trim adaptors and low-quality sequences. STAR (v2.5.3a) was then applied to align the trimmed reads to the reference genome (hg19, UCSC) with the parameters: `--outSAMstrandField intronMotif --outSAMattributes All --outFilterMismatchNmax 10 --seedSearchStartLmax 30 --outSAMtype BAM SortedByCoordinate` ([Dobin et al., 2013](#)). Then, subread (v1.5.2) was used to generate read counts data with the following parameters `featureCounts -p -t exon -g gene_id` ([Liao et al., 2013](#)). Significantly differentially expressed genes between treatments were analysed and selected using edgeR (v3.22.3) with false discovery rate (FDR)  $< 0.05$  ([Robinson et al., 2010](#)). Removal of unwanted variance (RUVs) package in R was applied to two different batches of transcriptome datasets to eliminate batch variance ([Risso et al., 2014](#)). APE was used to cluster the treatments ([Paradis et al., 2004](#)) followed by RUVs. Gene Ontology (GO) overrepresentation analyses were performed using clusterProfiler with the parameter `ont =`



"BP"(Biological Process), pAdjustMethod = "BH", pvalueCutoff = 0.01, and qvalueCutoff = 0.05 (Yu et al., 2012). Signalling Pathway Impact Analysis (SPIA) was carried out to identify the commonly perturbed pathways within the treatments using the SPIA R package (Tarca et al., 2008). Significantly perturbed pathways were visualized using Pathview package in R (Luo and Brouwer, 2013).

### Statistical analysis

Statistical analyses were carried out using GraphPad Prism 8.0 (GraphPad Software Inc., CA, USA). Student's t-test or ANOVA (one-way or two-way) was used when there were two or three groups to compare respectively. Post hoc "Bonferroni's multiple comparisons test" was performed when ANOVA results were significant. Statistically significant results were represented as  $p < 0.05$  (\*) or  $p < 0.01$  (\*\*),  $p < 0.001$  (\*\*\*), or  $p < 0.0001$  (\*\*\*\*); ns (not significant). All data were shown as mean  $\pm$  standard deviation (SD).

### FIGURE LEGENDS

**Figure 1:** Fractionation of Compound Kushen Injection. (A) Diagram illustrating the process of subtractive fractionation, reconstitution, and screening of fractionated compounds using three cell-based assays. (B) HPLC profile of the 9 purified and reconstituted major peaks (MJ) demonstrating nine major compounds. (C) HPLC profile of reconstituted fractions not containing the 9 major compounds (MN) showing the remaining peaks with no remaining major compounds.

**Figure 2:** XTT Cell viability assays of subtractive fractions in MDA-MB-231 cells at 24- and 48-hour time points treated with 2mg/mL of CKI and 2mg/mL equivalent concentrations of all other treating agents. (A) Suppression of cell viability from the following fractions: UT (untreated), MJ, MN, MJ+MN (combination of MJ and MN) and Syn\_CKI (synthetic CKI generated using nine major compounds). (B) Effect of 9 (N-1) subtractive fractions compared to CKI. (C) Effect of N-OmtOspc subtractive fraction and OmtOspc compared to CKI. (D) Effect of N-MacOmtOspc subtractive fraction and MacOmtOspc, compared to CKI. Statistically significant results relative to CKI treatment shown as  $p < 0.05$  (\*) or ns (not significant), all data were shown as mean  $\pm$  standard deviation (SD).

**Figure 3:** Cell cycle assay of subtractive fractions at 24- and 48-hour time point treatments. (A) Effect of 9 (N-1) subtractive fractions on cell cycle in MDA-MB-231 cells as determined by FACS PI cell cycle staining assay. (B) Effect of N-OmtOspc subtractive fraction and

OmtOspc on cell cycle in MDA-MB-231 cells as determined by FACS PI cell cycle staining assay. (C) Effect of N-MacOmtOspc subtractive fraction and MOO on cell cycle in MDA-MB-231 cells as determined by FACS PI cell cycle staining assay. (D) The representative histograms of cell cycle analysis by the treatments as compared to UT. Statistically significant results shown as  $p < 0.05$  (\*) or  $p < 0.01$  (\*\*),  $p < 0.001$  (\*\*\*), or  $p < 0.0001$  (\*\*\*\*). All data were shown as mean  $\pm$  standard deviation (SD).

**Figure 4:** Apoptosis and cytotoxicity assays of subtractive fractions at 24- and 48-hour time point treatments. Apoptotic effect of N-OmtOspc, N-MacOmtOspc subtractive fractions, OmtOspc and MacOmtOspc in (A) MDA-MB-231 cells, (B) HEK-293 cells, and (C) HFF cells as determined by FACS AnnexinV/PI assay. (D) Cytotoxic effect of CKI, N-OmtOspc and N-MacOmtOspc subtractive fractions was determined using Alamar Blue cytotoxicity assay. Statistically significant results shown as  $p < 0.05$  (\*) or  $p < 0.01$  (\*\*),  $p < 0.001$  (\*\*\*), or  $p < 0.0001$  (\*\*\*\*); ns (not significant). All data were shown as mean  $\pm$  standard deviation (SD).

**Figure 5:** Gene expression clustering and summary of differential gene expression (A) Clustering of treated samples based on gene expression, calculated as transcripts per million (TPM) using Ward's hierarchical cluster analysis (Ward.D2) method. Number of differentially expressed (DE) genes (FDR  $< 0.05$  according to edgeR) associated with each treatment was calculated using pair-wise comparison at (B) 24 hours and (C) 48 hours time point. Treatments were compared column versus row. Up-regulated genes are shown in shades of red and down-regulated genes are shown in shades of blue.

**Figure 6:** Over-representation analysis of GO functional annotation and KEGG pathway perturbation analysis. (A) Over-represented GO terms (Biological Process, BP=3) for DE genes identified from comparison of subtractive fraction treated cells against CKI treatment in order to show the relative change from depleting compounds. Gene ratio of each term calculated from Cluster Profiler was plotted based on the adjusted  $P$ -values. Top 5 most significant categories of GO terms were plotted by default. Colour gradient of adjusted  $P$ -values ranging from red to blue in order of increasing  $p$ -values (high to low significance). Number of identified genes in each treatment (numbers in parentheses) were shown in the bottom and the sizes of the dots correspond to the ratio of genes out of all significant DE gene from each treatment involved in the particular terms. (B) Identification of significantly perturbed pathways using SPIA ( $p < 0.05$ ) analysis. Eighty-six significantly perturbed pathways from twenty-two comparisons were found (Supplementary Figure10). Only 15

pathways most obviously linked to our phenotypes of cell viability, cell cycle and apoptosis were shown here. Positive (overall increase in gene expression for pathway) and negative (overall decrease in gene expression for pathway) perturbation accumulation values of the pathways were shown in red and blue respectively. Mean perturbation values of each pathway were shown in bar plot.

**Figure 7:** Differential gene expression profiles of all treatments for Cytokine-Cytokine Receptor pathway. The left panel shows comparison of subtractive fraction treated cells against CKI treatment and the right panel shows comparison of single compound subtractive fraction treated cells against the treatments for two and three compound subtractive fractions. Asterisks in green shows a subset of genes that had opposite changes in gene expression across treatments.

**Figure 8:** Differential gene expression profiles of all treatments for Cell Cycle pathway. The left panel shows comparison of subtractive fraction treated cells against CKI treatment and the right panel shows comparison of single compound subtractive fraction treated cells against the treatments for two and three compound subtractive fractions. Asterisks in green shows a subset of genes that have opposite changes in gene expression across treatments.

**Supplementary Figure 1:** HPLC profiles of 25 mixtures including CKI, MJ, MN, 9 (N-1), 4 (N-3) and 9 (N-2). 50  $\mu$ l of the samples at 1 mg/ml concentration was injected through the semi-preparative column to achieve the profiles.

**Supplementary Figure 2:** XTT Cell viability assays of subtractive fractions in MDA-MB-231 cells at 24- and 48-hour time points treated with 1 mg/ml or 2 mg/ml of CKI and 2 mg/ml equivalent concentrations of all other treating agents. (A) Suppression of cell viability from the following treatments: CKI, WRCKI-B (whole reconstituted CKI in buffer/vehicle control), and WRCKI-H (whole reconstituted CKI in milliQ H<sub>2</sub>O) (B) assessment of the interaction effects of single MJ compounds by the addition to the MN subtractive fraction. Single major compounds were dissolved in either MilliQ H<sub>2</sub>O or DMSO (Dimethyl sulfoxide). Effect of subtractive fractions (C) N-OmtOspcSpc, (D) N-MacAdeTri, (E) N-MtNmeSpr, (F) N-MtNme, (G) N-OmtSpc, (H) N-OspcSpc, (I) N-MacTri, (J) N-AdeTri, (K) N-MacAde, (L) N-MtSpr, (M) NmeSpr. Statistically significant results shown as  $p < 0.05$  (\*) or  $p < 0.01$  (\*\*)  $p < 0.001$  (\*\*\*), or  $p < 0.0001$  (\*\*\*\*); ns (not significant). All data were shown as mean  $\pm$  standard deviation (SD).

**Supplementary Figure 3:** XTT Cell viability assays of subtractive fractions N-OmtOspc and N-MacOmtOspc in MDA-MB-231 cells at 24- and 48-hour time points treated 4 different concentrations ranging from 0.25mg/mL to 2mg/mL of CKI and equivalent concentrations of two other agents. Statistically significant results shown as  $p < 0.05$  (\*) or  $p < 0.01$  (\*\*)  $p < 0.001$  (\*\*\*), or  $p < 0.0001$  (\*\*\*\*); ns (not significant). All data were shown as mean  $\pm$  standard deviation (SD).

**Supplementary Figure 4:** Representative plots of Annexin V and PI staining in (A) MDA-MB-231, (B) HEK-293, and (C) HFF.

**Supplementary Figure 5:** Multiple dimensional scaling (MDS) plot for samples based on expression profiles of all genes before the removal of unwanted variance (RUVs in R package).

**Supplementary Figure 6:** Multiple dimensional scaling (MDS) plot for samples based on expression profiles of all genes after the application of RUVs.

**Supplementary Figure 7:** Box plot for samples based on expression profiles of all genes before the application of RUVs.

**Supplementary Figure 8:** Box plot for samples based on expression profiles of all genes after the application of RUVs.

**Supplementary Figure 9:** Venn diagrams showing the number of overlapping DE genes between treatments at 24-hours and 48-hours.

**Supplementary Figure 10:** Identification of significantly perturbed pathways using SPIA ( $p < 0.05$ ) analysis. Eighty-six significantly perturbed pathways from twenty-two comparisons were found.

**Supplementary Figure 11:** DE genes from the following comparisons (CKI vs UT, CKI vs N-Mac, CKI vs N-Nme, CKI vs N-Omt and CKI vs N-Tri) shown in the "Cytokine-Cytokine Receptor Interaction pathway at 48-hours. Significant up- and down-regulated DE genes were coloured red and green respectively. Each coloured box was separated into five parts according

to treatments in this order: CKI vs UT, CKI vs N-Mac, CKI vs N-Nme, CKI vs N-Omt and CKI vs N-Tri. White or grey colours represented gene(s) that were not significantly differentially expressed by the treatments.

**Supplementary Figure 12:** DE genes from the following comparisons (CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc) shown in the Cytokine-Cytokine Receptor Interaction pathway at 48-hours. Significant up- and down-regulated DE genes were coloured red and green respectively. Each coloured box was separated into five parts according to this order: CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc. White or grey colours represented gene(s) that were not significantly differentially expressed by the treatments.

**Supplementary Figure 13:** DE genes from the following comparisons (CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc) shown in the Cytokine-Cytokine Receptor Interaction pathway at 24-hours. Significant up- and down-regulated DE genes were coloured red and green respectively. Each coloured box was separated into five parts according to this order: CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc. White or grey colours represented gene(s) that were not significantly differentially expressed by the treatments.

**Supplementary Figure 14:** DE genes from the following comparisons (CKI vs UT, CKI vs N-Mac, CKI vs N-Nme, CKI vs N-Omt and CKI vs N-Tri) shown in the Cell Cycle pathway at 48-hours. Significant up- and down-regulated DE genes were coloured with red and green respectively. Each coloured box was separated into five parts according to this order: CKI vs UT, CKI vs N-Mac, CKI vs N-Nme, CKI vs N-Omt and CKI vs N-Tri. White or grey colours represented gene(s) that were not significantly differentially expressed by the treatments.

**Supplementary Figure 15:** DE genes from the following comparisons (CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc) shown in the Cell Cycle pathway at 48-hour treatments. Significant up- and down-regulated DE genes were coloured red and green respectively. Each coloured box was separated into five parts according to this order: CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc. White or grey colours represented gene(s) that were not significantly differentially expressed by the treatments.

**Supplementary Figure 16:** DE genes from the following comparisons (CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc) shown in the Cell Cycle pathway at 24-hours. Significant up- and down-regulated DE genes were coloured red and green respectively. Each coloured box was separated into five parts according to this order: CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc. White or grey colours represented gene(s) that were not significantly differentially expressed by the treatments.

**Supplementary Figure 17:** Differential gene expression profiles of all treatments for TGF- $\beta$  signalling pathway: the left panel shows comparison of subtractive fraction treated cells against CKI treatment and the right panel shows comparison of single compound subtractive fraction treated cells against the treatments for two and three compound subtractive fractions.

**Supplementary Figure 18:** DE genes from the following comparisons (CKI vs UT, CKI vs N-Mac, CKI vs N-Nme, CKI vs N-Omt and CKI vs N-Tri) shown in the TGF- $\beta$  signalling pathway at 48-hours. Significant up- and down-regulated DE genes were coloured red and green respectively. Each coloured box was separated into five parts according to this order: CKI vs UT, CKI vs N-Mac, CKI vs N-Nme, CKI vs N-Omt and CKI vs N-Tri. White or grey colours represented gene(s) that were not significantly differentially expressed by the treatments.

**Supplementary Figure 19:** DE genes from the following comparisons (CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc) shown in the TGF- $\beta$  signalling pathway at 48-hour treatments. Significant up- and down-regulated DE genes were coloured red and green respectively. Each coloured box was separated into five parts according to this order: CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc. White or grey colours represented gene(s) that were not significantly differentially expressed by the treatments.

**Supplementary Figure 20:** DE genes from the following comparisons (CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc) shown in the TGF- $\beta$  signalling pathway at 24-hours. Significant up- and down-regulated DE genes were coloured red and green respectively. Each coloured box was separated into five parts according to this order: CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and

CKI vs MacOmtOspc. White or grey colours represented gene(s) that were not significantly differentially expressed by the treatments.

#### **ACKNOWLEDGEMENTS**

This work was supported by the special international corporation project of traditional Chinese medicine (GZYYGJ2017035) and The University of Adelaide, Zendong Australia China Centre for Molecular Chinese Medicine. The authors would like to thank Dr. Denis Scanlon and Associate Prof. Stephen Bell for the assistance with HPLC usage and Jue Zeng for her valuable discussions.

#### **Author contributions**

T.N.A, S.N, Z.Q and D.L.A designed the study, analysed the data and wrote the manuscript. T.N.A and S.N conducted the experiments. Y.H-L, J.C, H.S, T.P, and H.D assisted the experiments and analysed the data. A.J.Y and D.K assisted in writing the manuscript.

#### **Conflict of interest**

The authors declare no conflicts of interest.

#### **Declaration of transparency and scientific rigour**

This declaration acknowledges that this paper adheres to the principles for transparent reporting and scientific rigour of preclinical research recommended by funding agencies, publishers and other organisations engaged with supporting research.

**Table 1:** Summarized results of HPLC fractionation and treatments using three cell-based assays. N represents the number of all compounds contained in CKI, Mac = macrozamin, Ade = adenine, Tri = trifolirhizin, Nme = N-methylcytisine, Spr = Sophoridine, Mt = Matrine, Omt = Oxymatrine, Spc = Sophocarpine, and Ospc = Oxysophocarpine. Significant results of CKI treatment were calculated based on UT whereas those of other treatments were calculated based on CKI treatments. Statistically significant results were represented as  $p < 0.05$  (\*) or  $p < 0.01$  (\*\*),  $p < 0.001$  (\*\*\*), or  $p < 0.0001$  (\*\*\*\*).

HPLC Fractionation	Treatments	Proliferation Assay (MDA-MB-213)	Cell Cycle Assay (MDA-MB-213)	Apoptosis Assay in three cell lines		
				MDA-MB-231	HEK-293	HFF
9 known + small unknown (N)	Original CKI	*	*			
9 known major compounds	MJ					
Small unknown minor compounds	MN	*				
N-1	N-Mac					
	N-Ade					
	N-Tri					
	N-Nme					
	N-Spr					
	N-Mt					
	N-Omt					
	N-Ospc					
	N-Spc					
N-3	N-MacAdeTri					
	N-MtNmeSpr					
	N-OmtOspcSpc					
	N-MacOmtOspc	*	****	****	****	**
N-2	N-MacAde					
	N-MacTri					
	N-AdeTri					
	N-MtNme					
	N-MtSpr					
	N-NmeSpr					
	N-OmtOspc	*	****	****	****	**
	N-OmtSpc					
3 compounds	N-OspcSpc					
	MacAdeTri					
	MtNmeSpr					
	OmtOspcSpc					
2 compounds	MacOmtOspc					
	MacAde					
	MacTri					
	AdeTri					
	MtNme					
	MtSpr					
	NmeSpr					
	OmtOspc					
OmtSpc						
OspcSpc						



**Table 2:** Concentration of 9 major compounds in CKI (N) (Batch No:20170322), remaining major compounds in N-OmtOspc, and remaining major compounds in N-MacOmtOspc.

Mixtures	Compounds	Regression Line	Regression coefficient	concentration (mg/mL)(n=2)	% Contribution
CKI (N)	Macrozamin	$y = 6E-05x + 6E-05$	0.996	1.1 ± 0.03	4.4
	Adenine	$y = 0.021x + 0.0074$	0.994	0.09 ± 0.09	0.4
	N-methylcytisine	$y = 0.0937x + 0.042$	0.9895	0.17 ± 0.02	0.7
	Sophoridine	$y = 0.1443x + 0.2679$	0.987	0.4 ± 0.08	1.6
	Matrine	$y = 0.0132x + 0.7512$	0.993	1.26 ± 0.06	5.04
	Sophocarpine	$y = 0.0343x + 0.3974$	0.994	0.54 ± 0.02	2.2
	Oxysophocarpine	$y = 0.0371x - 0.0108$	0.999	1.1 ± 0.1	4.4
	Oxymatrine	$y = 0.0132x + 0.6226$	0.992	6.1 ± 0.09	24.4
	Trifolirhizin	$y = 0.0026x - 0.0002$	0.990	0.08 ± 0.002	0.3
	<b>Total</b>			<b>10.8</b>	<b>43.4</b>
N-OmtOspc	Macrozamin	$y = 6E-05x + 6E-05$	0.996	1.1 ± 0.04	4.4
	Adenine	$y = 0.021x + 0.0074$	0.994	0.3 ± 0.5	1.2
	N-methylcytisine	$y = 0.0937x + 0.042$	0.9895	0.03 ± 0.03	0.1
	Sophoridine	$y = 0.1443x + 0.2679$	0.987	0.3 ± 0.07	1.2
	Matrine	$y = 0.0132x + 0.7512$	0.993	1.7 ± 0.1	6.8
	Sophocarpine	$y = 0.0343x + 0.3974$	0.994	0.3 ± 0.01	1.2
	Trifolirhizin	$y = 0.0026x - 0.0002$	0.990	0.07 ± 0.01	0.3
	<b>Total</b>			<b>3.8</b>	<b>15.2</b>
N-MacOmtOspc	Adenine	$y = 0.021x + 0.0074$	0.994	0.06 ± 0.09	0.2
	N-methylcytisine	$y = 0.0937x + 0.042$	0.9895	0.02 ± 0.01	0.1
	Sophoridine	$y = 0.1443x + 0.2679$	0.987	0.04 ± 0.02	0.2
	Matrine	$y = 0.0132x + 0.7512$	0.993	1.7 ± 0.06	6.8
	Sophocarpine	$y = 0.0343x + 0.3974$	0.994	0.2 ± 0.02	0.8
	Trifolirhizin	$y = 0.0026x - 0.0002$	0.990	0.08 ± 0.003	0.3
	<b>Total</b>			<b>2.1</b>	<b>8.4</b>

\*Total alkaloid content in CKI = 26.5mg/ml based on manufacturer's assay.

**Table 3:** Summary of shared, differentially expressed (DE) genes across treatments. Similarity (%) calculated from total number of shared DE genes from all listed comparisons. To find the number of DE genes, CKI treatment was used as a baseline to compare all other fractionated treatments in order to emphasize the effect of depleted compounds and UT (untreated) was used as a base to calculate the DE genes for CKI treatment.

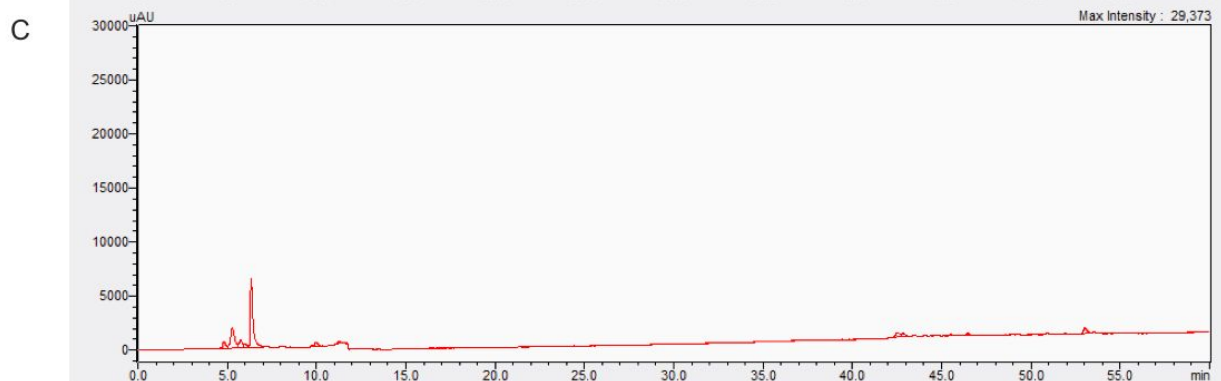
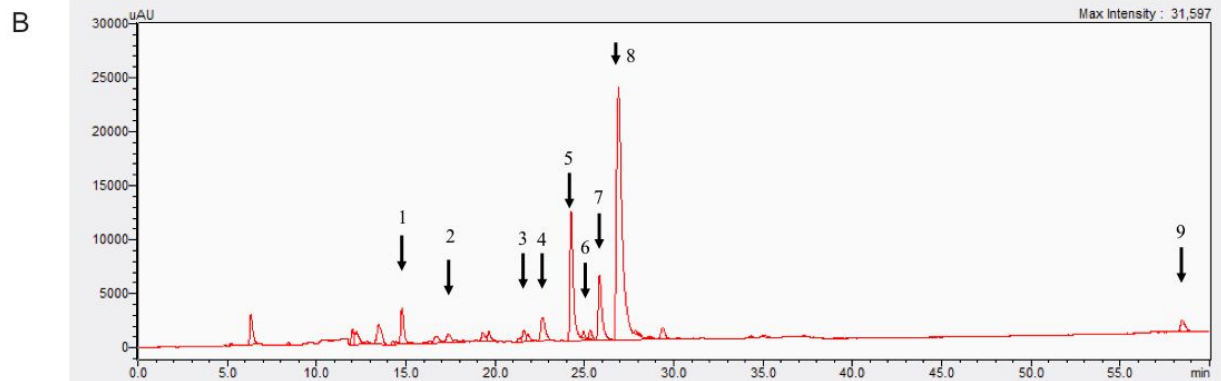
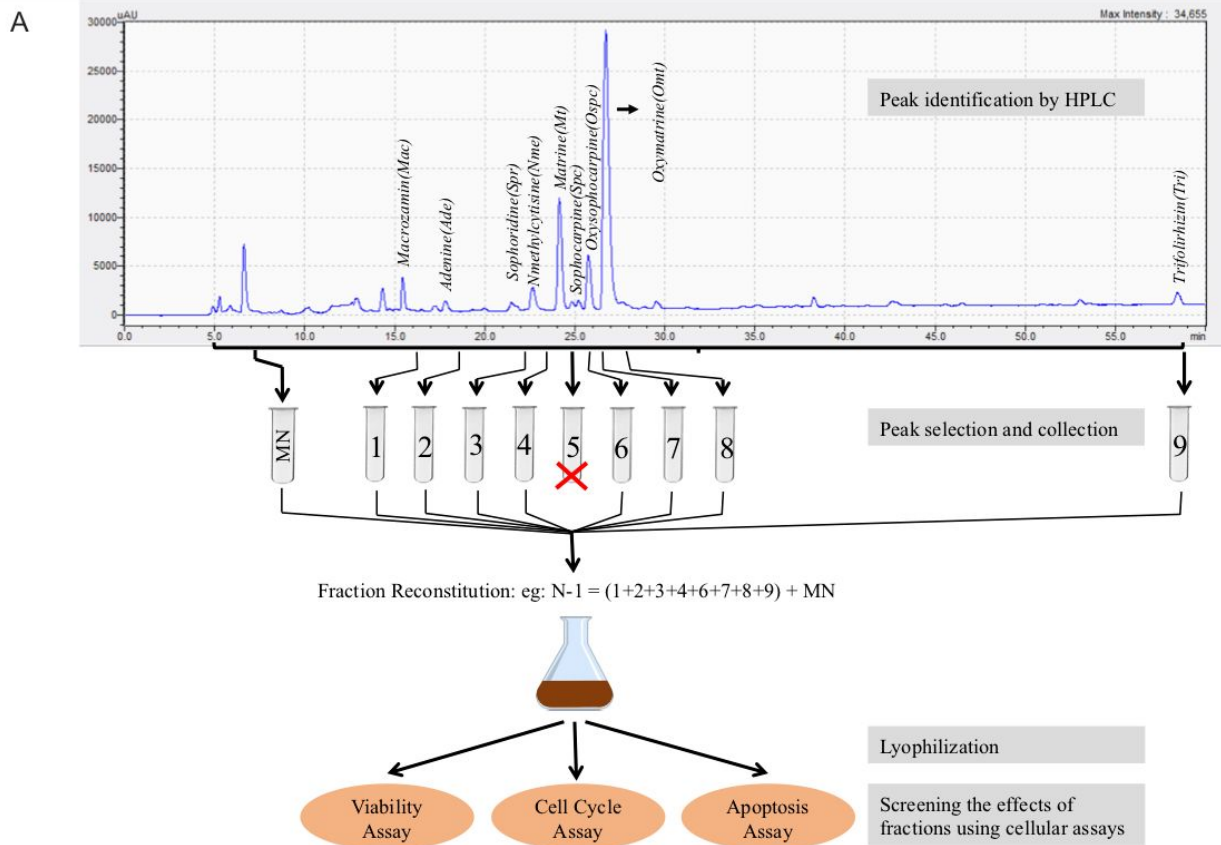
Treatments Comparisons	Similarity (%)	Time course (hours)
4 x (N-1)	71.3	48
4 x (N-1) and CKI	42.7	48
4 x (N-1), CKI and N-OmtOspc	30.4	48
4 x (N-1), CKI, N-OmtOspc and N-MacOmtOspc	24.6	48
4 x (N-1), CKI, OmtOspc, and MacOmtOspc	24.6	48
4 x (N-1), OmtOspc, and MacOmtOspc	33.7	48
4 x (N-1), N-OmtOspc and N-MacOmtOspc	37.7	48
CKI and N-OmtOspc	64.4	48
CKI and N-MacOmtOspc	63.9	48
CKI, N-OmtOspc and N-MacOmtOspc	50.1	48
CKI, N-OmtOspc, N-MacOmtOspc, OmtOspc, and MacOmtOspc	30.2	48
UT, OmtOspc, and MacOmtOspc	54.9	48
CKI and N-OmtOspc	56.0	24
CKI and N-MacOmtOspc	45.6	24
CKI, N-OmtOspc and N-MacOmtOspc	31.6	24
CKI, N-OmtOspc, N-MacOmtOspc, OmtOspc, and MacOmtOspc	13.3	24
UT, OmtOspc, and MacOmtOspc	39.1	24

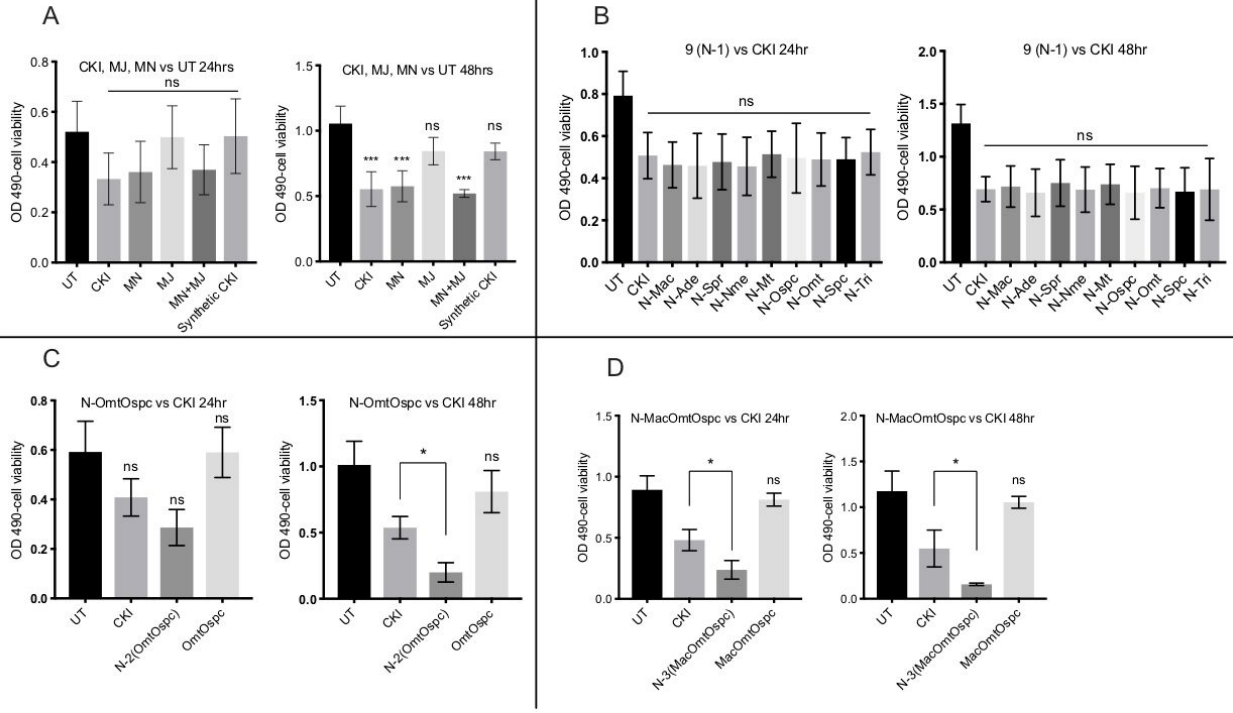
## Reference:

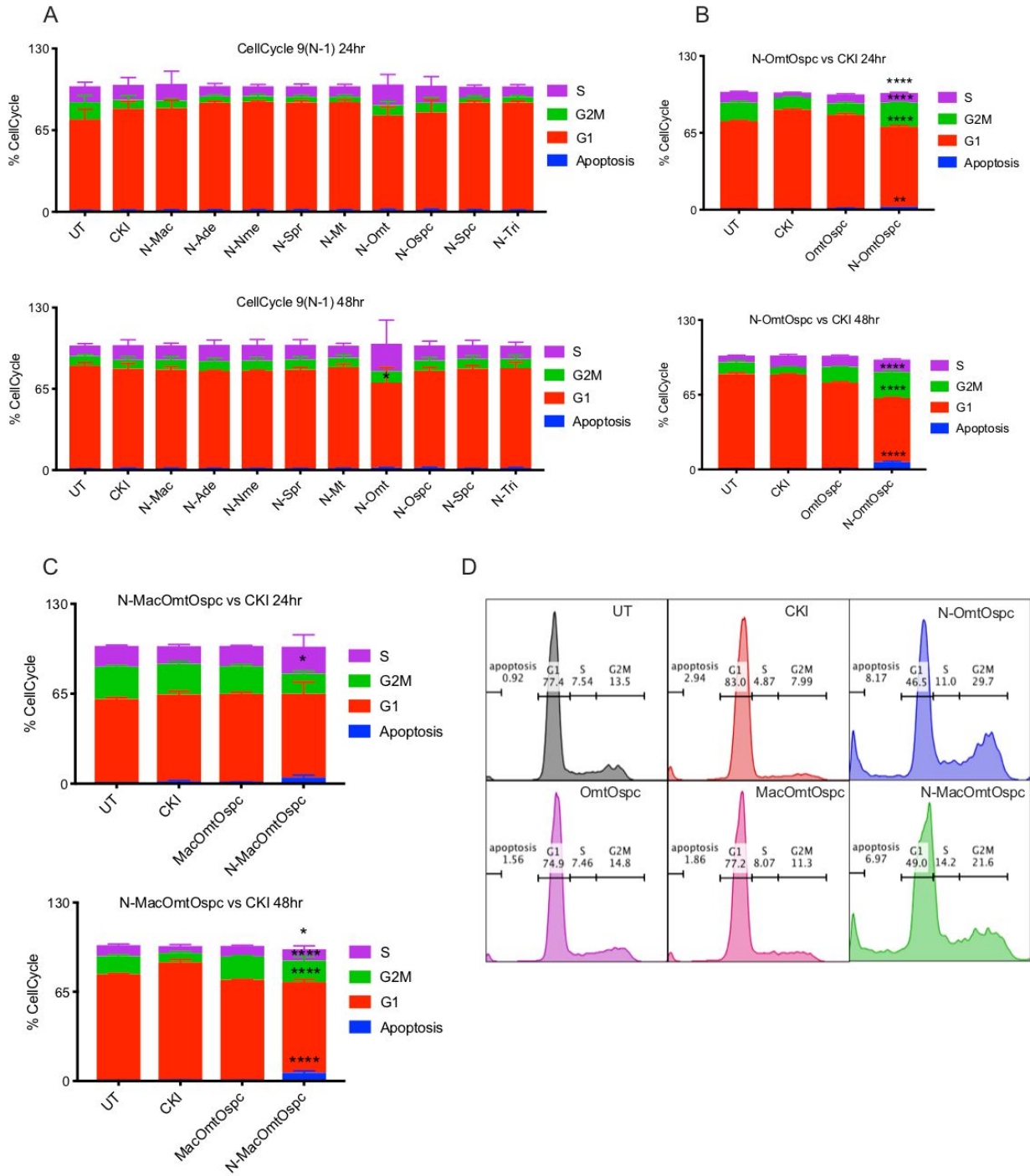
- Aung, T.N., Qu, Z., Kortschak, R.D., and Adelson, D.L. (2017). Understanding the effectiveness of natural compound mixtures in cancer through their molecular mode of action. *International journal of molecular sciences* 18, 656.
- Cai, Y., Xu, P., Yang, L., Xu, K., Zhu, J., Wu, X., Jiang, C., Yuan, Q., Wang, B., and Li, Y. (2016). HMGB1-mediated autophagy decreases sensitivity to oxymatrine in SW982 human synovial sarcoma cells. *Scientific reports* 6, 37845.
- Cheng, H., Xia, B., Zhang, L., Zhou, F., Zhang, Y.-X., Ye, M., Hu, Z.-G., Li, J., Li, J., and Wang, Z.-L. (2006). Matrine improves 2, 4, 6-trinitrobenzene sulfonic acid-induced colitis in mice. *Pharmacological research* 53, 202-208.
- Cui, J., Qu, Z., Harata-Lee, Y., Aung, T.N., Shen, H., Wang, W., and Adelson, D.L. (2018). Cell Cycle, Energy Metabolism and DNA Repair Pathways in Cancer Cells are Suppressed by Compound Kushen Injection. *bioRxiv*, 348102.
- Dai, Z., Wang, L., Wang, X., Zhao, B., Zhao, W., Bhardwaj, S.S., Ye, J., Yin, Z., Zhang, J., and Zhao, S. (2018). Oxymatrine induces cell cycle arrest and apoptosis and suppresses the invasion of human glioblastoma cells through the EGFR/PI3K/Akt/mTOR signaling pathway and STAT3. *Oncology reports* 40, 867-876.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15-21.
- Dong, Y., Xi, H., Yu, Y., Wang, Q., Jiang, K., and Li, L. (2002). Effects of oxymatrine on the serum levels of T helper cell 1 and 2 cytokines and the expression of the S gene in hepatitis B virus S gene transgenic mice: A study on the anti-hepatitis B virus mechanism of oxymatrine. *Journal of gastroenterology and hepatology* 17, 1299-1306.
- Gao, L., Wang, K.-X., Zhou, Y.-Z., Fang, J.-S., Qin, X.-M., and Du, G.-H. (2018). Uncovering the anticancer mechanism of Compound Kushen Injection against HCC by integrating quantitative analysis, network analysis and experimental validation. *Scientific reports* 8, 624.
- Harvey, A.L., Edrada-Ebel, R., and Quinn, R.J. (2015). The re-emergence of natural products for drug discovery in the genomics era. *Nature Reviews Drug Discovery* 14, 111.
- Jiang, J., Wu, F., Lu, J., Lu, Z., and Xu, Q. (1997). Anti-inflammatory activity of the aqueous extract from rhizoma smilacis glabrae. *Pharmacological research* 36, 309-314.
- Jin, Y., Yang, Q., Liang, L., Ding, L., Liang, Y., Zhang, D., Wu, B., Yang, T., Liu, H., and Huang, T. (2018). Compound kushen injection suppresses human acute myeloid leukaemia by regulating the Prdxs/ROS/Trx1 signalling pathway. *Journal of Experimental & Clinical Cancer Research* 37, 277.
- Li, H., Li, X., Bai, M., Suo, Y., Zhang, G., and Cao, X. (2015). Matrine inhibited proliferation and increased apoptosis in human breast cancer MCF-7 cells via upregulation of Bax and downregulation of Bcl-2. *International journal of clinical and experimental pathology* 8, 14793.
- Liao, Y., Smyth, G.K., and Shi, W. (2013). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923-930.
- Liu, W., Shi, J., Zhu, L., Dong, L., Luo, F., Zhao, M., Wang, Y., Hu, M., Lu, L., and Liu, Z. (2015). Reductive metabolism of oxymatrine is catalyzed by microsomal CYP3A4. *Drug design, development and therapy* 9, 5771.
- Luo, W., and Brouwer, C. (2013). Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics* 29, 1830-1831.

- Ma, Y., Gao, H., Liu, J., Chen, L., Zhang, Q., and Wang, Z. (2014). Identification and determination of the chemical constituents in a herbal preparation, Compound Kushen injection, by HPLC and LC-DAD-MS/MS. *Journal of Liquid Chromatography & Related Technologies* 37, 207-220.
- Paradis, E., Claude, J., and Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20, 289-290.
- Qu, Z., Cui, J., Harata-Lee, Y., Aung, T.N., Feng, Q., Raison, J.M., Kortschak, R.D., and Adelson, D.L. (2016). Identification of candidate anti-cancer molecular mechanisms of compound kushen injection using functional genomics. *Oncotarget* 7, 66003.
- Riccardi, C., and Nicoletti, I. (2006). Analysis of apoptosis by propidium iodide staining and flow cytometry. *Nature protocols* 1, 1458.
- Risso, D., Ngai, J., Speed, T.P., and Dudoit, S. (2014). Normalization of RNA-seq data using factor analysis of control genes or samples. *Nature biotechnology* 32, 896.
- Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139-140.
- Tarca, A.L., Draghici, S., Khatri, P., Hassan, S.S., Mittal, P., Kim, J.-S., Kim, C.J., Kusanovic, J.P., and Romero, R. (2008). A novel signaling pathway impact analysis. *Bioinformatics* 25, 75-82.
- Tu, H., Lei, B., Meng, S., Liu, H., Wei, Y., He, A., Zhang, W., and Zhou, F. (2016). Efficacy of compound kushen injection in combination with induction chemotherapy for treating adult patients newly diagnosed with acute leukemia. *Evidence-Based Complementary and Alternative Medicine* 2016.
- Wang, X., and Yang, R. (2003). Movement disorders possibly induced by traditional Chinese herbs. *European neurology* 50, 153-159.
- Wu, C., Huang, W., Guo, Y., Xia, P., Sun, X., Pan, X., and Hu, W. (2015). Oxymatrine inhibits the proliferation of prostate cancer cells in vitro and in vivo. *Molecular medicine reports* 11, 4129-4134.
- Wu, L., Wang, G., Liu, S., Wei, J., Zhang, S., Li, M., Zhou, G., and Wang, L. (2016). Synthesis and biological evaluation of matrine derivatives containing benzo- $\alpha$ -pyrone structure as potent anti-lung cancer agents. *Scientific reports* 6, 35918.
- Xu, G., Yao, L., Rao, S., Gong, Z., Zhang, S., and Yu, S. (2005). Attenuation of acute lung injury in mice by oxymatrine is associated with inhibition of phosphorylated p38 mitogen-activated protein kinase. *Journal of ethnopharmacology* 98, 177-183.
- Xu, W., Lin, H., Zhang, Y., Chen, X., Hua, B., Hou, W., Qi, X., Pei, Y., Zhu, X., and Zhao, Z. (2011). Compound Kushen Injection suppresses human breast cancer stem-like cells by down-regulating the canonical Wnt/b-catenin pathway. *J Exp Clin Cancer Res* 30, 103.
- Ying, X.-J., Jin, B., Chen, X.-W., Xie, J., Xu, H.-M., and Dong, P. (2015). Oxymatrine downregulates HPV16E7 expression and inhibits cell proliferation in laryngeal squamous cell carcinoma Hep-2 cells in vitro. *BioMed research international* 2015.
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: a journal of integrative biology* 16, 284-287.
- Yu, P., Liu, Q., Liu, K., Yagasaki, K., Wu, E., and Zhang, G. (2009). Matrine suppresses breast cancer cell proliferation and invasion via VEGF-Akt-NF- $\kappa$ B signaling. *Cytotechnology* 59, 219-229.
- Yuan, F., Chen, J., Wu, W.J., Chen, S.Z., Wang, X.D., Su, Z., and Huang, M. (2010). Effects of matrine and oxymatrine on catalytic activity of cytochrome P450s in rats. *Basic & clinical pharmacology & toxicology* 107, 906-913.

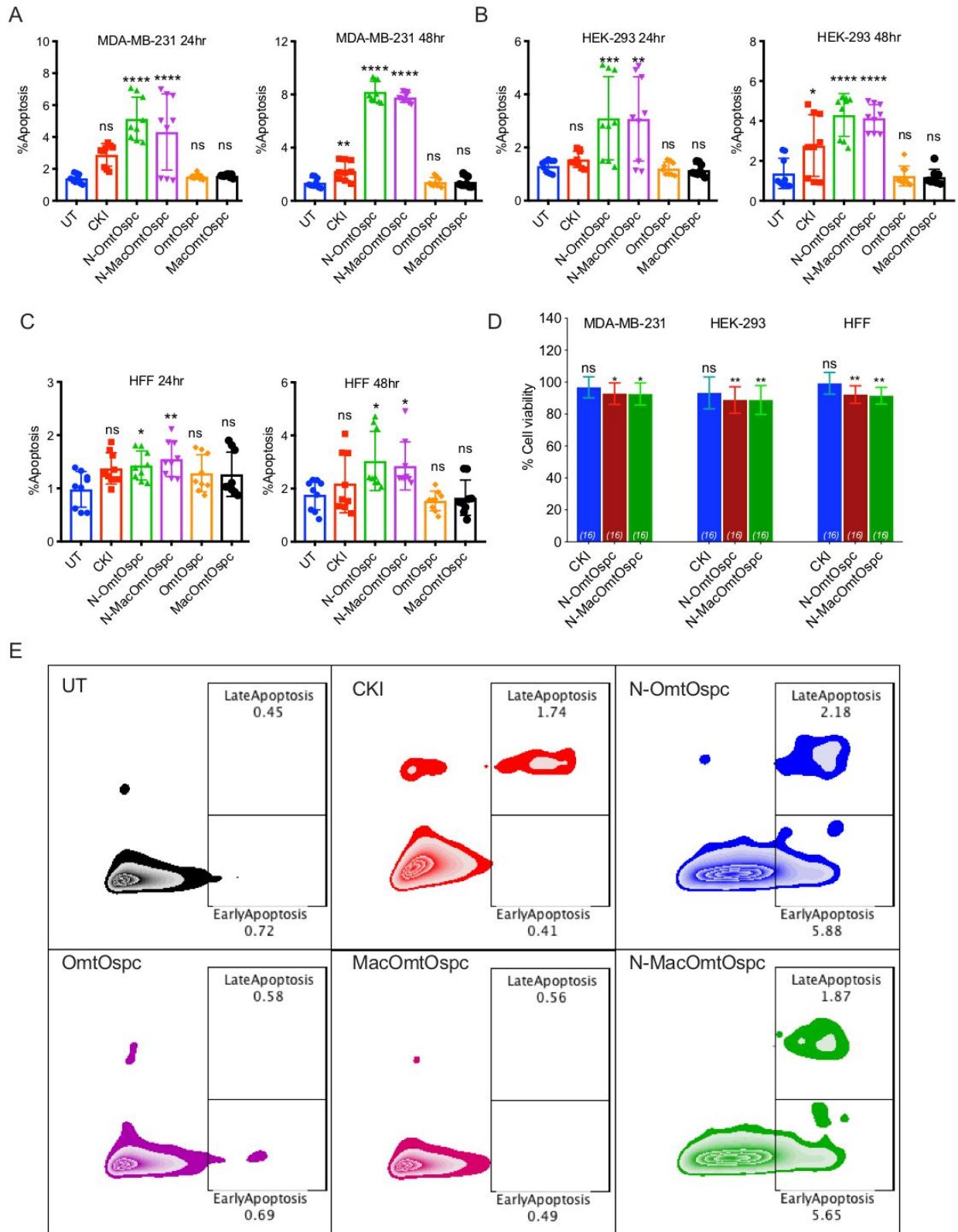
- Zhang, L., Zheng, Y., Deng, H., Liang, L., and Peng, J. (2014). Aloperine induces G2/M phase cell cycle arrest and apoptosis in HCT116 human colon cancer cells. *International journal of molecular medicine* 33, 1613-1620.
- Zhou, Y.J., Guo, Y.J., Yang, X.L., and Ou, Z.L. (2018). Anti-Cervical Cancer Role of Matrine, Oxymatrine and Sophora Flavescens Alkaloid Gels and its Mechanism. *Journal of Cancer* 9, 1357.



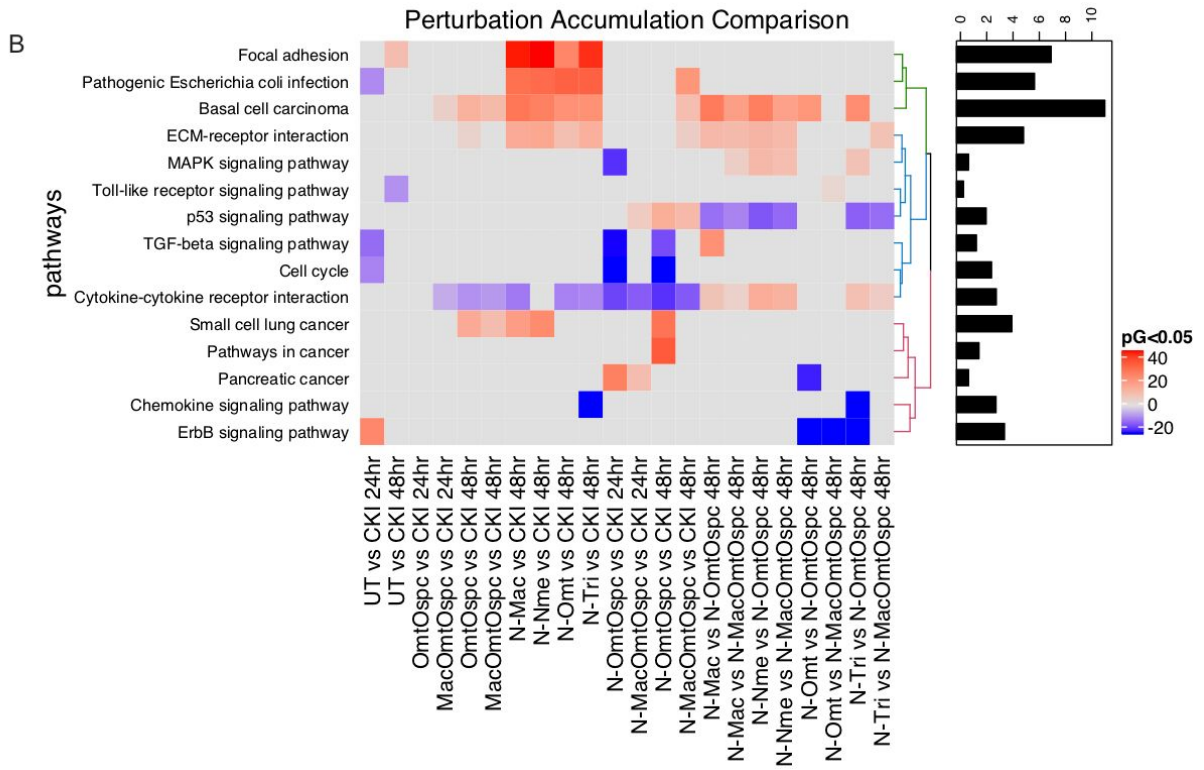
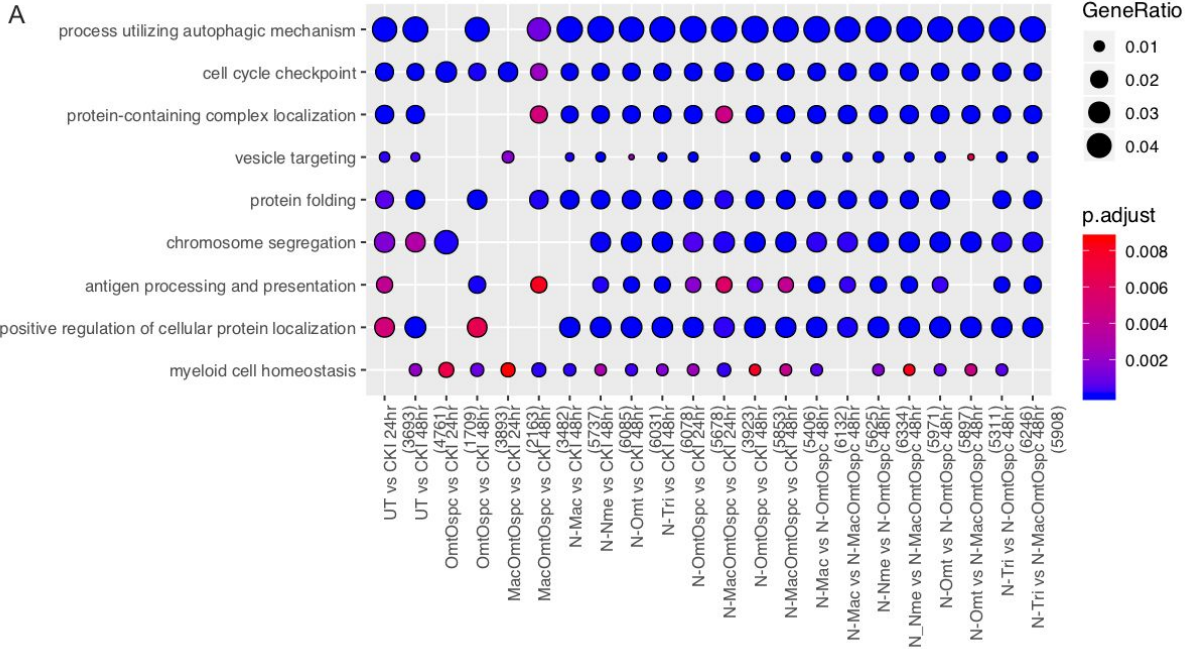


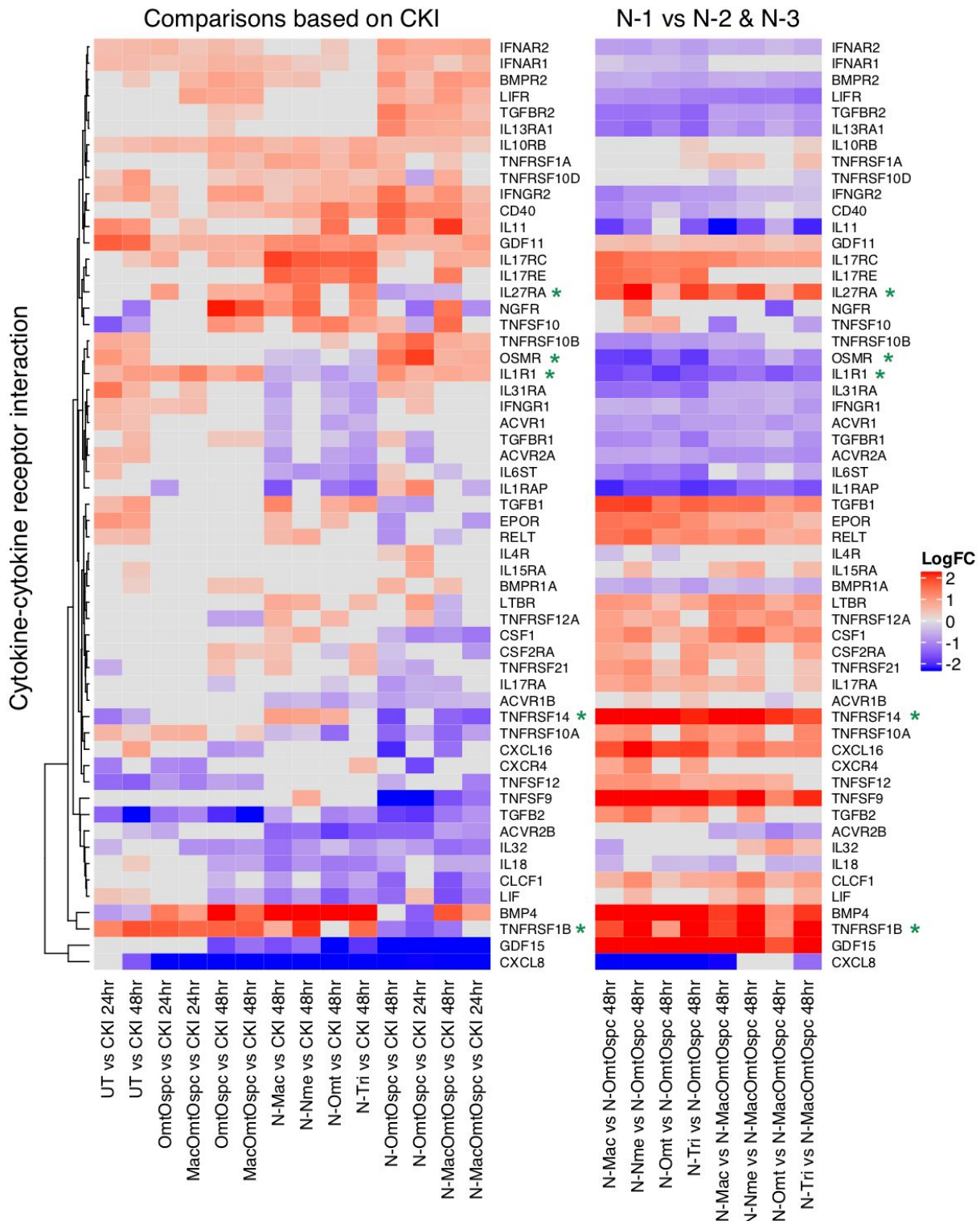




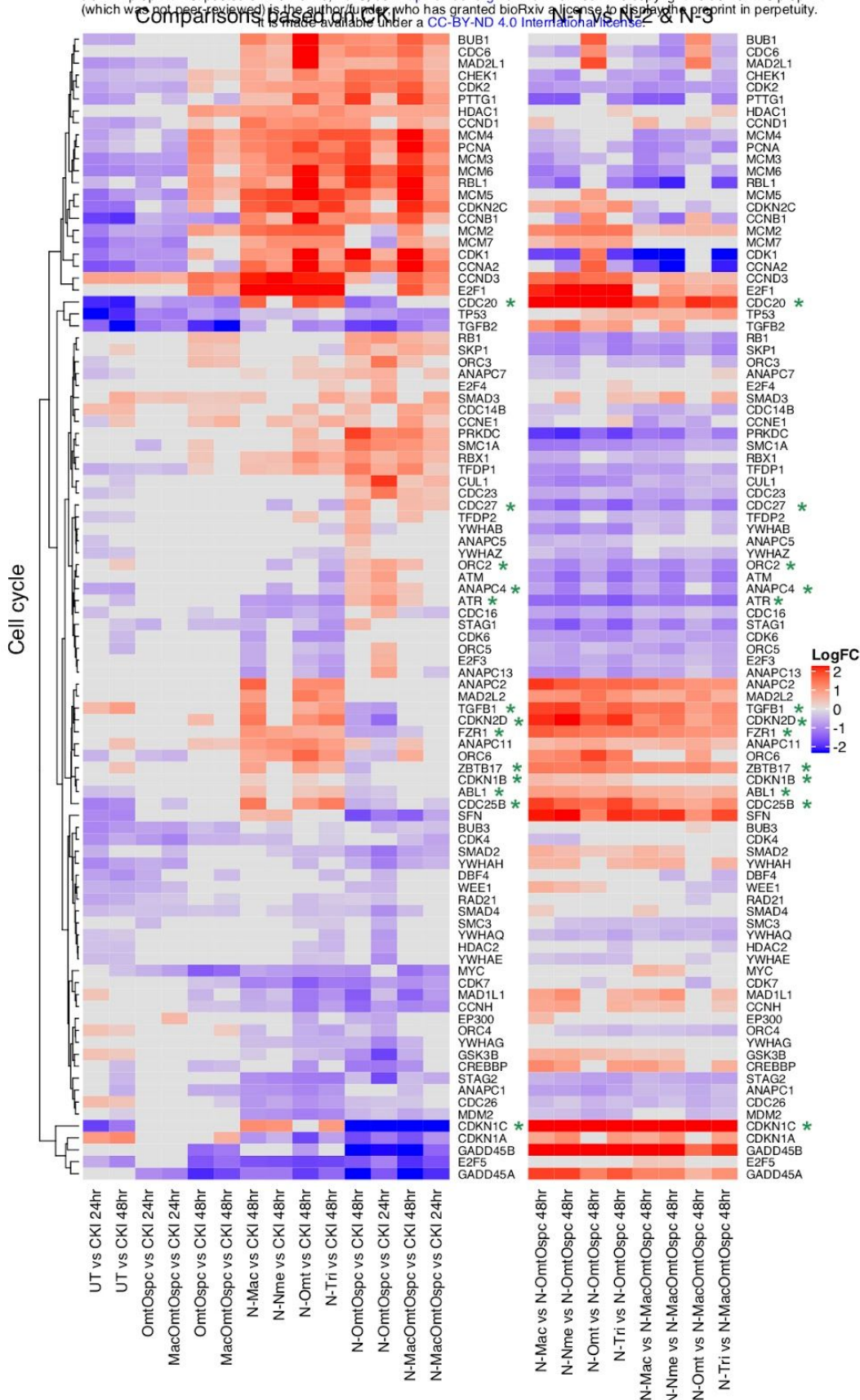








(which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).



# Chapter 7

## Conclusions and Future Directions

TCM or herbal medicines are widely applied therapeutic methods in Asian countries and becoming more popular in western countries. However, the extreme complexity of their components makes it difficult to explain the mechanisms of TCM. Currently, the dominant opinion is that TCM contains multiple bioactive compounds that perturb multiple targets or even pathways to generate a relatively strong effect. This therapeutic pattern makes it almost impossible to research mechanisms of TCM based on the 'one compound one target' mode for modern pharmaceuticals. The systems-biology approach provides an overall and network view for biological changes, which is probably a better way to conduct TCM research.

In this dissertation, we used CKI as our model TCM formula and applied systems biology and functional genomics methods to reveal its potential anticancer mechanisms from different aspects. Firstly, I described new methods to investigate the interactions between CKI and other chemotherapy agents and proved that -omics methods can not only explain observed interactions but also provide clues for potential interactions. Therefore, the results for this method are closer to the clinical application of TCM and more comprehensive than existing methods. Secondly, I applied transcriptome analysis to herbal compatibility research and provided a modern view of the underlying mechanisms used to explain the combination of Kushen and Baituling. Thirdly, I contributed to the investigation of anticancer effects of CKI as a whole. Based on transcriptome results from different cancer cell lines, the cell cycle, energy metabolism and DNA repair pathways were identified as primary pathways through which CKI exerts effects on cancer cells. In addition, these results were experimentally validated. Finally, through an approach combining analytical chemistry, cellular experiments and systems biology approaches, I contributed to work that revealed the effects and interactions of main components in CKI. Altogether, by generating a series of new methods, I and others successfully applied transcriptome analysis for the study of TCM and showed that systems biology methods are a powerful tool in TCM research. Through my work on understanding the interactions of plant extracts and the interactions of plant extracts with pharmaceutical cancer chemotherapy drugs, I

have significantly contributed to the understanding of TCM through the investigation of CKI's anticancer mechanisms and markedly improved the corpus of knowledge in this respect.

Many areas in this thesis can be explored in more depth. For example, transcriptomics is the only omics technique used in our study. We only selected limited target proteins for validation so there is still room for improvement in order to understand the effects of CKI at the protein level. Also, TCM is usually taken for extended periods to treat chronic disease or sub-acute health conditions, making it likely that long periods of use may also involve epigenetic changes. Therefore, combined experimental approaches using different -omics techniques and integrating the results across different levels of biological molecules may provide a more comprehensive view of the mechanisms of TCM.

It is also important to remember that all of our results and analyses are based only on cancer cells. The treatment of cancer not only relies on the interaction between cancer cells and medicines but is also related to tumor microenvironment and systemic immune responses. Although our results showed that CKI and Baituling may have effects on the immune system, these observations are by no means comprehensive and still require verification and additional research on immune cells. Therefore, experiments and systems biology analysis based on animal assays or clinical samples are likely to provide more definitive and practical results for translation of our research into practice.

There are also technical limitations to some of the approaches described in this thesis, the HPLC analysis and knockout methods in Chapter 6 are based on HPLC with a photodiode-array UV-Vis detector for 9 primary compounds in CKI. Although these compounds are considered to be the bioactive components in CKI at relatively high concentration, we cannot exclude the possibility that other minor compounds do not contribute to the effects of CKI. If UPLC with the high-resolution mass spectrometer or other analytical techniques could be applied and more compounds' structure in CKI could be described, future work will likely improve the approach based on knocking out similar compounds as a group.

TCM is a valuable resource for the pharmaceutical and healthcare industries. However, restricted by limited scientific knowledge because of its ancient origin, the principles of herbal medicine have been obfuscated by ambiguous ancient philosophy. Because it emphasizes

overall and systematic views, systems biology has many common points with TCM theory, which make it a useful tool for the modernization of TCM. Our work is just a beginning and introduction to the integration of fast developing omics techniques for TCM research.



# Appendix A

Supplementary Tables and Data.

All the supplementary tables and data for chapters 2, 3, 4, 5 and 6 can be obtained online:  
[https://drive.google.com/drive/folders/1m6KKwK\\_O8i96exUNz-c2UhsRnyay0CeL?usp=sharing](https://drive.google.com/drive/folders/1m6KKwK_O8i96exUNz-c2UhsRnyay0CeL?usp=sharing)

1. For chapter 2, there are three supplementary tables.
  - a. Supplementary Table 1: Mapping rates for RNA-seq data.
  - b. Supplementary Table 2: List of DE genes for different comparisons.
  - c. Supplementary Table 3: Gene list for groups based on type of regulation (Group A-D) and their over-represented GO terms (count > 4 and P-value < 0.05).
  
2. For chapter 3, there are five supplementary tables.
  - a. Supplementary Table 1: Mapping rates for each RNA-seq result.
  - b. Supplementary Table 2: Primer sequences for RT-qPCR target genes.
  - c. Supplementary Table 3: DE gene lists for different comparisons.
  - d. Supplementary Table 4: SPIA results of significantly perturbed KEGG pathways for three injections.
  - e. Supplementary Table 5: Significantly over-represented GO and KEGG terms (count > 4 and P-value < 0.05)
    - I. Sheet 1-2: GO and KEGG enrichment of DE genes shared by CKI compared to 'untreated' and Kushen compared to 'untreated'.
    - II. Sheet 3-4: GO and KEGG enrichment of DE genes for CKI compared to Kushen.
  
3. For chapter 4, there are four supplementary tables.
  - a. Table S1. RT-qPCR target genes and their primer sequences.
  - b. Table S2. Mapping rate of each cell line.
  - c. Table S3. List of DE genes in each cell line at each time point.
  - d. Table S4. The summary of the functional analysis of both separate datasets and shared datasets.
    - I. Sheet 1-4: GO enrichment of each cell line at two time points. Selection standard: cut off p value<0.01, cut off q value<0.01.

II. Sheet 5-8: KEGG enrichment of each cell line at two time points.

Selection standard: cut off p value<0.01, cut off q value<0.01.

III. Sheet 10-12: GO enrichment of each cell line at two time points.

Selection standard: cut off p value<0.01, cut off q value<0.01.

IV. Sheet 13: GO enrichment of shared genes by both cell lines. Selection standard: cut off p value<0.01.

V. Sheet 14: KEGG enrichment of shared genes by both cell lines.

Selection standard: cut off p value<0.01.

4. For chapter 6, there are one supplementary data and one supplementary table.
  - a. Data 1. Calibration curve for the concentrations of known compounds in CKI and reconstituted subtractive fractions.
  - b. Table 1. Summary of shared, differentially expressed (DE) genes across treatments. Similarity percentage calculated from total number of shared DE genes from all listed comparisons. To find the number of DE genes, CKI treatment was used as a baseline to compare all other fractionated treatments in order to emphasize the effect of depleted compounds and UT (untreated) was used as a base to calculate the DE genes for CKI treatment.

# Appendix B

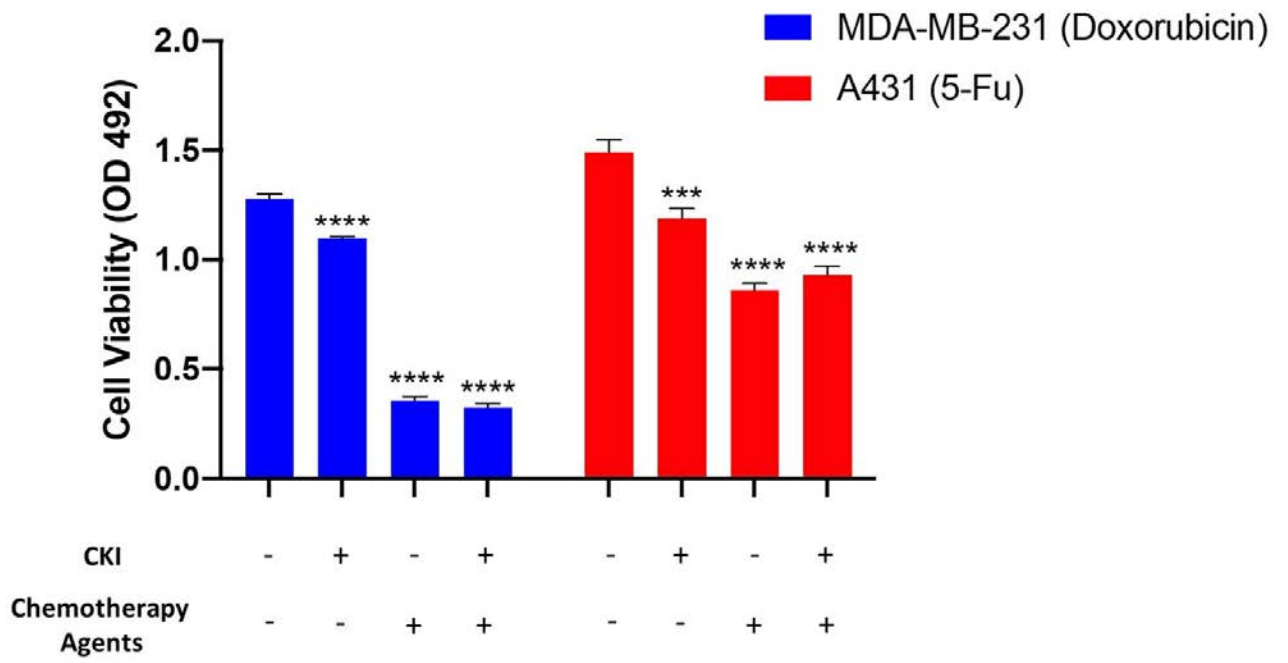
## Supplementary Figures for Chapter 2

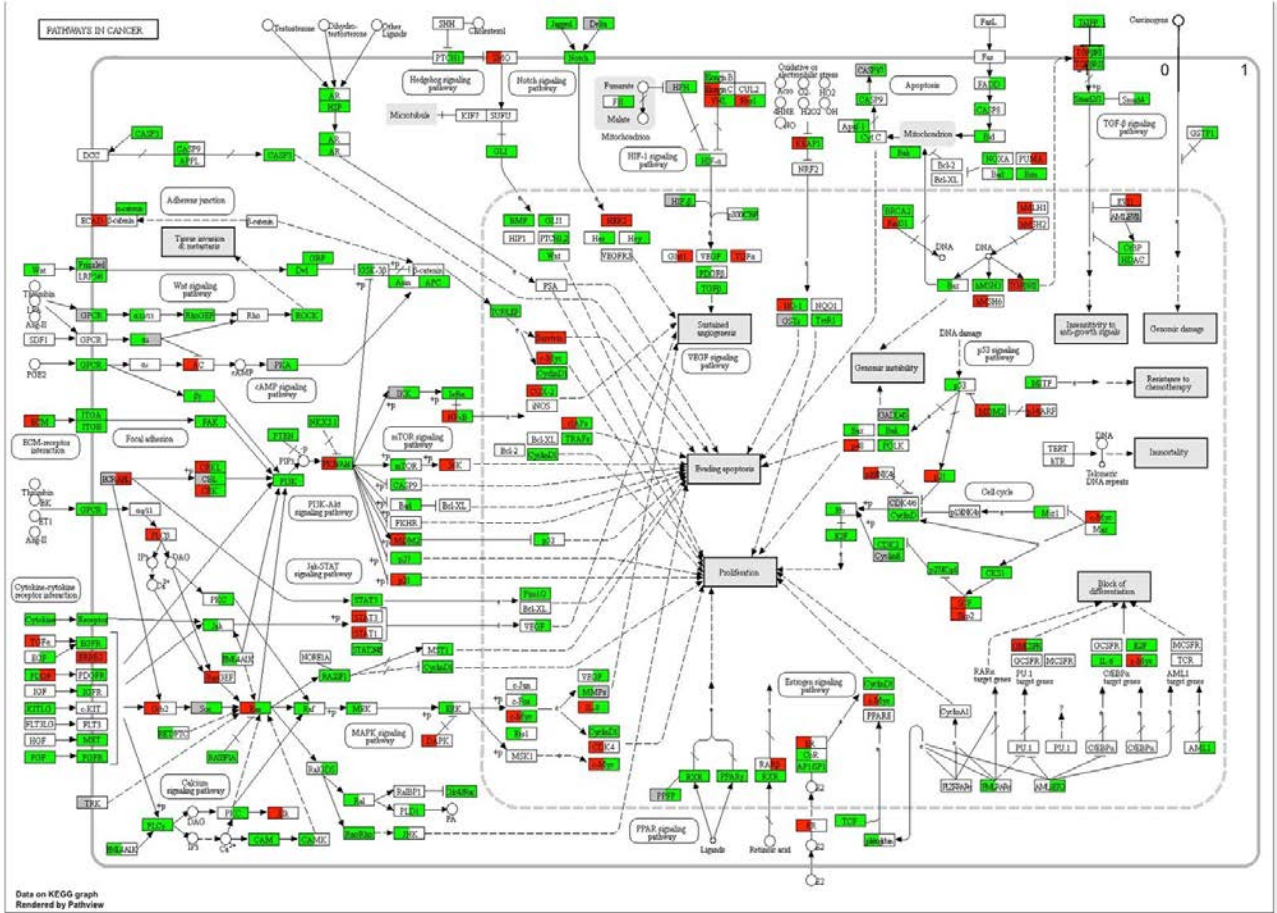
Supplementary Figure 1: Cell viability of cancer cells treated with different drug combinations for 48 hours (MDA-MB-231 cells with doxorubicin and A431 cells with 5-Fu). Results are represented as mean  $\pm$ SEM (n=9). Statistical analysis was performed by comparing each treatment to untreated (\*\*p < 0.001, \*\*\*\* p < 0.0001)

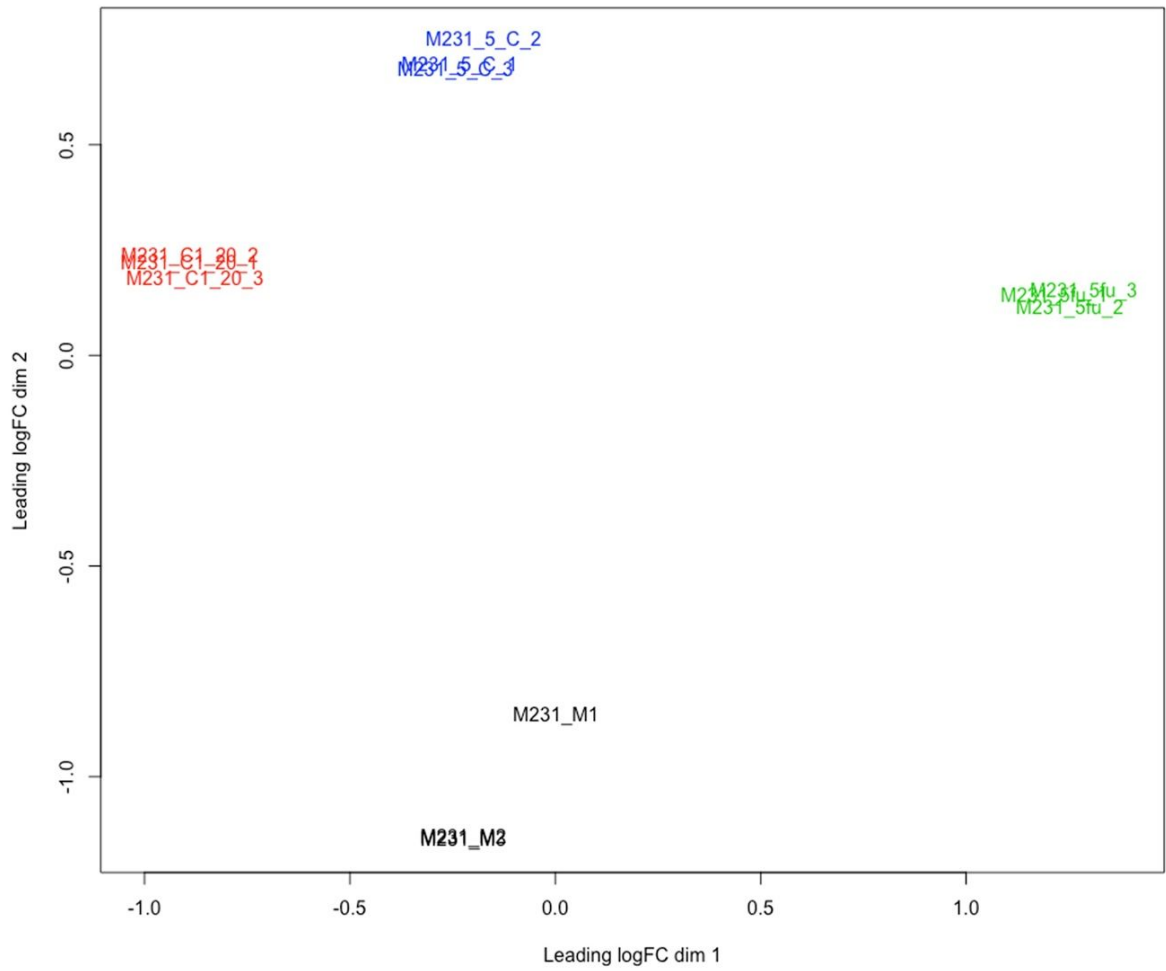
Supplementary Figure 2: Comparison of types of regulation for CKI with doxorubicin and 5-Fu in the “Pathways in cancer” pathway. Left half of the rectangle for each gene represents CKI with doxorubicin in A431 cells and the right half represents CKI with 5-Fu in MDA-MB-231 cells. Red and green colors indicate agonistic and antagonistic regulation, respectively.

Supplementary Figure 3: Multiple dimensional scaling (MDS) plot for MDA-MB-231 samples based on expression profiles of all genes (Untreated in black, CKI in red, 5-Fu in green and CKI+5-Fu in blue).

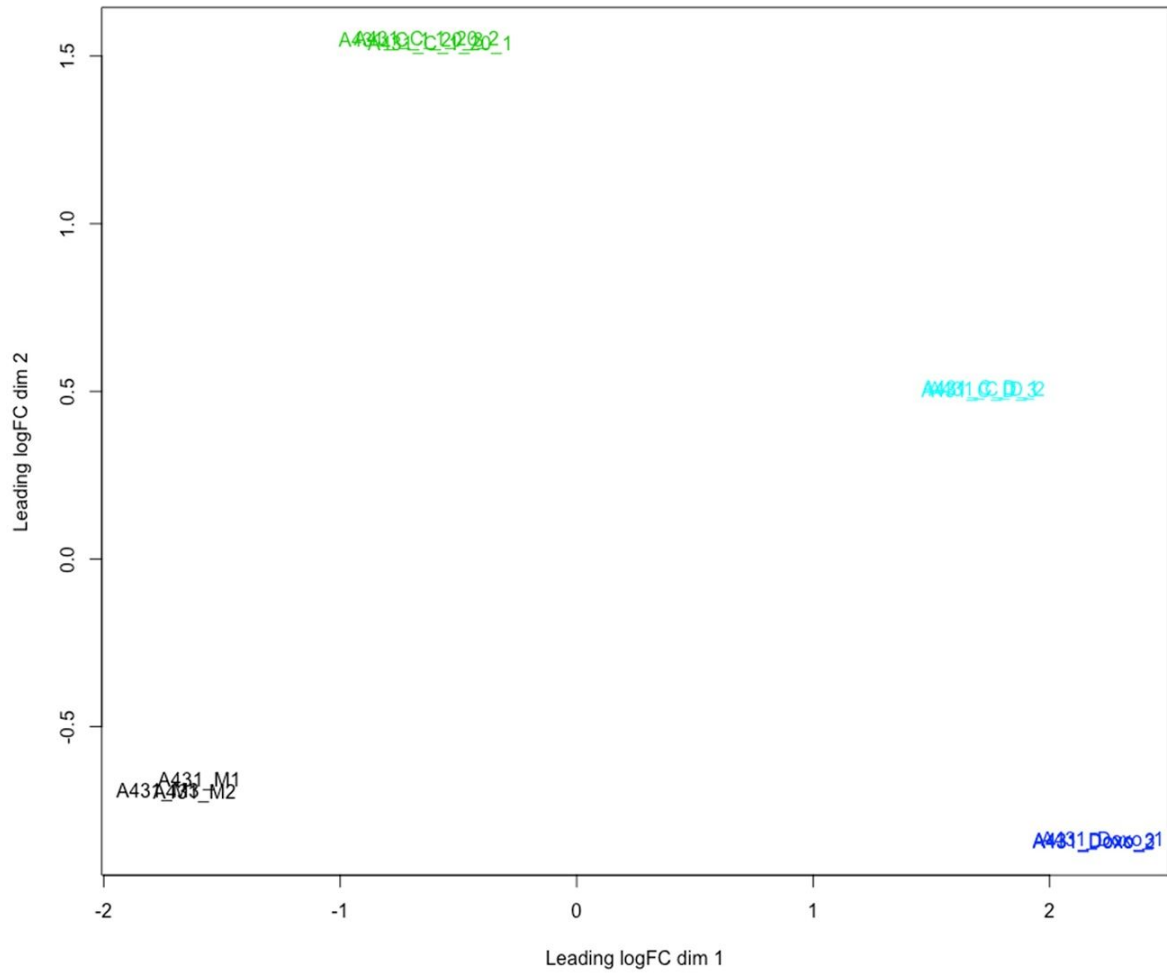
Supplementary Figure 4: Multiple dimensional scaling (MDS) plot for A431 samples based on expression profiles of all genes (Untreated in black, CKI in green, doxorubicin in blue, CKI+doxorubicin in cyan).











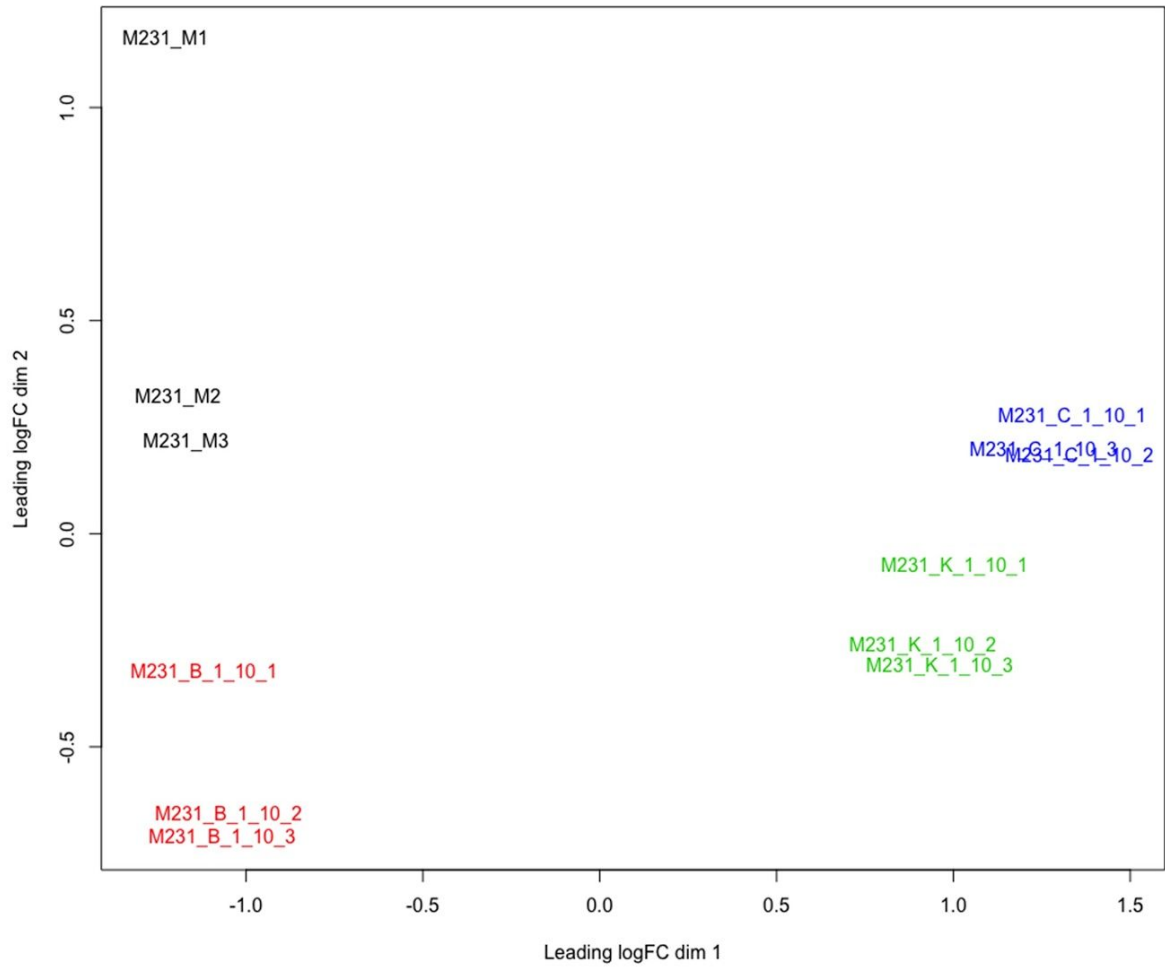
# Appendix C

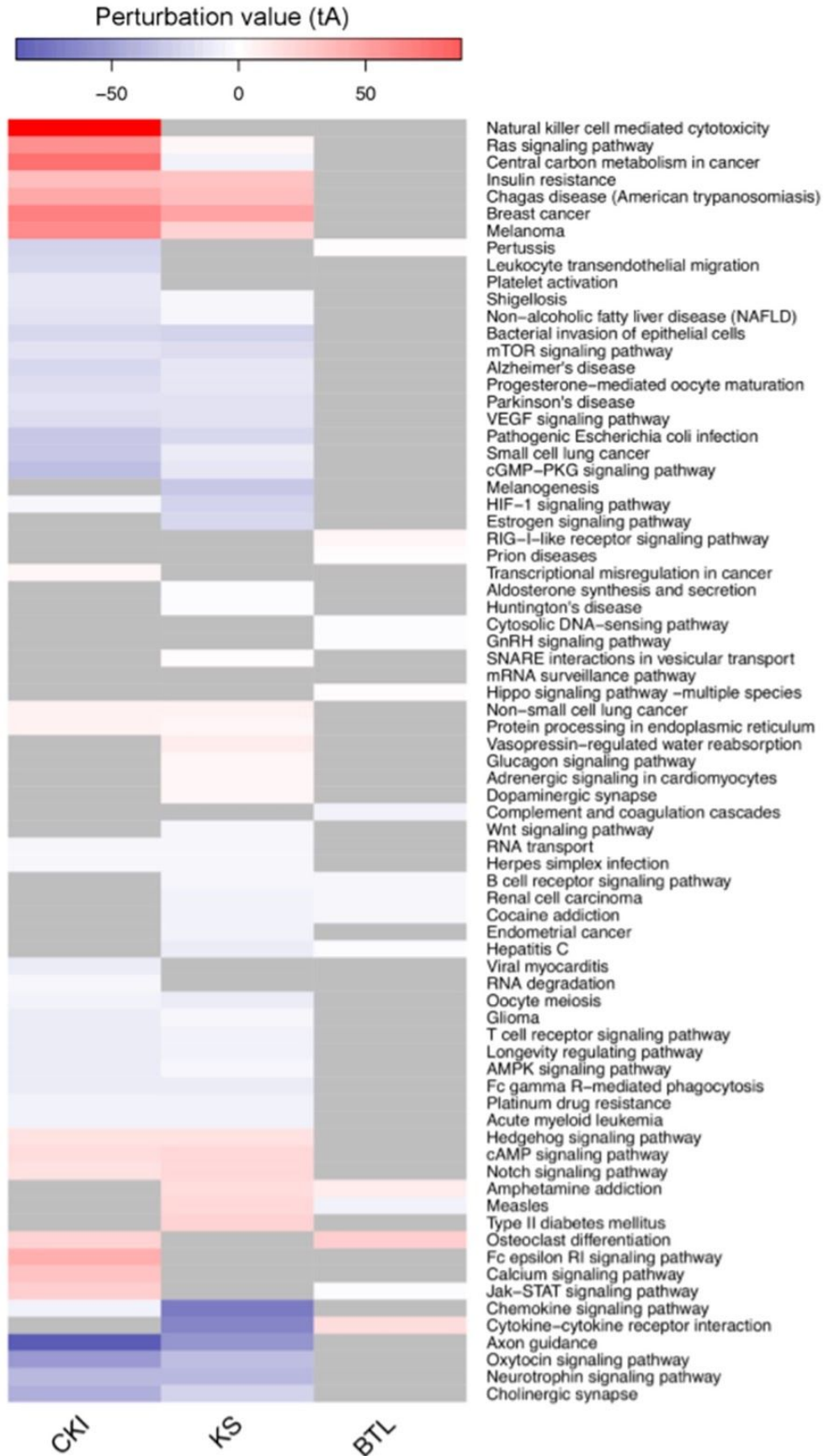
## Supplementary Figures for Chapter 3

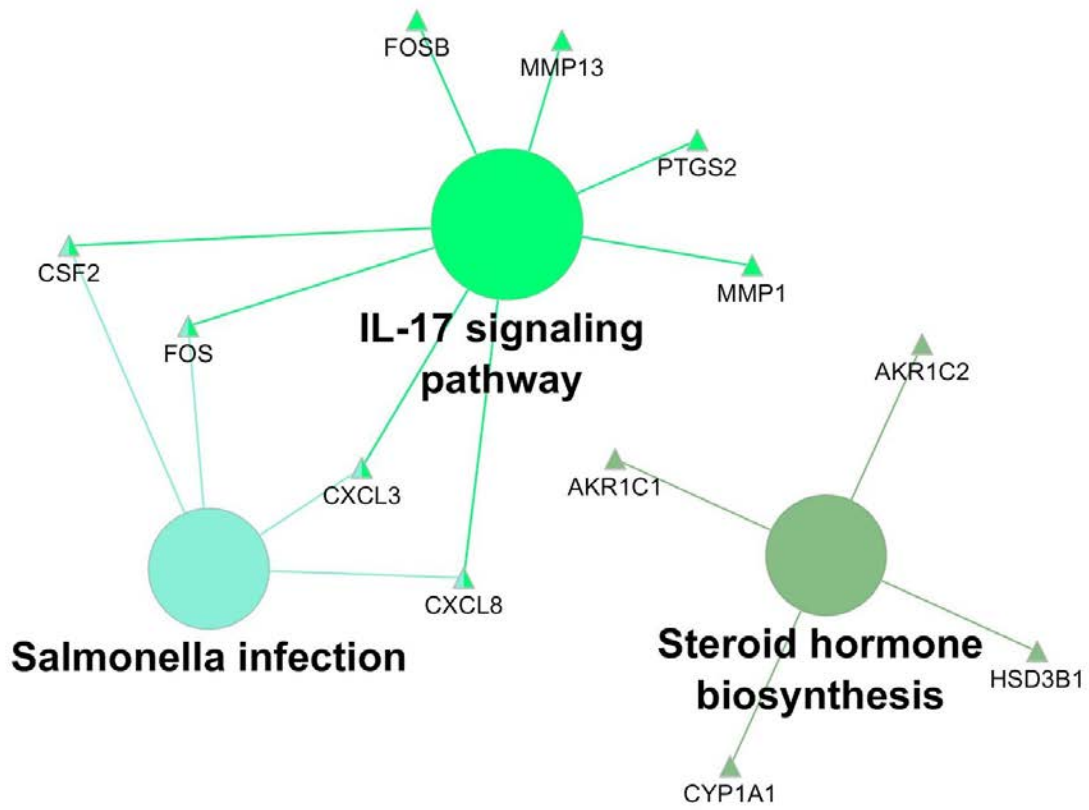
Supplementary Figure 1: Multiple dimensional scaling (MDS) plot for samples based on expression profiles of all genes (Untreated in black, Baituling in blue, Kushen in green and CKI in blue).

Supplementary Figure 2: Heatmap showing the perturbation value of significantly perturbed pathways only for one or two injections.

Supplementary Figure 3: Over-represented KEGG pathways and their contained genes showing shared DE genes between CKI (DE calculated by comparison to Kushen treated) and Kushen (DE calculated by comparison to untreated). Node size is proportional to the statistical significance of over-representation and genes are connected to their belonged pathways with edges.







# Appendix D

## Supplementary Figures for Chapter 4

Figure S1. MDS plot of the DE gene distribution of two cell lines under different conditions.

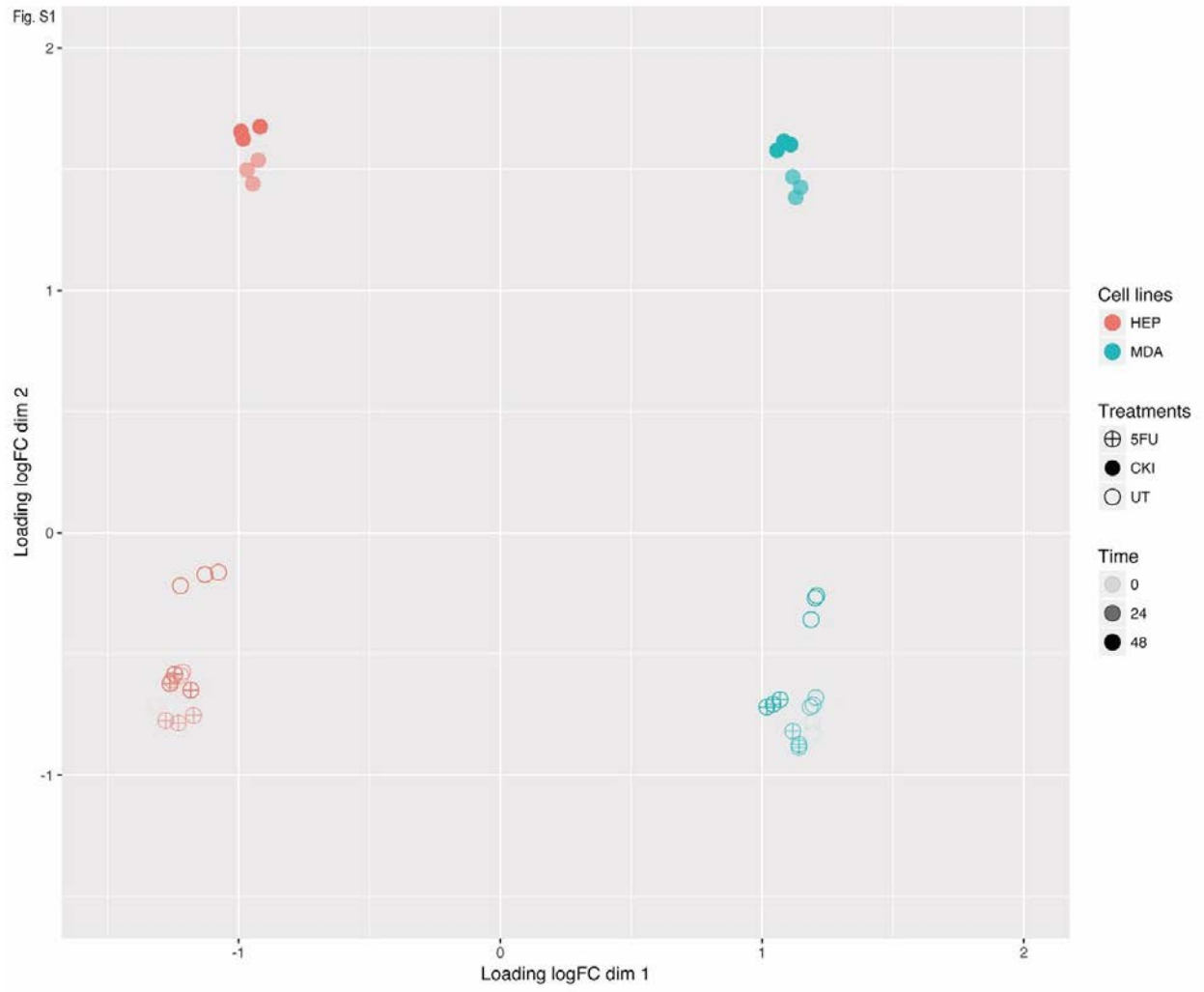
Figure S2. The GO semantic similarity analysis of each data set. (A-D) Each small square represents a GO biological process functions in level 3. Size of the squares positively correlates with the statistical significance of related biological process. Different colour is to distinguish biological process clusters that are described by the top shaded functional representatives.

Figure S3. DE genes distribution of two cell lines in the pathways in cancer. In the cell cycle pathway, each coloured box is separated into 4 parts, from left to right representing 24h CKI treated HepG2, 48h CKI treated Hep G2, 24h CKI treated MDA-MB-231 and 48h CKI treated MDA-MB-231.

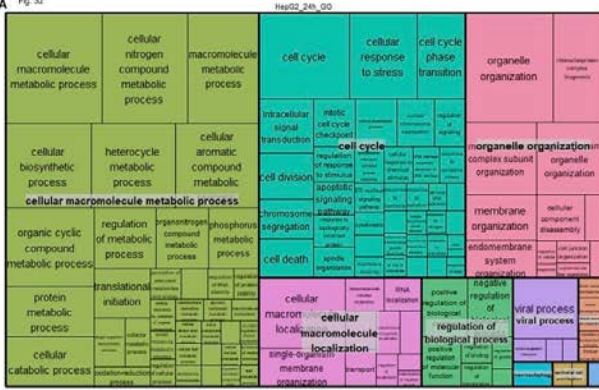
Figure S4. The heatmap of core genes of three cell lines. Heatmap revealing the expression fold changes of core genes in three cell lines at two time points. All the core genes can be separated into 3 clusters, namely consistently up/down regulated genes and uneven genes.



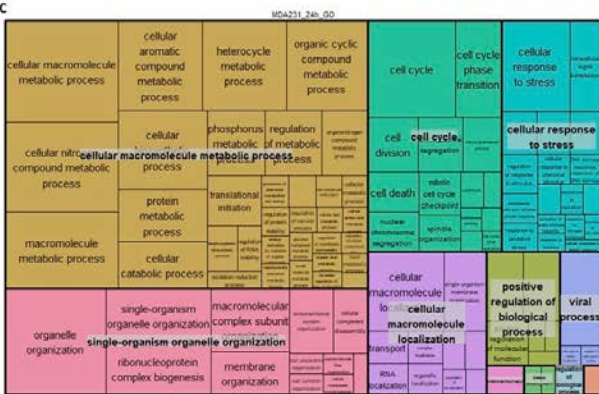
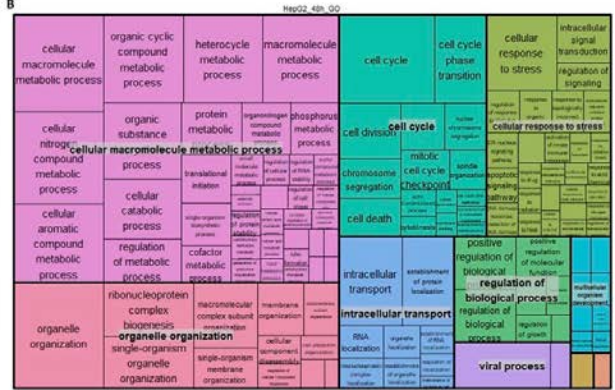
Fig. S1



A Fig 52



B



D

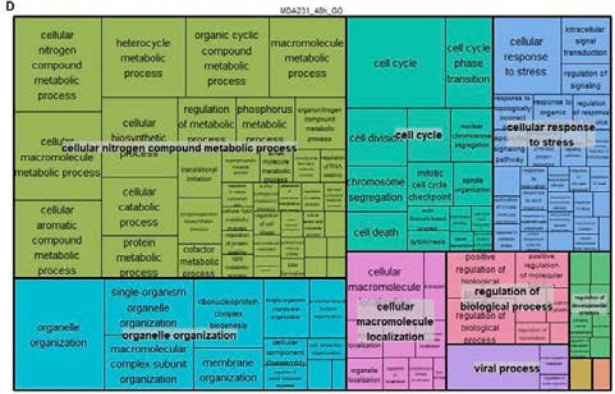
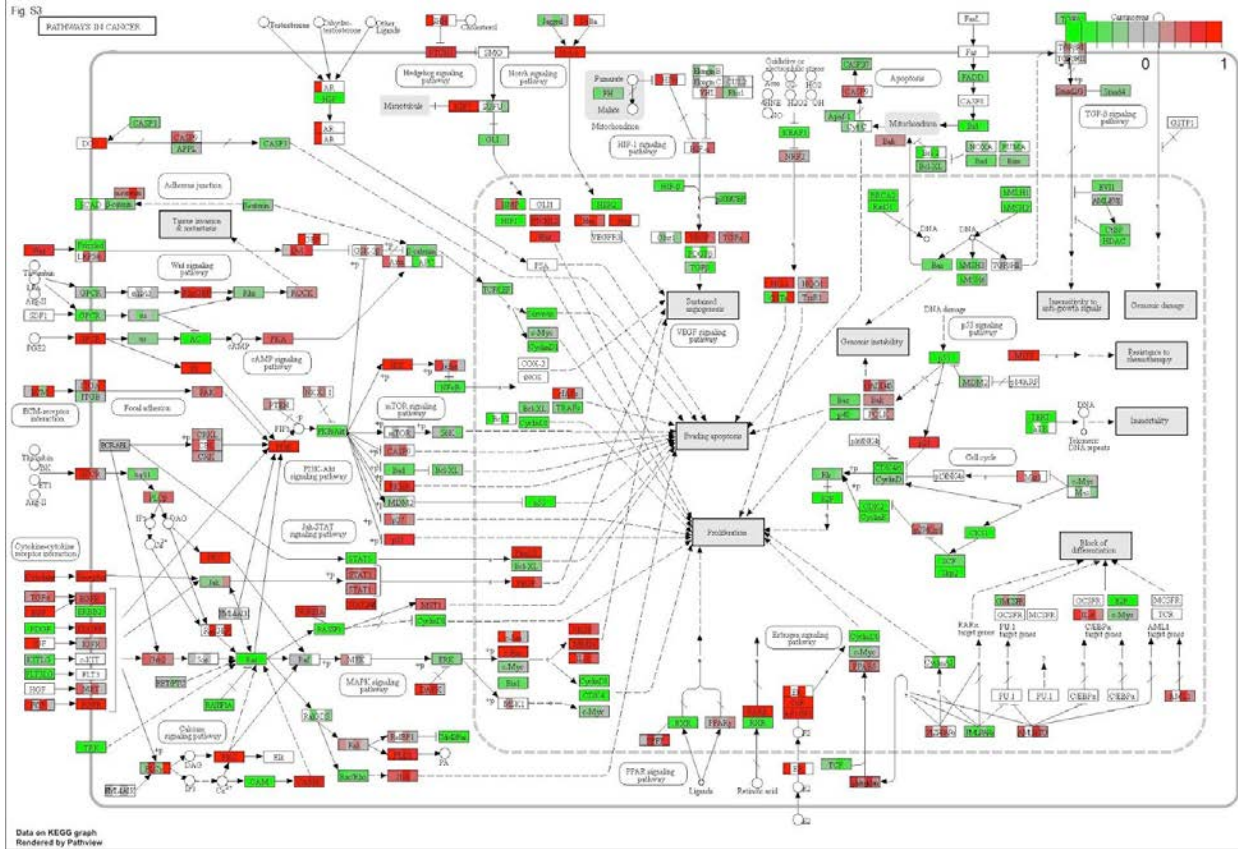
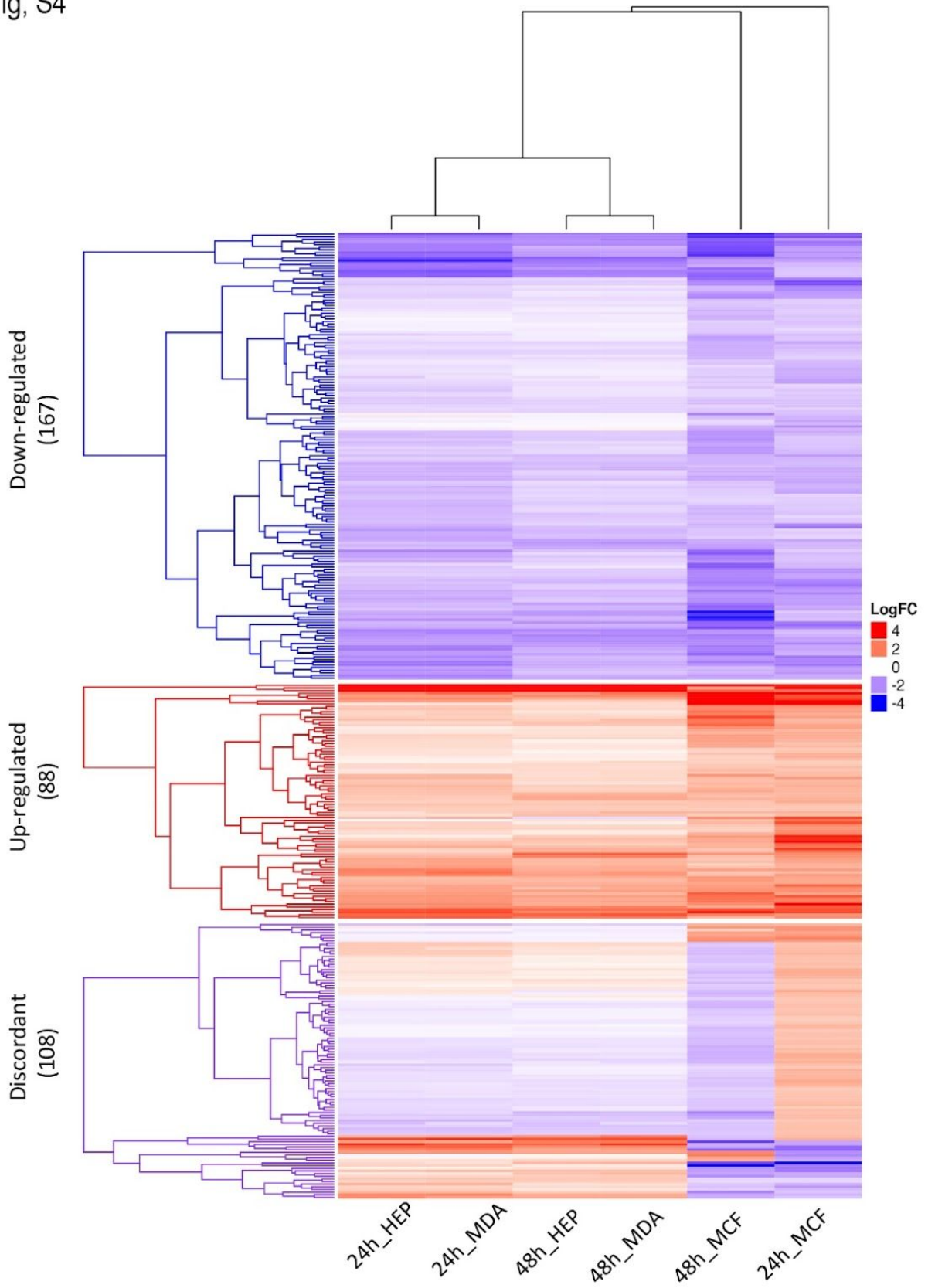


Fig S3



Data on KEGG graph  
Rendered by Pathview

Fig, S4



# Appendix E

## Supplementary Figures for Chapter 5

## Additional File1:

# Cell Cycle, Energy Metabolism and DNA Repair Pathways in Cancer Cells are Suppressed by Compound Kushen Injection

Jian Cui, Zhipeng Qu, Yuka Harata-Lee, Thazin Nwe Aung, Hanyuan Shen and David L Adelson

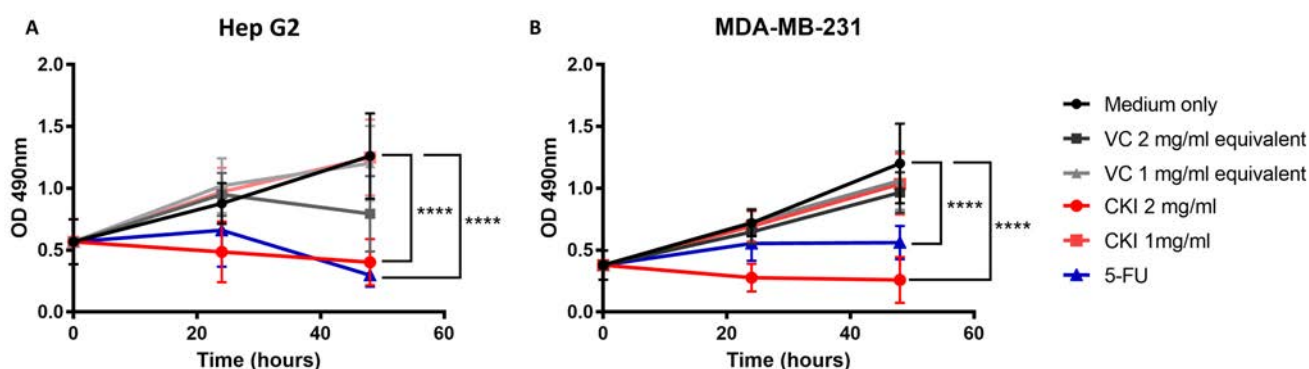
The University of Adelaide, School of Biological Sciences, Dept of Molecular and Biomedical Sciences

## 1 SUPPLEMENTARY DATA

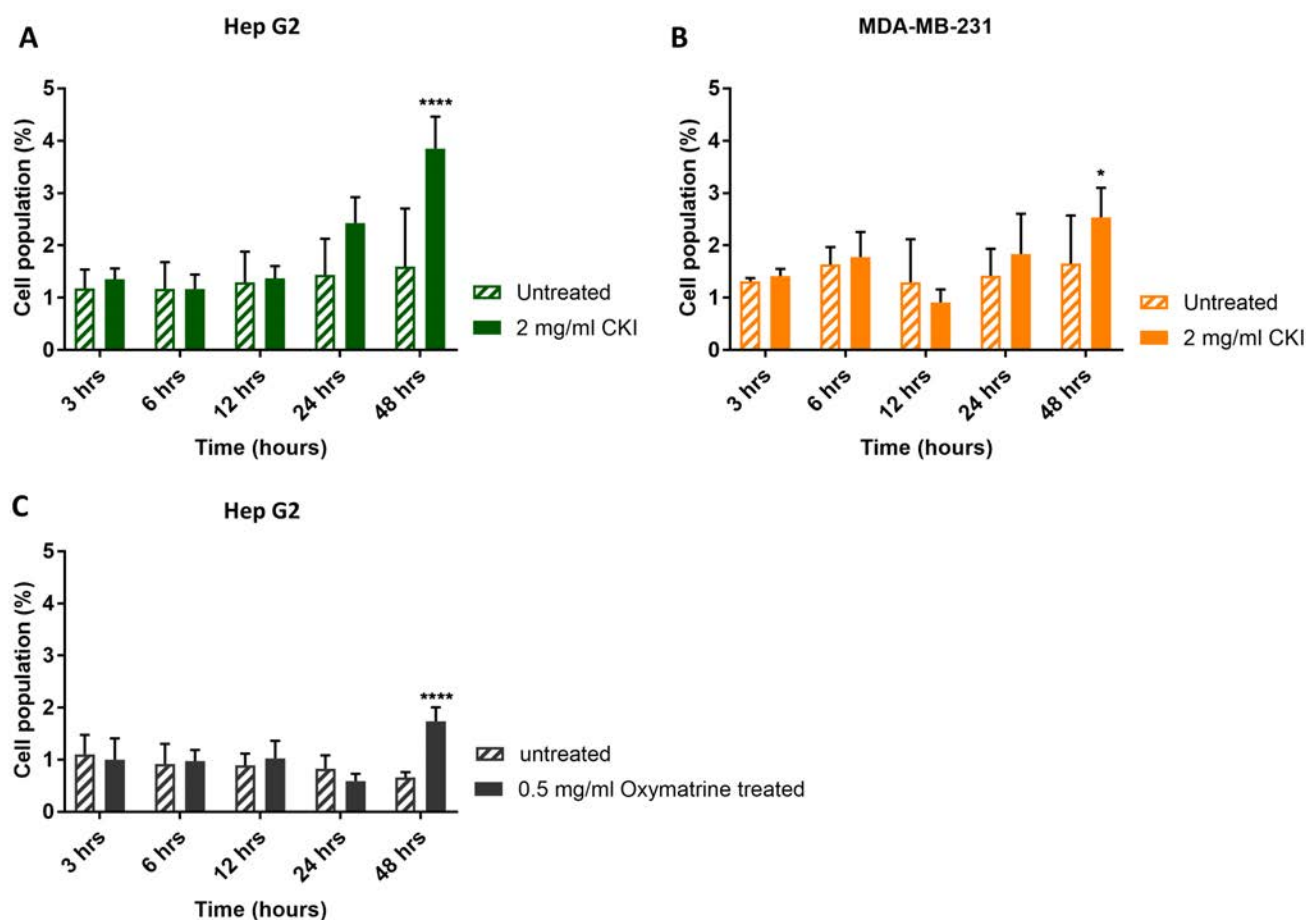
### 1.1 Methods

**XTT assay:** The wells of 96-well tray were seeded with  $4 \times 10^3$  cells per well for Hep G2 cells and  $8 \times 10^4$  cells per well for MDA-MB-231 cells in  $50 \mu\text{L}$  of medium and cultured overnight. On the following day,  $50 \mu\text{L}$  of either medium, CKI or 5-FU were added to the cells. Viability of the cells was measured at 0, 24 and 48 hours after the treatment by adding XTT:PMS (50:1; Sigma-Aldrich). After 4-hour incubation at  $37^\circ\text{C}$  optical density (OD) of each well was read at 490 nm. The background OD was also measured and the average was subtracted from the OD readings of appropriate wells.

### 1.2 Figures



**Figure S1. XTT assay result of Hep G2 and MDA-MB-231 cell lines.** The XTT assay measures levels of NADH and NADPH by producing a formazan dye product that can be detected at 490nm. **A.** XTT assay result for Hep G2 cells. The assay was carried out at three time points: 0, 24, and 48 hours. 5 treatment groups were used and compared,  $150 \mu\text{g/ml}$  5-FU as a positive control for a cytotoxic agent, 1 mg/ml and 2 mg/ml CKI as well as the corresponding concentration of vehicle control (VC). CKI has a clear effect on the amount of formazan dye produced indicating a significant and marked suppression in the production of NADH and NADPH. **B.** XTT result of MDA-MB-231 cells. This test is with a low concentration of 5-FU ( $20 \mu\text{g/ml}$ ). CKI has a clear and marked effect on the level of formazan dye produced indicating a significant and marked suppression in the production of NADH and NADPH. Statistical analyses were performed using two-way ANOVA comparing with untreated (\*\*\*\* $p < 0.0001$ ); bars show 1 standard deviation from the mean.



**Figure S2. Cell apoptosis assay.** **A.** Cell apoptosis in Hep G2 cells treated with CKI. The assay was carried out at 5 time points to detect apoptosis levels between untreated and 2 mg/ml CKI treated groups. From 3 to 12 hours, both groups maintained a baseline level of apoptosis. After 24 hour, apoptosis of CKI treated cells increased, with the difference attaining statistical significance at 48 hours. **B.** Cell apoptosis in MDA-MB-231 cells treated with CKI. From 3 to 24 hours, both groups show similar, if noisy results. By 48 hours apoptosis has increased and was statistically significantly different to the control. **C.** Cell apoptosis in Hep G2 cells treated with oxymatrine. We compare apoptosis levels between an untreated group and a group treated with 0.5 mg/ml oxymatrine. From 3 to 24 hours we observed a baseline level of apoptosis. By 48 hours apoptosis in the oxymatrine treated group is significantly greater than in the control group. Statistical analyses were performed using two-way ANOVA comparing with untreated (\*\*\*\* $p < 0.0001$ ); bars show 1 standard deviation from the mean.

# Appendix F

## Supplementary Figures for Chapter 6



Supplementary Fig. 1: HPLC profiles of 25 mixtures including CKI, MJ, MN, 9 (N-1), 4 (N-3), and 6 (N-2). Numbers represent compounds 1: Mac, 2: Ade, 3: Nme, 4: Spr, 5: Mt, 6: Spc, 7: Ospc, 8: Omt and 9: Tri. 50  $\mu$ l of the samples at 1 mg/ml concentration was injected through the semi-preparative column to achieve the profiles.

Supplementary Fig. 2: XTT Cell viability assays of subtractive fractions in MDA-MB-231 cells at 24- and 48-hour timepoints treated with 1 mg/ml or 2 mg/ml of CKI and 2 mg/ml equivalent concentrations of all other treating agents. (a) Suppression of cell viability from the following treatments: CKI, WRCKI-B (whole reconstituted CKI in buffer/vehicle control), and WRCKI-H (whole reconstituted CKI in milliQ H<sub>2</sub>O) (b) assessment of the interaction effects of single MJ compounds by the addition to the MN subtractive fraction. Single major compounds were dissolved in either MilliQ H<sub>2</sub>O or Dimethyl sulfoxide. Effect of subtractive fractions (c) N-OmtOspcSpc, (d) N-MacAdeTri, (e) N-MtNmeSpr, (f) N-MtNme, (g) N-OmtSpc, (h) N-OspcSpc, (i) N-MacTri, (j) N-AdeTri, (k) N-MacAde, (l) N- MtSpr, (m) NmeSpr. Statistically significant results shown as  $p < 0.05$  (\*) or  $p < 0.01$  (\*\*)  $p < 0.001$  (\*\*\*), or  $p < 0.0001$  (\*\*\*\*); ns (not significant). All data were shown as mean  $\pm$  SD.

Supplementary Fig. 3: XTT Cell viability assays of subtractive fractions N-OmtOspc and N-MacOmtOspc in MDA-MB-231 cells at 24- and 48-hour timepoints treated 4 different concentrations ranging from 0.25 mg/ml to 2 mg/ml of CKI and equivalent concentrations of two other agents. Statistically significant results shown as  $p < 0.05$  (\*) or  $p < 0.01$  (\*\*)  $p < 0.001$  (\*\*\*), or  $p < 0.0001$  (\*\*\*\*); ns (not significant). All data were shown as mean  $\pm$  SD.

Supplementary Fig. 4: Representative plots of Annexin V and PI staining in (A) MDA-MB-231, (B) HEK-293, and (C) HFF.

Supplementary Fig. 5: Multiple dimensional scaling (MDS) plot for samples based on expression profiles of all genes before the removal of unwanted variance (RUVs) in R package.

Supplementary Fig. 6: MDS plot for samples based on expression profiles of all genes after the application of RUVs.

Supplementary Fig. 7: Box plot for samples based on expression profiles of all genes before the application of RUVs.

Supplementary Fig. 8: Box plot for samples based on expression profiles of all genes after the application of RUVs.

Supplementary Fig. 9: Venn diagrams showing the number of overlapping DE genes between treatments at 24-hours and 48-hours.

Supplementary Fig. 10: Identification of significantly perturbed pathways using SPIA ( $pG < 0.05$ ) analysis. Eighty-six significantly perturbed pathways from twenty-two comparisons were found.

Supplementary Fig. 11: DE genes from the following comparisons (CKI vs UT, CKI vs N-Mac, CKI vs N-Nme, CKI vs N-Omt and CKI vs N-Tri) shown in the "Cytokine-Cytokine Receptor Interaction pathway at 48-hours. Significantly up- and down-regulated DE genes were coloured red and green respectively. Each coloured box was separated into five parts according to treatments in this order: CKI vs UT, CKI vs N-Mac, CKI vs N-Nme, CKI vs N-Omt and CKI vs N-Tri. White or grey colours represented gene(s) that were not significantly differentially expressed by the treatments.

Supplementary Fig. 12: DE genes from the following comparisons (CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc) shown in the Cytokine-Cytokine Receptor Interaction pathway at 48-hours. Significantly up- and down-regulated DE genes were coloured red and green respectively. Each coloured box was separated into five parts according to this order: CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc. White or grey colours represented gene(s) that were not significantly differentially expressed by the treatments.

Supplementary Fig. 13: DE genes from the following comparisons (CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc) shown in the Cytokine-Cytokine Receptor Interaction pathway at 24-hours. Significantly up- and down-regulated DE genes were coloured red and green respectively. Each coloured box was separated into five parts according to this order: CKI vs UT, CKI vs N-OmtOspc, CKI vs

N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc. White or grey colours represented gene(s) that were not significantly differentially expressed by the treatments.

Supplementary Fig. 14: DE genes from the following comparisons (CKI vs UT, CKI vs N-Mac, CKI vs N-Nme, CKI vs N-Omt and CKI vs N-Tri) shown in the Cell Cycle pathway at 48-hours. Significantly up- and down-regulated DE genes were coloured with red and green respectively. Each coloured box was separated into five parts according to this order: CKI vs UT, CKI vs N-Mac, CKI vs N-Nme, CKI vs N-Omt and CKI vs N-Tri. White or grey colours represented gene(s) that were not significantly differentially expressed by the treatments.

Supplementary Fig. 15: DE genes from the following comparisons (CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc) shown in the Cell Cycle pathway at 48-hour treatments. Significantly up- and down-regulated DE genes were coloured red and green respectively. Each coloured box was separated into five parts according to this order: CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc. White or grey colours represented gene(s) that were not significantly differentially expressed by the treatments.

Supplementary Fig. 16: DE genes from the following comparisons (CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc) shown in the Cell Cycle pathway at 24-hours. Significantly up- and down-regulated DE genes were coloured red and green respectively. Each coloured box was separated into five parts according to this order: CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc. White or grey colours represented gene(s) that were not significantly differentially expressed by the treatments.

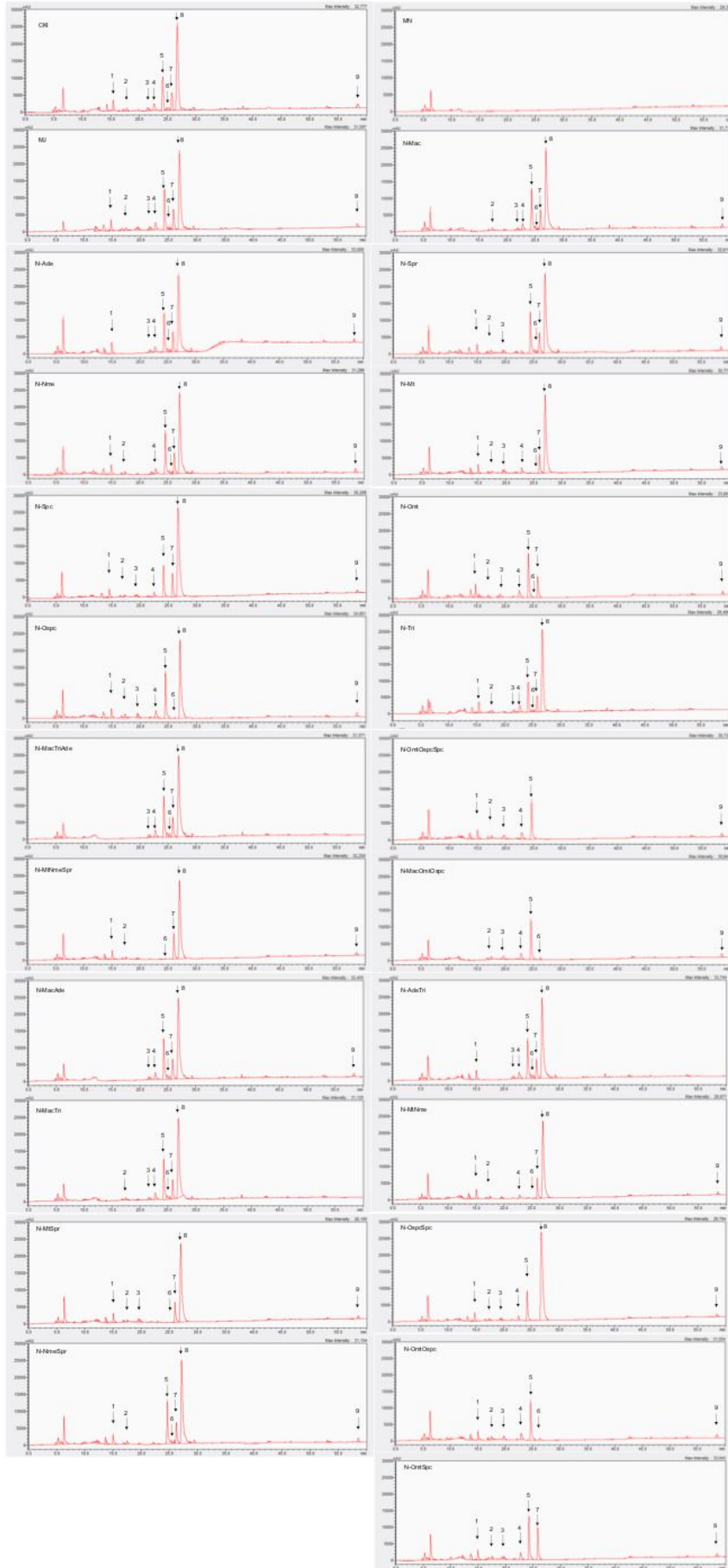
Supplementary Fig. 17: Differential gene expression profiles of all treatments for TGF- $\beta$  signalling pathway: the left panel shows comparison of subtractive fraction treated cells against CKI treatment and the right panel shows comparison of single compound subtractive fraction treated cells against the treatments for two and three compound subtractive fractions.

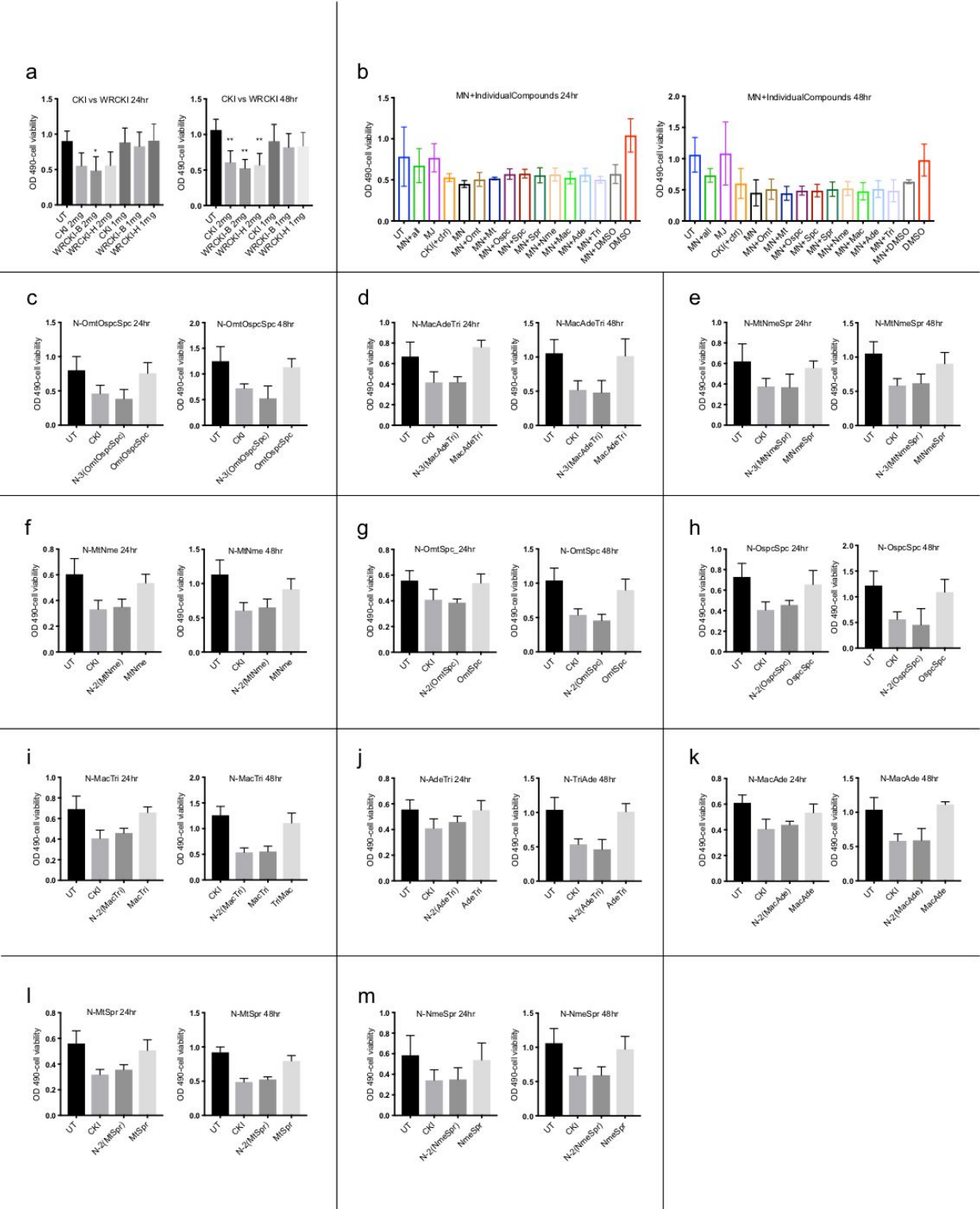
Supplementary Fig. 18: DE genes from the following comparisons (CKI vs UT, CKI vs N-Mac, CKI vs N-Nme, CKI vs N-Omt and CKI vs N-Tri) shown in the TGF- $\beta$  signalling pathway at

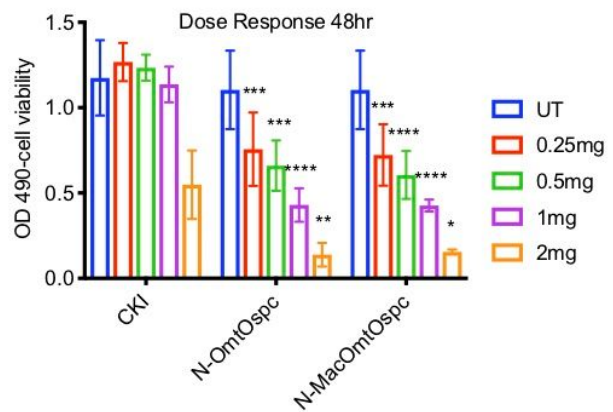
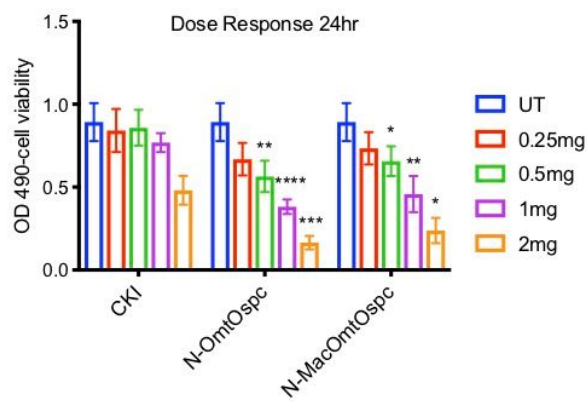
48-hours. Significantly up- and down-regulated DE genes were coloured red and green respectively. Each coloured box was separated into five parts according to this order: CKI vs UT, CKI vs N-Mac, CKI vs N-Nme, CKI vs N-Omt and CKI vs N-Tri. White or grey colours represented gene(s) that were not significantly differentially expressed by the treatments.

Supplementary Fig. 19: DE genes from the following comparisons (CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc) shown in the TGF- $\beta$  signalling pathway at 48-hour treatments. Significantly up- and down-regulated DE genes were coloured red and green respectively. Each coloured box was separated into five parts according to this order: CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc. White or grey colours represented gene(s) that were not significantly differentially expressed by the treatments.

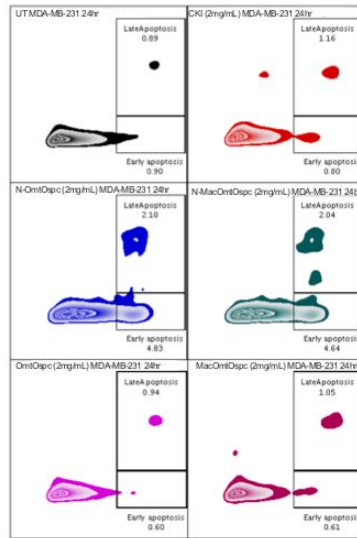
Supplementary Fig. 20: DE genes from the following comparisons (CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc) shown in the TGF- $\beta$  signalling pathway at 24-hours. Significantly up- and down-regulated DE genes were coloured red and green respectively. Each coloured box was separated into five parts according to this order: CKI vs UT, CKI vs N-OmtOspc, CKI vs N-MacOmtOspc, CKI vs OmtOspc and CKI vs MacOmtOspc. White or grey colours represented gene(s) that were not significantly differentially expressed by the treatments.



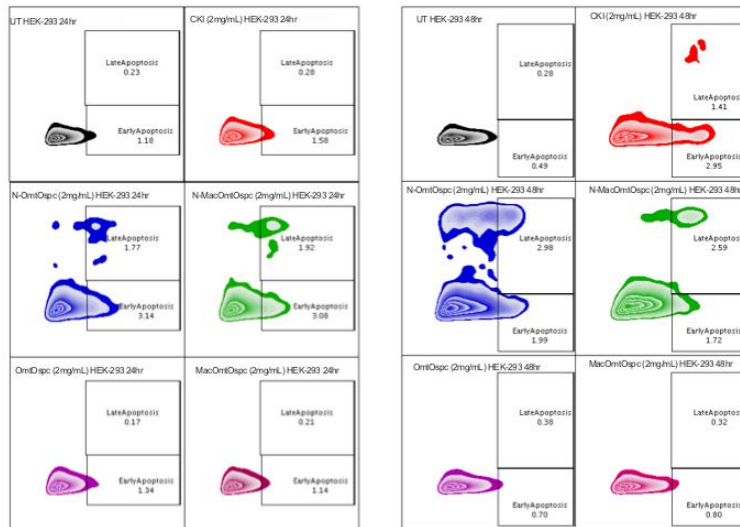




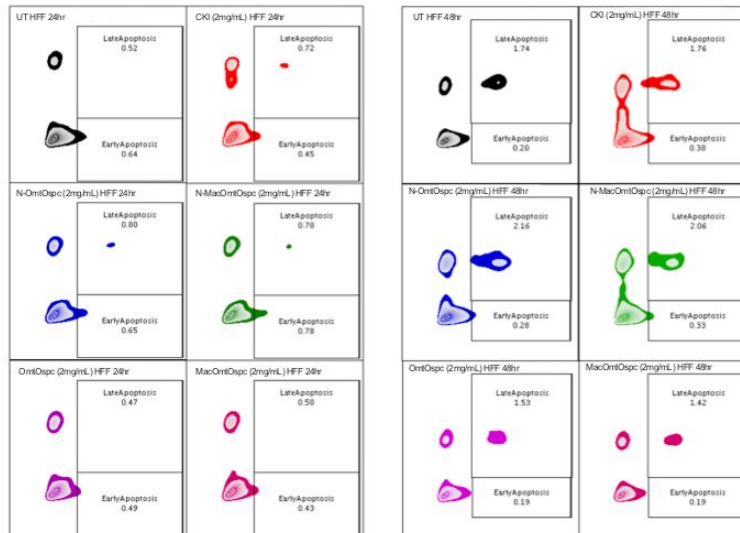
a



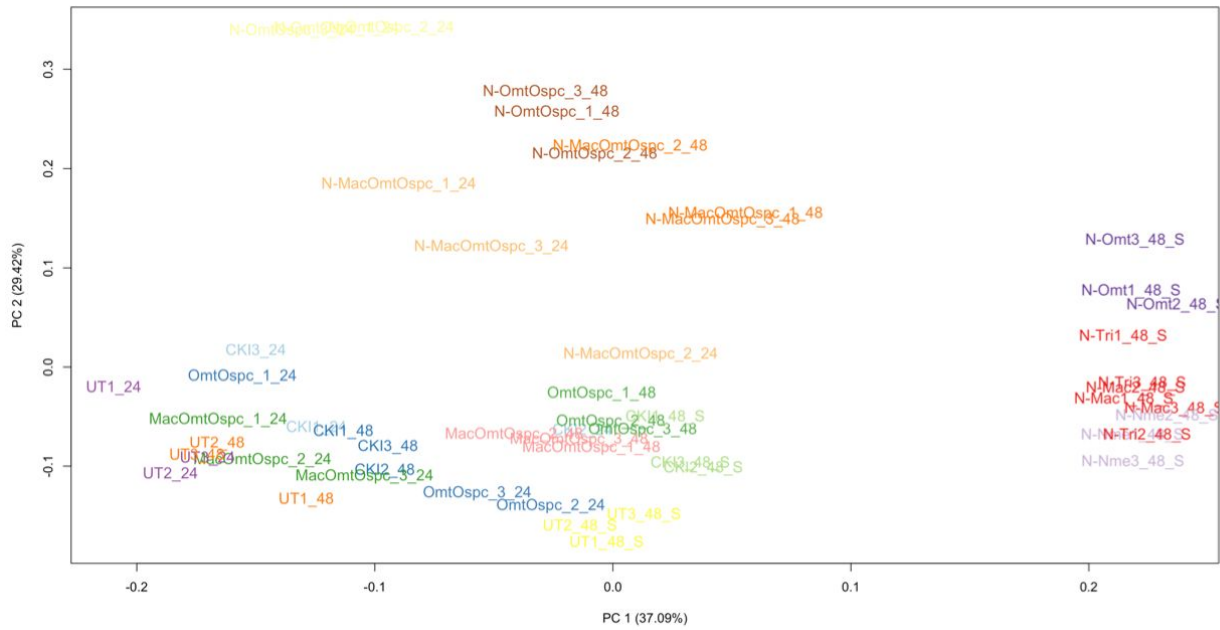
b



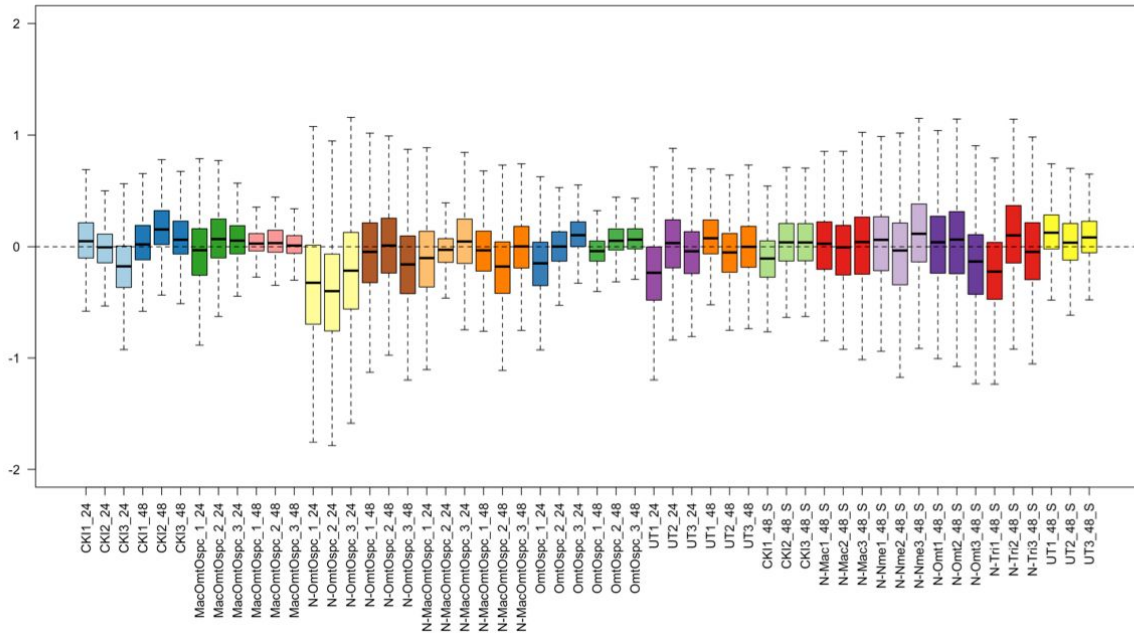
c

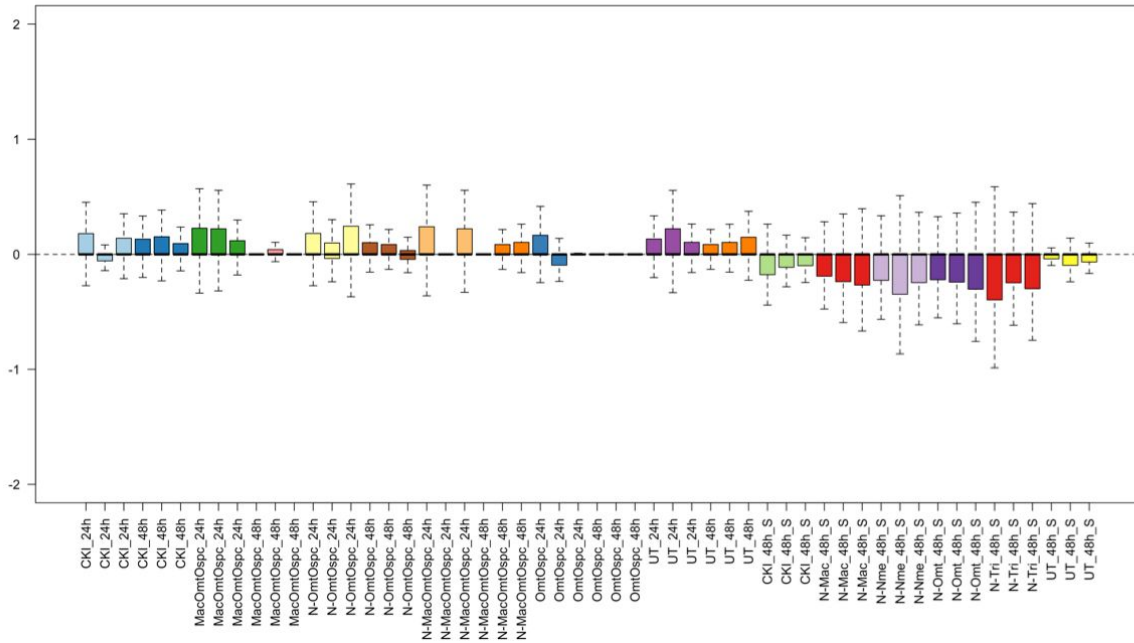


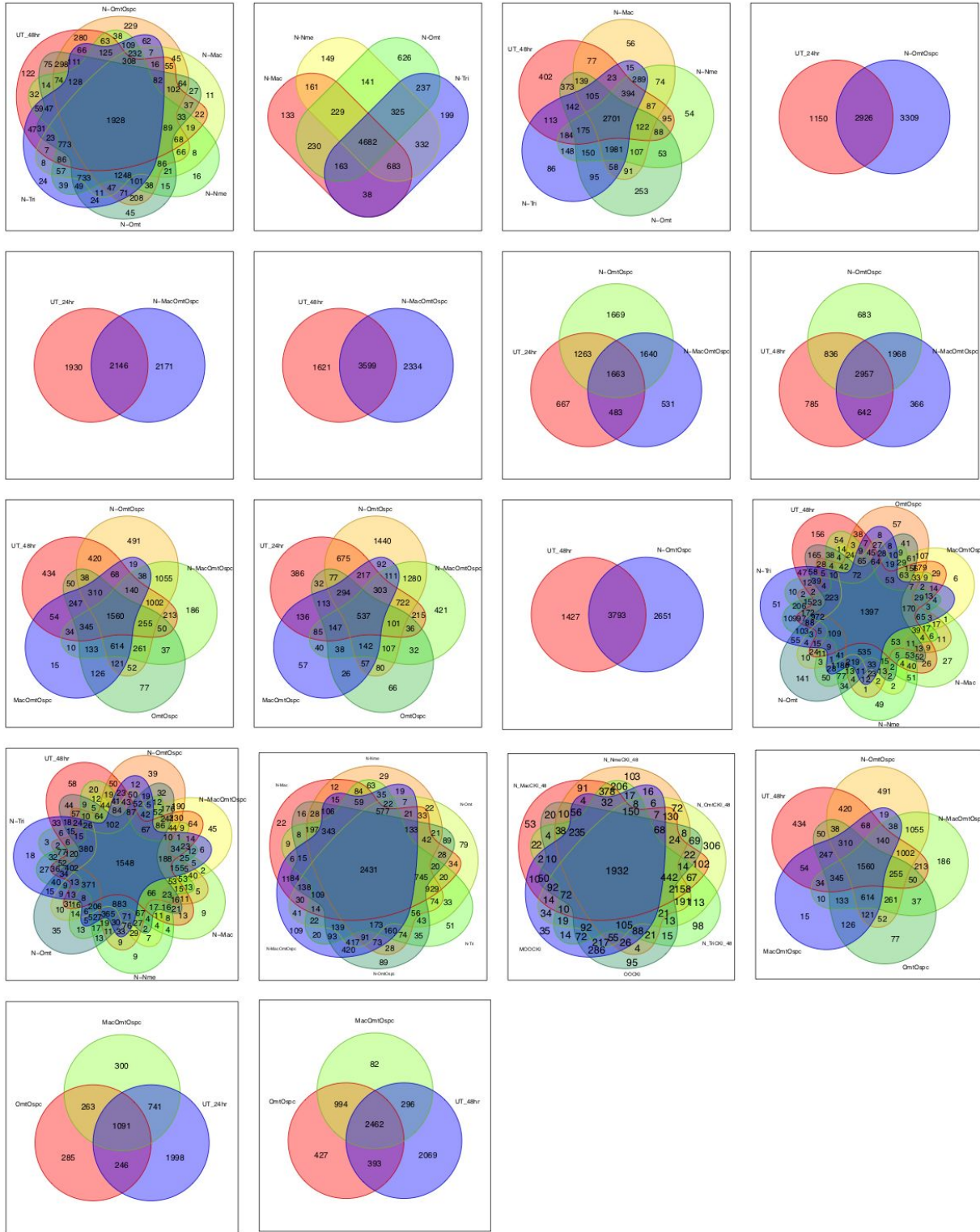




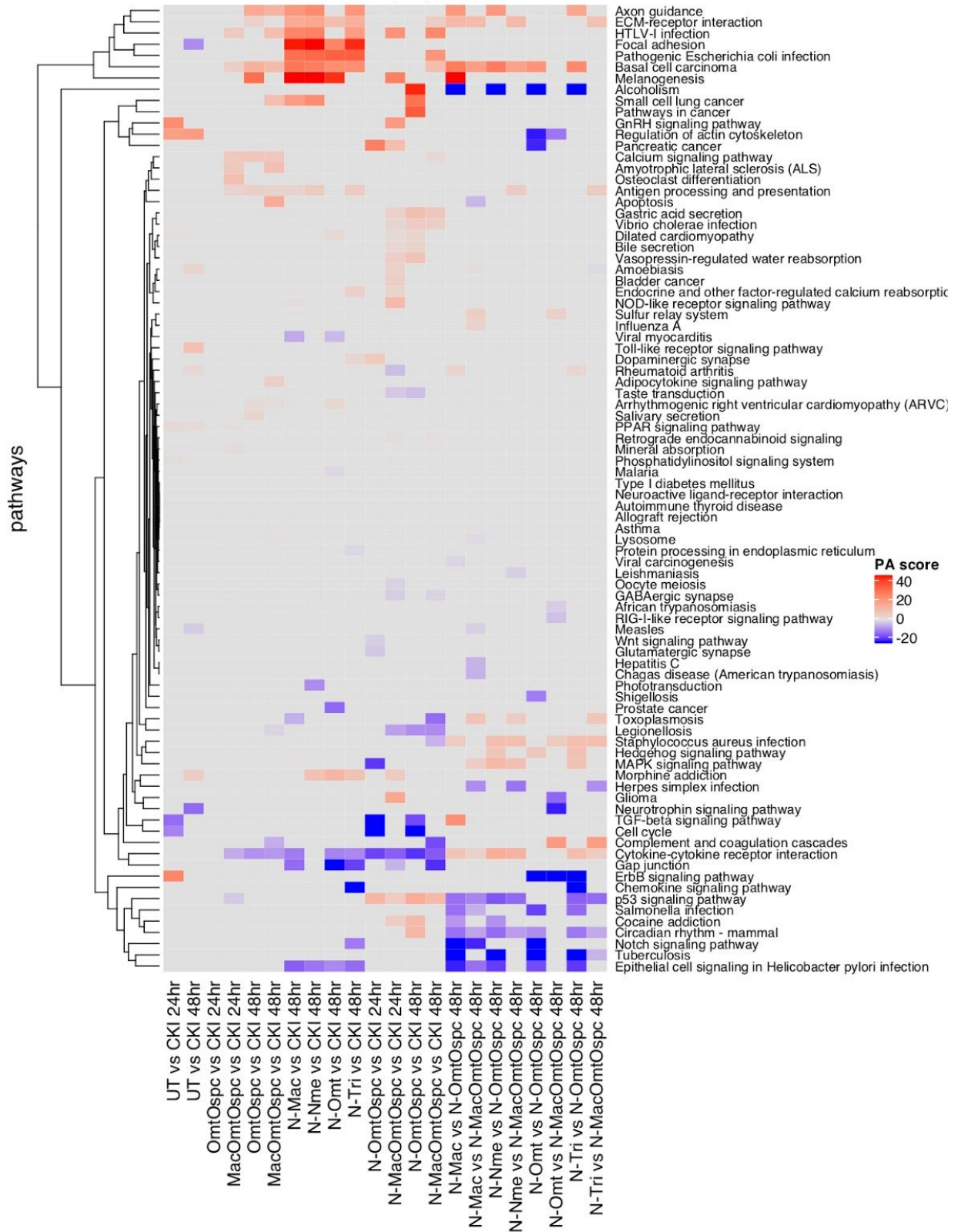








### Perturbation Accumulation(PA) between Treatments



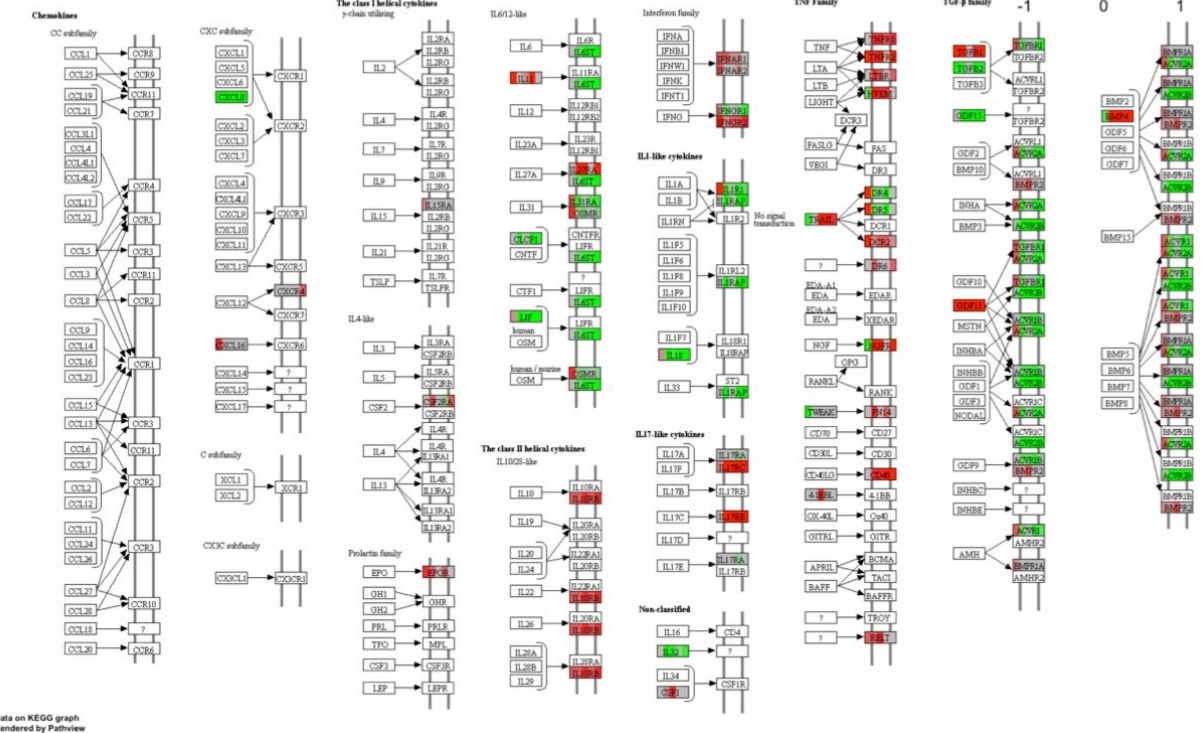




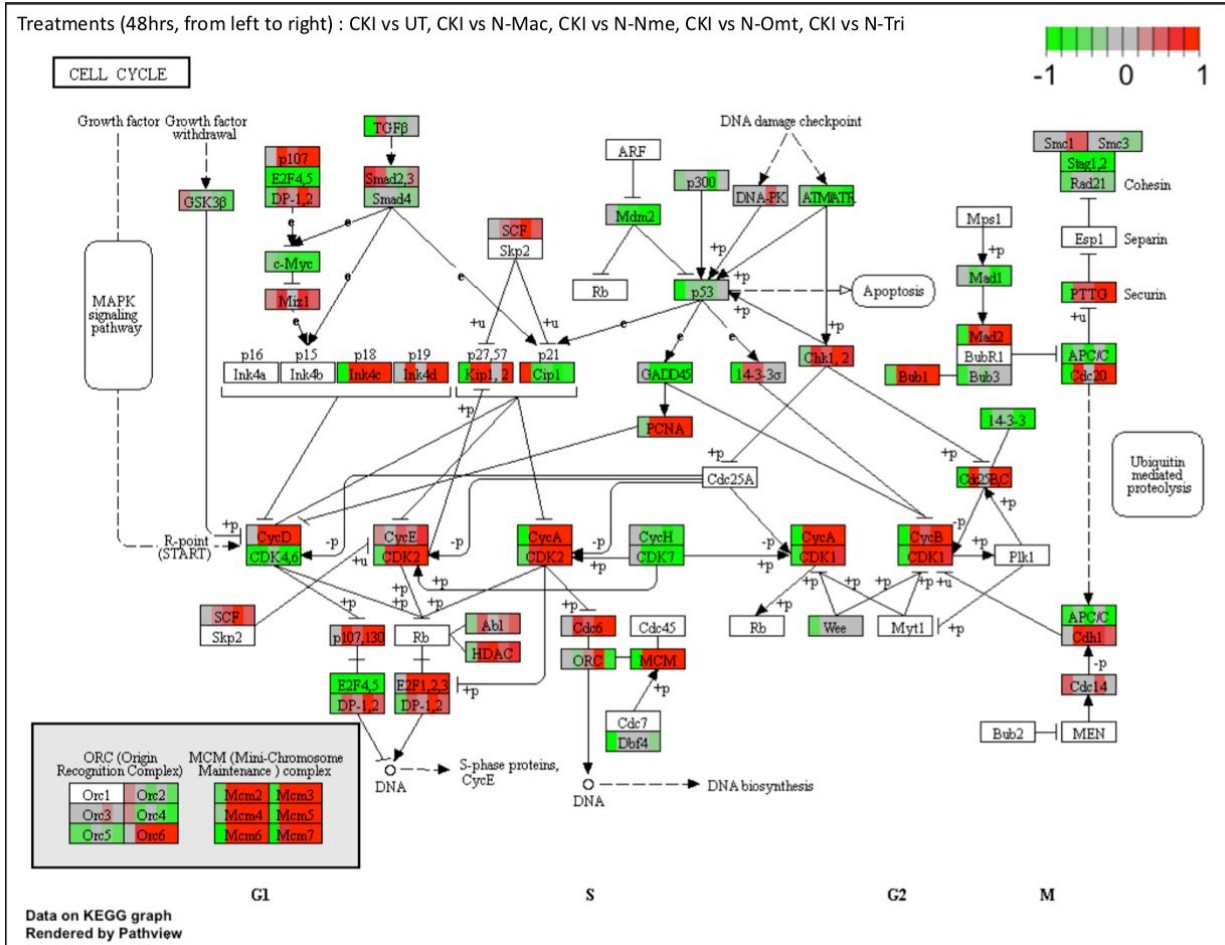


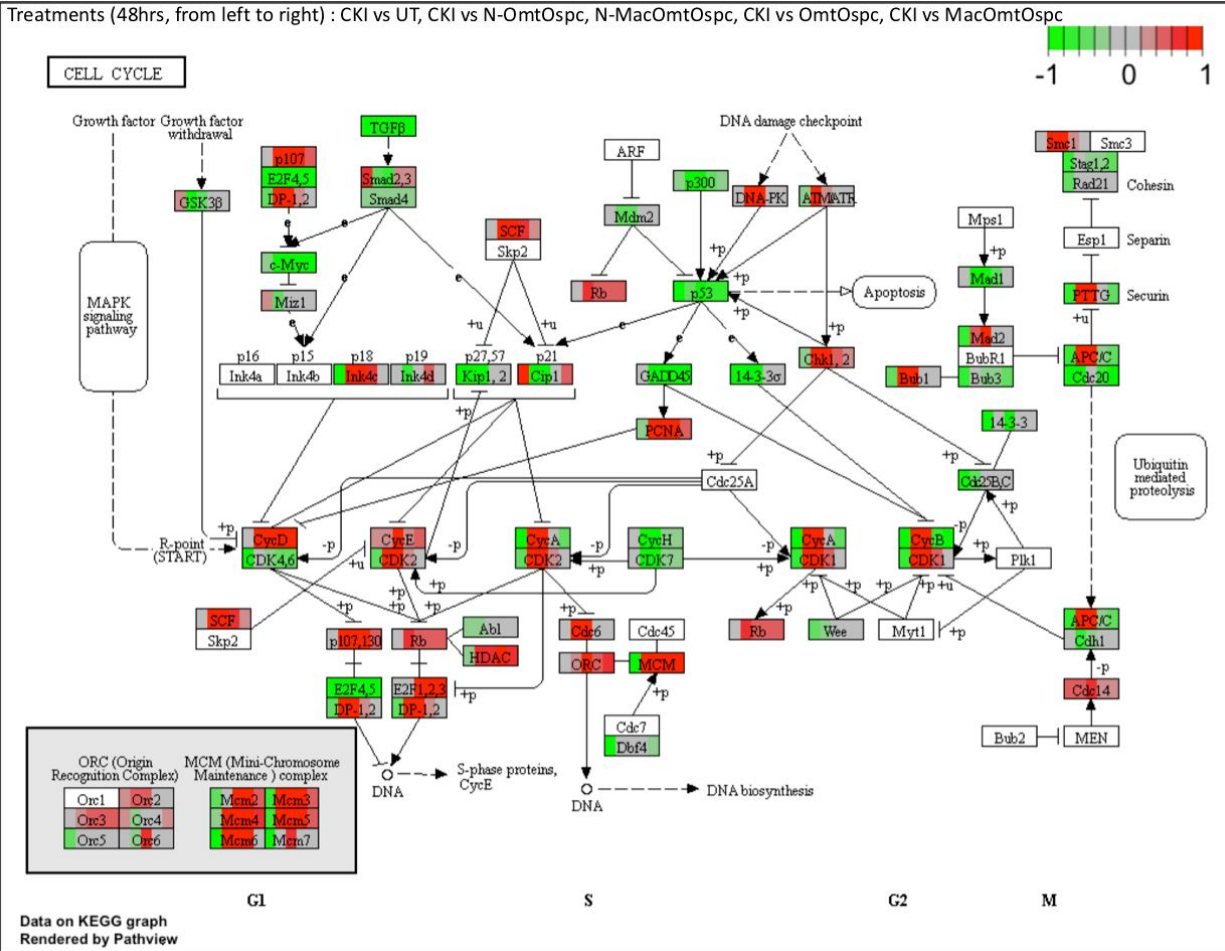
Treatments (24hrs, from left to right) : CKI vs UT, CKI vs N-OmtOspc, N-MacOmtOspc, CKI vs OmtOspc, CKI vs MacOmtOspc

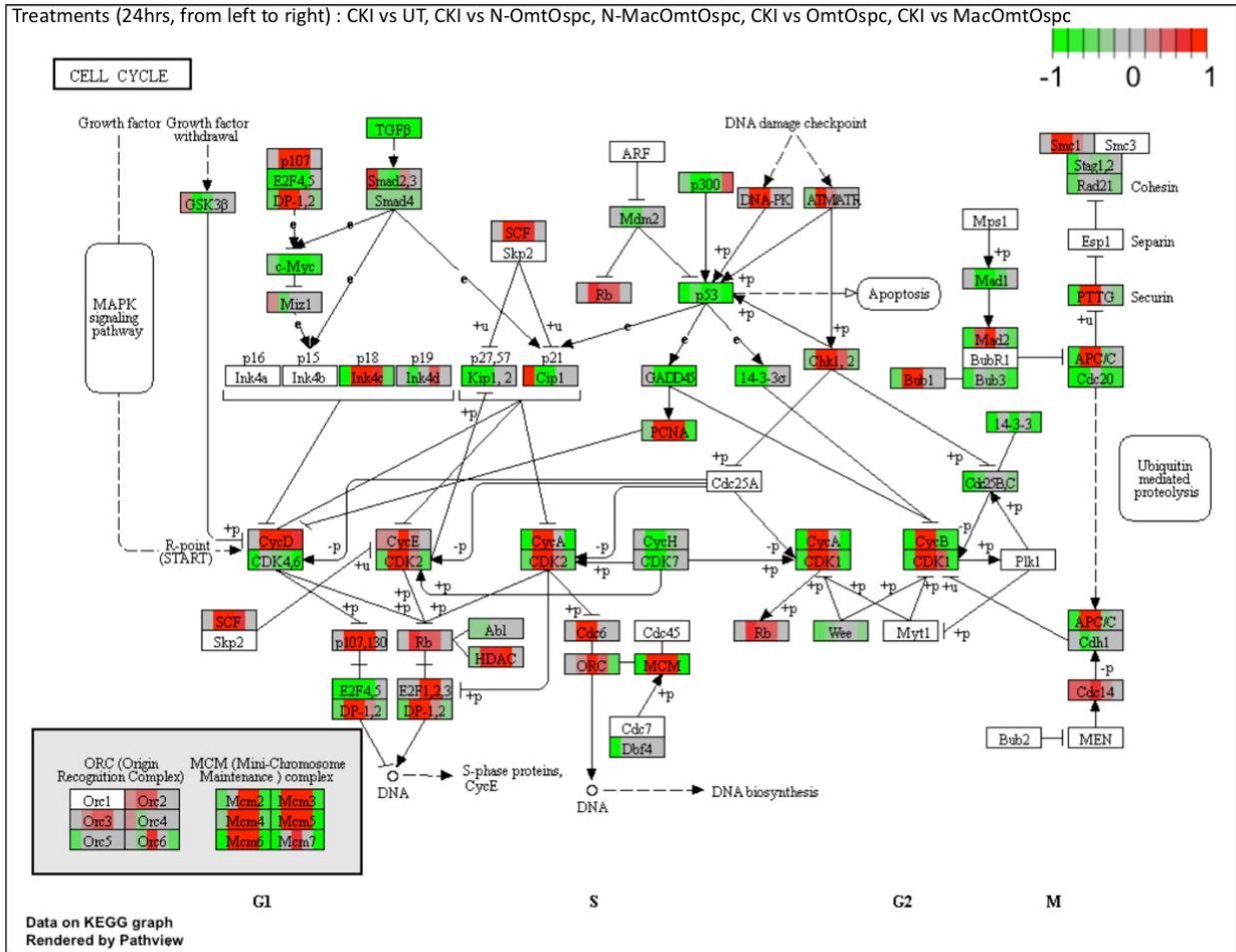
CYTOKINE-CYTOKINE RECEPTOR INTERACTION

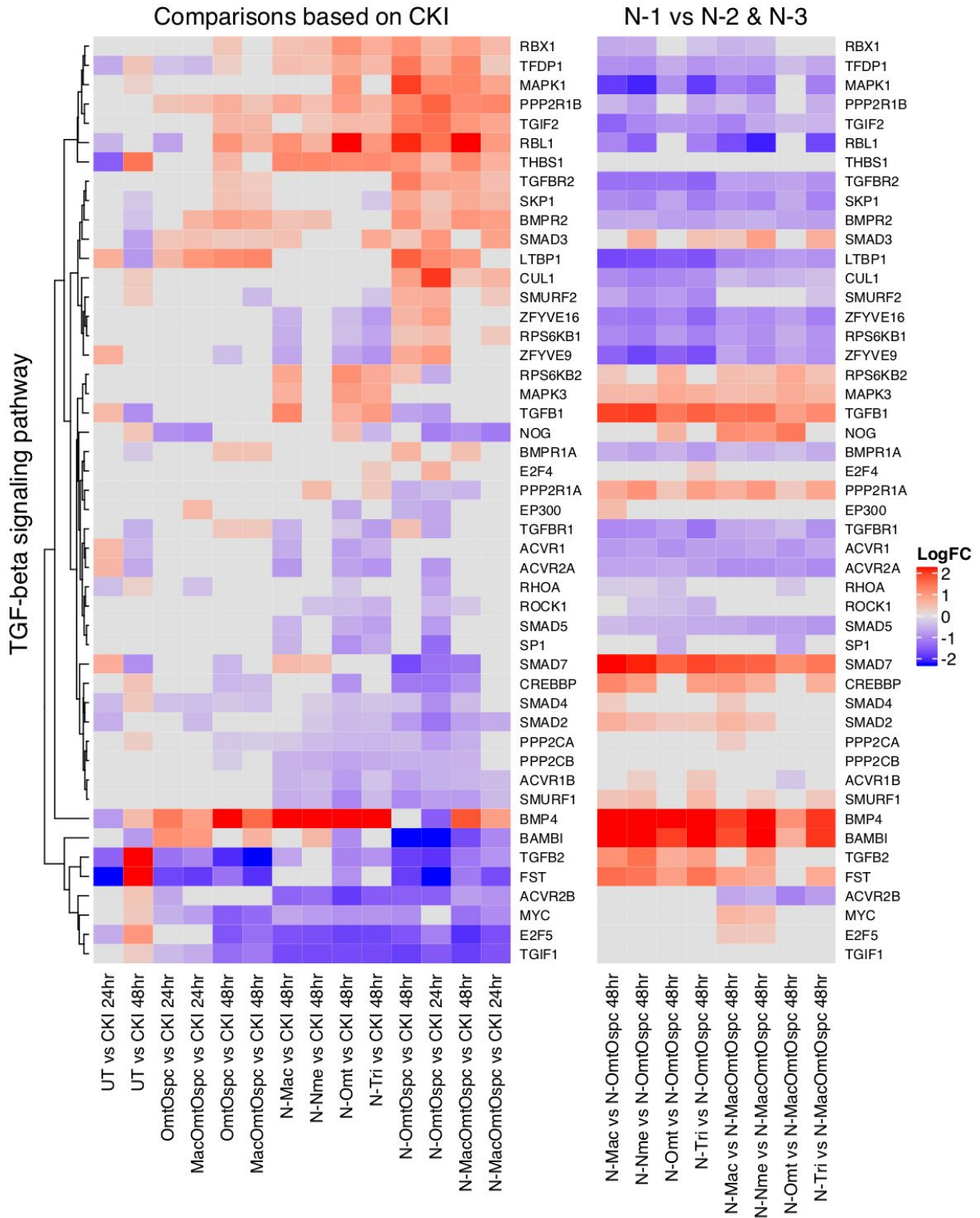


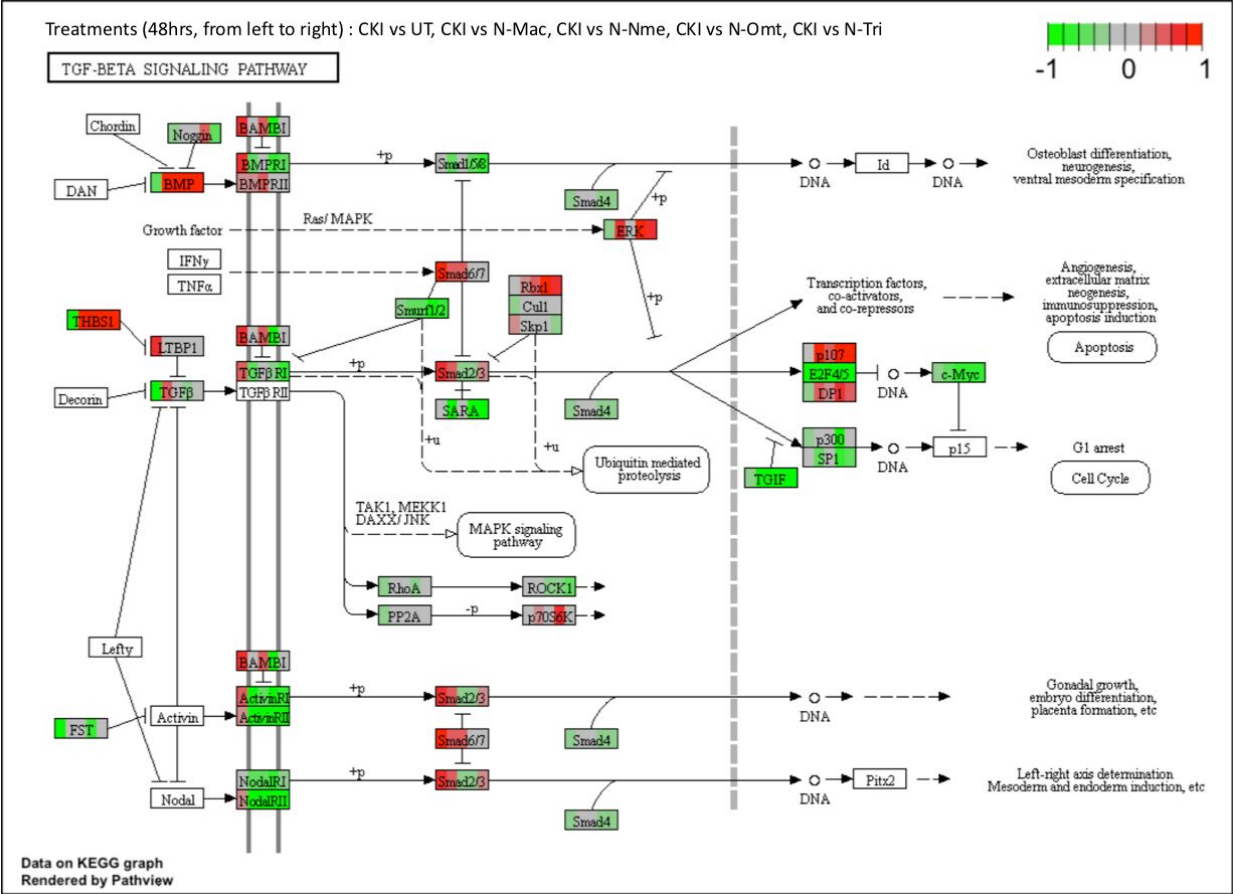
Treatments (48hrs, from left to right) : CKI vs UT, CKI vs N-Mac, CKI vs N-Nme, CKI vs N-Omt, CKI vs N-Tri



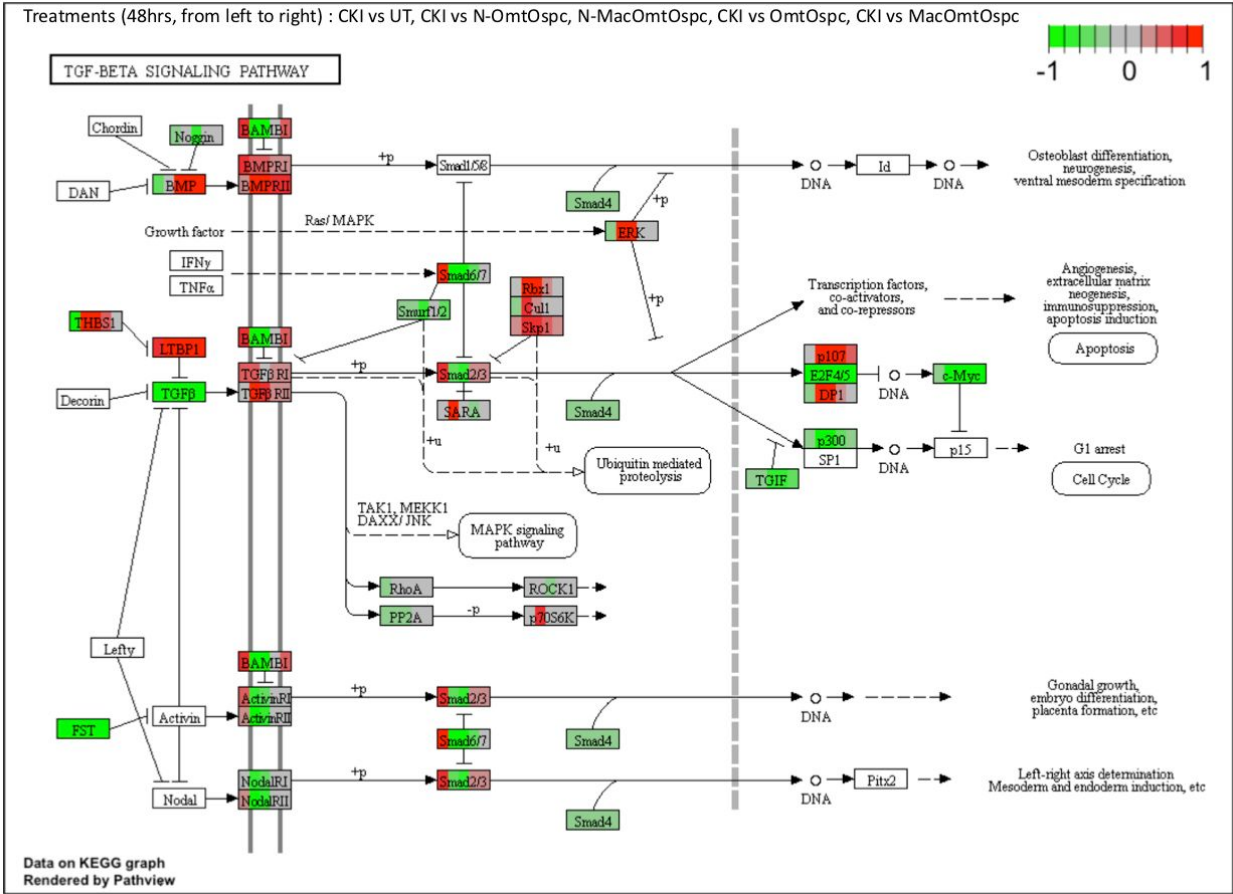








Treatments (48hrs, from left to right) : CKI vs UT, CKI vs N-OmtOspc, N-MacOmtOspc, CKI vs OmtOspc, CKI vs MacOmtOspc



Treatments (24hrs, from left to right) : CKI vs UT, CKI vs N-OmtOspc, N-MacOmtOspc, CKI vs OmtOspc, CKI vs MacOmtOspc

