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[1]†, Nourmohammadi S[2]†, Qu Z [1], Harata-Lee Y [1], Cui J[1], Shen HY[1], Yool AJ[2], Pukala T[3], Hong Du[4], Kortschak RD[1], and Adelson DL*[1].

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

(2) Adelaide Medical School, University of Adelaide, Adelaide, South Australia, 5005.

(3) School of Physical Sciences, University of Adelaide, Adelaide, South Australia, 5005.

(4) 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**

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 most 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 transcriptomics 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 cancer patients (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 4mg/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 cervical cancer, prostate cancer, synovial sarcoma, and hepatocellular carcinoma (Wu et al., 2015;Cai 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;Chan et al., 2014). 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 cytisine 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

Cell viability

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 (Figure2A).

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 Figure2B).

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 the number of viable cells significantly more than CKI 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, from N-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 (Figure 2C, D).

Cell cycle

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 any of the N-1 treatments compared to CKI (Figure 3A). However, N-OmtOspc treatment significantly altered the cell cycle for MDA-MB-231 cells and induced significant higher apoptosis from 0.25mg/mL through 2mg/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 3[c]).

Apoptosis

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 2mg/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 2mg/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 (Figure 3F and Supplementary Figure 3A).

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**<0.01) at 48-hour and showed no significant apoptotic induction in HFF (Figure 3A, B and Supplementary Figure 3B, C).

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.

Cytotoxicity

Because of the significant decreased viability accompanied by increased apoptosis triggered by subtractive fractions, cytotoxicity tests were carried out for HEK-293 and HFF cells using N-OmtOspc and N-MacOmtOspc subtractive fractions at concentrations equivalent to CKI 2mg/mL . N-OmtOspc and N-MacOmtOspc at equivalent concentration to CKI 2mg/mL were significantly cytotoxic to both non-cancerous cell lines (Figure 3G).

Overall, these results indicate 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 kill 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 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 4A 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 4B). 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 4B) 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 Figure 8, 9 and Supplementary Table3). 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 CKI treatment was compared to the four (N-1) treatments 42.8% of DE genes were shared, and when CKI was compared to N-OmtOspc and N- MacOmtOspc treatments, 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 identify over-represented genes in these shared genes. The only common function enriched across all comparisons was for "cell cycle checkpoint". This confirmed earlier results (Qu et al., 2016) and was consistent with the phenotype data for CKI.

Subtracted fractionation altered pathways

We also used 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 (SuppFig 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 5). By comparing the pathway gene expression perturbation scores

(pG) between treatments we can see that three specific observations can 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-hour, 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 hypothesise 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 causes more cell death than CKI.

More detailed examination of some of these interactions within significantly perturbed pathways highlights the gene-specific changes in expression for some key regulators of of inflammation and the cell-cycle (Figure 6). 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 6A) 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 6B) 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)), TGFB1 (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 and induce apoptosis in cancer cell lines (Xu et al., 2011;XU et al., 2013;Qu et al., 2016). CKI arrest of the cell cycle 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, oxymatrine was previously shown to arrest the cell cycle and induce apoptosis in primary human myeloma cells through caspase-dependent pathway (Wu et al., 2014) and inhibit the proliferation of laryngeal squamous cell carcinoma Hep-2 cells (Ying et al., 2015). As shown in this report, oxymatrine or oxysophocarpine or combined OmtOspc treatment caused no significant change in cell viability, the cell cycle or apoptosis, in agreement with prior work showing that oxymatrine and oxysophocarpine showed no significant effect on apoptosis, cell cycle or cell proliferation in HCT116 human colon cancer cells (Zhang et al., 2014). However, at the level of gene expression, gene ontology analysis indicated that “cell cycle checkpoint” was significantly

enriched in genes expressed in cells treated with all fractionated mixtures or mixtures of Omt and Ospc. Our results demonstrate 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" 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(Tu et al., 2016). 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.

The paradoxical result that removal of OmtOspc caused a striking increase in apoptosis is most simply explained by a model that assumes an integreation of effects of multiple compounds on many targets. 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. Anrea J 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 37ºC with 5% CO_2.

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), 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 a reparative C_{18} column (5 µm, 250 x 10 mm, Phenomenex) with the following mobile phases: 0.01 M ammonium acetate (adjusted to pH 8.0, solvent A) and acetonitrile + 0.09% trifluroacetic 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 carried out using MilliQ water and buffered with 10mM Hepes pH 6.8-7.0 (Gibco, Life technologies, USA). 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. 10uL 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_2O + 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,

Diamosnsil, 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 Data1.

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 50ul of medium were plated in 96 wells trays overnight prior to drug treatments in triplicate. Cells were subsequently treated with 50ul of drug mixtures to provide final concentrations of 0.25, 0.5, 2 and 2mg/mL in 100uL. Cell viability was then measured at 24 and 48 hours after drug treatment by the addition of 50 ul 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 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, $4x10^5$ 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 was 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 $2.5x10^3$ cells per well in triplicate. CKI and fractionated mixtures to produce a final concentration of 1mg/mL and 2mg/mL were added to each well and after 24 hours of incubation, viable cells were measured using the AlamarBlue assay (Molecular Probes, Eugene, OR). Mercuric chloride (Sigma-Aldrich , St. Louis, MO) (5µM) 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 $2x10^5$ cells/mL overnight prior to drug treatments. On the following day, CKI (to give a final concentration of 2mg/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 Foundatin (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 ± 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 ± standard deviation (SD).

Figure 3: Cell cycle, apoptosis and cytotoxicity assays 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) Apoptotic effect of N-OmtOspc, N-MacOmtOspc subtractive fractions, OmtOspc and MacOmtOspc in MDA-MB-231 cells as determined by FACS AnnexinV/PI assay. (E) Apoptotic effect of N-OmtOspc and N-MacOmtOspc subtractive fractions, OmtOspc and MacOmtOspc in HEK-293 cells as determined by FACS AnnexinV/PI assay. (F) Apoptotic effect of N-OmtOspc and N-MacOmtOspc subtractive fractions, OmtOspc and MacOmtOspc in HFF cells as determined by FACS AnnexinV/PI assay. (G) Cytotoxic effect of CKI, N-OmtOspc and N-MacOmtOspc subtractive fractions were 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 ± standard deviation (SD).

Figure 4: 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 5: 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 gens out of all significant DE gene from each treatment involved in the particular terms. (B) Identification of significantly perturbed pathways using SPIA (pG< 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 6: Differential gene expression profiles of all treatments for two highly perturbed pathways: (A) Cytokine-Cytokine Receptor pathway and (B) Cell Cycle pathway. In each 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 1: HPLC profiles of 25 mixtures including CKI, MJ, MN, 9 (N-1), 4 (N-3) and 9 (N-2). 50uL of the samples at 1mg/mL concentration was injecteded 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 1mg/mL or 2mg/mL of CKI and 2mg/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_2O) (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_2O 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 ± 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 ± 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 overlaping DE genes between treatments at 24-hours and 48-hours.

Supplementary Figure 10: Identification of significantly perturbed pathways using SPIA (pG< 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-ß 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-ß 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-ß 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-ß 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.

Author contributions

T.N.A, S.N, Z.Q and D.L.A designed the study, analysed the data and wrote the manuscript. Y.H-L, J.C, H.S, T.P, 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	*	*			
9known 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					
	N-OspcSpc					
3 compounds	MacAdeTri					
	MtNmeSpr					
	OmtOspcSpc					
	MacOmtOspc					
2 compounds	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)	% Contribution
CKI (N)	Macrozamin	y = 0.0192x + 0.2501	0.9879	0.3890	1.467924528
	Adenine	y = 0.021x + 0.0074	0.9935	0.1602	0.604528302
	N-methylcytisine	y = 0.0937x + 0.042	0.9895	0.1892	0.713962264
	Sophoridine	y = 0.1443x + 0.2679	0.9872	0.4685	1.767924528
	Matrine	y = 0.0132x + 0.7512	0.9929	1.4612	5.513962264
	Sophocarpine	y = 0.0343x + 0.3974	0.994	0.5725	2.160377358
	Oxysophocarpine	y = 0.0371x - 0.0108	0.9994	1.0108	3.814339623
	Oxymatrine	y = 0.0132x + 0.6226	0.992	6.1290	23.12830189
	Trifolirhizin	y = 0.0026x - 0.0002	0.9903	0.0807	0.304528302
	Total			**10.4611**	**39.47584906**
N-OmtOspc	Macrozamin	y = 0.0192x + 0.2501	0.9879	0.1580	0.596226415
	Adenine	y = 0.021x + 0.0074	0.9935	0.7887	2.976226415
	N-methylcytisine	y = 0.0937x + 0.042	0.9895	0.0062	0.023396226
	Sophoridine	y = 0.1443x + 0.2679	0.9872	0.2378	0.897358491
	Matrine	y = 0.0132x + 0.7512	0.9929	1.4993	5.657735849
	Sophocarpine	y = 0.0343x + 0.3974	0.994	0.2858	1.078490566
	Trifolirhizin	y = 0.0026x - 0.0002	0.9903	0.0696	0.262641509
	Total			**3.0454**	**11.49207547**
N-MacOmtOspc	Adenine	y = 0.021x + 0.0074	0.9935	0.1579	0.595849057
	N-methylcytisine	y = 0.0937x + 0.042	0.9895	0.1550	0.58490566
	Sophoridine	y = 0.1443x + 0.2679	0.9872	0.3742	1.412075472
	Matrine	y = 0.0132x + 0.7512	0.9929	1.3276	5.009811321
	Sophocarpine	y = 0.0343x + 0.3974	0.994	0.2383	0.899245283
	Trifolirhizin	y = 0.0026x - 0.0002	0.9903	0.0816	0.307924528
	Total			**2.3346**	**8.809811321**

*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.

Chan, K., Zhang, H., and Lin, Z. (2014). "Treatments used in complementary and alternative medicine," in *Side Effects of Drugs Annual*. Elsevier), 889-898.

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.

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.

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 fromrhizoma smilacis glabrae. *Pharmacological research* 36**,** 309-314.

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.

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, W., Wu, C., Zi, F., Li, Y., Yang, L., He, J., He, D., Yang, Y., Zhao, Y., and Han, X. (2014). "Oxymatrine Induces Apoptosis in Human Myeloma Cells By Cell Cycle Arrest and Activation of Caspase-Dependent Apoptosis". Am Soc Hematology).

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, J., Zhao, S., Wang, A., and Wang, W. (2013). Clinical Observation of Compound Kushen Injection Combined With FOLFOX Regimen in Treatment of Advanced Gastric Cancer [J]. *Chinese Archives of Traditional Chinese Medicine* 8**,** 084.

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.

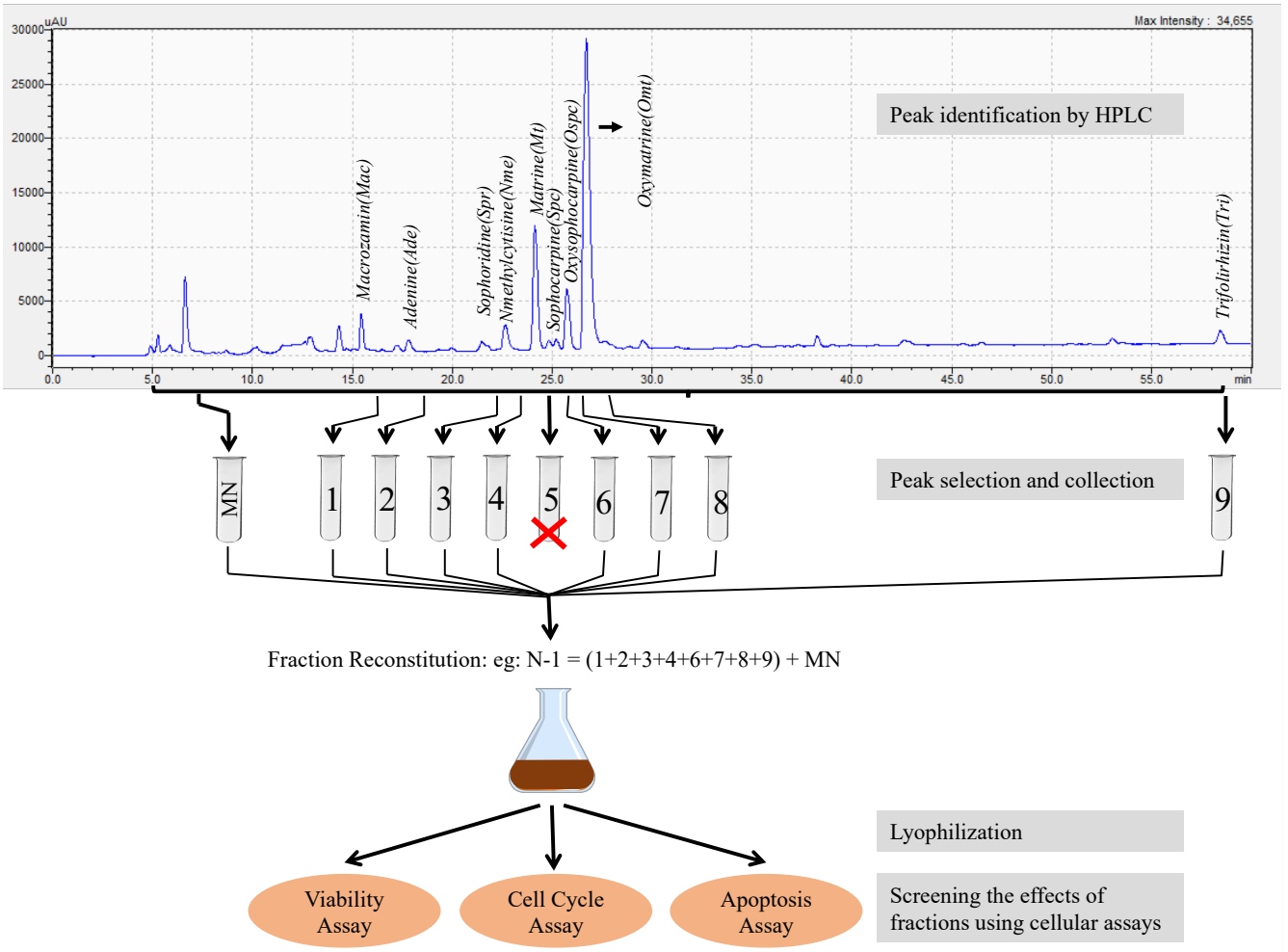
Ying, X.-J., Jin, B., Chen, X.-W., Xie, J., Xu, H.-M., and Dong, P. (2015). Oxymatrine downregulates HPV16E7 expression and inhibits cell proliferation in laryngeal squamous cell carcinoma Hep-2 cells in vitro. *BioMed research international* 2015.

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.

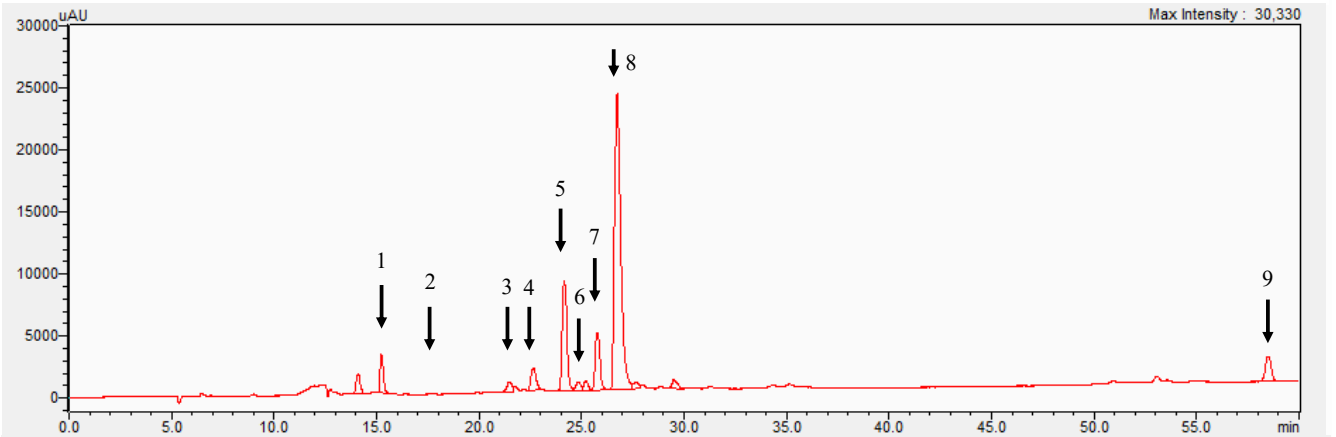
Zhang, L., Zheng, Y., Deng, H., Liang, L., and Peng, J. (2014). Aloperine induces G2/M phase cell cycle arrest and apoptosis in HCT116 human colon cancer cells. *International journal of molecular medicine* 33**,** 1613-1620.

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.

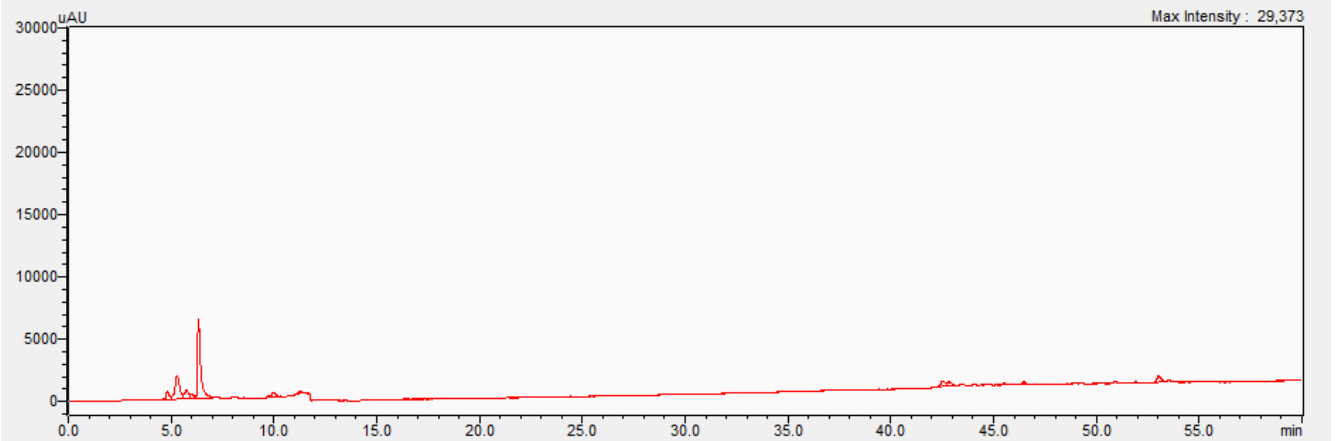
A

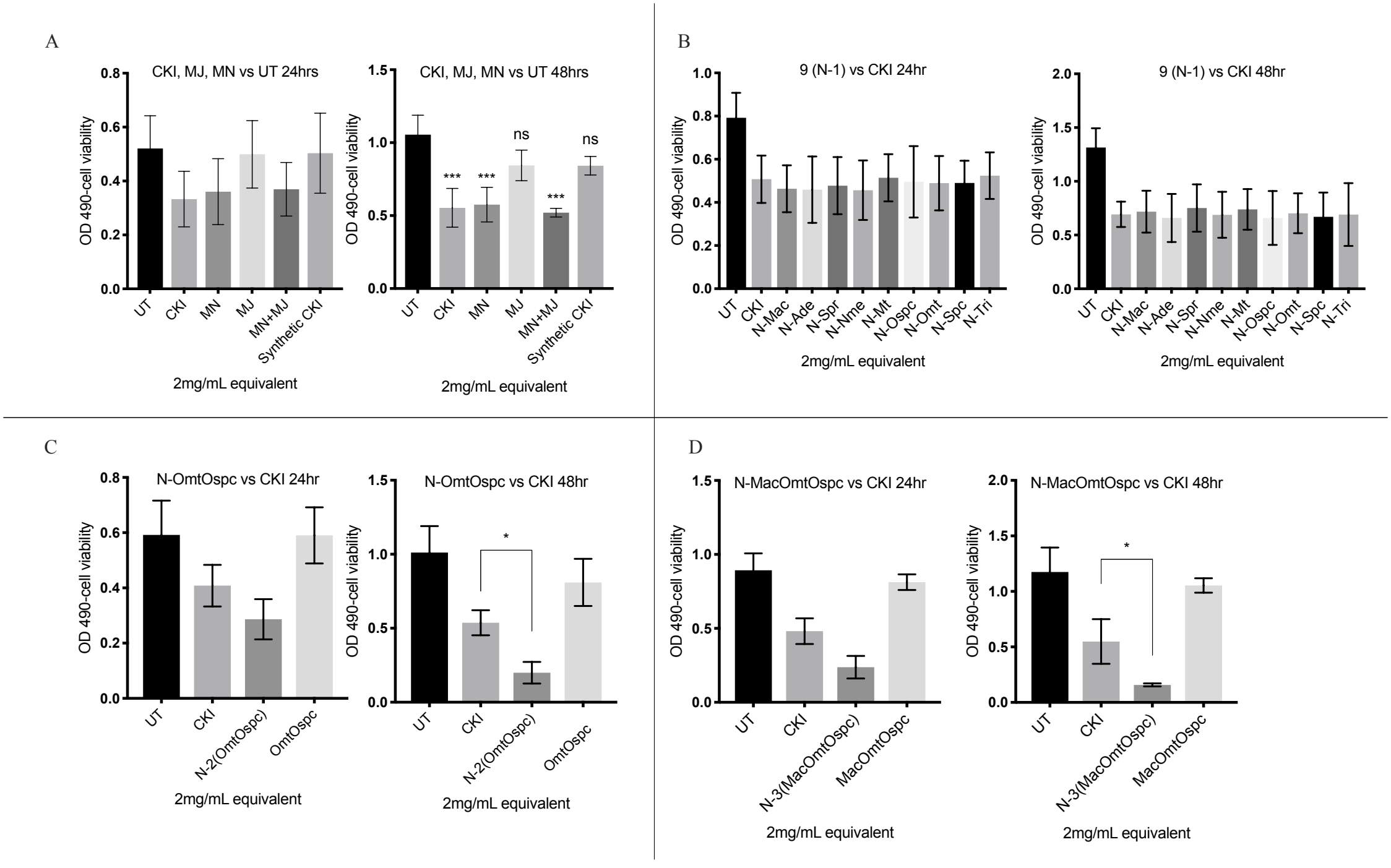

uAU
Max Intensity : 34,655
Macrozamin(Mac)
Adenine(Ade)
Sophoridine(Spr)
Nmethylcytisine(Nme)
Matrine(Mt)
Sophocarpine(Spc)
Oxysophocarpine(Ospc)
Oxymatrine(Omt)
Trifolirhizin(Tri)
Peak identification by HPLC
MN
1
2
3
4
5
6
7
8
9
Peak selection and collection
Fraction Reconstitution: eg: N-1 = (1+2+3+4+6+7+8+9) + MN
Lyophilization
Viability Assay
Cell Cycle Assay
Apoptosis Assay
Screening the effects of fractions using cellular assays

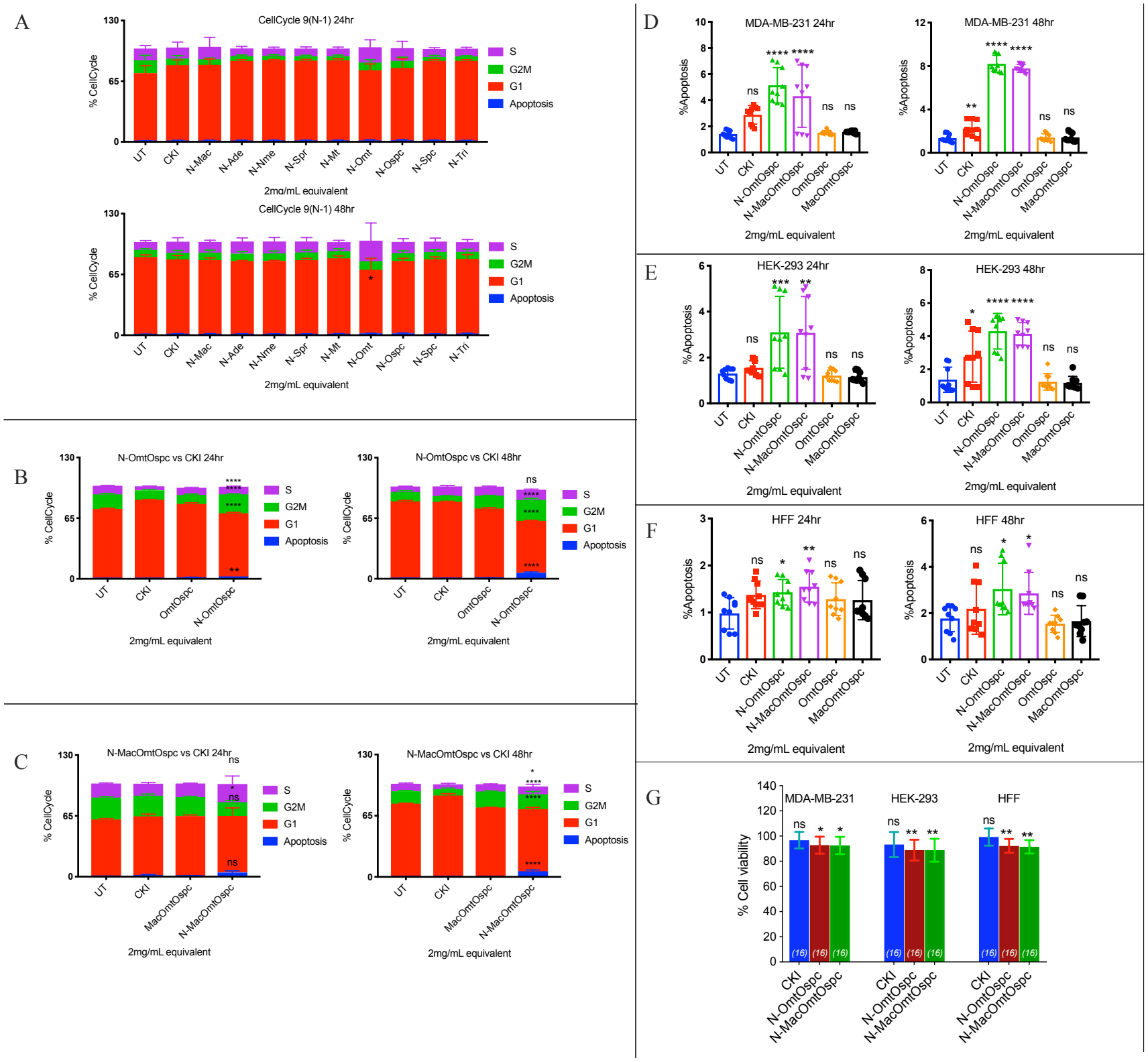
B


uAU
Max Intensity : 30,330
1
2
3
4
5
6
7
8
9
min

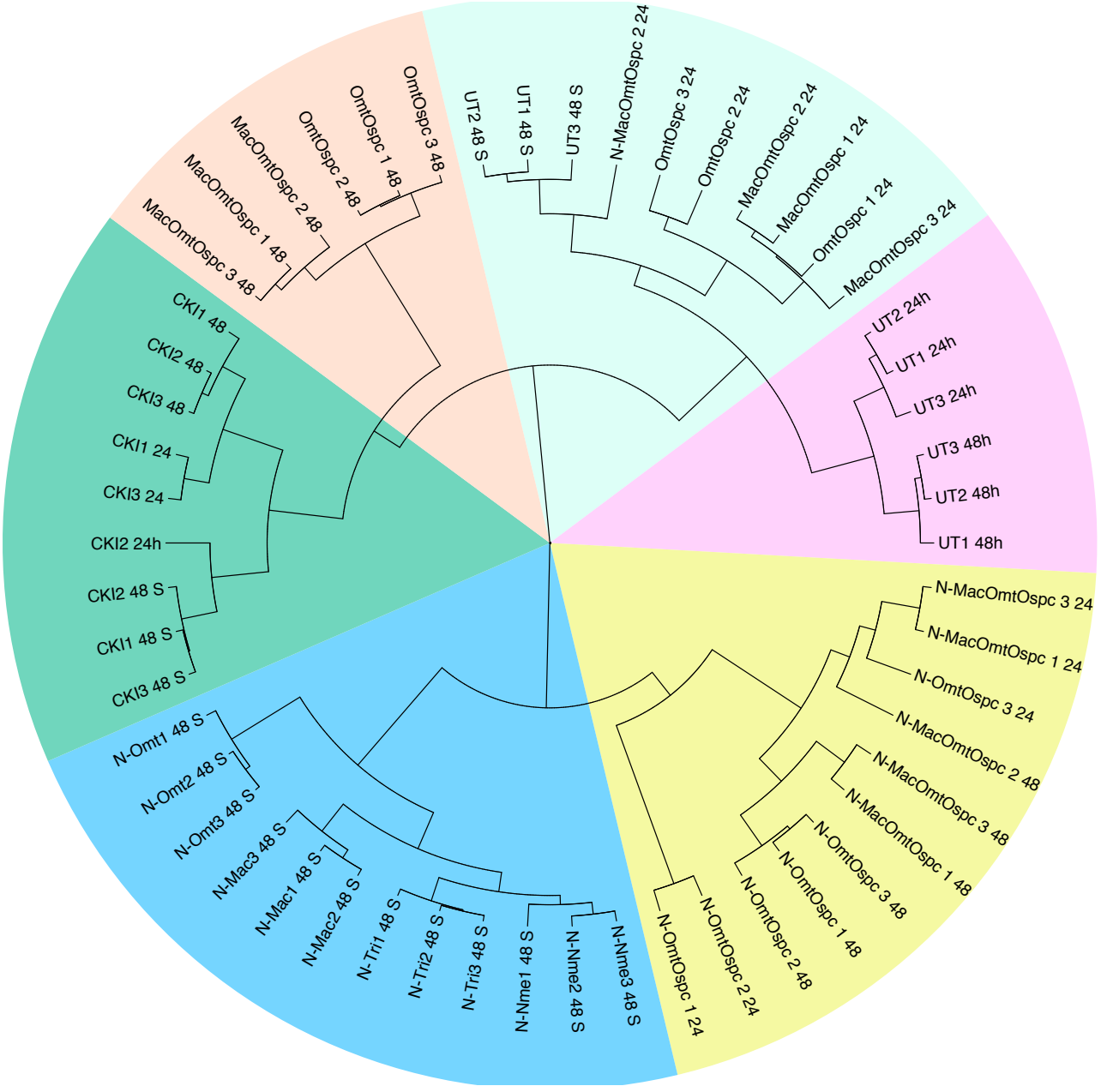
C


uAU
Max Intensity : 29,373
min


A
CKI, MJ, MN vs UT 24hrs
CKI, MJ, MN vs UT 48hrs
OD 490-cell viability
0.0
0.2
0.4
0.6
0.8
0.5
1.0
1.5
***
ns
UT
CKI
MN
MJ
MN+MJ
Synthetic CKI
2mg/mL equivalent
B
9 (N-1) vs CKI 24hr
9 (N-1) vs CKI 48hr
2.0
N-Mac
N-Ade
N-Spr
N-Nme
N-Mt
N-Ospc
N-Omt
N-Spc
N-Tri
C
N-OmtOspc vs CKI 24hr
N-OmtOspc vs CKI 48hr
*
N-2(OmtOspc)
OmtOspc
D
N-MacOmtOspc vs CKI 24hr
N-MacOmtOspc vs CKI 48hr
N-3(MacOmtOspc)
MacOmtOspc


A
CellCycle 9(N-1) 24hr
% CellCycle
130
65
0
S
G2M
G1
Apoptosis
UT
CKI
N-Mac
N-Ade
N-Nme
N-Spr
N-Mt
N-Omt
N-Ospc
N-Spc
N-Tri
2mg/mL equivalent
CellCycle 9(N-1) 48hr
B
N-OmtOspc vs CKI 24hr
N-OmtOspc vs CKI 48hr
OmtOspc
N-OmtOspc
C
N-MacOmtOspc vs CKI 24hr
N-MacOmtOspc vs CKI 48hr
MacOmtOspc
N-MacOmtOspc
D
MDA-MB-231 24hr
MDA-MB-231 48hr
%Apoptosis
N-OmtOspc
N-MacOmtOspc
OmtOspc
MacOmtOspc
E
HEK-293 24hr
HEK-293 48hr
F
HFF 24hr
HFF 48hr
G
MDA-MB-231
HEK-293
HFF
% Cell viability
(16)

A



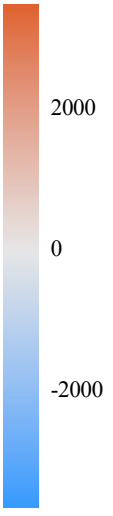
B

Pairwise comparison of DE genes 24-hour

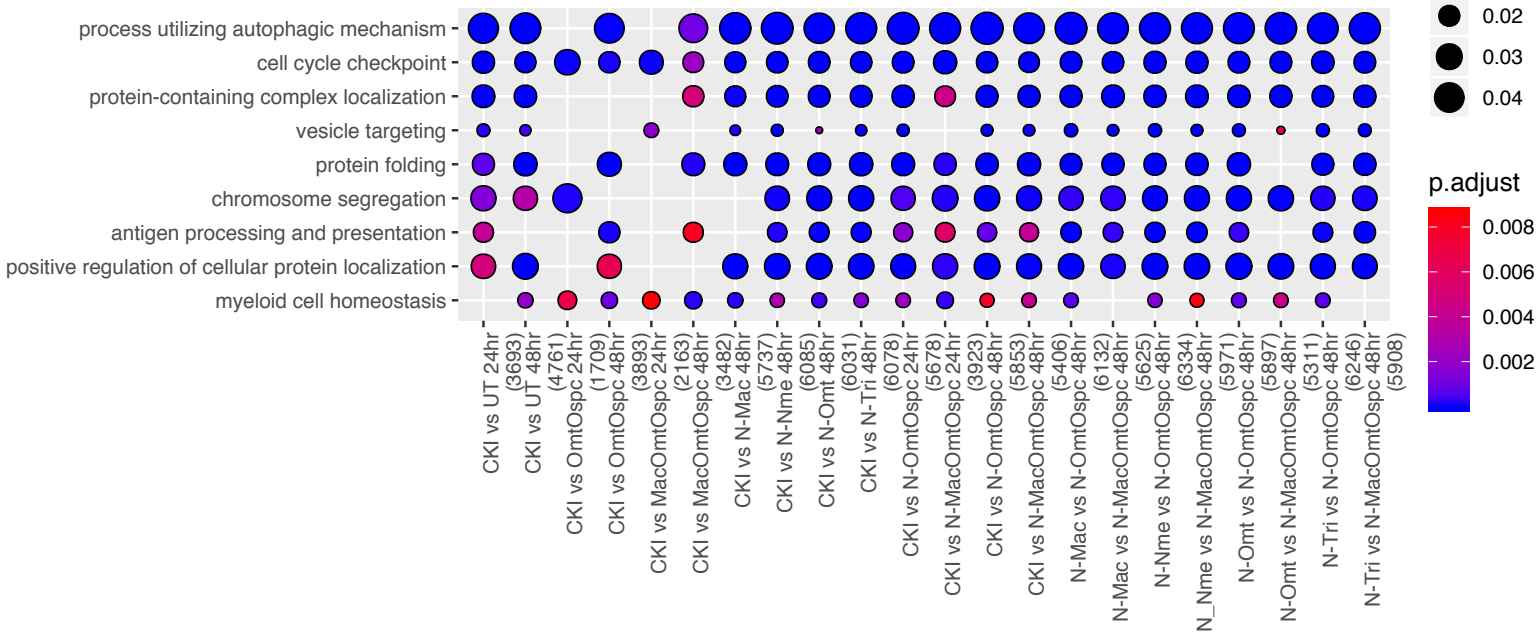
	UT 24hr	CKI 24hr	N-OmtOspc 24hr	N-MacOmtOspc 24hr	OmtOspc 24hr	MacOmtOspc 24hr
UT 24hr	0	2061	3247	2597	1533	957
CKI 24hr	-2015	0	3107	2197	1019	1257
N-OmtOspc 24hr	-3250	-3128	0	2815	3241	3167
N-MacOmtOspc 24hr	-2427	-2120	-2709	0	2225	2024
OmtOspc 24hr	-1305	-866	-3160	-2244	0	0
MacOmtOspc 24hr	-855	-1138	-3149	-2216	0	0

Pairwise comparison of DE genes 48-hour

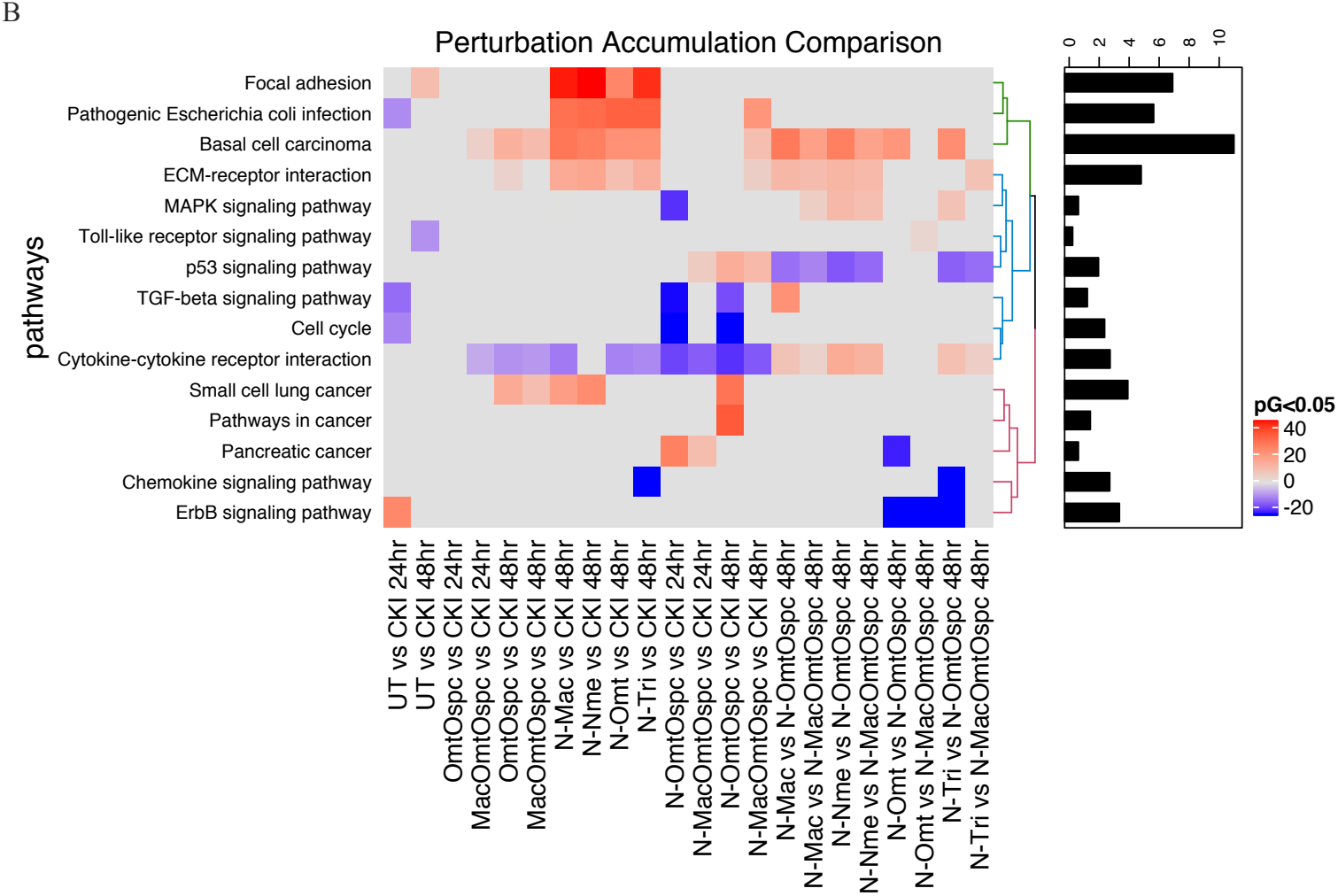
	UT 48hr	CKI 48hr	N-Mac 48hr	N-Nme 48hr	N-Omt 48hr	N-Tri 48hr	N-OmtOspc 48hr	N-MacOmtOspc 48hr	OmtOspc 48hr	MacOmtOspc 48hr
UT 48hr	0	2638	3330	3424	3440	3437	3272	3007	2159	1704
CKI 48hr	-2682	0	3206	3454	3332	3435	3197	2937	2170	1969
N-Mac 48hr	-3214	-3113	0	59	1230	121	3177	2855	2705	2782
N-Nme 48hr	-3275	-3248	-158	0	1634	7	3306	3102	2961	2985
N-Omt 48hr	-3382	-3301	-1395	-1832	0	1762	3016	2720	2988	3048
N-Tri 48hr	-3254	-3224	-221	-7	-1431	0	3286	3097	2917	2936
N-OmtOspc 48hr	-3404	-3247	-3577	-3668	-3461	-3571	0	1577	2898	2937
N-MacOmtOspc 48hr	-3048	-2996	-3318	-3452	-3098	-3377	-1555	0	2153	2914
OmtOspc 48hr	-2201	-2106	-2884	-3165	-3046	-3059	-2769	-2070	0	1
MacOmtOspc 48hr	-1743	-1865	-2979	-3170	-3083	-3081	-2836	-2859	0	0

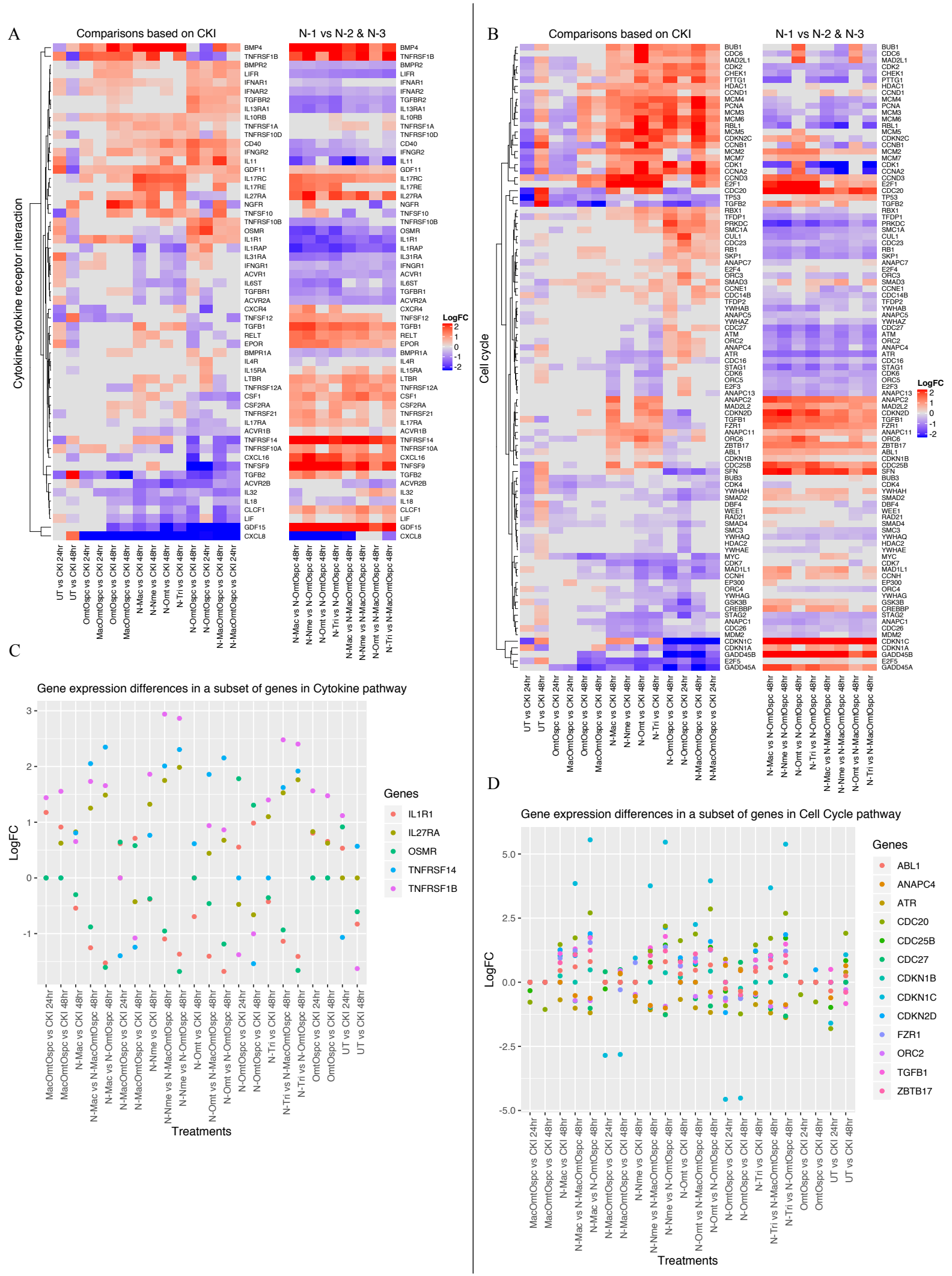


A


process utilizing autophagic mechanism
cell cycle checkpoint
protein-containing complex localization
vesicle targeting
protein folding
chromosome segregation
antigen processing and presentation
positive regulation of cellular protein localization
myeloid cell homeostasis
CKI vs UT 24hr (3693)
CKI vs UT 48hr (4761)
CKI vs OmtOspc 24hr (1709)
CKI vs OmtOspc 48hr (3893)
CKI vs MacOmtOspc 24hr (2163)
CKI vs MacOmtOspc 48hr (3482)
CKI vs N-Mac 48hr (5737)
CKI vs N-Nme 48hr (6085)
CKI vs N-Omt 48hr (6031)
CKI vs N-Tri 48hr (6078)
CKI vs N-OmtOspc 24hr (5678)
CKI vs N-MacOmtOspc 24hr (3923)
CKI vs N-OmtOspc 48hr (5853)
CKI vs N-MacOmtOspc 48hr (5406)
N-Mac vs N-OmtOspc 48hr (6132)
N-Mac vs N-MacOmtOspc 48hr (5625)
N-Nme vs N-OmtOspc 48hr (6334)
N_Nme vs N-MacOmtOspc 48hr (5971)
N-Omt vs N-OmtOspc 48hr (5897)
N-Omt vs N-MacOmtOspc 48hr (5311)
N-Tri vs N-OmtOspc 48hr (6246)
N-Tri vs N-MacOmtOspc 48hr (5908)
0.02
0.03
0.04
p.adjust
0.008
0.006
0.004
0.002

B


Perturbation Accumulation Comparison
pathways
Focal adhesion
Pathogenic Escherichia coli infection
Basal cell carcinoma
ECM-receptor interaction
MAPK signaling pathway
Toll-like receptor signaling pathway
p53 signaling pathway
TGF-beta signaling pathway
Cell cycle
Cytokine-cytokine receptor interaction
Small cell lung cancer
Pathways in cancer
Pancreatic cancer
Chemokine signaling pathway
ErbB signaling pathway
UT vs CKI 24hr
UT vs CKI 48hr
OmtOspc vs CKI 24hr
MacOmtOspc vs CKI 24hr
OmtOspc vs CKI 48hr
MacOmtOspc vs CKI 48hr
N-Mac vs CKI 48hr
N-Nme vs CKI 48hr
N-Omt vs CKI 48hr
N-Tri vs CKI 48hr
N-OmtOspc vs CKI 24hr
N-MacOmtOspc vs CKI 24hr
N-OmtOspc vs CKI 48hr
N-MacOmtOspc vs CKI 48hr
N-Mac vs N-OmtOspc 48hr
N-Mac vs N-MacOmtOspc 48hr
N-Nme vs N-OmtOspc 48hr
N-Nme vs N-MacOmtOspc 48hr
N-Omt vs N-OmtOspc 48hr
N-Omt vs N-MacOmtOspc 48hr
N-Tri vs N-OmtOspc 48hr
N-Tri vs N-MacOmtOspc 48hr
0
2
4
6
8
10
pG<0.05
40
20
0
-20


A
Comparisons based on CKI
N-1 vs N-2 & N-3
Cytokine-cytokine receptor interaction
LogFC
B
Comparisons based on CKI
N-1 vs N-2 & N-3
Cell cycle
LogFC
C
Gene expression differences in a subset of genes in Cytokine pathway
LogFC
Treatments
Genes
IL1R1
IL27RA
OSMR
TNFRSF14
TNFRSF1B
D
Gene expression differences in a subset of genes in Cell Cycle pathway
LogFC
Treatments
Genes
ABL1
ANAPC4
ATR
CDC20
CDC25B
CDC27
CDKN1B
CDKN1C
CDKN2D
FZR1
ORC2
TGFB1
ZBTB17