Multi-omics analysis in mouse primary cortical neurons reveals complex positive and negative biological interactions between constituent compounds in Centella asiatica

Steve Chamberlin* [1,2,3], Jonathan A Zweig[2,3], Cody J Neff[2,3], Luke Marney[3,4], Jaewoo Choi[3,4], Liping Yang[3,4] , Claudia S Maier[3,4], Amala Soumyanath[2,3], Shannon McWeeney[1,5], Nora E Gray[2,3]

[1]Department of Medical Informatics and Clinical Epidemiology, Oregon Health & Science University, Portland, OR; [2] Department of Neurology, Oregon Health & Science University, Portland, OR; [3] BENFRA Botanical Dietary Supplements Research Center, Oregon Health & Science University (OHSU) Portland, OR; [4] Department of Chemistry, Oregon State University, Corvallis, OR; [5] Knight Cancer Institute, Oregon Health & Science University, Portland, OR;

Corresponding author: chambest@ohsu.edu

Abstract

Background: A water extract of the Ayurvedic plant *Centella asiatica* (CAW) improves cognitive function in mouse models of aging and Alzheimer's disease, and affects dendritic arborization, mitochondrial activity and oxidative stress in mouse primary neurons. Triterpenes (TT) and caffeoylquinic acids (CQA) are constituents associated with these bioactivities of CAW although little is known about how interactions between these compounds contribute to the plant's therapeutic benefit.

Methods: Mouse primary cortical neurons were treated with CAW, or equivalent concentrations of four TT combined, eight CQA combined, or these twelve compounds combined (TTCQA). Treatment effects on the cell transcriptome (18,491 genes) and metabolome (192 metabolites) relative to vehicle control were evaluated using RNAseq and metabolomic analyses respectively.

Results: Extensive differentially expressed genes (DEGs) were seen with all treatments, as well as evidence of interactions between compounds. Notably many DEGs seen with TT treatment were not observed in the TTCQA condition, possibly suggesting CQA reduced the effects of TT. Moreover, additional gene activity seen with CAW as compared to TTCQA indicate the presence of additional compounds in CAW that further modulate TTCQA interactions. Weighted Gene Correlation Network Analysis (WGCNA) identified 4 gene co-expression modules altered by treatments that were associated with extracellular matrix organization, fatty acid metabolism, cellular response to stress and stimuli, and immune function. Compound interaction patterns were seen at the eigengene level in these modules. Interestingly, in metabolomics analysis, the TTCQA treatment saw the highest number of changes in individual metabolites (20), followed by CQA (15), then TT (8) and finally CAW (3). WGCNA analysis found two metabolomics modules with significant eigenmetabolite differences for TT and CQA, and possible compound interactions at this level.

Conclusions: Four gene expression modules and two metabolite modules were altered by the four types of treatments applied. This methodology demonstrated the existence of both negative and positive interactions between TT, CQA and additional compounds found in CAW on the transcriptome and metabolome of mouse primary cortical neurons.

Keywords:

A. Introduction

Traditional and complementary medicines are often sought to improve or maintain health, and it is estimated that 50% of the population in industrialized countries use these approaches [1]. In the elderly, these interventions are primarily used to promote resilience [2] since many of these are botanical extracts with purported anti-aging effects [3-6]. However, claims are controversial and few clinical trials have been done with these treatments [7].

The study of botanical extracts for healthcare is challenging, in part due to the complex mix of compounds found in these plants. Unlike single-chemical drugs or isolated natural products, botanical extracts may contain hundreds of compounds [8]. Many traditional practioners believe that these botanical mixtures are more clinically effective than isolated compounds due to interactions between the constituents. Interactions can occur between individual compounds, groups or fractions of compounds, constituents from different parts of the plant or even between compounds derived from different species as seen in complex traditional herbal formulas [9]. These mixtures can interact with biological systems in a way that is additive, synergistic, or antagonistic [9-12]. Fraction-based methods have been developed to evaluate synergy between compounds in whole plant extracts using a combination of mass spectrometry, isolation of natural products and synergy assays [13]. In a whole organism, the nature of these interactions can vary depending on the compounds, and be considered pharmacokinetic or pharmacodynamic. Pharmacokinetic interactions affect the bioavailability of a constituent in an organism through changes in its metabolism or cellular transport induced by another constituent. Pharmacodynamic interactions occur when two compounds affect each others' binding to targets, or have similar or opposing biological activity mediated through interactions at different points in molecular pathways [10, 12].

While synergy studies have traditionally utilized targeted assays, leveraging multi-omic analysis in pre-clinical models offers an effective way to understand multiple bioactivities and mechanisms of action that are modulated by this mix of compounds, including understanding the interactions between the compounds. For the current project we propose to evaluate two untargeted molecular domains, transcriptomics and metabolomics, in an *in vitro* model system, to broadly study compound interactions in the biological activity of compounds within a water extract (CAW) of the botanical *Centella asiatica*. This plant has a history of use in several systems of traditional medicine, including traditional Chinese medicine and Ayurvedic medicine [14], for treatment of various conditions including cognitive impairment [15]. Current in vitro, in vivo and clinical research also supports the neurological benefits of *Centella asiatica* [16-28]. CAW can evoke particularly potent neuroprotective effects both *in vitro* and *in vivo* that may be mediated by effects on dendritic arborization [25, 28], increased synaptic density [21, 25, 28], mitochondrial biogenesis [20, 25, 28], and activation of endogenous antioxidant mechanisms [20, 25, 28].

Two classes of compounds have been associated with the health promoting bioactivity of CAW: pentacyclic triterpenes (TT) and mono- and di-caffeoylquinic acids

(CQA) [29] although the contribution of each compound group to the overall effects of CAW, and the mechanisms by which the compounds can elicit those effects, have not been fully elucidated. The goal of this project is to begin to determine the contributions to bioactivity of these two compound classes separately or combined as well as in the complete CAW extract.

In this study, mouse primary cortical neurons were treated with either CAW, groups of TT or CQA compounds separately, or the combination of the two (TTCQA). All compounds were applied at concentrations equivalent to their presence in CAW. Transcriptomic and metabolomic data was collected from these treated cells. This design allowed us to assess the effects of, and nature of interactions between, the TT and CQA compounds, as well as interactions between these compounds and other unknown compounds in CAW, at a biomolecular level.

B. Results

B.1 Analysis of culture media post-incubation

Neuron culture media was collected and frozen after the 48 hour treatment period at the time of cell harvest. The cultured media was then extracted with OstroPlate® solid-phase extraction and concentrations of selected phytochemicals measured using liquid chromatography coupled to multiple reaction monitoring mass spectrometry (LC-MRM-MS; **Supp Table ST1**). In many of the incubations, compound concentrations were somewhat higher than the pre-incubation time zero starting value, possibly due to evaporation of solvent during incubation (**Supp Table ST1**). Interestingly post-incubation concentration of CQA metabolites was ten-fold lower in CAW incubations compared to incubations containing CQA alone or TTCQA. This suggests a rapid degradation of CQAs in the presence of other CAW constituents. CQA metabolites are known to undergo chemical changes in non-acidic solutions [30, 31]. Triterpene-glycoside content (MS and AS) was similar in all cultures containing these compounds. Triterpene aglycones (MA and AA) were present in higher concentrations post-incubation in CAW than the TT or TTCQA cultures, potentially arising from precursors present in CAW, as they were considerably higher than the time 0 values (**Supp Table ST1**). Differences in response to CAW compared to the compounds may in part have been due to these concentration differences.

B.2 CAW and its constituent compounds induce extensive gene expression changes

Thirty nine out of 40 of the culture samples were of sufficient quality for RNA-seq processing, displaying a high level of inter-neuronal networking and neuro-spheres (n=8 per treatment except TTCQA which had n=7). RNA extracted from the samples was of high quality with an average RNA Integrity Number (RIN) of 9.9. For the metabolomics processing we had 50 cell culture samples of sufficient quality (n=10 per treatment). To investigate both differentially expressed genes (DEG) and differentially abundant metabolites (DAM), we compared each of the four treatment groups to the vehicle control (CAW vs Ctrl, TT vs Ctrl, CQA vs Ctrl and TTCQA vs Ctrl). RNA-seq genes with expression less than 10 counts in any of the samples were removed, leaving 18,491

genes for analysis. Our in-house library provided 192 metabolites for analysis. A false discovery rate cutoff of 0.05 was used for both DEG and DAM analyses. While there was a high degree of treatment related gene expression change in the neurons (CAW 946 upregulated, 1,721 downregulated; CQA 33 upregulated, 165 downregulated; TT 514 upregulated, 1.246 downregulated; TTCQA 303 upregulated, 678 downregulated), based on fold change **(Supp Table ST2)**, the same was not true for changes in metabolites. In contrast the metabolomics analysis found only 3 downregulated metabolites for CAW, 8 upregulated and 7 downregulated metabolites for CQA, 3 upregulated and 5 downregulated metabolites for TT, and 12 upregulated and 8 downregulated metabolites for TTCQA **(Supp Table ST3)**.This difference in magnitude of DEGs as compared to DAMs was largely because the full transcriptome (18,491) was evaluated, but only a select number of metabolites (192) was assessed. When compared by the percentage of effected genes or metabolites, relative to the total number tested, the difference is not as dramatic (transcriptome: 21.8% of the 18,491 genes had a change in expression; metabolome: 16.1% of the 192 metabolites had a change in abundance).

A principal component analysis (PCA) was used to assess the effect of each treatment relative to control, and also to compare the treatments to each other, using both the transcriptomic and metabolomic data. This work was done with the PCA tools R package [32]. Metabolites and genes in the lowest decile of variability across samples were removed. There was not a strong separation of samples seen between treatments for both data types, as would be expected since some of these treatments contain the same compounds (**Fig 1A).** However, for the independent treatment vs control analyses, there is greater separation of samples as the number of differentially expressed metabolites between the groups being compared increases **(Fig 1C).** The same is generally true with the transcriptomic data **(Fig 1B).** This confirmed that the treatments were having some effect.
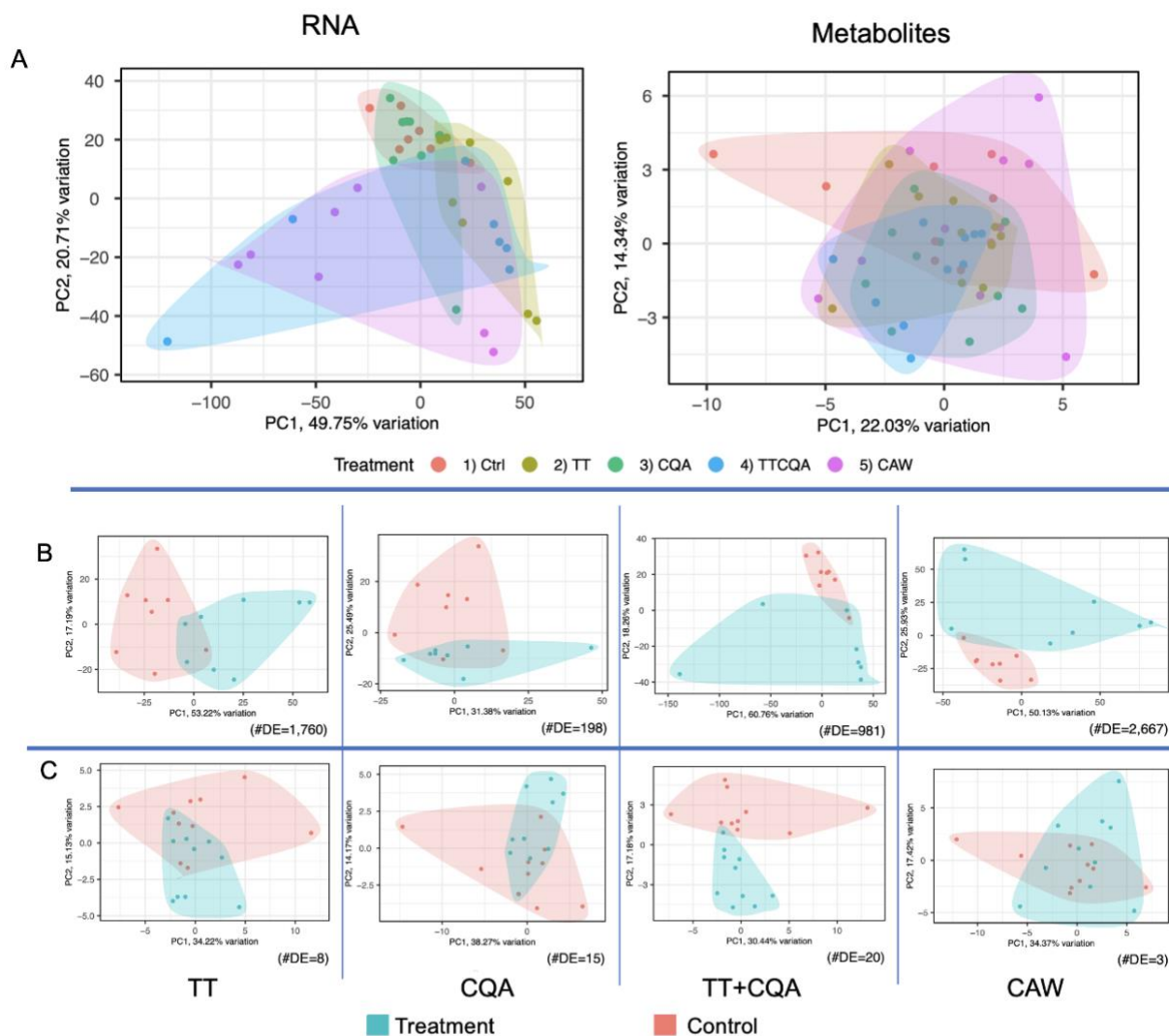
Figure 1. PCA plots for the two untargeted molecular domains of metabolomics and transcriptomics. **A.** All treatments plotted together, including the vehicle control, for either metabolic or RNA-seq data. **B.** RNA-seq data, each treatment compared individually to control. **C.** Metabolic data, each treatment compared individually to control. For **B** and **C,** the number of differentially expressed metabolites or genes (#DE) is shown in the lower right corner of each plot. The plots are also ordered left to right by increasing number of CA compounds in each treatment.

B.3 Complex interactions were seen between CAW compounds, both activating and inactivating genes.

Transcriptomic analysis revealed many interactions were apparent between constituent compounds in CAW. Only 17% of the DEGs that are seen in the TT and CQA separate treatments are also seen in the combined TTCQA treatment. (**Supp Figure SF1**), However, there are 420 new DEGs seen in this combination, of which only 263 are seen in the full CAW treatment indicating the potential 'silencing' of 157 DEGs by interactions with unknown additional compounds in CAW. Interestingly, 177 of the TT DEGs not seen in the TTCQA treatment return in the full CAW treatment, and there are
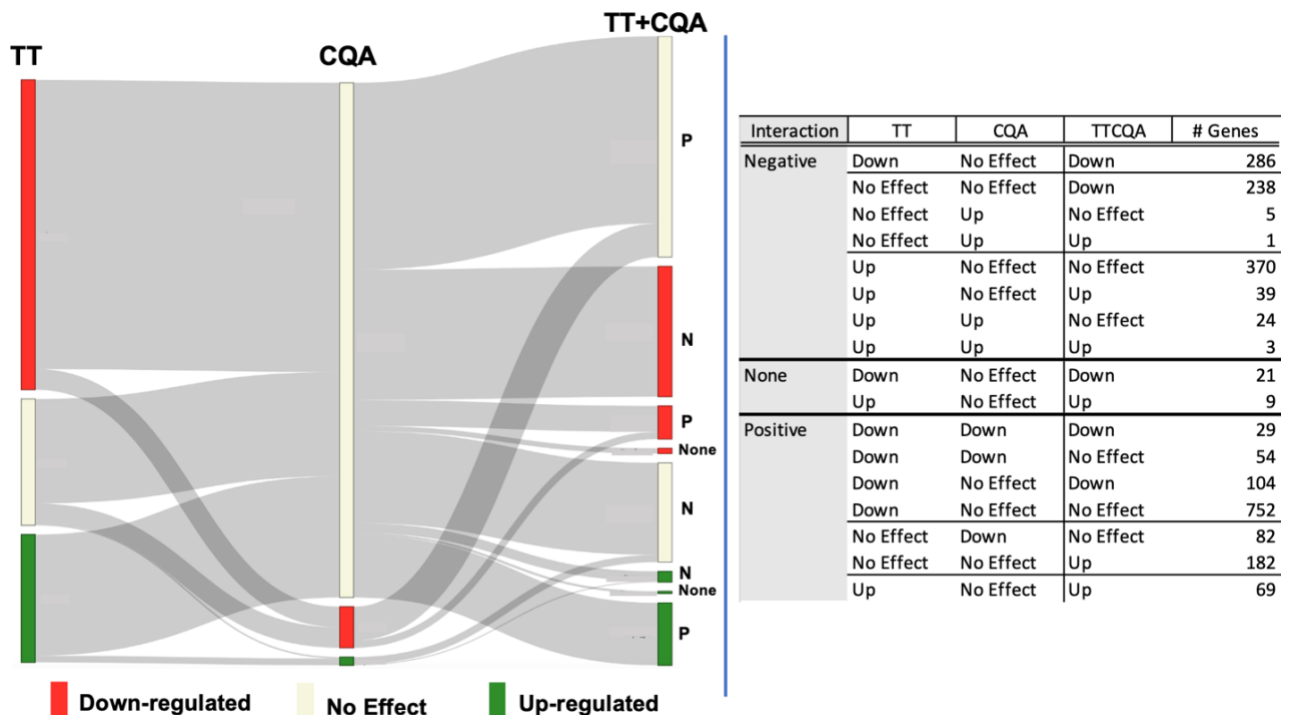
only 26 DEGs in common to all four treatments **(Supp Figure SF1A).** In the metabolomic data, only two metabolites downregulated by a constituent treatment (CQA) were seen in the full CAW treatment **(Supp Figure SF1B).**

A more detailed classification of the different types of gene expression interactions (additive, synergistic, antagonistic) from the differential gene expression model between TT and CQA can be seen in **Figure 2 (See Methods D.7).**

For this analysis, a little less than half of the genes examined had a negative interaction between TT and CQA (N=966). A substantial number of these genes had a significant effect from the TT treatment (N=728), while very few had an effect from CQA (N=33). There was an increased downregulation, in the TTCQA treatment, in 524 of these genes, 238 of which did not have a significant change from either TT or CQA (considered a negative interaction for the sake of this analysis). Other interactions reduced or neutralized significant upregulated effects seen in either TT or CQA (N=443).

We saw a larger number of genes with a positive interaction between the TT and CQA treatments (N=1,272). Most of these were either downregulated by TT, CQA or both and had their downregulation either diminished (N=154) or neutralized (N=888) in the TTCQA treatment. Interestingly there were 182 genes that were not affected by either the TT or CQA treatment that were significantly upregulated by TTCQA mixture.

Finally, there were 30 genes that did not have any interaction between the TT and CQA treatment, all of which were either upregulated (N=21) or downregulated (N=9) by only TT (**Fig 2**). A table of the parameter estimates for each treatment can be found in **(Supp Table ST4)** for the TT and CQA gene expression interaction classification.

Interaction	TT	CQA	TTCQA	# Genes
Negative	Down	No Effect	Down	286
	No Effect	No Effect	Down	238
	No Effect	Up	No Effect	5
	No Effect	Up	Up	1
	Up	No Effect	No Effect	370
	Up	No Effect	Up	39
	Up	Up	No Effect	24
	Up	Up	Up	3
None	Down	No Effect	Down	21
	Up	No Effect	Up	9
Positive	Down	Down	Down	29
	Down	Down	No Effect	54
	Down	No Effect	Down	104
	Down	No Effect	No Effect	752
	No Effect	Down	No Effect	82
	No Effect	No Effect	Up	182
	Up	No Effect	Up	69

Down-regulated   No Effect   Up-regulated

Figure 2. Effect on gene expression of the TT and CQA compounds administered separately and interactions between these groups observed in the TTCQA treated samples. The figure on the left shows a mapping of the interactions listed in the table on the right. Any gene with significant expression changes, relative to control, seen with any of the TT, CQA or TTCQA treatments is represented in the figure and table (N=2,268 unique genes). The relative number of genes and the expression status is represented within each of the three columns labeled TT, CQA or TTCQA (red=significant downregulation, green=significant upregulation, beige=no significant expression changes seen with this treatment). The TT and CQA compound 'Interaction effect' for the combined group is shown next to the TTCQA column ('P"=positive, or synergistic, 'N'=negative, or antagonistic, 'none'=additive, or no interaction). The gray ribbons show the numbers of genes interacting between TT and CQA, the interaction status, and the final expression status in the TTCQA treatment. The table on the right shows the number of genes in each category. As an example, a positive interaction has a higher fold change value in the TTCQA treatment than the sum of the two fold change values from TT and CQA, and a negative interaction has a lower TTCQA fold change than the sum of TT and CQA fold changes. So a positive interaction could still result in a downregulated or unregulated TTCQA gene and a negative interaction could result in an upregulated or unregulated TTCQA gene.

B.4 Multiple modules capturing a diverse variety of compound interactions point to distinct cellular mechanisms affected by CAW and its constituent compounds in the transcriptomic data.

Weighted Correlation Gene Network Analysis (WCGNA) was used to find groups of genes expressed in a correlated manner across the 39 samples, irrespective of DEG status, indicating probable participation in a common biological function. The soft thresholding power was set at 12 to create scale-free topology in the weighted co-expression network. The 18,491 genes clustered into 14 modules (**See Methods D.8**). A least squares linear method with a post-hoc Tukey multiple comparison test was used to identify treatment differences, relative to control, within these modules (FDR cutoff of 0.05). Eight of these modules showed significant eigengene differences between at least one of the treatments and the control. To functionally classify these modules, we performed an over-representation analysis using a hypergeometric test to see if the number of genes from a module mapped to a pathway was greater than would be expected by random chance. Genes in each module were used to enrich *Mus musculus* pathways from the Reactome Pathway Database using the ReactomePA package in R [33]. Genes in five of the eight modules significantly enriched molecular pathways in Reactome. Sample eigengene distributions for these five modules are shown in **Figure 3** by treatment group.

The size of the five modules varied from 256 to 2,490 genes. Module 1 (M1) was enriched for Extracellular Matrix Organization and Collagen Biosynthesis pathways. CAW was the only treatment with a significant effect on this eigengene (70% up-regulated DEG's) (**Fig 3, Module 1**). Fatty Acid Metabolism pathways were associated with Module 2 (M2), and three treatments show a significant effect (TT, TTCQA and CAW; 80% down-regulated DEG's) (**Fig 3, Module 2**). Module 3 (M3) enriched pathways included Cellular Response to Stress and Stimuli and was significantly impacted by the TT treatment (61% up-regulated DEG's)(**Fig 3, Module 3**). Module 4

(M4) enriched pathways are primarily involved in Immune System functions. Here again only the TT treatment showed a significant impact (90% down-regulated DEG's). (**Fig 3, Module 4**). Enriched pathways for Module 5 (M5) are mostly associated with Electron Transport and Mitochondrial Biogenesis, only affected by the CQA treatment (91% down-regulated DEG's) (**Fig 3, Module 5**). The pathways mentioned here were the highest level. For a more detailed description of smaller enriched pathways see **Supp Table ST5**.

Module 1 of **Figure 3** also shows an associated heatmap of individual gene expression levels for this module with groups defined by hierarchical clustering. The two highest level clusters show a reversal of expression patterns associated with the treatment of significant effect (CAW), with an upregulated cluster for the treatment group showing a mostly downregulated pattern in the control group, and vice versa. For heatmaps associated with Modules 2-5 see **Supp Fig SF6**. The number of module genes in the enriched pathways were generally small in our analysis, compared to the total number in the module (M1=6%, M2=2%, M3=11%, M4=21%, M5=4%,). It is not unusual for the fraction of genes related to the primary biological function of the module to be less than 20% of all genes in that module. While not all genes in the module were associated with the enriched function(s), genes that were not in these pathways but had a strong relationship to the module could be (guilt by association) [34]. To evaluate the strength of the relationship of both enriched and non-enriched genes to the entire module, we assessed both module membership (MM) and intramodular connectivity (IC). MM, or kME, is the correlation between individual module gene expression with the eigenvalue of that module. IC, or kWithin, is the gene degree, or number of neighboring connections within the module, in the weighted gene co-expression network that we constructed. The MM of the genes in enriched pathways were evaluated using a T-test with Bonferroni correction and found that genes within four out of the five modules (M1, M2, M3, M4) had significantly higher MM scores than genes not found in enriched pathways. Interestingly, the MM of M5 was significantly lower than non-enriched pathway MM in this module, having a weak negative correlation (r=-.19 average) with the eigengene.

Next, the treatment DEGs contribution to each module were evaluated. While not all treatment DEGs were found in the five modules, the significant treatments, with respect to the eigengenes, had the highest number of significant DEGs in the module. CAW DEGs comprised 28% of the genes in M1, where all other treatment DEGs were less than 1% of this module. For M2, TT DEGs were 30% of the total module genes, TTCQA DEGs 23%, CAW DEGs 38%, and CQA DEGs were less than 1%. TT DEGs in M3 were 22% of total genes and none of the other treatments were more than 10%. In M4, TT DEGs comprised 63% of the module and none of the other treatment DEGs were more than 7% of the module. Treatment DEG contributions were lower in M5, but the significant eigengene treatment CQA still had the highest percentage of the module at 9%.

The MM measure was significantly higher for the DEGs associated with significant eigengene treatments, compared to non DEGs in the module, in two modules, M2 and M4 (M2: TT .54 vs .18, TTCQA .52 vs .22, and CAW .58 vs .12; M4: TT .68 vs .31; Bonferroni adjusted t-test). No other significant treatment DEG's in the other three modules were found to have significantly higher MM measures.

Each module displays a distinctly different pattern of CAW compound interaction at the eigengene level **(Fig 3, all modules)**. M1 (Extracellular Matrix Organization) shows an increasing effect as the number of compounds increases, reaching significance with the full CAW treatment. M2 (Fatty Acid Metabolism) shows the effect of the TT treatment carried through to the full CAW treatment. CQA does not have an effect in this module, and does not seem to diminish the TT effect, nor do the additional unknown compounds in CAW. The significant effect of TT seen in M3 (Cellular Response to Stress) seems to be diminished by CQA, and even further diminished by the unknown compounds. The same pattern is seen with the TT treatment in M4 (Immune Function), but here TT has a downregulating effect whereas it had an upregulating effect in M3. And in M5 (Electron Transport and Mitochondrial Biogenesis) only the CQA treatment had a significant effect, which seems to be negated through interactions with the TT compounds, and not further impacted by interactions with unknown compounds in CAW.

Information about the individual genes in these five modules, and the pathways they enrich, can be found in **Supp Table ST5.**
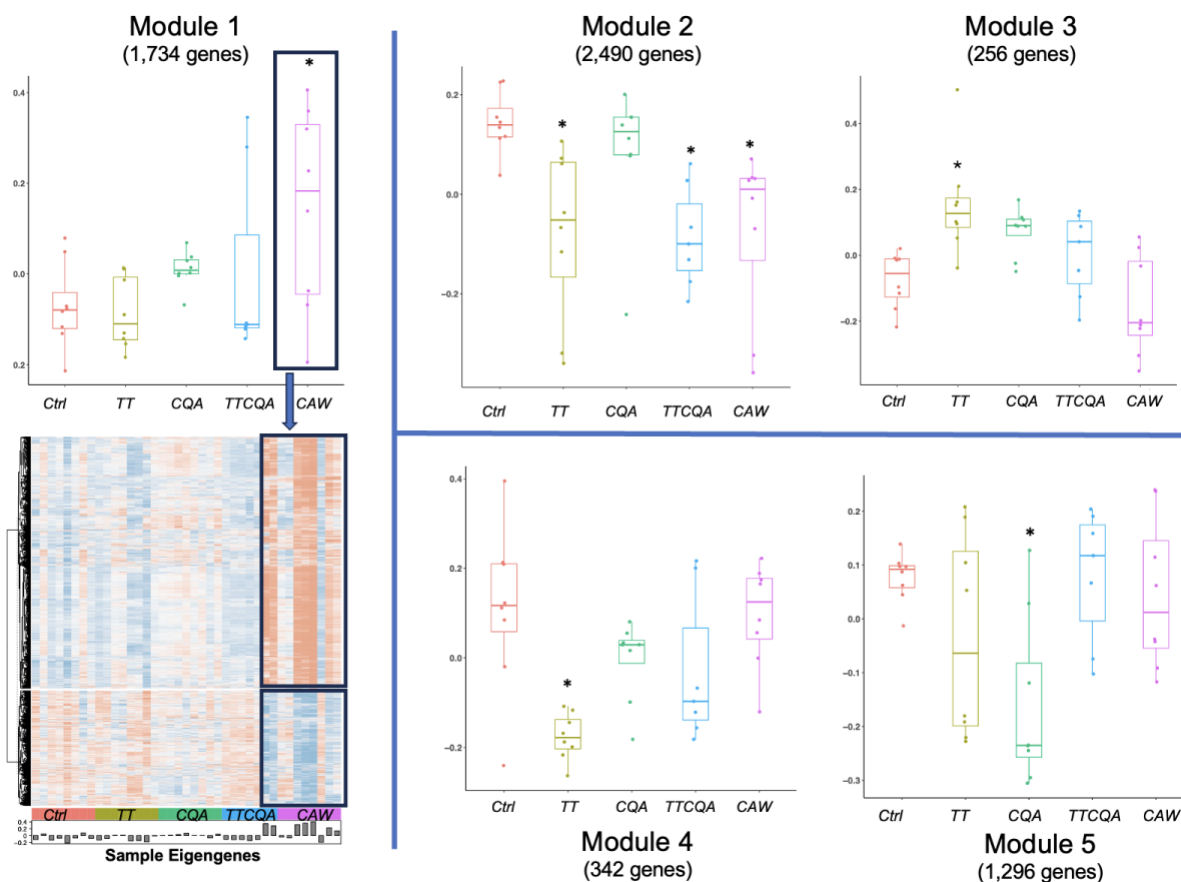
Figure 3. Sample eigengene distributions for the five gene co-expression modules with significant eigengene differences between at least one treatment and control. Boxplots for module eigengene distributions by each treatment group (* p value<=.05). Module 1=Extracellular Matrix Organization and Collagen Biosynthesis , Module 2=Fatty Acid Metabolism, Module 3=Cellular Response to Stress and Stimuli, Module 4=Immune System, Module 5=Electron Transport and Mitochondrial Biogenesis. Module 1 also shows a heat map of individual gene expression levels (blue=low, red=high). The highlighted section represents samples in the significant treatment for this module. See **Supp Figure SF6** for heatmaps for Modules 2-5.

<u>B.5 TT treatment has the greatest effect on functions associated with metabolite co-abundance and this effect is diminished by interactions with other compounds.</u>

A modified protocol for WCGNA, with normalized metabolite abundance values, was used to find groups of metabolites that were co-expressed across the 50 samples (10 from each treatment plus control) [35]. Using a soft thresholding power of 12 to construct the co-abundance network, the 192 metabolites clustered into 6 modules (**See Methods D.8**). Two of the six modules had significant differences between at least one of the treatments and control (FDR cutoff of 0.05). A Tukey multiple comparison test was employed to identify the specific treatment differences in each of these two modules **(Fig 4A).** Metabolites from both modules significantly enriched 31 Reactome *Mus musculus* pathways for Module 1 and 345 pathways for Module 2 **(Supp Table ST6 for pathway information),** using a hypergeometric test**.**

The majority of metabolites in Module 1 (M1) were amino acids **(Fig 4C, M1),** which was reflected in several of the 31 enriched pathways ('branched-chain amino acid catabolism', 'phenylalanine and tyrosine metabolism')**.** Most of the metabolites in Module 2 (M2) were either nucleotides or amino acids **(Fig 4C, M2)**. This was also reflected in several of the 345 enriched pathways for this module ('metabolism of amino acids and derivatives','metabolism of nucleotides')**.**

Information about individual metabolites in both modules can be found in **Supp Table ST7**.

There were no enriched pathways in common between those associated with metabolite module M1 and the pathways associated with the five transcriptomic modules described in **Results B.4**. However, there were common enriched pathways between metabolite module M2 and four of the transriptomic modules (transcriptomic M1, M3, M4, M5). Some of these pathways include 'signaling by PDGF' for M1, 'cellular responses to stress and stimuli' for M3, 'homeostasis' for M4 and 'respiratory electron transport' for M5. For a complete list of the common pathways see **Supp Table ST8.**
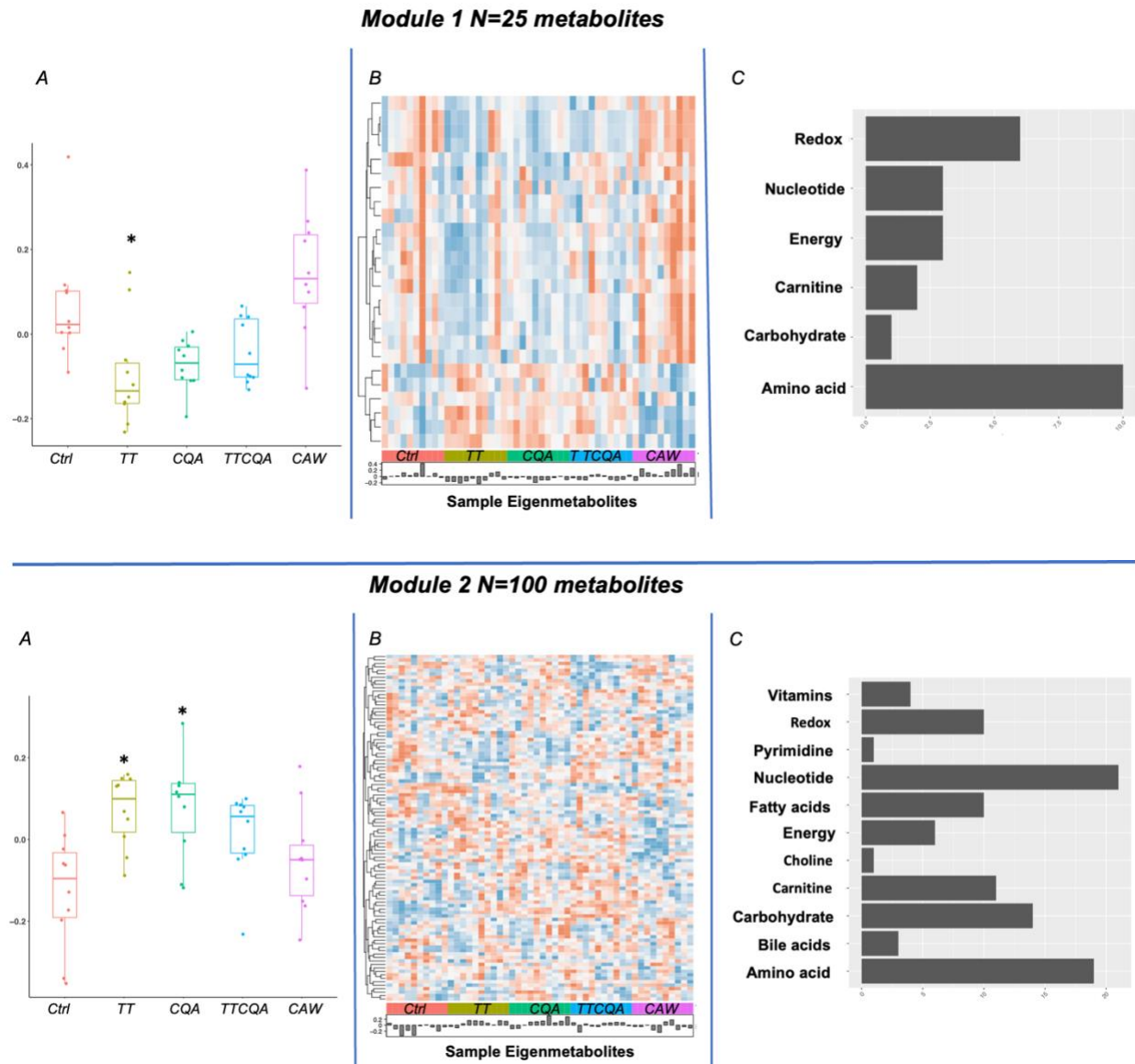
Figure 4. Metabolite co-abundance modules with significant eigenmetabolite differences between at least one treatment and control. A. Boxplots for module eigenmetabolite distributions by each treatment group (* adj p-value<=.05). **B.** Heatmap of individual scaled metabolite abundance for the module (red=high abundance, blue=low abundance). Samples are ordered by treatment group, labeled on the horizontal axis. Histogram shows eigenmetabolite values for each sample. **C.** Metabolite category counts for all metabolites in the module.

B.6 Integration of transcriptome and metabolome shows little overlap between TT, CQA and TTCQA in significant gene activity, but commonality in pathways affected

In the previous section it was noted that there were several common biological pathways enriched with both genes and metabolites from the WGCNA modules

generated for both omics domains. To focus more specifically on how an individual treatment impacts genes and metabolites that are functionally related to each other within a certain biological function (pathway), we performed an integration analysis as described in **Methods D.9.** This integration approach considered all DEG's for a specific treatment, then used a computational framework to find metabolites, from both the DAM's and WGCNA modules, that were also impacted by the same treatment and were statistically associated with the DEG's. This framework used network methods and prior knowledge of biological functions (pathways).

The 1,760 DEG's and 133 DAM's (8 individual DAMs, 25 from Module 1 and 100 from Module 2) identified following TT treatment were used as seeds to construct the composite network as described in **Methods D.9**. This network was constructed without the necessity of additional connector genes/proteins to contain the maximum number of seeds possible. However, there were genes in the metabolite-gene network that were not differentially expressed by TT. There were 6 communities detected in the composite network that contained both a gene and metabolite seed, for a total of 37 metabolites and 42 genes in close functional relationship to each other (**See Supp Table ST9)**. Reactome pathway enrichment, done at the individual community level, using both seed and non-seed genes, identified a total of 29 pathways containing at least one DEG for the TT treatment overall (**See Supp Table ST10)**. An example of the network derived from one TT community from the composite network is in **Figure 5**. Individual network figures for each community containing both seed metabolites and genes are in **Supp Figure SF2.**

The 198 DEG's and 115 DAM's (15 individual DAMs 100 from Module 2) affected by CQA treatment were used as seeds. In this case, connector genes/proteins were used for maximum seed inclusion in the composite network, which yielded 5 communities meeting our criteria of containing both a seed metabolite and gene. These communities contained a total of 23 seed metabolites and 26 seed genes (**See Supp Table ST11),** and these DEGs contributed to enrichment of 74 pathways (**See Supp Table ST10)**. Module networks figures are in **Supp Figure SF4**.

The 981 DEG's and 20 individually DAM's (no Module metabolites) altered with TTCQA treatment were used as seeds. Again, connector genes/proteins were used to construct the composite network, yielding 2 communities of interest, containing a total of 2 seed metabolites and 24 seed genes (**See Supp Table ST12)** contributing to the enrichment of 90 pathways (**See Supp Table ST10)**. Module networks figures are in **Supp Figure SF5**.

No composite network communities built using CAW seed genes and metabolites met the necessary criteria.

Next, pathway and gene overlap was evaluated between treatments using the composite network module level data, as described above. Comparisons were done between TT and CQA, TTCQA and TT, and TTCQA and CQA. There were very few genes in common between any of these treatments (GPT for TT and CQA, CYP4F15 and ACOT5 for TTCQA and TT, and none for TTCQA and CQA). However, while there was only one pathway in common between TT and CQA, there were 6 pathways in common between TTCQA and TT and 34 for TTCQA and CQA.
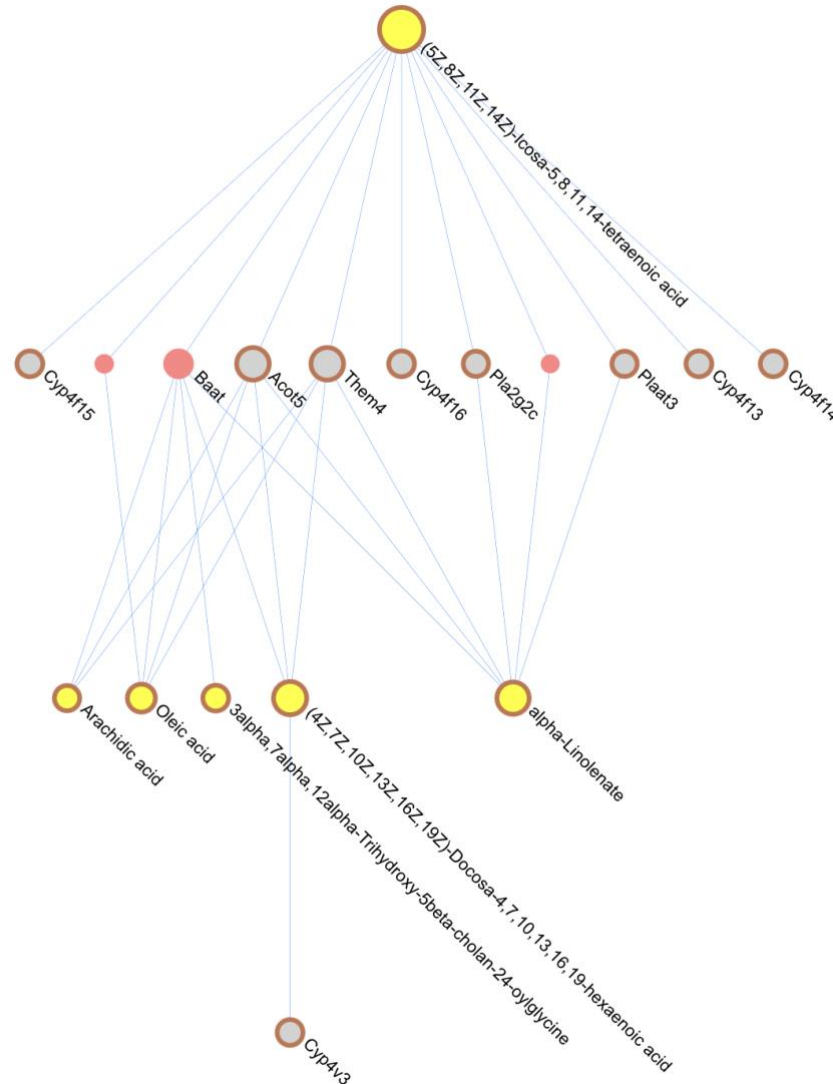
Figure 5. TT Integration Community Network Example. Network is derived from seed genes and metabolites in one community of the TT integrated composite network. Seed genes and metabolites have red circles, genes in gray and metabolites in yellow. This network enriched mostly fatty acid metabolism pathways.

B.7 Functional relationships of DEGs for all treatment groups is not by random chance

A network approach was used to determine the overall functional similarity of the genes differentially expressed by each treatment. First a protein-protein interaction (PPI) network was constructed using the STRING Database [36]. Protein interactions with a physical interaction score >= 400 (scale 0-1000), which is a moderate to high confidence in the interaction, were kept and then the DEG's from each treatment was mapped onto this network and the average network shortest path between the DEGs within each treatment separately was calculated. Permutation testing (1000

permutations) in the PPI network was then conducted by randomly distributing the same number of genes that were differentially expressed in each treatment. This was done to see if the average shortest path distance between the DEGs for a specific treatment was closer than would be expected by random distribution. DEGs associated with all four treatments were significantly closer in the STRING PPI than would be expected by chance, as was also the case with the DEGs found in the five co-expression modules.

C. <u>Discussion</u>

Herbal or botanical healthcare products generally consist of complex mixtures in the form of powders or extracts of single or multiple herbs. The effects of the product may be due to the bioactivities of individual compounds present in the mixture, but may also result from complex interactions between the different components. Studying interactions that may occur between multiple components of a botanical mixture presents a significant challenge. Several methods exist to understand these interactions including the use of isobolograms of bioactivity vs proportion of the components being tested, and biochemometric approaches where the activity of mixtures including or missing specific components is evaluated [13, 37-42]. In the present study integration of transcriptomic and metabolomic data, as well as separate analyses for the two omics domains, was applied to evaluate interactions between components of *Centella asiatica* water extract, CAW, which contains a complex set of components as previously reported [29, 43]. This study compared different subsets of CAW constituent compounds, as well as the complete CAW extract in a nested compound design.

While the earlier studies used targeted analyses of specific effects, here the effects of the nested subsets of compounds were compared in an untargeted fashion, to identify broad patterns of activity and interactions for further research. For example, previous *in vitro* synergy experiments have generally evaluated a single biological activity in a targeted fashion, such as cell viability or an anti-microbial effect, often in a pairwise fashion between 2 compounds [39] and previous systems biology *in silico* methods are reliant on computational predictions [38]. To our knowledge, this is the first *in vitro* approach using systems biology methods to study compound interaction effects on bioactivity in a complex plant extract with a nested compound design with two untargeted molecular domains.

At the individual gene level, the TT treatment stood out in the gene expression results, with the most activated genes of the three select compound treatments (TT, CQA, TTCQA), and the most activated genes that were still retained in the full CAW treatment, as compared to those retained from the CQA and TTCQA treatments in CAW. This would be consistent with the body of previous research focusing on TTs as the main bioactive compound in *Centella asiatica,* with activity demonstrated for many health conditions including neurodegenerative, dermatological and others [14]. However, in this study there were many complex interactions reflected in the overlap of activated genes between TTs and other treatments. A large portion of gene activity seen with either the TT or CQA treatments was not seen when the two groups were combined or in the complete CAW extract, with 66% of the TT and CQA gene effects disappearing with TTCQA, and 56% of these effects disappearing with CAW.

Interestingly, 10% of gene expression lost on combining compounds in TTCQA is restored with the full CAW treatment, which suggests that the additional compounds within CAW are modulating these effects. Based on the post-incubation cultured media analysis, this return of gene activity in CAW could in part be modulated by the degradation of CQA compounds, which appear to inhibit the TT activity in the TTCQA treatment. Since there was no apparent degradation of either TT or CQA compounds seen in the TTCQA treatment (the post-incubation concentrations remained relatively unchanged for CQAs in TTCQA), it could be assumed that the TT inactivation was through some other type of compound inhibition. There was also large amount of new gene activity unique to TTCQA and CAW that was not observed in TT or CQA groups alone suggesting the presence of interactions between TT and CQA (in the case of TTCQA) and/or the involvement of additional compounds in the case of CAW for interactions with these genes. Futher studies are needed to identify whether DEG's unique to CAW could be the result of interactions between TT or CQA and other compounds within the extract.

Continuing with individual genes, the combination TTCQA treatment inactivated many genes whose expression was altered by TT. Most of this interaction is positive on genes downregulated by TT, with no effect by CQA alone. But there is also a large number of upregulated TT genes neutralized by negative interaction with CQA (**Supp Figure SF1 and Fig 2)**.

Some of these inactivated genes explain the therapeutic mechanisms of TT. The effect of TT on these genes not only has a therapeutic effect on neurodegenerative diseases, but also endocrine, dermatological, cardiovascular, digestive, respiratory, gynecological, and rheumatoid diseases (asiatic acid impact on Akt, mTOR, NF-kB, BDNF, CPT-1, SOX2, BCL2, IL-18, CASP-3, NLRP3 modulates the therapeutic effect for some neurological, digestive, dermatological, endocrine, cardiovascular, and respiratory conditions. Asiaticoside impact on NF-kB, MAPK, BCL2, TLR4, TRAF6, IL18, IL10 modulates the therapeutic effect for some neurological, endocrine, cardiovascular, digestive, and respiratory conditions. And madecassoside impact on MAPK, TLR2, IL10 modulates the therapeutic effect for some neurological, endocrine, dermatological and rheumatoid conditions.) [14, 44]. This finding could suggest a therapeutic rationale for <u>not</u> combining TT and CQA compounds ie using isolated TT compounds. While TT compounds are more widely studied, recent research has also shown health benefits from the CQA compounds, such as ameliorating cognitive impairment in an Alzheimers mouse model [26].

When evaluating higher level function using gene co-expression modules affected by each treatment, TT again showed the greatest effect across both the transcriptomics modules and metabolomics modules. However, there were very different patterns of interaction with TT and other compounds for these different modules. In Module 2, which was enriched for Fatty Acid Metabolism and had a dominant gene downregulation effect from TT, there was virtually no interaction with CQA or unknown compounds. This module effect was still seen with TTCQA and CAW. One of the pathways within fatty acid metabolism was arachidonic acid metabolism. Downregulation of this pathway could explain some of the anti-inflammatory and immune modulating effects of both TT and CAW [14]. With both Module 3 and Module 4, the significant effect of TT, seen with eigengenes (**Fig 3A**) appears to be diminished by CQA and then further diminished by

unknown compounds in CAW. Both of the overall functions associated with these modules through the enrichment analysis (cellular response to stress and stimuli, immune functions) have been previously researched for CA. While the pattern of compound interaction is the same in these two modules, the effect of that interaction is different. For Module 3 (cellular response to stress and stimuli), the effect is overall an upregulation of genes, which seems to be neutralized by interactions. The reverse is true for Module 4 (immune functions). However, both triterpene compounds and extracts of CA, water or ethanol, have shown effects in these functions in previous research [14]. But most of these were *in vivo* studies. The metabolomic Module 1 had the same pattern of interaction and differential abundance as the transcriptomic Module 4, but the metabolomic Module 2 has a pattern of interaction distinct from any transcriptomic module. For this module, both TT and CQA have relatively equivalent significant effects that are apparently neutralized through their interaction and then further reduced by interactions with unknown compounds.

The significant effect for CAW seen in transcriptomic Module 1 (extracellular matrix organization) appears to be mediated through interactions of all of the compounds tested in this experiment since no significant effect is seen until all compounds are together in the full CAW extract, with possibly a strong contribution from unknown compounds (**Fig 3, Module 1)**. Collagen formation sub-pathways, within extracellular matrix organization are dominant in this module. Interestingly, TT have been found to reduce the deposition of the extra-cellular matrix and are thus useful in treating liver, pulmonary and other fibrotic conditions [45, 46]. However, a methanol extract *of Centella asiatica* has been found to stimulate collagen and extra-cellular matrix formation, which is beneficial in certain dermatological conditions, as well as wound healing [47] [14] [44]. TT compounds alone have shown benefit in treating dermatological conditions, but this effect appears to be modulated through other mechanisms besided collagen formation [14].

The integration analysis done for this experiment was metabolomics-centric, since we had a much smaller number of metabolites of interest than genes. Finding genes associated with these metabolites helped to focus in on specific functions for the compound interaction analysis. This is likely why some of the functionality seen with the gene co-expression analysis is not reflected in this integration. Generally, functions seen here are more closely reflected in the metabolite co-abundance findings, except for fatty acid metabolism.

Distinctly different functions between the TT and CQA treatments were apparent at the pathway level based on this integration analysis. Some of the top pathways associated with DEGs in the TT integration included metabolism of amino acids, nucleotides and fatty acids. Top pathways associated with DEGs in the CQA integration included chromatin organization, epigenetic regulation of gene expression and DNA repair. There were also functional variations across network community groups within a treatment (**See Supp Tables ST9, ST11, ST12).** Many of the same pathways seen with CQA are also seen in TTCQA, such as chromatin organization, epigenetic regulation of gene expression and DNA repair. Interestingly, there is almost no overlap between these three treatments at the DEG level, but a substantial amount of overlap at the pathway level, particularly between TTCQA and CQA, suggesting that some functions are retained in the TTCQA combination, but mediated through different genes.

The PPI network analysis is another 'guilt by association' approach which considers two genes in a PPI to have a more likely functional relationship the shorter the path between them [48]. There is also some evidence that network proximity of DEG's in a PPI is associated with therapeutic synergy of the associated compounds [49]. This analysis revealed that none of the collections of altered genes were distributed by random chance. This would be expected for collections of compounds with structural similarity, such as the four compounds in the TT treatment of the eight compounds in the CQA treatment. But it is more surprising to see this phenomenon maintained with in the combination of the these two sets of compounds (TTCQA), and especially in the large number of strucuturally diverse compounds seen with CAW. It is possible that the lack of random distribution seen in CAW is mediated through compound interactions, and even possibly driven by the two therapeutic classes of compounds.

This study revealed some compelling patterns for further investigation. Future analyses could address some of the limitations of this current study. The study sample sizes were relatively small and may not have been able to detect smaller significant effects, particularly with the metabolomics differential analyses. Also, the individual experiments were not able to collect sufficient cells from the cultures to perform the two omics analyses on the same batch of neuronal cultures. This particularly impacts our methodological approach for the omics integration. Using a knowledge-based approach has limitations due to incomplete annotation of biological functions. The two independent co-expression/abundant analyses were data-driven and likely a more complete picture of functional molecular relationships, although still subject to limitations from the use of known pathway data for the enrichment analysis. And even the co-expression/abundance analyses are based on the assumption of a linear relationship between biological molecules, which is not always the case. Some sample variablity can be seen in the eigengenes and eigenmetabolites in **Figures 3, 4A.** While the two omics experiments were done in separate batches, the samples within each omics domain were from the same culture batch, so this would not have been a source of variation. This could be related to interactions at the cellular transport level. Since this was an *in-vitro* experiment some of the findings may not translate at an organism level. Some compounds present in these treatments may not be aborbed when given orally, and, in general, most pharmacokinetic mechanisms are not accounted for in this type of experiment. Future studies should explore the *in-vivo* absorption of these compounds and also examine metabolomic and transcriptomic changes in the brains of mice treated with TT, CQA, TTCQA and CAW.

D. <u>Methods</u>

<u>D.1 Mouse primary cortical neuron cell cultures</u>

Mouse primary cortical neurons were isolated and cultured (37° C) following a previously published protocol [50]. Briefly, six well plates were coated with poly l-lysine (PLL; Sigma), three days prior to primary neuron harvest. After one day, the PLL was removed and the plates were washed with double distilled water . Following the wash, plating media (Minimal Essential Media, Fetal Bovine Serum 4.6%, Glucose 0.55%, Antibiotic-antimycotic: 10,000 U/mL Penicilin 10,000 U/mL Streptomycin and 25 µg of

amphotericin B/ml) was added to each well for the remaining two days. Then, neurons were isolated from the cortices of C57BL6 mouse embryos at embryonic days 16-18, incubated in a mixture of HBSS and trypsin (2.5%) at 37° C, then dissociated. Neurons were plated onto six well plates (1 million/well) and incubated (37° C) for 3 hours after which the plating media was replaced with supplemented neurobasal media (with B27 1:50 dilution, Glutamax 1:100 dilution, Antibiotic-antimycotic: 10,000 U/mL Penicilin 10,000 U/mL Streptomycin and 25 µg of amphotericin B/ml). Neurons were then cultured for five days at 37° C before receiving experimental treatments.

D.2 Treatments:
   Centella asiatica dried plant material was produced by Oregon's Wild Harvest (OWH). A water extract (CAW) was prepared by reflux extraction of the dried plant, filtering and lyophilization[20-22, 51, 52]. The CAW treatment was prepared from this dried extract (50 µg/mL in 0.025% v/v aqueous methanol vehicle in each well). The triterpene (TT) and caffeoylquinic acid (CQA) content in the CAW was determined using liquid chromatography coupled to multiple reaction monitoring mass spectrometry (LC-MRM-MS) as previously described [29]. The triterpene (TT) treatment was prepared (0.025% v/v aqueous methanol vehicle final concentration in the cell culture) with madecassoside (1.7945 µg/mL), asiaticoside (0.7376 µg/mL), madecassic acid (0.0377 µg/mL), and asiatic acid (0.021 µg/mL), concentrations equivalent to that in the CAW 50 µg/mL solution (LC-MRM-MS verified concentrations). The caffeoylquinic acid (CQA) solution was prepared in the same fashion using chlorogenic acid (0.3749 µg/mL), neo-chlorogenic acid (0.1695 µg/mL) , crypto-chlorogenic acid (0.1492 µg/mL), 3,4-dicaffeoylqinic acid (0.1146 µg/mL), 3,5-dicaffeoylqinic acid (0.0884 µg/mL), 4,5-dicaffeoylqinic acid (0.0975 µg/mL), 1,3-dicaffeoylqinic acid (0.1291 µg/mL), and 1,5-dicaffeoylqinic acid (0.1946 µg/mL), (LC-MRM-MS verified concentrations). And the triterpene and caffeoylqinic acid (TTCQA) solution was prepared using the 12 compounds that were used in the TT and CQA treatments, prepared in the same fashion. Control vehicle consisted of aqueous methanol (.025% v/v). These are the stock treatment 'theoretical values'. These stock treatment solutions were then quantifed using LC-MRM-MS. These are the expected time zero values and can be found in the first four lines of **Supp Table ST1**.
   CAW or compound treatments were initiated after five days of incubation and lasted for 48 hours. There were four compound treatments (CAW, TT, CQA, TTCQA) and one vehicle control treatment. After the 48 hour treatment period, an aliquot of medium from each well was collected for analysis of post-incubation cultured media compound concentrations and cells were harvested by trypsinization (0.25% trypsin in media, 37° C, 5 min). Detached cells were then centrifuged and the resulting cell pellet was flash frozen and stored at -80C until analysis. Eight samples for each experimental condition were prepared for RNA sequencing and ten samples were prepared for each condition for metabolomics analyses.

D.3. Analysis of compound concentrations in the media after 48h of treatment
For quantifying 12 CA phytochemical marker compounds (3 mono-CQA, 5 di-CQAs, and 4 TTs) in media, a SPE method was developed that removes phospholipids and proteins (using Ostro® Protein Precipitation and Phospholipid Removal Plate (Waters

Corp.)) prior to LC-MRM-MS analysis. Each media sample was thawed and a 100 µL aliquot was transferred into Eppendorf tube, that contained 300 µL methanol containing 0.1% formic acid and 5 µL digoxin-d3 (10 µg/mL) as an internal standard. Media sample was loaded into a Ostro plate well and 300 µL acetonitrile containing 0.1% formic acid was added and mixed by pipetting up and down 3 times. The plate was placed onto a vacuum manifold and the elution solvent was drawn into glass inserts of the collectin plate. The 1-mL glass inserts containing the eluate were placed into a 2-mL micro-centrifuge tubes, and dried in Speed-vac for approximately 1 hour for obtaining solid residues. For reconstitution, 50 µL of 70% methanol containing 0.1% formic acid were added to the glass inserts and resuspended by pipetting up and down 3 times. Samples were centrifuged at 13,000 rpm at 4 °C for 10 min, and each supernatant was transferred to an LC-MS vial and stored at -20°C until LC-MRM-MS analysis using a Waters Xevo TQXS system connected ot a Waters UPLC I-class. The MS method details have been described by us previously [29].

D.4 RNA_seq

Cell pellets pellets were resuspended in RLT-ß-mercaptoethanol (QIAGEN) lysis buffer and quick frozen for RNA extraction at the Oregon Health & Science University (OHSU) Gene Profiling Shared Resource . RNA quality was assessed at the GPSR and assigned an RNA Integrity Number (RIN). Only samples with RIN>=8.0 were considered for sequencing, which occurred at the OHSU Massively Parallel Sequencing Shared Resource facility. This was performed on a HiSeq 2500 (Illumina) and aligned to the mm38 mouse genome with the Star aligner [53]. Subsequent QA/QC was performed using the MultiQC package [54].

D.5 Metabolomics

Cells for control and treatments (approximately 1 million each) were thawed on ice and transferred to Precellys Lysing Kit bead blender tube with pre-supplied beads (2 mL Tissue Homogenizing Mixed Beads Kit (CKMix), Item No. 10409, Cayman Chemicals, Michigan, USA). Extraction solvents (500 µL, cold 80% methanol in water containing 12-[[(cyclohexylamino)carbonyl]amino]-dodecanoic acid (CUDA, 20 ppm, 10 µL) were added to each vial and homogenized. The homogenates were incubated for 1 hr at -20°C, then centrifuged at 10,000 rpm at 4 °C for 10 min. Each supernatant was transferred to a new vial and dried using a Speed-vac. Each residue was reconstituted in a cold 50% aqueous acetonitrile, vortexed for 20 seconds, and centrifuged at 10,000 rpm at 4 °C for 10 min. Each supernatant was transferred to an LC-MS vial and stored at -80°C until LC-MS analysis. Quality control (QC) samples were generated by pooling 10 µL aliquots from each sample extract and were analyzed with samples. Blanks contained only extraction solvent.

The LC-MS/MS metabolomics workflow was described by us previously (Choi et al., J Am Heart Assoc. 2019;8:e012809) and entails two chromatographic methods: (1) Reverse-phase liquid chromatography was carried out using ACE Excel C18-PFP (1.7 µm, 2.1 mm x 100 mm). The sample injection volume was 5 µL, and the flow rate was 0.5 mL/min. The mobile phases consisted of water (A) and acetonitrile (B), both

containing 0.1% formic acid. The gradient was as follows: initial hold at 5%B for 3 min, then increase to 80% B in 13 min, held until 16 min, then shift to 5% B and hold for equilibration from 16.5 min to 20.5 min. The column oven temperature was maintained at 30°C. (2) Hydrophilic interaction liquid chromatography, the method was identical to that previously reported (Choi et al. J Am Heart Assoc. 2019;8:e012809). For both chromatographic methods ESI-MS/MS data were acquired in positive and negative ionizaton mode using a quadrupole time-of-flight MS system (AB Sciex tripleTOF 5600). Individual metabolites were annotated using an existing in-house libraries library based on retention time, exact mass and MS/MS of 650 compounds compiled in the Mass Spectrometry Metabolite Library of Standards (MSMLS, IROA Technologies, Bolton, MA, USA). Raw data files were imported and processed using PeakView (ver. 1.2, Sciex) and MultiQuant (ver. 3.0.2, Sciex) software. Metabolites were verified using chromatographic retention time (error <10%), accurate mass (error <10 ppm), MS/MS fragmentation (score >70), and isotopic pattern (error <20%). The RP-C18 LC method yielded 117 metabolites (58 metabolites in positive ion mode and 59 in negative ion mode), and the HILIC method resulted in 234 metabolites (135 metabolites in positive ion mode and 119 in negative ion mode). The metabolites that were assigned in both methods were evaluated and the metabolites detected with the lower coefficients variation (CV) value in the QC samples were kept. This evaluation resulted in 192 high confident unique metabolites (178 metabolites (<20%), 10 metabolites (<30%), and 4 metabolites (>30%) in QC CV values, Supp Fig SF.xxxx).

D.6 Differential analyses for gene expression and metabolite abundance

For differential gene expression the limma-voom package was used, a modification of the limma package designed for RNA-seq count data [55]. Genes were removed from the data if they did not have a transcript count of at least 10 in each sample. Next, a $\log_2$ transformation of counts per million (lcpm) was performed, followed by a trimmed mean of M-values (TMM) normalization. Heteroscedasticity was controlled with the voom function to meet the assumptions of a linear method. Differential calculations were done between each of the four treatments (TT, CQA, TTCQA, CAW) and the control (FDR cutoff 0.05).

For metabolite abundance the MetaboAnalyst package was used [56]. The data was median normalized, $\log_2$ transformed and then pareto scaled (**Supp Fig SF3**). T-tests were performed to compare each treatment to the control for all 192 metabolites (FDR cutoff 0.05).

D.7 TT and CQA Gene Expression Interaction Classification

To assess the interaction between the TT and CQA treatments on normalized gene expression data, parameter estimates were derived from a least squares linear method with the control group as reference, using the limma R package [55]. For each group, the effects (parameter estimates) of each treatment were assessed relative to control. Next, the sum of the effects of TT and CQA treatments were contrasted to the effect of the TTCQA treatment using the contrasts.fit eBayes functions in limma. Any gene found to be differentially expressed by any of these three treatments was examined for

interaction (N=2,268). We included genes that might not have been differentially expressed by either one or two of the other treatments to see if an interaction might create or inhibit a significant expression.

If the TTCQA effect on the expression of a particular gene was greater than the sum of the TT and CQA effects, the interaction was classified as positive, or synergistic. If the TTCQA effect was less than the sum, the interaction was classified as negative, or antagonistic. If they were equal, the interaction was classified as additive, or no interaction [57]. Since these coefficients represent $\log_2$ fold changes, negative values represent downregulation and positive values represent upregulation. So, any negative interaction that decreases this coefficient value would potentially be increasing downregulation, and positive interactions, as well as additive, would be increasing upregulation.

D.8 Weighted Gene Correlation Network Analysis (WGCNA gene co-expression and metabolite co-abundance)

The) WGCNA package in R [58] was used for both RNA-seq and metabolomic data [35]. Both data types were normalized for the creation of this network by variance stabilizing transformation. To create a scale free network, Pearson correlation was first used to construct an adjacency matrix for all gene or metabolite pairs using a soft thresholding power of 12 for both data types. The matrix was transformed to a topological overlap matrix (TOM) which was then converted to a dissimilarity matrix (1-TOM). Hierarchical clustering was applied to the final matrix to identify clusters of genes or metabolites with similar expression or abundance profiles. Minimum node size was 5 for metabolomics data and 30 for RNA-seq data. Module eigengenes and eigenmetabolites (first principal components) were calculated using the moduleEigengenes of WGCNA. Module membership for each gene or metabolite was calculated as the correlation between the normalized expression or abundance values with the module eigengene or eigenmetabolite, and was calculated with the signedKME in WGCNA. Intramodular connectivity was the degree of each gene or metabolite in the adjacency matrix described above, calculated with the intramodularConnectivity function in WGCNA.

Following the construction of these modules in each of the two omics domains, we then tested for differences between each treatment and control seen at the eigengene or eigenmetabolite level.

### D.9 Integration of transcriptomics and metabolomics data.

A modification of a protocol previously used to integrate transcriptomic and metabolomic data for the study of neurodegenerative disease [59] was used to get a more comprehensive functional perspective for comparing the four treatments. This protocol was modified to fit the limitations of the data. Since this project was a repeated experiment design, with the transcriptome coming from different samples than the metabolome, rather than a split design with both omics coming from the same samples, a 'data driven' approach was not appropriate. A 'data driven' approach would correlate expression and abundance levels (or eigenvalues) across domains to create

correlation-based networks for further analyses. Instead, a 'knowledge-based' approach was used, with existing molecular pathway information, to construct networks within and between individual genes and metabolites. These networks and subsequent analyses were done using OMICSNET software [60].

Genes and metabolites associated with significant treatment effects, relative to control, were used as seeds to construct the 'knowledge-based' networks in OMICSNET. Analyses were done separately for each treatment. Any gene differentially expressed by a treatment was used as a seed. Any individual metabolite affected by a treatment or member of a metabolite module whose eigenmetabolite was affected by a treatment was also used as a seed. Due to the sparse number of individual metabolites affected, we also included co-expressed metabolites. Seed genes were used to construct a primary protein-protein interaction (PPI) network using the STRING database [36]. Genes and proteins are used interchangeably here. Connector proteins (non-seed) were added if necessary to create a more fully connected network. Seed metabolites were used to first create a primary metabolite-protein network using the KEGG pathway database [61]. A secondary PPI was created using the proteins in the primary metabolite-protein network using STRING, adding connector proteins if necessary. Finally, a composite network was created by linking the three independent network layers by common nodes (genes/proteins). To control the network size, the 'Minimum network' function was used in OMICSNET to create the smallest network possible which still contained as many seeds as possible.

Once an integrated network was created, tightly clustered communities of nodes were detected using the Walktrap algorithm [62]. Only communities that contained both seed metabolites and seed genes were further analyzed. These communities identified close functional relationships between genes and metabolites both affected by the same treatment. Subsequent Reactome pathway overrepresentation analysis was done using all genes in a community (seeds and connectors) using ReactomePA [33]. A comparison of pathways impacted by each treatment was then done to identify overlapping and distinct functions using the integrated omics data (**Fig 3**).
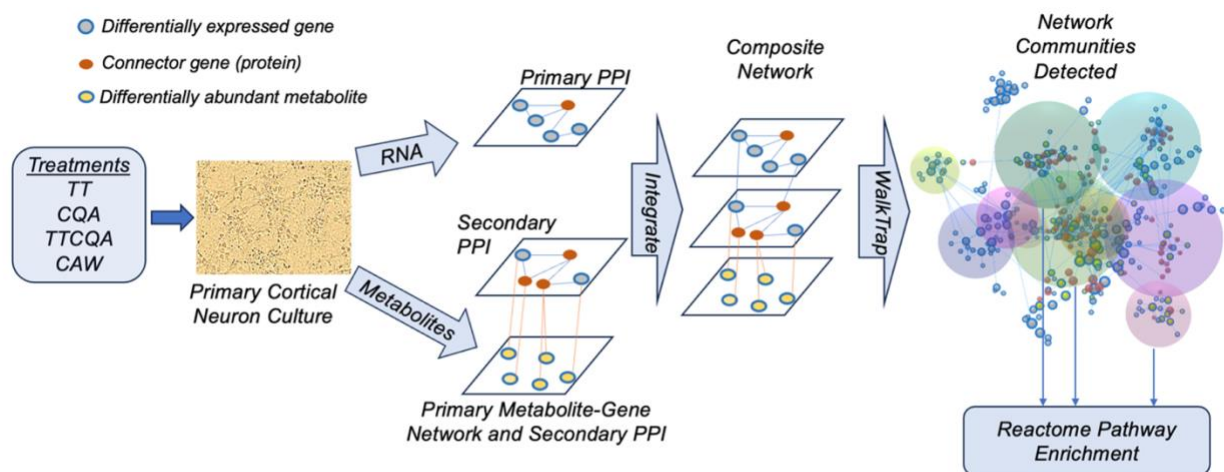
Figure 3. Omics integration methodology. Individual seed networks are shown in the center of the figure. For metabolite seeds two networks are shown, a primary metabolite-gene network and a secondary PPI created from the genes in the primary network. Seed genes are shown in this secondary PPI if present, but are not used to construct this network. For the seed genes, only one primary PPI network is created. The red connector (non-seed) genes (proteins) are added to both PPI's, if necessary, to increase connectivity in the overall network for community detection that will better contain seeds of interest.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/S, Figures SF1-SF6 ; Tables ST1-ST12.

Data Availability Statement: The original contributions presented in this study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author(s) for RNAseq files.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.
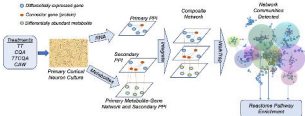
References

1. Organization, W.H. *Ageing and Health*. 2022; Available from: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health.

2. AARP/NCCAM, *Complementary and Alternative Medicine: What People Aged 50 and Older Discuss with Their Health Care Providers*, in *Survey Report*. 2011.

3. Nascimento, A.L., et al., *Functional aspects of the use of plants and animals in local medical systems and their implications for resilience.* J Ethnopharmacol, 2016. **194**: p. 348-357.

4. Ward, L. and G.M. Pasinetti, *Recommendations for Development of Botanical Polyphenols as "Natural Drugs" for Promotion of Resilience Against Stress-Induced Depression and Cognitive Impairment.* Neuromolecular Med, 2016. **18**(3): p. 487-95.

5. Ho, Y.S., K.F. So, and R.C. Chang, *Anti-aging herbal medicine--how and why can they be used in aging-associated neurodegenerative diseases?* Ageing Res Rev, 2010. **9**(3): p. 354-62.

6. Ismail, Z., et al., *Usage of traditional medicines among elderly and the prevalence of prednisolone contamination.* The Malaysian journal of medical sciences : MJMS, 2005. **12**(2): p. 50-55.

7. Xu, Q.Q., et al., *Chinese Herbal Medicine for Vascular Dementia: A Systematic Review and Meta-Analysis of High-Quality Randomized Controlled Trials.* J Alzheimers Dis, 2018. **62**(1): p. 429-456.

8. Enke, C.G. and L.J. Nagels, *Undetected components in natural mixtures: how many? What concentrations? Do they account for chemical noise? What is needed to detect them?* Anal Chem, 2011. **83**(7): p. 2539-46.

9. van Vuuren, S. and A. Viljoen, *Plant-based antimicrobial studies--methods and approaches to study the interaction between natural products.* Planta Med, 2011. **77**(11): p. 1168-82.

10. Efferth, T. and E. Koch, *Complex interactions between phytochemicals. The multi-target therapeutic concept of phytotherapy.* Curr Drug Targets, 2011. **12**(1): p. 122-32.

11. Wagner, H. and G. Ulrich-Merzenich, *Synergy research: approaching a new generation of phytopharmaceuticals.* Phytomedicine, 2009. **16**(2-3): p. 97-110.

12. Caesar, L.K. and N.B. Cech, *Synergy and antagonism in natural product extracts: when 1 + 1 does not equal 2.* Nat Prod Rep, 2019. **36**(6): p. 869-888.

13. Junio, H.A., et al., *Synergy-directed fractionation of botanical medicines: a case study with goldenseal (Hydrastis canadensis).* J Nat Prod, 2011. **74**(7): p. 1621-9.

14. Sun, B., et al., *Therapeutic Potential of Centella asiatica and Its Triterpenes: A Review.* Front Pharmacol, 2020. **11**: p. 568032.

15. Shinomol, G.K., Muralidhara, and M.M. Bharath, *Exploring the Role of "Brahmi" (Bacopa monnieri and Centella asiatica) in Brain Function and Therapy.* Recent Pat Endocr Metab Immune Drug Discov, 2011. **5**(1): p. 33-49.

16.   Gray, N.E., et al., *Centella asiatica: phytochemistry and mechanisms of neuroprotection and cognitive enhancement.* Phytochemistry Reviews, 2017: p. 1-34.

17.   Veerendra Kumar, M.H. and Y.K. Gupta, *Effect of Centella asiatica on cognition and oxidative stress in an intracerebroventricular streptozotocin model of Alzheimer's disease in rats.* Clinical & Experimental Pharmacology & Physiology, 2003. **30**(5-6): p. 336-342.

18.   Veerendra Kumar, M.H. and Y.K. Gupta, *Effect of different extracts of Centella asiatica on cognition and markers of oxidative stress in rats.* Journal of ethnopharmacology, 2002. **79**(2): p. 253-260.

19.   Gupta, Y.K., M.H. Veerendra Kumar, and A.K. Srivastava, *Effect of Centella asiatica on pentylenetetrazole-induced kindling, cognition and oxidative stress in rats.* Pharmacology, Biochemistry & Behavior, 2003. **74**(3): p. 579-585.

20.   Gray, N.E., et al., *Centella asiatica modulates antioxidant and mitochondrial pathways and improves cognitive function in mice.* J Ethnopharmacol, 2016. **180**: p. 78-86.

21.   Gray, N.E., et al., *Centella asiatica increases hippocampal synaptic density and improves memory and executive function in aged mice.* Brain Behav, 2018. **8**(7): p. e01024.

22.   Gray, N.E., et al., *Centella asiatica attenuates hippocampal mitochondrial dysfunction and improves memory and executive function in beta-amyloid overexpressing mice.* Mol Cell Neurosci, 2018. **93**: p. 1-9.

23.   Soumyanath, A., Zhong YP, Henson E, Wadsworth T, Bishop J, Gold BG, Quinn JF., , *Centella asiatica Extract Improves Behavioral Deficits in a Mouse Model of Alzheimer's Disease: Investigation of a Possible Mechanism of Action.* Int J Alzheimers Dis,, 2012: p. 381974.

24.   Han, J., et al., *Neuroprotective effect of 3,5-di-O-caffeoylquinic acid on SH-SY5Y cells and senescence-accelerated-prone mice 8 through the up-regulation of phosphoglycerate kinase-1.* Neuroscience, 2010. **169**(3): p. 1039-45.

25.   Gray, N.E., et al., *Centella asiatica - Phytochemistry and mechanisms of neuroprotection and cognitive enhancement.* Phytochem Rev, 2018. **17**(1): p. 161-194.

26.   Matthews, D.G., et al., *Caffeoylquinic Acids in Centella asiatica Reverse Cognitive Deficits in Male 5XFAD Alzheimer's Disease Model Mice.* Nutrients, 2020. **12**(11).

27.   Matthews, D.G., et al., *Centella Asiatica Improves Memory and Promotes Antioxidative Signaling in 5XFAD Mice.* Antioxidants (Basel), 2019. **8**(12).

28.   Gray, N.E., et al., *Centella asiatica attenuates Aβ-induced neurodegenerative spine loss and dendritic simplification.* Neurosci Lett, 2017. **646**: p. 24-29.

29.   Yang, L., et al., *Quantification of Caffeoylquinic Acids and Triterpenes as Targeted Bioactive Compounds of Centella asiatica in Extracts and Formulations by Liquid Chromatography Mass Spectrometry.* J Chromatogr Open, 2023. **4**.

30.   Wianowska, D. and M. Gil, *Recent advances in extraction and analysis procedures of natural chlorogenic acids.* Phytochemistry Reviews, 2019. **18**(1): p. 273-302.

31.   Xue, M., et al., *Stability and Degradation of Caffeoylquinic Acids under Different Storage Conditions Studied by High-Performance Liquid Chromatography with Photo Diode Array Detection and High-Performance Liquid Chromatography with Electrospray Ionization Collision-Induced Dissociation Tandem Mass Spectrometry.* Molecules, 2016. **21**(7).

32.   Blighe, K., and A Lun, *"PCAtools: everything Principal Components Analysis."*. 2019.

33. Yu, G. and Q.-Y. He, *ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization.* Molecular BioSystems, 2016. **12**(2): p. 477-479.

34. van Dam, S., et al., *Gene co-expression analysis for functional classification and gene-disease predictions.* Brief Bioinform, 2018. **19**(4): p. 575-592.

35. Pei, G., L. Chen, and W. Zhang, *WGCNA Application to Proteomic and Metabolomic Data Analysis.* Methods Enzymol, 2017. **585**: p. 135-158.

36. Szklarczyk, D., et al., *The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets.* Nucleic Acids Res, 2021. **49**(D1): p. D605-d612.

37. Gilbert, B. and L.F. Alves, *Synergy in plant medicines.* Curr Med Chem, 2003. **10**(1): p. 13-20.

38. Wang, X., et al., *A systems biology approach to uncovering pharmacological synergy in herbal medicines with applications to cardiovascular disease.* Evid Based Complement Alternat Med, 2012. **2012**: p. 519031.

39. Lin, J., H. Jiang, and X. Ding, *Synergistic combinations of five single drugs from Centella asiatica for neuronal differentiation.* Neuroreport, 2017. **28**(1): p. 23-27.

40. Britton, E.R., et al., *Biochemometrics to Identify Synergists and Additives from Botanical Medicines: A Case Study with Hydrastis canadensis (Goldenseal).* J Nat Prod, 2018. **81**(3): p. 484-493.

41. Caesar, L.K., et al., *Integration of Biochemometrics and Molecular Networking to Identify Antimicrobials in Angelica keiskei.* Planta Med, 2018. **84**(9-10): p. 721-728.

42. Kellogg, J.J., et al., *Biochemometrics for Natural Products Research: Comparison of Data Analysis Approaches and Application to Identification of Bioactive Compounds.* J Nat Prod, 2016. **79**(2): p. 376-86.

43. Alcazar Magana, A., et al., *Integration of mass spectral fingerprinting analysis with precursor ion (MS1) quantification for the characterisation of botanical extracts: application to extracts of Centella asiatica (L.) Urban.* Phytochem Anal, 2020. **31**(6): p. 722-738.

44. Park, K.S., *Pharmacological Effects of Centella asiatica on Skin Diseases: Evidence and Possible Mechanisms.* Evid Based Complement Alternat Med, 2021. **2021**: p. 5462633.

45. Wei, L., et al., *Asiatic acid attenuates CCl(4)-induced liver fibrosis in rats by regulating the PI3K/AKT/mTOR and Bcl-2/Bax signaling pathways.* Int Immunopharmacol, 2018. **60**: p. 1-8.

46. Dong, S.-H., et al., *Asiatic acid ameliorates pulmonary fibrosis induced by bleomycin (BLM) via suppressing pro-fibrotic and inflammatory signaling pathways.* Biomedicine & Pharmacotherapy, 2017. **89**: p. 1297-1309.

47. Yao, C.H., et al., *Wound-healing effect of electrospun gelatin nanofibres containing Centella asiatica extract in a rat model.* J Tissue Eng Regen Med, 2017. **11**(3): p. 905-915.

48. Valdeolivas, A., et al., *Random walk with restart on multiplex and heterogeneous biological networks.* Bioinformatics, 2019. **35**(3): p. 497-505.

49. Tang, Y.-C. and A. Gottlieb, *SynPathy: Predicting Drug Synergy through Drug-Associated Pathways Using Deep Learning.* Molecular Cancer Research, 2022. **20**(5): p. 762-769.

50. Kaech, S. and G. Banker, *Culturing hippocampal neurons.* Nat Protoc, 2006. **1**(5): p. 2406-15.

51.   Houghton, P.J.R., A. , *Handbook for the fractionation of natural extracts.* 1998, London: Chapman and Hall.

52.   Soumyanath, A., Zhong YP, Henson E, Wadsworth T, Bishop J, Gold BG, Quinn JF., , *Centella asiatica Extract Improves Behavioral Deficits in a Mouse Model of Alzheimer's Disease: Investigation of a Possible Mechanism of Action. .* Int J Alzheimers Dis,, 2012: p. 381974.

53.   Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.

54.   Ewels, P., et al., *MultiQC: summarize analysis results for multiple tools and samples in a single report.* Bioinformatics, 2016. **32**(19): p. 3047-3048.

55.   Ritchie, M.E., et al., *limma powers differential expression analyses for RNA-sequencing and microarray studies.* Nucleic Acids Res, 2015. **43**(7): p. e47.

56.   Pang, Z., et al., *MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights.* Nucleic Acids Res, 2021. **49**(W1): p. W388-w396.

57.   Law, C.W., et al., *A guide to creating design matrices for gene expression experiments.* F1000Res, 2020. **9**: p. 1444.

58.   Langfelder, P. and S. Horvath, *WGCNA: an R package for weighted correlation network analysis.* BMC Bioinformatics, 2008. **9**(1): p. 559.

59.   Horgusluoglu, E., et al., *Integrative metabolomics-genomics approach reveals key metabolic pathways and regulators of Alzheimer's disease.* Alzheimers Dement, 2021.

60.   Zhou, G., et al., *OmicsNet 2.0: a web-based platform for multi-omics integration and network visual analytics.* Nucleic Acids Research, 2022. **50**(W1): p. W527-W533.

61.   Kanehisa, M., et al., *KEGG: integrating viruses and cellular organisms.* Nucleic Acids Res, 2021. **49**(D1): p. D545-d551.

62.   Pons, P. and M. Latapy. *Computing Communities in Large Networks Using Random Walks*. 2005. Berlin, Heidelberg: Springer Berlin Heidelberg.
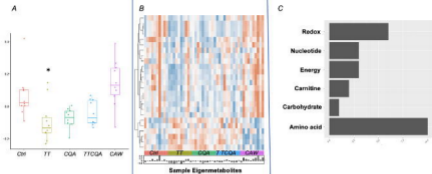
Figure 3. Omics integration methodology. Individual seed networks are shown in the center of the figure. For metabolite seeds two networks are shown, a primary metabolite seed network and a secondary PPI created from the genes in the primary network. Seed genes are shown in the secondary PPI if present, but are not used to construct this network. For the seed genes, only one primary PPI network is created. The red connector (non-seed) genes (proteins) are added to both PPI, if necessary, to increase connectivity in the overall network for community detection that will better contain seeds of interest.

Figure 5. TT Integration Community Network Example. Network is derived from seed genes and metabolites in one community of the TT integrated composite network. Seed genes and metabolites have red circles, genes in gray and metabolites in yellow. This network enriched mostly fatty acid metabolism pathways.

Figure 4. Metabolite co-abundance modules with significant eigenmetabolite differences between at least one treatment and control. A. Boxplots for module eigenmetabolite distributions by each treatment group (* adj p-value<=.05). **B.** Heatmap of individual scaled metabolite abundance for the module (red=high abundance, blue=low abundance). Samples are ordered by treatment group, labeled on the horizontal axis. Histogram shows eigenmetabolite values for each sample. **C.** Metabolite category counts for all metabolites in the module.

Figure 3. Sample eigengene distributions for the five gene co-expression modules with significant eigengene differences between at least one treatment and control. Boxplots for module eigengene distributions by each treatment group (* value<=.05). Module 1=Extracellular Matrix Organization and Collagen Biosynthesis , Module 2=Fatty Acid Metabolism, Module 3=Cellular Response to Stress and Stimuli, Module 4=Immune System, Module 5=Electron Transport and Mitochondrial Biogenesis. Module 1 also shows a heat map of individual gene expression levels (blue=low, red=high). The highlighted section represents samples in the significant treatment for this module. See **Supp Figure SF6** for heatmaps for Modules 2-5.

Interaction	TT	CQA	TTCQA	# Genes
Negative	Down	No Effect	Down	286
	No Effect	Down	Down	238
	Down	Down	Down	5
	No Effect	Up	Up	1
	Up	No Effect	Up	370
	Up	Up	Up	39
	Up	Up	No Effect	24
	Up	Up	Down	3
None	Down	No Effect	No Effect	21
	Up	No Effect	No Effect	54
Positive	Down	Down	Down	29
	Down	Down	No Effect	9
	Down	Down	Up	44
	Down	No Effect	Down	104
	Down	No Effect	No Effect	752
	No Effect	Down	No Effect	82
	No Effect	Down	Up	182
	No Effect	No Effect	Up	69

🟥 Down-regulated 🟨 No Effect 🟩 Up-regulated

Figure 2. Effect on gene expression of the TT and CQA compounds administered separately and interactions between these groups observed in the TTCQA treated samples. The figure on the left shows a mapping of the interactions listed in the table on the right. Any gene with significant expression changes, relative to control, seen with any of the TT, CQA or TTCQA treatments is represented in the figure and table (N=2,268 unique genes). The relative number of genes and the expression status is represented within each of the three columns labeled TT, CQA or TTCQA (red=significant downregulation, green=significant upregulation, beige=no significant expression changes seen with this treatment). The TT and CQA compound 'Interaction effect' for the combined group is shown next to the TTCQA column ('P'=positive, or synergistic, 'N'=negative, or antagonistic, 'none'=additive, or no interaction). The gray ribbons show the numbers of genes interacting between TT and CQA, the interaction status, and the final expression status in the TTCQA treatment. The table on the right shows the number of genes in each category. As an example, a positive interaction has a higher fold change value in the TTCQA treatment than the sum of the two fold change values from TT and CQA, and a negative interaction has a lower TTCQA fold change than the sum of TT and CQA fold changes. So a positive interaction could still result in a downregulated or unregulated TTCQA gene and a negative interaction could result in an upregulated or unregulated TTCQA gene.

Figure 1. PCA plots for the two untargeted molecular domains of metabolomics and transcriptomics. **A.** All treatments plotted together, including the vehicle control, for either metabolic or RNA-seq data. **B.** RNA-seq data, each treatment compared individually to control. **C.** Metabolic data, each treatment compared individually to control. For **B** and **C**, the number of differentially expressed metabolites or genes (#DE) is shown in the lower right corner of each plot. The plots are also ordered left to right by increasing number of CA compounds in each treatment.