7.13 Uncovering Bioactive Natural Products Via Biochemometric Methodologies

Joshua J Kellogg[a] **and Nadja B Cech**, Department of Chemistry & Biochemistry, University of North Carolina at Greensboro, Greensboro, NC, United States

7.13.1 Introduction

Natural products represent a tremendously rich resource for drug discovery paradigms or as dietary supplements to improve human health.[1,2] In their unrefined form, natural products consist of complex mixtures of many structurally diverse components.[3] The identities of bioactive constituents are (a) often not known, and (b) likely to differ depending on the biological activity evaluated.[4–6] The central challenge facing natural product researchers is unraveling the complexities of such complex mixtures and assigning structures and biological activities to the components. With this chapter, we review various strategies that can be employed towards this goal and discuss their advantages and disadvantages. We summarize the traditional approach of bioassay guided fractionation and describe how newer methodologies (biochemometrics strategies) that integrate metabolite profiles (metabolomics data) with biological activity evaluation can be employed to enhance the bioassay guided fractionation process.

7.13.2 Bioassay-Guided Fractionation

The traditional approach to discovering bioactive constituents from complex natural product mixtures is a range of related techniques generally referred to as "bioassay-guided fractionation." With bioassay guided fractionation, extracts are simplified and separated using various chromatographic techniques, the resulting (sub)fractions are screened for biological activity, and the process is repeated iteratively until a single bioactive constituent has been isolated and identified (Fig. 1).[7,8] Bioassay-guided fractionation has remained the "gold standard" of natural product discovery for the last few decades, and has resulted in the discovery of many bioactive compounds, including paclitaxel,[9] artemisinin,[10] and vincristine/vinblastine.[11]

Despite its significance in natural product discovery, it is not uncommon with bioassay guided fraction to experience a loss of activity or failure to isolate a single bioactive constituent from the complex mixture.[12] The bioactive principles are sometimes minor constituents,[13] and bioassay-guided fractionation tends to be biased towards the most abundant (or highly responsive) components in a mixture.[1,14] This bias has also led to instances of re-isolation of already known active compounds. The problem of re-isolation of known compounds is typically addressed using various "dereplication" approaches, which often rely on comparing spectroscopic data against libraries and databases containing the experimental data of known compounds.[15–18]

7.13.3 A Note on In Vitro Data

Any attempt to identify or characterize bioactive principles from a complex mixture axiomatically requires biological evaluation of the extract(s), fraction(s), or isolated compounds. The most common assessment of biological activity is an in vitro assay that analyzes cellular response to a stimulus, or inhibition/activation of an isolated protein/enzyme.[19–21] In vitro assays have been used to evaluate complex mixtures for potential lead compounds to address an array of health conditions, including chronic diseases such as diabetes, cancer, and obesity,[22,23] infectious diseases (e.g., bacterial, fungal, and parasitic infections),[24,25] and to promote overall health and wellness.[26,27] Cell-based assays have been used to provide insight into potential interactions between natural products and conventional pharmaceuticals,[28–30] investigate how natural products interact with the gut microbiome,[31–33] and to model how nutritional matrices could impact the pharmacodynamics and pharmacokinetics of natural product constituents.[34] Furthermore, in vitro data have the potential to provide mechanistic or structure-function insight, which can be imperative in guiding future evaluation via mathematical or statistical modeling experiments, as well as in vivo or clinical studies.[35]

[a]Current address: Department of Veterinary and Biomedical Sciences, Pennsylvania State University, University Park, PA, United States.

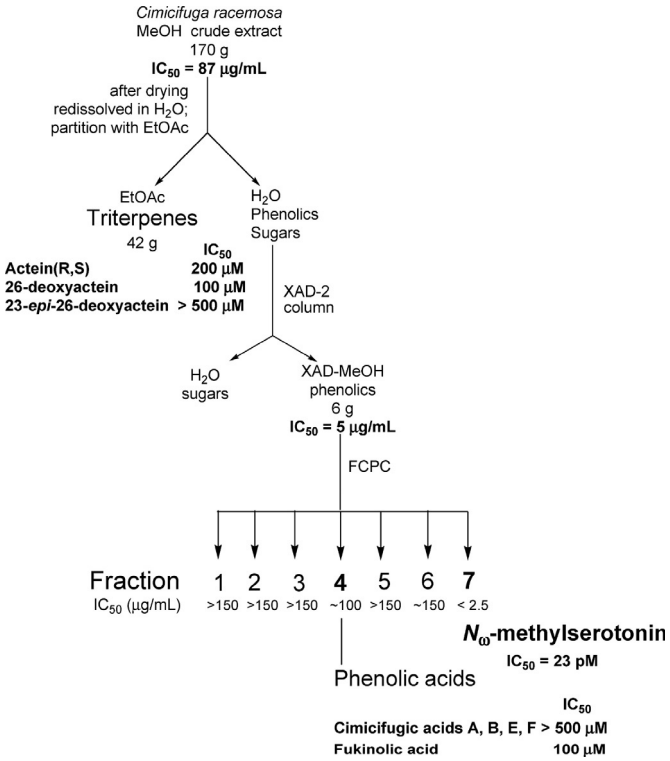


Fig. 1 Representative scheme for bioassay-guided fractionation employed for discovery of bioactive natural products. Bioassay-guided fractionation is an iterative process that involves repeated fractionation and bioassaying until a single bioactive entity (i.e., pure compound) is isolated from the mixture. In this case, three rounds of separation, with 5-HT7 receptor binding evaluation each round, yielded the sertotogenic Nω-methylserotonin. Reproduced with permission from Powell, S. L.; Gödecke, T.; Nikolic, D.; Chen, S.-N.; Ahn, S.; Dietz, B.; Farnsworth, N. R.; van Breemen, R. B.; Lankin, D. C.; Pauli, G. F.; Bolton, J. L. In Vitro Serotonergic Activity of Black Cohosh and Identification of Nω-Methylserotonin as a Potential Active Constituent. *J. Agric. Food Chem.* **2008**, *56*(24), 11718–11726, American Chemical Society.

However, it is worth noting that in vitro assay data does not always translate to activity in the whole organism. This is due to the increased complexity of whole biological systems (i.e., organisms vs single cells/isolated proteins), and the potential effects of other biological factors (e.g., cytosol composition, microbiota, etc.) that are not accounted for in most in vitro studies. García-Contreras et al.[36] found that reaction medium composition plays a role in enzyme activity, and thus researchers must take care extrapolating conclusions from in vitro studies investigating enzymatic inhibition by natural products, as there could be augmentation of activity and efficacy when scaling up to a more complex model system.[36] Another challenge is that many of the purported health benefits of botanical natural products (e.g., "maintain joint health and flexibility", "support proper immune function", or "enhance adaptability to stress") are difficult to model with an in vitro assay. Successful bioassay-guided fractionation of natural products mixtures necessitates the availability of a relevant, robust and translatable biological assay.

7.13.4 Biochemometrics Modeling

Metabolomics represents an inclusive approach to analyze the entirety of measure metabolites from a complex mixture and is a rapidly emerging field in a number of disciplines, including natural product chemistry and drug discovery.[37–41] Metabolomics studies can utilize data from a variety of analytical instrumentation, though mass spectrometry (MS) and nuclear magnetic resonance spectroscopy (NMR) are the most common.[42,43] Metabolomic profiling results in the generation of large datasets comprising potentially major and minor components from the mixture and require tools to parse the data and provide meaningful analysis. The most commonly employed statistical tool for comparing chemical composition of complex mixtures is principal component analysis (PCA).[44] However, PCA analysis only takes into consideration the chemical data.[43]

There has been increased interest in integrating the chemical fingerprints obtained from metabolomics studies of natural products with bioactivity data and statistically modeling the changing behavior as composition alters across fractions, extracts, or taxa. These methods, collectively termed "biochemometrics",[43,45] have emerged as means to improve the efficiency of discovery of new bioactive natural product metabolites and overcome some of the limitations of traditional bioassay-guided fractionation.[46,47] To merge two disparate datasets into a single model, statistical methods are required. Several different statistical approaches have been employed for this purpose, including Pearson correlations,[13,48,49] partial least squares (PLS),[43,50–52] PLS-discriminant analysis (PLS-DA, OPLS-DA)[53–55] and hierarchical cluster analysis (HCA).[56]

Biochemometrics has been successfully applied by different groups to identify bioactive principles from natural product mixtures. For example, a recent study investigating the anti-tuberculosis (*Mycobacterium tuberculosis* H37Rv) activity of volatile

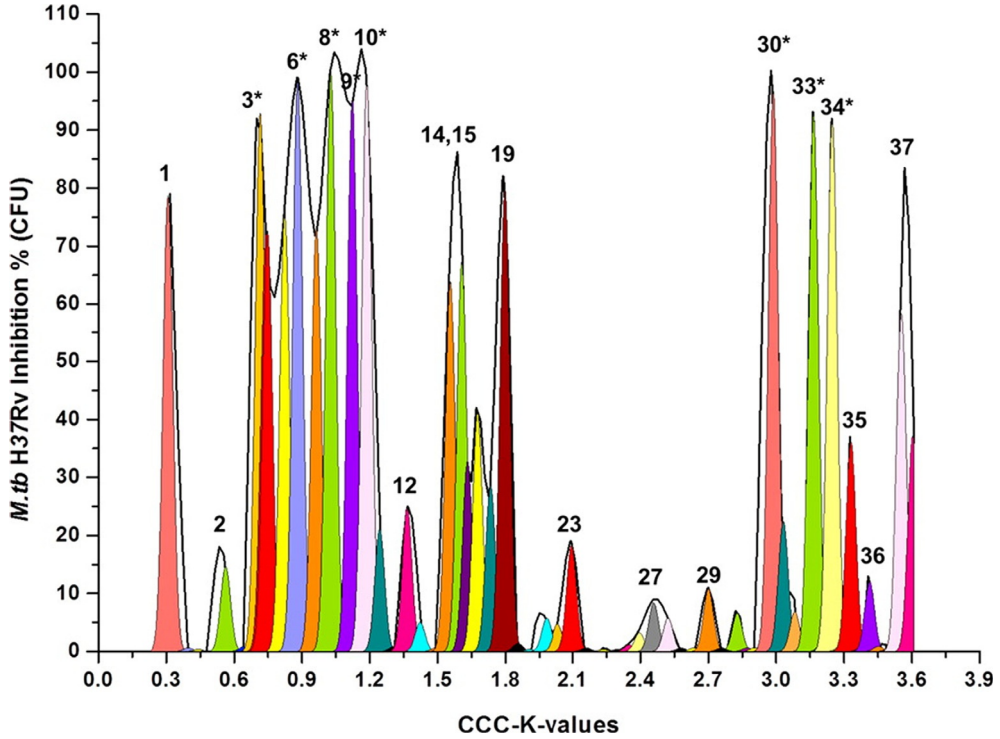


Fig. 2 A deconvoluted "biochromatogram" of fractions of *E. citriodora* essential oil (HR-HSCCC K-values on *x*-axis) and their corresponding percentage inhibition against *Mycobacterium tuberculosis* H37Rv in a gaseous contact assay (*y*-axis). The biochromatogram distinguished 37 peaks of significant activity. Reproduced with permission from Ramos Alvarenga, R. F.; Wan, B.; Inui, T.; Franzblau, S. G.; Pauli, G. F.; Jaki, B. U. Airborne Antituberculosis Activity of *Eucalyptus citriodora* Essential Oil. *J. Nat. Prod.* **2014**, *77* (3), 603–610, American Chemical Society.

components the essential oil of *Eucalyptus citriodora* used Pearson correlation to identify 37 volatile anti-tuberculosis compounds (Fig. 2). The monoterpene citronellol demonstrated significant inhibition of tuberculosis in the airborne stage, and the observed activity was attributed to mixtures of several compounds (citronellol, (*S*)-citronella and eucalyptol).[57]

Partial least squares (PLS) has become one of the dominant statistical modeling approaches for biochemometrics studies. Biochemometric analyses are frequently multi-dimensional models, and can be difficult to decipher. To this end, a number of visualization techniques have been developed to interpret PLS models, with the S-plot and the selectivity ratio being the two dominant metrics.[43]

Abdelmohsen et al. (2014) employed OPLS-DA to model potential bioactive metabolites from *Actinokineospora* sp. strain EG49, an actinomycete cultivated from the sponge *Spheciospongia vagabunda*. Using the source of the metabolite (broth supernatant vs. cell) as the dependent variables, the model generated an S-plot (Fig. 3), in which the upper and lower extremes reveal metabolites that most strongly correlate and covary with the biological activity, suggesting they are responsible for the observed activity. This study yielded two new O-glycosylated angucyclines, named actinosporins A and B, which were isolated and subsequently demonstrated activity against *Trypanosoma brucei brucei*.[58] S-plots were also used to identify a neuroprotective agent against Parkinson's disease. The S-plot, from an OPLS model of *Alpinia oxyphylla* fractions and associated neuroprotective activities, revealed oxyphylla A as a potentially powerful neuroprotectant.[59] Finally, studying the stipe of *Ganoderma sinense*, the S-plot approach was used to uncover five cytotoxic compounds to prioritize for follow-up studies.[53]

A complimentary approach for highlighting significant PLS model features from the independent dataset is the selectivity ratio, which provides a quantitative metric for each variable and its ability to distinguish between different groups. The selectivity ratio is defined as the ratio between the explained and residual variance of the spectral variables (features) of a target-projection from the PLS model. With selectivity ratio modeling, features (spectral variables) that have high selectivity ratio have are more strongly associated with bioactive (rather than nonbioactive) samples.[43,60,61] A recent study compared the use of the S-plot and selectivity ratio analysis when PLS datasets were employed to identify antimicrobial constituents from two endophytic fungi (*Alternaria* sp. and *Pyrenochaeta* sp.). For *Pyrenochaeta* sp., both the S-plot and selectivity ratio identified macrosphelide A as the dominant antimicrobial component. However, interpreting the PLS model from *Alternaria* sp., the S-plot suggested alternariol monomethyl ether as the main bioactive principle, while the selectivity ratio yielded altersetin as the responsible compound. Follow-up assaying confirmed altersetin as the more potent bioactive component, with a minimum inhibitory concentration of 0.59 μM against a sensitive strain of *Staphylococcus aureus* (compared to 275 μM for alternariol monomethyl ether).[43] In a pharmacokinetic study of *Camellia sinensis* (green tea), a PLS model was used to predict which constituents were responsible for the in vitro inhibition of intestinal UDP-glucuronosyltransferases (UGTs). The selectivity ratio predicted (−)-epicatechin gallate (ECG) as a major contributor to inhibition of UGTs (Fig. 4).[62] This prediction was validated with follow up assays of isolated ECG.

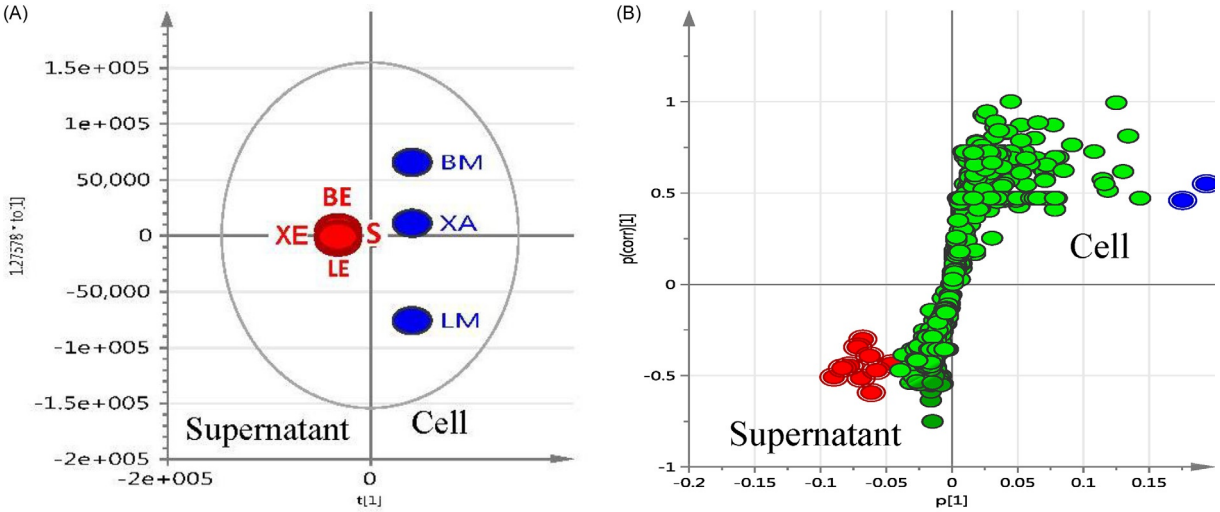
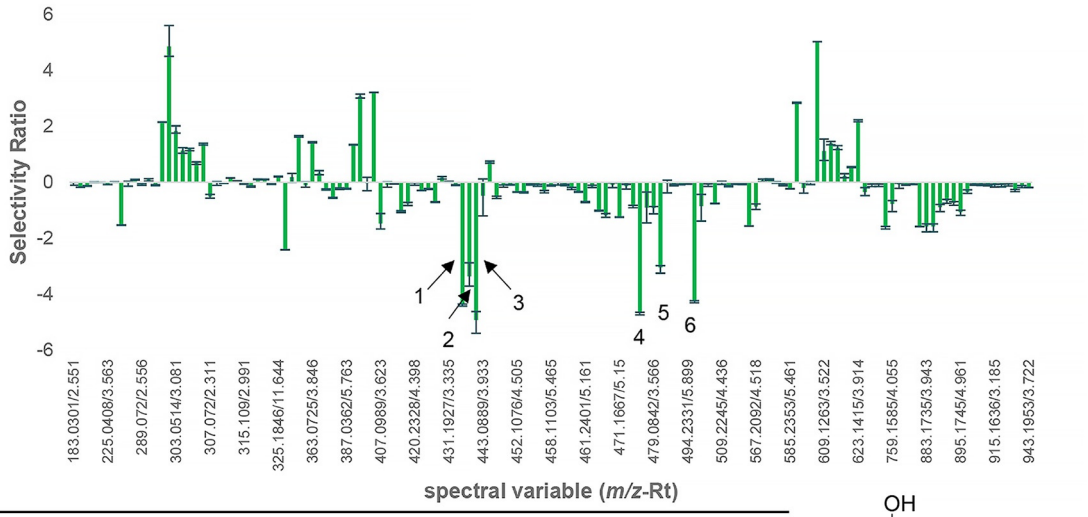


Fig. 3 Orthogonal partial least square-discriminant analysis (OPLS-DA) (A) and resulting S-plot (B) of supernatant broth (*red*) vs. cellular (*blue*) extracts from cultures of Actinokineospora sp. strain EG49. The S-plot's upper (*blue dots*) and lower (*red dots*) termini represent unique secondary metabolites that differentiated the two classes, and was used to identify actinosporins A and B. Reproduced with permission from Abdelmohsen, U. R.; Cheng, C.; Viegelmann, C.; Zhang, T.; Grkovic, T.; Ahmed, S.; Quinn, R. J.; Hentschel, U.; Edrada-Ebel, R. Dereplication Strategies for Targeted Isolation of New Antitrypanosomal Actinosporins A and B From a Marine Sponge Associated-*Actinokineospora* sp. EG49. *Mar. Drugs* **2014**, *12* (3), 1220, MDPI.



	Ion (*m/z*)[a]		Molecular formula	Δ (ppm)	Tentative identification
1	443.0869	$[M-H]^-$	$C_{22}H_{17}O_{10}$	+ 0.9	^{13}C isotope peak of ECG
2	504.0890	$[M+ACN+Na-2H]^-$	$C_{24}H_{19}NO_{10}Na$	+ 2.7	Acetonitrile-sodium adduct of ECG
3	477.0597	$[M+Cl]^-$	$C_{22}H_{18}O_{10}Cl$	+ 0.8	Chloride adduct of ECG
4	441.0832	$[M-H]^-$	$C_{22}H_{17}O_{10}$	+ 1.1	Deprotonated molecular ion of ECG
5	442.0864	$[M-H]^-$	$C_{22}H_{17}O_{10}$	+ 1.8	^{13}C isotope peak ECG
6	487.0886	$[M+FA-H]^-$	$C_{23}H_{19}O_{12}$	+ 1.0	Formic acid adduct of ECG

[a] ACN, acetonitrile; FA, formic acid.

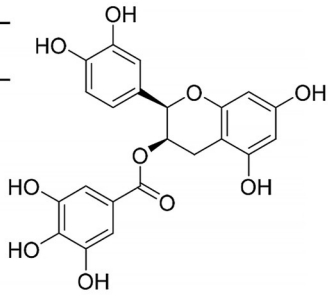


Fig. 4 Biochemometric analysis of the inhibition of UDP-glucuronosyltransferase (UGT) activity by green tea constituents. (A) Mass spectrometry-based metabolomics was correlated with bioactivity to generate a partial least squares (PLS) model, and the selectivity ratio for each feature detected in the green tea extract. The more negative values on the *y*-axis represent more significant contribution to the observed UGT inhibitory activity. (B) All six detected ions represented isotope peaks and adduct ions of (–)-epicatechin gallate (ECG) molecular ion. Reproduced with permission from Tian, D.-D.; Kellogg, J. J.; Okut, N.; Oberlies, N. H.; Cech, N. B.; Shen, D. D.; McCune, J. S.; Paine, M. F. Identification of Intestinal UDP-Glucuronosyltransferase Inhibitors in Green Tea (*Camellia sinensis*) Using a Biochemometric Approach: Application to Raloxifene as a Test Drug Via In Vitro to In Vivo Extrapolation. *Drug Metab. Dispos.* **2018**, *46*, 552–560, American Society for Pharmacology and Experimental Therapeutics (ASPET).

7.13.5 Structural Annotation

Biochemometric methodologies rely on the "raw" spectral data provided by the analytical instrumentation as independent variable input. For mass spectrometry-based metabolomics, these data take the form of the *m/z* value of the analyte, which can be paired with retention time if an LC-MS system is employed. For NMR metabolomics, the spectroscopic variables represent binned peak signals for the nucleus of interest (traditionally ^{1}H or ^{13}C) and depends on the resolution of the acquisition instrument.[63–65] However, this does not provide any structural information about the putative bioactive metabolites, especially if they are unknown.

Attempts to better annotate and elucidate chemical structural information combined biochemometrics analyses with metabolomics tools that provide structural insight into the potential bioactives. One such approach utilizes the Global Natural Product Social Molecular Networking (GNPS) approach, in which MS/MS fragmentation data is mined from metabolomics datasets to provide a network of structurally-similar metabolites.[16] GNPS has proven to be a powerful tool in annotation and de-replication of mass spectrometry-based datasets.[66,67] For example, a recent study combined biochemometrics analysis of *Angelica keiskei* with molecular networking to tentatively annotate putative bioactive metabolites (Fig. 5). The analysis identified a set of chalcone analogs as being potentially active, and subsequent isolation efforts yielded two known components as well as one compound that

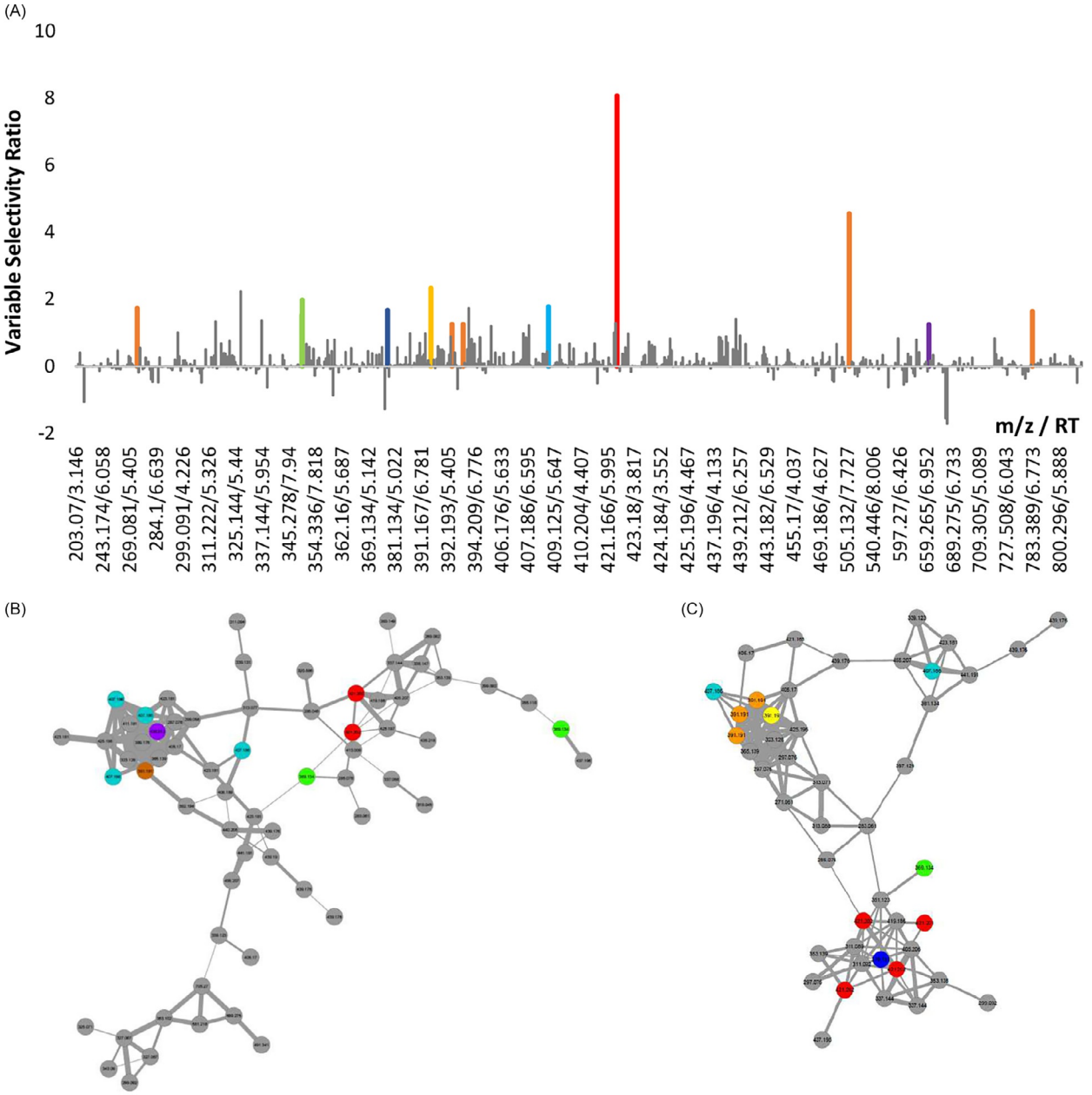


Fig. 5 PLS model of antimicrobial activity from Angelica keiskei fractions yielded a selectivity ratio plot (A), and was combined with selected molecular networks of second-stage (B) and third-stage (C) fractions. Metabolites correlated with bioactivity (higher magnitude selectivity ratio scores, (A)) were color coded to correspond with the molecular network analysis (B and C). Molecular networks indicated that a particular class of compounds, chalcone derivatives, was responsible for A. keiskei's antimicrobial activity. Reproduced with permission from Caesar, L. K.; Kellogg, J. J.; Kvalheim, O. M.; Cech, R. A.; Cech, N. B. Integration of Biochemometrics and Molecular Networking to Identify Antimicrobials in *Angelica keiskei*. *Planta Med.* **2018**, *84* (9–10), 721–728, Thieme.

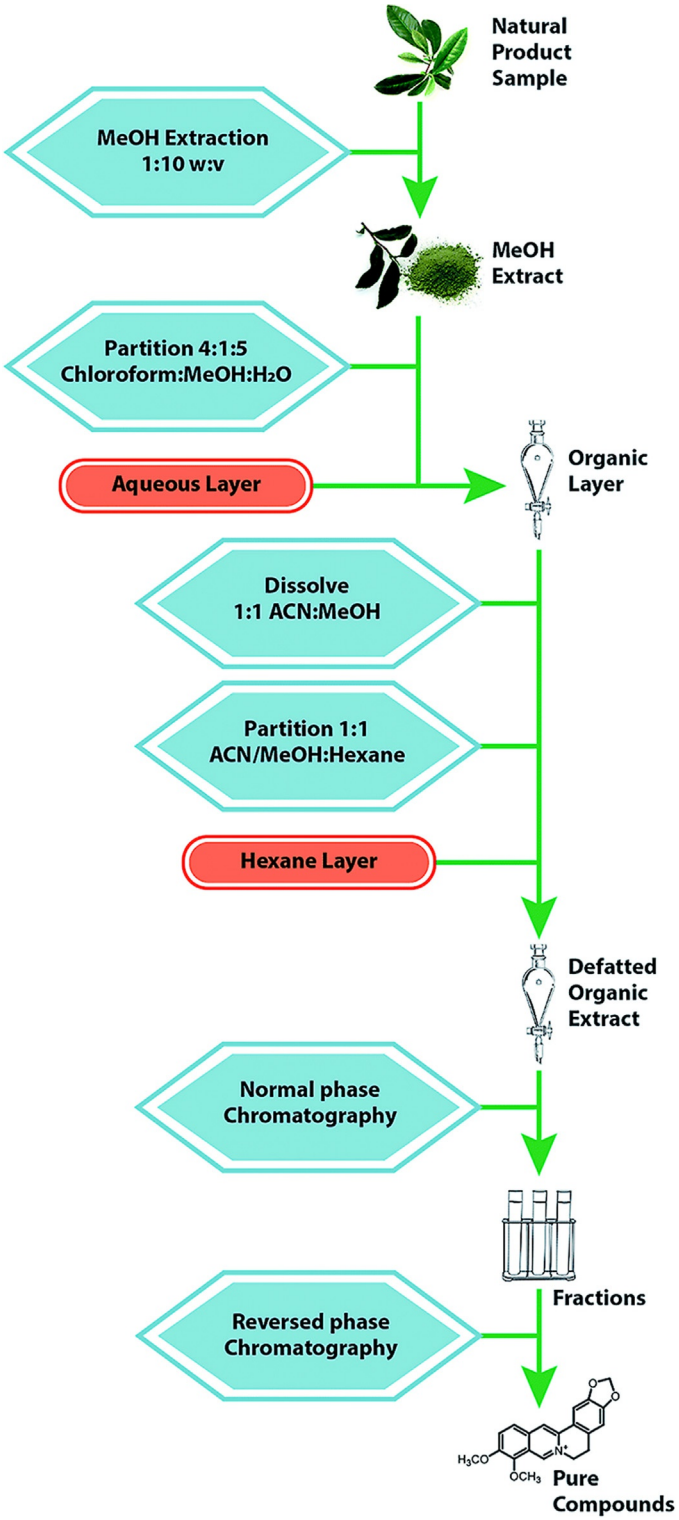


Fig. 6 A representative natural product extraction and fractionation scheme. The sample material first undergoes solvent extraction, after which it is subjected to liquid-liquid partitioning to produce an organic extract. This residue is solubilized and then fractionated with a normal-phase or open column chromatography system, and if further purification is necessary, the fractions are separated via reverse-phase HPLC to yield purified compounds. Reproduced with permission from Kellogg, J. J.; Paine, M. F.; McCune, J. S.; Oberlies, N. H.; Cech, N. B. Selection and Characterization of Botanical Natural Products for Research Studies: A NaPDI Center Recommended Approach. *Nat. Prod. Rep.* **2019**, *36* (8), 1196–1221, Royal Society of Chemistry.

had no previously reported antimicrobial activity.[51] The GNPS approach has been adapted further to create "bioactive molecular networking," where the two analyses are harmonized to directly visualize bioactive metabolites in the context of their structural relationship to one another. Nothias et al. developed a biochemometric workflow to map a "bioactivity score" within a molecular network, and applied this approach to discover antiviral compounds from *Euphorbia dendroides*. This was achieved after a traditional bioassay-guided fractionation procedure was unable to discover the bioactive candidate molecules.[48]

7.13.6 Confirmation of Activity

Biochemometric modeling and interpretation is a statistical approach to prioritize features as putative bioactive metabolites from a more complex mixture. While efforts to identify and quantify constituents of natural product extracts often rely on purified standards,[68] this approach is viable only when standards are available. The alternative, when standards of the compounds of interest are not commercially available, is proceeding with an isolation scheme for the compounds from the complex starting material.[7] A representative review of isolation procedures and techniques is available elsewhere.[69] Generally, a common approach for natural product separation and isolation of constituents is a combination of liquid-liquid partitioning and column chromatography approaches (Fig. 6). A natural product is macerated with an organic solvent, typically methanol, and then partitioned between an immiscible binary solvent system (e.g., chloroform and water or ethyl acetate/water), followed by a "de-fatting" step using hexane-organic solvent partitioning to remove lipid-soluble components.[70] This methodology is inherently flexible and could easily be modified with acid/base conditions or substituting other solvents as the system demands. The extracted residue is dried (typically under nitrogen or reduced pressure) and undergoes normal phase or open column chromatography to provide a first round of fractions. If necessary, fractions are subfractionated using one or more rounds of reversed phase preparative HPLC. This can be a laborious process, and the isolation of even milligrams quantities of a pure isolated compound could take weeks to months of work, and it is possible that it may be ultimately unsuccessful.[71,72] This resource- and time-intensive is another argument in favor of using biochemometric approaches to guide isolation efforts.

7.13.7 Discussion

The identification of bioactive constituents from complex mixtures, with minimal reliance on multiple bioactivity-guided isolation steps remains a challenge for natural product discovery programs. The adaptation of statistical models to integrate metabolomic datasets with corresponding in vitro assay data, represents an important step to improve the efficiency and productivity of discovery efforts. And the combination of complimentary methods (e.g. biochemometrics and molecular networking) has the potential to provide new tools to accelerate discovery. As both the analytical and statistical capabilities of researchers improve, these techniques will drive secondary metabolite discovery and application, and new arenas for combining orthogonal datasets will aid in identifying biological roles of these small molecules.

References

1. Weller, M. G. A Unifying Review of Bioassay-Guided Fractionation, Effect-Directed Analysis and Related Techniques. *Sensors* **2012**, *12* (7), 9181–9209.
2. Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs From 1981 to 2014. *J. Nat. Prod.* **2016**, *79* (3), 629–661.
3. Enke, C. G.; Nagels, L. J. 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), 2539–2546.
4. Cech, N. B. Mass Spectrometry for Natural Products Research: Challenges, Pitfalls, and Opportunities. *LCGC North America* **2013**, *31* (11), 938–947.
5. Alvarez-Zapata, R.; Sánchez-Medina, A.; Chan-Bacab, M.; García-Sosa, K.; Escalante-Erosa, F.; García-Rodríguez, R. V.; Peña-Rodríguez, L. M. Chemometrics-Enhanced High Performance Liquid Chromatography-Ultraviolet Detection of Bioactive Metabolites From Phytochemically Unknown Plants. *J. Chromatogr. A* **2015**, *1422*, 213–221.
6. Simmler, C.; Kulakowski, D.; Lankin, D. C.; McAlpine, J. B.; Chen, S.-N.; Pauli, G. F. Holistic Analysis Enhances the Description of Metabolic Complexity in Dietary Natural Products. *Adv. Nutr.* **2016**, *7* (1), 179–189.
7. Bucar, F.; Wube, A.; Schmid, M. Natural Product Isolation–How to Get From Biological Material to Pure Compounds. *Nat. Prod. Rep.* **2013**, *30* (4), 525–545.
8. Teasdale, M. E.; Shearer, T. L.; Engel, S.; Alexander, T. S.; Fairchild, C. R.; Prudhomme, J.; Torres, M.; Le Roch, K.; Aalbersberg, W.; Hay, M. E.; Kubanek, J. Bromophycoic Acids: Bioactive Natural Products From a Fijian Red Alga *Callophycus sp.*. *J. Org. Chem.* **2012**, *77* (18), 8000–8006.
9. Wani, M. C.; Taylor, H. L.; Wall, M. E.; Coggon, P.; McPhail, A. T. Plant Antitumor Agents. VI. Isolation and Structure of Taxol, a Novel Antileukemic and Antitumor Agent From *Taxus brevifolia*. *J. Am. Chem. Soc.* **1971**, *93* (9), 2325–2327.
10. Tu, Y. The Discovery of Artemisinin (Qinghaosu) and Gifts From Chinese Medicine. *Nat. Med.* **2011**, *17*, 1217.
11. Noble, R. L. The Discovery of the Vinca Alkaloids—Chemotherapeutic Agents Against Cancer. *Biochem. Cell Biol.* **1990**, *68* (12), 1344–1351.
12. Qiu, F.; Cai, G.; Jaki, B. U.; Lankin, D. C.; Franzblau, S. G.; Pauli, G. F. Quantitative Purity–Activity Relationships of Natural Products: The Case of Anti-Tuberculosis Active Triterpenes From Oplopanax Horridus. *J. Nat. Prod.* **2013**, *76* (3), 413–419.
13. Inui, T.; Wang, Y.; Pro, S. M.; Franzblau, S. G.; Pauli, G. F. Unbiased Evaluation of Bioactive Secondary Metabolites in Complex Matrices. *Fitoterapia* **2012**, *83* (7), 1218–1225.
14. Prince, E. K.; Pohnert, G. Searching for Signals in the Noise: Metabolomics in Chemical Ecology. *Anal. Bioanal. Chem.* **2010**, *396* (1), 193–197.
15. El-Elimat, T.; Figueroa, M.; Ehrmann, B. M.; Cech, N. B.; Pearce, C. J.; Oberlies, N. H. High-Resolution MS, MS/MS, and UV Database of Fungal Secondary Metabolites as a Dereplication Protocol for Bioactive Natural Products. *J. Nat. Prod.* **2013**, *76* (9), 1709–1716.
16. Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T. Sharing and Community Curation of Mass Spectrometry Data With Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34* (8), 828.

17. Hoffmann, T.; Krug, D.; Hüttel, S.; Müller, R. Improving Natural Products Identification Through Targeted LC-MS/MS in an Untargeted Secondary Metabolomics Workflow. *Anal. Chem.* **2014**, *86* (21), 10780–10788.
18. Gaudêncio, S. P.; Pereira, F. Dereplication: Racing to Speed up the Natural Products Discovery Process. *Nat. Prod. Rep.* **2015**, *32* (6), 779–810.
19. Balouiri, M.; Sadiki, M.; Ibnsouda, S. K. Methods for In Vitro Evaluating Antimicrobial Activity: A Review. *J. Pharm. Anal.* **2016**, *6* (2), 71–79.
20. van Breemen, R. B. Development of Safe and Effective Botanical Dietary Supplements. *J. Med. Chem.* **2015**, *58*, 8360–8372.
21. Brusotti, G.; Cesari, I.; Dentamaro, A.; Caccialanza, G.; Massolini, G. Isolation and Characterization of Bioactive Compounds From Plant Resources: The Role of Analysis in the Ethnopharmacological Approach. *J. Pharm. Biomed. Anal.* **2014**, *87*, 218–228.
22. Dinic, J.; Podolski-Renic, A.; Stankovic, T.; Bankovic, J.; Pesic, M. New Approaches With Natural Product Drugs for Overcoming Multidrug Resistance in Cancer. *Curr. Pharm. Des.* **2015**, *21* (38), 5589–5604.
23. Wang, K.; Bao, L.; Zhou, N.; Zhang, J.; Liao, M.; Zheng, Z.; Wang, Y.; Liu, C.; Wang, J.; Wang, L. Structural Modification of Natural Product Ganomycin I Leading to Discovery of a α-Glucosidase and HMG-CoA Reductase Dual Inhibitor Improving Obesity and Metabolic Dysfunction In Vivo. *J. Med. Chem.* **2018**, *61* (8), 3609–3625.
24. Cretton, S.; Dorsaz, S.; Azzollini, A.; Favre-Godal, Q.; Marcourt, L.; Ebrahimi, S. N.; Voinesco, F.; Michellod, E.; Sanglard, D.; Gindro, K. Antifungal Quinoline Alkaloids From Waltheria Indica. *J. Nat. Prod.* **2016**, *79* (2), 300–307.
25. Dong, M.; Pfeiffer, B.; Altmann, K.-H. Recent Developments in Natural Product-Based Drug Discovery for Tuberculosis. *Drug Discov. Today* **2017**, *22* (3), 585–591.
26. Dietz, B. M.; Hajirahimkhan, A.; Dunlap, T. L.; Bolton, J. L. Botanicals and Their Bioactive Phytochemicals for Women's Health. *Pharmacol. Rev.* **2016**, *68* (4), 1026–1073.
27. Harvey, A. L.; Edrada-Ebel, R.; Quinn, R. J. The Re-emergence of Natural Products for Drug Discovery in the Genomics Era. *Nat. Rev. Drug Discov.* **2015**, *14* (2), 111.
28. Gufford, B. T.; Chen, G.; Vergara, A. G.; Lazarus, P.; Oberlies, N. H.; Paine, M. F. Milk Thistle Constituents Inhibit Raloxifene Intestinal Glucuronidation: A Potential Clinically Relevant Natural Product–Drug Interaction. *Drug Metab. Dispos.* **2015**, *43* (9), 1353–1359.
29. Li, G.; Huang, K.; Nikolic, D.; Van Breemen, R. B. High-Throughput Cytochrome P450 Cocktail Inhibition Assay for Assessing Drug-Drug and Drug-Botanical Interactions. *Drug Metab. Dispos.* **2015**, *43* (11), 1670–1678. https://doi.org/10.1124/dmd.115.065987.
30. Posadzki, P.; Watson, L.; Ernst, E. Herb–Drug Interactions: An Overview of Systematic Reviews. *Br. J. Clin. Pharmacol.* **2013**, *75* (3), 603–618.
31. Greiner, A. K.; Papineni, R. V. L.; Umar, S. Chemoprevention in Gastrointestinal Physiology and Disease. Natural Products and Microbiome. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2014**, *307* (1), G1–G15.
32. Possemiers, S.; Bolca, S.; Verstraete, W.; Heyerick, A. The Intestinal Microbiome: A Separate Organ Inside the Body With the Metabolic Potential to Influence the Bioactivity of Botanicals. *Fitoterapia* **2011**, *82* (1), 53–66.
33. Tian, Y.; Cai, J.; Gui, W.; Nichols, R. G.; Koo, I.; Zhang, J.; Anitha, M.; Patterson, A. D. Berberine Directly Affects the Gut Microbiota to Promote Intestinal Farnesoid X Receptor Activation. *Drug Metab. Dispos.* **2019**, *47* (2), 86.
34. Van Loo-Bouwman, C. A.; Naber, T. H. J.; Minekus, M.; van Breemen, R. B.; Hulshof, P. J. M.; Schaafsma, G. Food Matrix Effects on Bioaccessibility of β-Carotene Can Be Measured in an In Vitro Gastrointestinal Model. *J. Agric. Food Chem.* **2014**, *62* (4), 950–955.
35. Paine, M. F.; Shen, D. D.; McCune, J. S. Recommended Approaches for Pharmacokinetic Natural Product-Drug Interaction Research: A NaPDI Center Commentary. *Drug Metab. Dispos.* **2018**, *46* (7), 1041–1045. https://doi.org/10.1124/dmd.117.079962.
36. García-Contreras, R.; Vos, P.; Westerhoff, H. V.; Boogerd, F. C. Why In Vivo May Not Equal In Vitro—New Effectors Revealed by Measurement of Enzymatic Activities Under the Same In Vivo-Like Assay Conditions. *FEBS J.* **2012**, *279* (22), 4145–4159.
37. Yang, J. Y.; Sanchez, L. M.; Rath, C. M.; Liu, X.; Boudreau, P. D.; Bruns, N.; Glukhov, E.; Wodtke, A.; De Felicio, R.; Fenner, A. Molecular Networking as a Dereplication Strategy. *J. Nat. Prod.* **2013**, *76* (9), 1686–1699.
38. Allard, P.-M.; Genta-Jouve, G.; Wolfender, J.-L. Deep Metabolome Annotation in Natural Products Research: Towards a Virtuous Cycle in Metabolite Identification. *Curr. Opin. Chem. Biol.* **2017**, *36*, 40–49.
39. Olawode, E. O.; Tandlich, R.; Cambray, G. ^{1}H-NMR Profiling and Chemometric Analysis of Selected Honeys From South Africa, Zambia, and Slovakia. *Molecules* **2018**, *23* (3), 578.
40. Depke, T.; Franke, R.; Brönstrup, M. Clustering of MS2 Spectra Using Unsupervised Methods to Aid the Identification of Secondary Metabolites From *Pseudomonas aeruginosa*. *J. Chromatogr. B* **2017**, *1071*, 19–28.
41. Guo, X.; Long, P.; Meng, Q.; Ho, C.-T.; Zhang, L. An Emerging Strategy for Evaluating the Grades of Keemun Black Tea by Combinatory Liquid Chromatography-Orbitrap Mass Spectrometry-Based Untargeted Metabolomics and Inhibition Effects on α-Glucosidase and α-Amylase. *Food Chem.* **2018**, *246*, 74–81.
42. Kellogg, J. J.; Graf, T. N.; Paine, M. F.; McCune, J. S.; Kvalheiam, O. M.; Oberlies, N. H.; Cech, N. B. Comparison of Metabolomic Approaches for Evaluating the Variability of Complex Botanical Preparations: Green Tea (*Camellia sinensis*) as a Case Study. *J. Nat. Prod.* **2017**, *80*, 1457–1466.
43. Kellogg, J. J.; Todd, D. A.; Egan, J. M.; Raja, H. A.; Oberlies, N. H.; Kvalheim, O. M.; Cech, N. B. Biochemometrics for Natural Product Research: Comparison of Data Analysis Approaches and Application to Identification of Bioactive Compounds. *J. Nat. Prod.* **2016**, *79* (2), 376–386.
44. Rajalahti, T.; Kvalheim, O. M. Multivariate Data Analysis in Pharmaceutics: A Tutorial Review. *Int. J. Pharm.* **2011**, *417* (1–2), 280–290.
45. Martens, H.; Bruun, S. W.; Adt, I.; Sockalingum, G. D.; Kohler, A. Pre-Processing in Biochemometrics: Correction for Path-Length and Temperature Effects of Water in FTIR Bio-Spectroscopy by EMSC. *J. Chemometr.* **2006**, *20* (8–10), 402–417.
46. Calderon, A. I. Editorial:Combination of Mass Spectrometry and Omics/Chemometrics Approaches to Unravel Bioactives in Natural Products Mixtures. *Comb. Chem. High Throughput Screen.* **2017**, *20* (4), 278.
47. Roberts, G. K.; Gardner, D.; Foster, P. M.; Howard, P. C.; Lui, E.; Walker, L.; van Breemen, R. B.; Auerbach, S. S.; Rider, C. Finding the Bad Actor: Challenges in Identifying Toxic Constituents in Botanical Dietary Supplements. *Food Chem. Toxicol.* **2019**, *124*, 431–438.
48. Nothias, L.-F.; Nothias-Esposito, M.; da Silva, R.; Wang, M.; Protsyuk, I.; Zhang, Z.; Sarvepalli, A.; Leyssen, P.; Touboul, D.; Costa, J.; Paolini, J.; Alexandrov, T.; Litaudon, M.; Dorrestein, P. C. Bioactivity-Based Molecular Networking for the Discovery of Drug Leads in Natural Product Bioassay-Guided Fractionation. *J. Nat. Prod.* **2018**, *81* (4), 758–767.
49. Richards, L. A.; Oliveira, C.; Dyer, L. A.; Rumbaugh, A.; Urbano-Munoz, F.; Wallace, I. S.; Dodson, C. D.; Jeffrey, C. S. Shedding Light on Chemically Mediated Tri-Trophic Interactions: A ^{1}H-NMR Network Approach to Identify Compound Structural Features and Associated Biological Activity. *Front. Plant Sci.* **2018**, *9*, 1155.
50. Britton, E. R.; Kellogg, J. J.; Kvalheim, O. M.; Cech, N. B. Biochemometrics to Identify Synergists and Additives From Botanical Medicines: A Case Study With Hydrastis Canadensis (Goldenseal). *J. Nat. Prod.* **2018**, *81* (3), 484–493.
51. Caesar, L. K.; Kellogg, J. J.; Kvalheim, O. M.; Cech, R. A.; Cech, N. B. Integration of Biochemometrics and Molecular Networking to Identify Antimicrobials in *Angelica keiskei*. *Planta Med.* **2018**, *84* (9-10), 721–728.
52. Kvalheim, O. M.; Chan, H.-y.; Benzie, I. F. F.; Szeto, Y.-t.; Tzang, A. H.-c.; Mok, D. K.-w.; Chau, F.-t. Chromatographic Profiling and Multivariate Analysis for Screening and Quantifying the Contributions From Individual Components to the Bioactive Signature in Natural Products. *Chemom. Intel. Lab. Syst.* **2011**, *107* (1), 98–105.
53. Chan, K.-M.; Yue, G. G.-L.; Li, P.; Wong, E. C.-W.; Lee, J. K.-M.; Kennelly, E. J.; Lau, C. B.-S. Screening and Analysis of Potential Anti-Tumor Components From the Stipe of Ganoderma Sinense Using High-Performance Liquid Chromatography/Time-of-Flight Mass Spectrometry With Multivariate Statistical Tool. *J. Chromatogr. A* **2017**, *1487*, 162–167.
54. Chagas-Paula, D.; Zhang, T.; Da Costa, F.; Edrada-Ebel, R. A Metabolomic Approach to Target Compounds From the Asteraceae Family for Dual COX and LOX Inhibition. *Metabolites* **2015**, *5* (3), 404–430.
55. Wen, C.; Wang, D.; Li, X.; Huang, T.; Huang, C.; Hu, K. Targeted Isolation and Identification of Bioactive Compounds Lowering Cholesterol in the Crude Extracts of Crabapples Using UPLC-DAD-MS-SPE/NMR Based on Pharmacology-Guided PLS-DA. *J. Pharm. Biomed. Anal.* **2018**, *150*, 144–151.
56. Patras, A.; Brunton, N. P.; Downey, G.; Rawson, A.; Warriner, K.; Gernigon, G. Application of Principal Component and Hierarchical Cluster Analysis to Classify Fruits and Vegetables Commonly Consumed in Ireland Based on In Vitro Antioxidant Activity. *J. Food Compos. Anal.* **2011**, *24* (2), 250–256.

57. Ramos Alvarenga, R. F.; Wan, B.; Inui, T.; Franzblau, S. G.; Pauli, G. F.; Jaki, B. U. Airborne Antituberculosis Activity of *Eucalyptus citriodora* Essential Oil. *J. Nat. Prod.* **2014**, *77* (3), 603–610.
58. Abdelmohsen, U. R.; Cheng, C.; Viegelmann, C.; Zhang, T.; Grkovic, T.; Ahmed, S.; Quinn, R. J.; Hentschel, U.; Edrada-Ebel, R. Dereplication Strategies for Targeted Isolation of New Antitrypanosomal Actinosporins A and B From a Marine Sponge Associated-*Actinokineospora* sp. EG49. *Mar. Drugs* **2014**, *12* (3), 1220.
59. Li, G.; Zhang, Z.; Quan, Q.; Jiang, R.; Szeto, S. S. W.; Yuan, S.; Wong, W.-t.; Lam, H. H. C.; Lee, S. M.-Y.; Chu, I. K. Discovery, Synthesis, and Functional Characterization of a Novel Neuroprotective Natural Product From the Fruit of Alpinia Oxyphylla for Use in Parkinson's Disease Through LC/MS-Based Multivariate Data Analysis-Guided Fractionation. *J. Proteome Res.* **2016**, *15* (8), 2595–2606.
60. Rajalahti, T.; Arneberg, R.; Berven, F. S.; Myhr, K.-M.; Ulvik, R. J.; Kvalheim, O. M. Biomarker Discovery in Mass Spectral Profiles by Means of Selectivity Ratio Plot. *Chemom. Intel. Lab. Syst.* **2009**, *95* (1), 35–48.
61. Rajalahti, T.; Arneberg, R.; Kroksveen, A. C.; Berle, M.; Myhr, K.-M.; Kvalheim, O. M. Discriminating Variable Test and Selectivity Ratio Plot: Quantitative Tools for Interpretation and Variable (Biomarker) Selection in Complex Spectral or Chromatographic Profiles. *Anal. Chem.* **2009**, *81* (7), 2581–2590.
62. Tian, D.-D.; Kellogg, J. J.; Okut, N.; Oberlies, N. H.; Cech, N. B.; Shen, D. D.; McCune, J. S.; Paine, M. F. Identification of Intestinal UDP-Glucuronosyltransferase Inhibitors in Green Tea (*Camellia sinensis*) Using a Biochemometric Approach: Application to Raloxifene as a Test Drug Via In Vitro to In Vivo Extrapolation. *Drug Metab. Dispos.* **2018**, *46*, 552–560.
63. Clark, T. N.; Berrué, F.; Calhoun, L.; Boland, P.; Kerr, R.; Johnson, J. A.; Gray, C. A. A Comparison Between the Application of NMR and LC-HRMS Based Metabolomics on the Discovery of Natural Products From Endophytic Fungi. *Planta Med.* **2016**, *82* (S 01), P690.
64. Clendinen, C. S.; Stupp, G. S.; Ajredini, R.; Lee-McMullen, B.; Beecher, C.; Edison, A. S. An Overview of Methods Using ^{13}C for Improved Compound Identification in Metabolomics and Natural Products. *Front. Plant Sci.* **2015**, *6* (611).
65. Robinette, S. L.; Brüschweiler, R.; Schroeder, F. C.; Edison, A. S. NMR in Metabolomics and Natural Products Research: Two Sides of the Same Coin. *Acc. Chem. Res.* **2012**, *45* (2), 288–297.
66. Allard, P.-M.; Péresse, T.; Bisson, J.; Gindro, K.; Marcourt, L.; Pham, V. C.; Roussi, F.; Litaudon, M.; Wolfender, J.-L. Integration of Molecular Networking and In-Silico MS/MS Fragmentation for Natural Products Dereplication. *Anal. Chem.* **2016**, *88* (6), 3317–3323.
67. Duncan, K. R.; Crüsemann, M.; Lechner, A.; Sarkar, A.; Li, J.; Ziemert, N.; Wang, M.; Bandeira, N.; Moore, B. S.; Dorrestein, P. C. Molecular Networking and Pattern-Based Genome Mining Improves Discovery of Biosynthetic Gene Clusters and Their Products From *Salinispora* species. *Chem. Biol.* **2015**, *22* (4), 460–471.
68. Napolitano, J. G.; Gödecke, T.; Lankin, D. C.; Jaki, B. U.; McAlpine, J. B.; Chen, S.-N.; Pauli, G. F. Orthogonal Analytical Methods for Botanical Standardization: Determination of Green Tea Catechins by qNMR and LC–MS/MS. *J. Pharm. Biomed. Anal.* **2014**, *93*, 59–67.
69. Kellogg, J. J.; Paine, M. F.; McCune, J. S.; Oberlies, N. H.; Cech, N. B. Selection and Characterization of Botanical Natural Products for Research Studies: A NaPDI Center Recommended Approach. *Nat. Prod. Rep.* **2019**, *36* (8), 1196–1221.
70. El-Elimat, T.; Raja, H. A.; Graf, T. N.; Faeth, S. H.; Cech, N. B.; Oberlies, N. H. Flavonolignans From *Aspergillus iizukae*, a Fungal Endophyte of Milk Thistle (*Silybum marianum*). *J. Nat. Prod.* **2014**, *77* (2), 193–199.
71. Graf, T. N.; Wani, M. C.; Agarwal, R.; Kroll, D. J.; Oberlies, N. H. Gram-Scale Purification of Flavonolignan Diastereoisomers From *Silybum marianum* (Milk Thistle) Extract in Support of Preclinical In Vivo Studies for Prostate Cancer Chemoprevention. *Planta Med.* **2007**, *73* (14), 1495–1501.
72. Powell, S. L.; Gödecke, T.; Nikolic, D.; Chen, S.-N.; Ahn, S.; Dietz, B.; Farnsworth, N. R.; van Breemen, R. B.; Lankin, D. C.; Pauli, G. F.; Bolton, J. L. In Vitro Serotonergic Activity of Black Cohosh and Identification of Nω-Methylserotonin as a Potential Active Constituent. *J. Agric. Food Chem.* **2008**, *56* (24), 11718–11726.