



Published in final edited form as:

Nat Prod Rep. 2019 August 14; 36(8): 1196–1221. doi:10.1039/c8np00065d.

Selection and Characterization of Botanical Natural Products for Research Studies: A NaPDI Center Recommended Approach

Joshua J. Kellogg^a, Mary F. Paine^b, Jeannine S. McCune^c, Nicholas H. Oberlies^a, Nadja B. Cech^a

^aDepartment of Chemistry & Biochemistry, University of North Carolina at Greensboro, Greensboro, North Carolina, USA. nbcech@uncg.edu

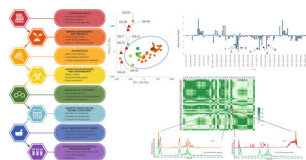
^bDepartment of Pharmaceutical Sciences, College of Pharmacy, Washington State University, Spokane, Washington, USA.

^cDepartment of Population Sciences, City of Hope, Duarte, California, USA.

Abstract

Dietary supplements, which include botanical (plant-based) natural products, constitute a multi-billion-dollar industry in the US. Regulation and quality control for this industry is an ongoing challenge. While there is general agreement that rigorous scientific studies are needed to evaluate the safety and efficacy of botanical natural products used by consumers, researchers conducting such studies face a unique set of challenges. Botanical natural products are inherently complex mixtures, with composition that differs depending on myriad factors including variability in genetics, cultivation conditions, and processing methods. Unfortunately, many studies of botanical natural products are carried out with poorly characterized study material, such that the results are irreproducible and difficult to interpret. This review provides recommended approaches for addressing the critical questions that researchers must address prior to *in vitro* or *in vivo* (including clinical) evaluation of botanical natural products. We describe selection and authentication of botanical material and identification of key biologically active compounds, and compare state-of-the-art methodologies such as untargeted metabolomics with more traditional targeted methods of characterization. The topics are chosen to be of maximal relevance to researchers, and are reviewed critically with commentary as to which approaches are most practical and useful and what common pitfalls should be avoided.

Graphical Abstract



Conflicts of interest

There are no conflicts to declare.

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here].
See DOI: [10.1039/x0xx00000x](https://doi.org/10.1039/x0xx00000x)

1. Introduction

More than 50% of Americans acknowledge taking dietary supplements,^{1, 2} with sales more than tripling since passage of the Dietary Supplement Health and Education Act in 1994.^{3, 4} In many cases, these dietary supplements are prepared from botanical (plant based) material, and as such are referred to as ‘botanical natural products.’ Countless studies have been devoted to the scientific evaluation of the safety and/or efficacy of botanical natural products. Investigators involved in such studies face a unique set of challenges (Table 1). The research methodology for evaluating safety and efficacy of conventional (pharmaceutical) drugs operates with the assumption that the product being tested is a single compound of known purity, identity, and concentration. Natural products differ from their pharmaceutical counterparts in that they are typically complex mixtures, for which the identities and quantities of components present are not fully known. The composition of these natural products can vary depending on the method of preparation or source material used.⁵ Such variability can impact the interpretation of *in vitro*, non-clinical *in vivo*, and/or clinical studies.

To enable studies to be as robust and as widely applicable as possible, selecting, characterizing, and ensuring continued quality and consistency of a botanical natural product is an essential task. The ideal characteristics of a botanical natural product used for research studies include that it is authenticated (of known identity), well-characterized in terms of potentially active constituents, and stable (Table 1). At present, many studies are conducted with botanical natural products that lack one or more of these characteristics. Herein, we seek to provide guidance for effective selection of botanical natural products prior to *in vitro*, non-clinical *in vivo*, or clinical evaluation. We discuss the relative merits of potential analytical approaches, including state-of-the-art metabolomics techniques, and recommend specific, effective and practical guidelines (Figure 1). The recommendations we provide are informed by the coauthors’ experience for several decades researching botanical natural products, culminating in our role as investigators participating in the Center of Excellence for Natural Product Drug Interaction Research (NaPDI Center). The National Center for Complementary and Integrative Health established the NaPDI Center to address the need for robust methods to study the potential for natural products to precipitate clinically significant pharmacokinetic interactions with conventional medications (natural product-drug interactions).⁶ While the focus of the NaPDI Center is specifically on the evaluation of interactions between conventional drugs and natural products, the recommendations provided herein are more broadly applicable to the selection of botanical natural products for any *in vitro* or *in vivo* investigation (including studies to evaluate efficacy or toxicity).

2. Terminology relevant to botanical natural products

“Natural products” are a broad range of substances that can come from a variety of sources, including bacteria, fungi, marine organisms, and plants. The term can be used to describe both complex mixtures and single isolated compounds that come from these mixtures. For this publication, we focus specifically on botanical natural products, i.e. natural products derived from plants. Such botanical natural products are available for use by consumers in a variety of forms, including capsules containing raw or extracted material, extracts, teas

(typically raw plant material that is extracted in hot water prior to use), tinctures (ethanolic extracts), and traditional formulations such as powders used in Traditional Chinese Medicine or Ayurvedic practices. In this manuscript, we use “botanical natural products” as an umbrella term to describe complex plant-based preparations. Terms used elsewhere include: supplements, herbal medicines, herbal drugs, herbs, botanical preparations, nutraceuticals, phytomedicines, and botanical medicines. Additional definitions for terms used in this manuscript are provided in Table 1S (Electronic Supplemental Information, ESI).

From a regulatory perspective, distinctions are often made between various botanical natural products depending on how they are used by consumers. The regulatory approaches and classifications of botanical natural products are complex and differ from country to country, and a detailed comparison is beyond the scope of the current work.⁷ The majority of botanical natural products consumed in the US (including the examples discussed in this manuscript) are regulated either as foods or dietary supplements. From a scientific perspective, the way in which the products are classified and regulated is less important than the material from which they are prepared, and much of the information presented herein is relevant to botanical natural products broadly defined, regardless of their regulatory classification.

3. Journal and Government Agency Guidelines for Characterization of Natural Product Study Material

Several agencies provide guidance for the type of information that should be obtained about a natural product prior to its use for research purposes. The United States Food and Drug Administration (FDA) recommendation regarding botanical drug clinical trials specifies that the investigational new drug application contains “a chemical identification for the active constituents or characteristic markers in the drug substance, if possible”.⁸ The National Center for Complementary and Integrative Health (NCCIH), a component of the National Institutes of Health, funds a large portion of the research conducted on botanical natural products in the United States. NCCIH has established a “Natural Product Integrity Policy”, which requires that researchers provide information about the identity, extraction solvent, characterization (i.e., metabolite profile), stability, standardization, and storage of all natural products that will be used in NCCIH funded studies.⁹ Some individual journals also provide specifications for how complex botanical natural products should be characterized before data on these products can be published.^{10–13} Journal editors have a great deal of influence in this regard, since their publishing requirements cross geopolitical borders.¹⁴

4. Literature Research

Before beginning botanical natural product studies, it is useful to obtain information from the peer-reviewed literature regarding: 1. identities of major metabolites, especially those that may be biologically active; 2. potential *in vitro* or *in vivo* targets and activities; and 3. current knowledge gaps. A recent publication by the NaPDI Center addresses useful strategies for critically evaluating botanical natural product literature to identify gaps in the knowledge base.⁶ Due to the complexity of natural product constituents, it is important to note that reported chemical structures may contain errors or inconsistencies. Thus, the

literature review by an organic or natural product chemist trained in structure elucidation is recommended.

Another important topic to research during sourcing of botanical natural products is information regarding traditional and consumer usage of the product to be studied. The investigator will wish to answer questions such as: What are the most common species used medicinally? Which plant parts (for example, roots, leaves, berries) are typically used? What types of preparations (capsule, tincture, tea, compress, topical versus internal) are employed? Some information of this nature may be available in the peer-reviewed literature, but it is often necessary to seek other sources. Common products used by consumers can be evaluated via national surveys, such as the National Institute of Health's Office of Dietary Supplement's Dietary Supplement Label Database,¹⁵ National Health and Nutrition Examination Survey (NHANES),¹⁶ consumer quality reports^{17, 18} as well as industry or retail sales reports.^{4, 19, 20}

5. Obtaining Botanical Natural Product Study Material

Prior to conducting *in vitro*, non-clinical *in vivo*, and/or clinical studies, it is important that researchers identify an *authentic* natural product (i.e. correct assignment of genus and species) that is available in sufficient quantity to conduct the studies. In the case of clinical studies, this product should be selected to closely resemble (or be identical to) the commercial product(s) used by the target population. Where is such a product obtained? One option is to purchase this material from a commercial distributor. Alternately, researchers may opt to obtain raw material and prepare a natural product themselves or to employ a contract laboratory for formulation.²¹ Finally, several suppliers provide high quality reference materials of botanicals for research purposes. Here we discuss the importance of voucher specimens and compare the relative merits of various sources of botanical material.

5.1 Voucher specimens

Under optimal circumstances, a voucher specimen of any botanical natural product to be studied is collected at the same time and from the same lot that the study material is obtained. The voucher specimen consists of an intact, dried sample of the plant material, including the flower when possible, and as many parts as can reasonably be collected, i.e. roots, stem, leaves, flowers and/or seeds.²² The voucher specimen is used for taxonomic identification of the study material by a trained botanist or otherwise qualified individual. Herbaria have existed for centuries,²³ and these vouchers serve as historical records of plant specimens, and modern uses of these go beyond taxonomy, including areas like conservation biology.²⁴ Thus, a pressed and dried sample of the voucher is deposited in a regional or national herbarium, where it is catalogued and stored for future reference. Herbarium vouchers are essential for preserving a record of the original sample tested and provide lasting, public access to that material in perpetuity. Authentication and retention of voucher specimens is required by major natural product journals.^{10–12, 25} For example, for the *Journal of Natural Products*, authors who “purchase dried ‘herbal remedies’ or other materials from companies must make provision for their proper deposit in a herbarium or other permanent repository, for access by future workers”.¹⁰ Another example is

Phytochemistry, which mandates that samples, “must also include a reference to voucher specimen(s) and voucher number(s) of the plants or other material examined” as well as “the name and address of the authority who identified each non-cultivated plant investigated.” Similar guidelines are in place for other major natural product journals.^{11, 12, 25}

5.2 Contracting an independent laboratory to prepare study material

One way to obtain quality study material is to contract a laboratory to prepare the botanical natural product study material from raw materials to pre-determined specifications. This approach has been employed by other research groups.^{21, 26, 27} A major advantage of employing a contract laboratory to produce study materials is that the researchers can be responsible for ensuring the quality, consistency, and rigor with which the material is prepared. However, contracting a laboratory to prepare a formulation can be expensive, and the production of such formulations is time-consuming and requires access to specialized facilities and expertise. Even more importantly, a botanical natural product prepared by a contract laboratory may not reflect commercial products being consumed by the general public. Thus, many researchers opt instead to conduct studies with commercial products (the same products used by consumers). This was the approach that we took with the NaPDI Center’s studies of both green tea [*Camellia sinensis* (L.) Kuntze (Theaceae)]²⁸ and goldenseal [*Hydrastis canadensis* L. (Ranunculaceae)].²⁹

5.3 Purchasing study material from a commercial supplier

Many botanical natural products can be purchased from online distributors, health food stores, grocery stores. The advantage of purchasing commercial botanical natural products for study is that it is possible to select material that directly reflects what is being used by consumers (with the caveat that there may be a great deal of variability among commercial products). In addition, for clinical studies, commercial products have the advantage of being “food grade” meaning that they can be administered to human subjects. On the other hand, it is difficult to control quality and ensure correct identity of botanical natural products obtained from commercial sources. Commercial botanical natural products typically do not come with associated voucher specimens. Additionally, adulteration or mis-identification of commercial natural products (both intentional and unintentional) is common.^{30–32} For example, in the process of selecting commercial products of goldenseal for clinical studies by the NaPDI Center, we identified several preparations sold as goldenseal (*Hydrastis canadensis*) that were a different botanical, Chinese goldenthread (*Coptis chinensis*).²⁹

5.4 Optimal number of samples to analyze

When a study will be conducted with commercial products, the question arises as to how many samples should be evaluated before one is selected for study material. The goal is to select a sample set that is representative of the variability in products used by the target population. How can this be accomplished? Purchasing every potential commercial botanical natural product would theoretically be an ideal strategy. However, this is not always feasible due to constraints of time and cost. If a botanical natural product has a relatively small commercial footprint, it is possible to acquire all available products at a given time. As an illustrative example, the 35 commercial goldenseal samples selected for the second NaPDI Center study represented an estimated >90% of the readily available commercial products.²⁹

However, when a botanical natural product is found in hundreds or potentially thousands of products, the logistical and potentially financial hurdles of sampling a great number of samples must be taken into account during this planning stage.

For analyses of commercial samples, there has not been a conventional minimum value set on the number of samples needed to yield a robust analysis. Previous studies on botanical natural products have sampled as few as five products,³³ with many sampling 10–18 samples;^{34–37} while some procure over 75 commercial products.³⁸ We elected to select a minimum of 30 commercial samples for each of the NaPDI Center's studies on green tea (34 products)²⁸ and goldenseal (35 products).²⁹ Importantly, these samples were selected based on sales data to represent the products most widely purchased for use by US consumers (our target population). The result was a sample set that captured variability among the products most widely consumed and could be realistically handled within constraints of processing and analysis time.

5.5 Purchasing authenticated botanical reference materials

Several sources provide authenticated botanical reference materials. The National Institute of Standards and Technology (NIST, <https://www.nist.gov>) offers a number of botanical natural products as verified reference materials available for testing and authentication purposes, as does the U.S. Pharmacopeia (USP, <https://www.usp.org>). Additionally, phytochemical reference suppliers like ChromaDex (Irvine, CA, <https://chromadex.com>) provide different grades of botanical reference material. A major advantage of these reference materials is that they have often been characterized already, and information is available about known genetic markers and the presence and identity of known metabolites. Some sources, such as ChromaDex, may also provide information regarding their internal voucher specimens. Even if the intention is to use a commercial product for the final studies, it is advisable, where possible, to obtain an authenticated reference material. The composition of this reference material can then be compared to that of commercial botanical natural products to verify identity and compare quality.^{28, 29}

5.6 Collecting cultivated or wild plant material

Certain botanicals are unavailable as authenticated references, and references must be obtained by other means. In some cases, it is possible to contract third-party vendors to obtain vouchered botanical samples. Another option is to purchase raw or minimally processed plant material from a nursery or farm, or, in the case of a botanical that grows wild, to harvest it from its native environment (sometimes termed “wild crafting”). In these cases, a voucher specimen should be collected at the time of harvest. For wild-harvested plants, permission of the landowner or overseeing agency should be sought prior to collection.³⁹ Researchers seeking to collect botanical material for research purposes should also be aware of regulations that govern intellectual property related to that material. In 1992, the United Nations, through the U.N. Convention on Biodiversity, enacted recognition of traditional knowledge and the natural resources that go with that knowledge. The Convention on Biological Diversity codified enforcement of ecologically sound harvesting practices and sharing of any intellectual or economic benefits that may arise as a result of the utilization of that traditional knowledge.^{40–42} Even though the United States did not ratify

this treaty, many funding agencies, particularly those of the National Institutes of Health, have chosen to adopt the principles of this convention.⁴³ In addition, during the 10th conference of parties of the Convention on Biological Diversity,⁴³ the Nagoya Protocol was adopted, which is a legally binding protocol for access to genetic resources and benefit sharing.⁴⁴ Again, since the United States is not a party to the Convention, this creates a gray area. However, this is another example where journal editors have a great degree of influence, as some journals (for example, *Planta Medica*) now require verification that study materials were acquired in accordance with the Nagoya Protocol, as it is interpreted by the country from where the sample was collected. This latter point is important, as individual countries have interpreted how to handle the Nagoya Protocol differently. Ultimately, collecting a plant specimen is not a trivial matter, especially when done outside the United States, and we recommend working with local authorities and seeking guidance and consultation from individuals with relevant experience before doing so.

6. Authentication

Once a botanical sample (or series of samples) has been collected for analysis, an important step is “authentication,” i.e. the definitive determination of the species represented by the sample.^{13, 45, 46} Botanical samples can be authenticated by multiple means, including examination of voucher samples, DNA barcoding, and both untargeted and targeted characterization of chemical constituents. Whenever possible, more than one technique should be applied to confirm the identity of a botanical sample. However, the type of processing that a sample has been subjected to will determine which techniques are effective for authentication, as described in the following section.

6.1 Authentication by microscopic and morphological characteristics

Traditional methods used to authenticate botanical samples rely on examining the morphological and microscopic characteristics of representative voucher specimens. For example, macroscopic observations surrounding the arrangement (e.g., alternate, opposite, whorled) and shape (e.g., elliptic, ovate, cordate) of the leaves, the shape of inflorescence (e.g., cruciform, head, spurred) as well as its type (e.g., raceme, umbel, spike), the classification of the root system (e.g., taproot, fibrous root) can provide insight into the genus and species of the plant material. Similarly, microscopic examination of the histological characteristics of plant parts (e.g., stems, roots and rhizomes, bark, leaves, flowers, seeds, wood) can also aid in positive identification. These techniques, whether independently or in combination, have formed a baseline of approaches for botanical natural product quality control, and they continue to serve as essential methods used in pharmacopeia worldwide.⁴⁷ However, accurate microscopic and morphological analysis requires a trained expert, and the robustness of these procedures is also reliant on the presence (and expert knowledge) of diagnostic characteristics that enable one natural product to be separated from another. Challenges arise for closely related species and plant parts that share similar morphological structures yet could have differing biological activities.

Studies conducted using natural product extracts, botanical supplements, and commercial material are subject to several obstacles to authentication. When using commercial suppliers as material sources, voucher specimens are not always readily available. Identification by morphology or microscopy is not suitable for extracted natural product materials, and it may be difficult to determine the identity to the species level. Thus, to address the need for additional, more robust methods, researchers developed complementary techniques of differentiating botanical natural products.

6.2 Authentication by DNA Barcoding

DNA barcoding is a genetic technique used to identify species using specific differences in small regions of DNA.⁴⁸ Barcoding is an effective tool in authentication of botanical natural products that are comprised of fresh, dried, or powdered material, where intact DNA is still present.^{46, 49, 50} However, the DNA barcoding approach is more difficult when applied to botanical extracts. The manufacturing process for botanical natural products may involve heat treatment, oxidation, extraction, distillation, filtration, pressing, encapsulation, spray drying and/or UV light exposure, and these often lead to removal or degradation of DNA. DNA barcoding is not feasible for processed botanical products where the DNA is either not present or potentially highly degraded, or where there are two or more species present (either mixtures of active botanicals or the use of botanical “filler” such as rice).⁵⁰ Unfortunately, DNA barcoding is occasionally used to characterize processed botanical natural products without recognition of these limitations. For example, in 2015, the New York State Attorney General investigated potential adulteration of commercial products. This investigation employed DNA barcoding methodology to suggest that only five of 24 samples actually contained the purported botanical natural product,⁵¹ and as a result many of the products were subjected to a forced recall. However, the methods and results of the study were suspect. Among other concerns, there were reports of cross-contamination of the samples, and the investigation focused on botanical extracts, not whole plant material, such that intact DNA may not have been present even if the correct botanical natural products were used.⁵² This controversy may have been avoided with the use of complementary analytical approaches suitable for characterizing extracts (such as described in 6.3 and 6.4).

Although beyond the scope of this review, there are other factors that make DNA barcoding of plants challenging, such as hybridization and polyploidy, slow mutation rates compared with speciation rates, as well as lineages that show rapid and recent divergence rates. In some species of plants, these factors can influence the discriminatory power of standard plant barcoding markers.^{53–58} Despite these limitations, the use of two core plant DNA barcode regions *rbcL* (ribulose-bisphosphate carboxylase) and *matK* (maturase K) as well as two supplementary regions, *trnH-psbA* (chloroplast intergeneric spacers) and internal transcribed spacer (ITS) have been used for species discrimination in identifying plants via barcoding.⁵⁸ Given all these complexities and controversies, it is recommended that DNA barcoding of botanical materials be conducted in consultation with experts.

6.3 Authentication by targeted analysis of “marker compounds”

Information about the identities of known small molecule constituents (metabolites) present in a botanical extract (known as “chemotaxonomy”) can be extremely useful in the

authentication process. For most botanicals sold commercially as dietary supplements, there is a broad literature base that describes known metabolites and quantities in which they are likely to be present. When these constituents are unique to the botanical of interest, they are often referred to as “marker compounds.” Information about the chemical composition of a given plant is available in the primary peer reviewed literature, and in the form of “monographs” that summarize the chemical content of a particular botanical. Monographs are published by several sources, including the German Commission E,⁵⁹ US Pharmacopeia,⁶⁰ and *Tyler’s Herbs of Choice*.⁶¹ To make use of these monographs, the sample under question is analyzed by a method appropriate for detecting and quantifying the key marker compounds (see Section 9, *Identification, Structure Elucidation, and Quantitative Analysis of Bioactive Compounds in the Selected Study Material*), and their identities and concentrations are compared with literature values. The most common analytical techniques for analyzing botanical constituents are ultraviolet/visible (UV/VIS) spectrophotometry, mass spectrometry (MS), and Nuclear Magnetic Resonance (NMR) spectroscopy. Because of the complicated nature of botanical samples, it is also typical to perform chromatographic separation prior to analysis by spectrometric approaches. These chromatographic separations are most commonly carried out using gas chromatography (GC), high-performance thin-layer chromatography (HPTLC), high-performance liquid chromatography (HPLC), or a modern advancement on HPLC, termed ultrahigh performance liquid chromatography (UHPLC) prior to analysis by spectrometric approaches. The relative merits of different separation and analysis approaches for botanical natural products have been previously described.^{62–65} It is advised to use multiple approaches to characterize a single sample such that orthogonal data can be collected and compared.

6.4 Authentication by “chemical fingerprints”

In addition to specifically analyzing individual marker compounds, researchers seeking to authenticate botanical natural products often collect a “chemical fingerprint” that represents a more comprehensive profile of metabolites produced by the plant. This chemical fingerprint can be compared between two or more samples to assess similarity or difference (Figure 2).⁶⁶ The most common techniques used to measure chemical fingerprints are gas or liquid chromatography, coupled to either spectrophotometric or mass spectrometric detection.^{62, 65} The resulting data can be viewed as a chromatogram, which displays detector response as a function of time (Figure 2), such that each “peak” in the chromatogram corresponds to a chemical constituent (or multiple constituents) from the sample. Quantitative NMR (qNMR) is another technique which provides chemical fingerprinting capable of distinguishing species and authenticating botanical origins.⁶⁷

The relative intensities, retention times, and spectroscopic data obtained by chromatographic analysis can be used to putatively identify a sample. Small molecules are more stable than DNA, and they typically survive the processing involved with preparing botanical natural products. Thus, it is possible to authenticate many samples based on their metabolite profiles that may not be characterizable by DNA barcoding. However, it is important to note that the “fingerprint” of which metabolites are present differs depending on the extraction procedure as well as analytical technique used to collect the data and the parameters of the specific analysis.⁶⁸ Thus, it is necessary to include authenticated reference material for comparison

in the same analysis with the sample under question and processed by the same extraction procedure and solvent system (see *Purchasing Authenticated Reference Materials, Section 5.5*) (Figure 2).

6.5 Authentication by untargeted metabolomics

Often, it is of interest to compare more than two botanical natural product samples to each other and assess their relative differences and similarities. The technique, currently referred to as “untargeted metabolomics,” can be very useful for this purpose. The “metabolome” is defined as the complete set of small molecules produced by a biological sample. Metabolomics techniques rely on measuring as many of these small molecules as possible (although it is never possible to measure the entire metabolome, due to limitations in analytical methods). Metabolomics has been employed to characterize the relationships between the metabolome of a given natural product and corresponding genotype, origin, quality, or other biotic or abiotic attributes.^{28, 69–71} Two main analytical techniques are currently employed for metabolomic studies – MS and NMR. The relative advantages of these approaches have been discussed elsewhere.^{28, 72–74} The ability of untargeted metabolomics analyses to simultaneously profile a large fraction of the phytochemical make-up of the botanical, and offer detailed characterization of the samples in question has led metabolomics to be a very useful tool in the authentication of botanical natural products.

The process of data collection for a metabolomics analysis is essentially the same as that described for chemical fingerprinting, except that it is typical to analyze scores to hundreds of samples in a single analysis. This need to collect large datasets requires careful attention to experimental design, because run-times are often long, and drift in instrument response can cause artifacts in the data. Quality control checks throughout the run, as well as attention to sample order, blanks, and replication, are critically important.^{75, 76}

6.5.1 Metabolomics data analysis.—The major challenge in metabolomics studies is not in data collection, but in interpreting the very large datasets that are generated. The process known as “chemometrics” refers to the statistical approaches used to analyze metabolomics datasets. Various approaches are used to simplify and compare the data (reviewed in Kellogg et al. (2016)).⁷⁷ Most common among these is Principal Component Analysis (PCA), in which a dataset is transformed into a set of orthogonal variables (components) that account for the greatest degree of variability in the data. An example of how PCA can be employed to compare botanical samples is shown in Figure 3. Visual analysis of an untargeted metabolomic PCA scores plot of various commercial green tea preparations (Figure 3A) yielded distinct clusters of loose-leaf green tea samples (green symbols), green tea supplements (blue symbols), and a non-green tea sample (orange). Targeted metabolomics, using only “marker compounds” found in green teas, produced less distinct separations (Figure 3B).²⁸

7. Selecting a “Representative” Sample

Faced with a range of commercial samples purported to have been prepared from a given botanical natural product, how does the researcher select a representative product to advance to *in vitro*, non-clinical *in vivo* or clinical studies? Numerous factors influence the

composition of a botanical natural product. For example, with green tea, cultivar, geography, elevation, sun exposure, climate, harvest time, harvest location (where on the plant the leaves were removed), as well as processing, fermentation, and manufacturing methodologies all led to discernable differences in final phytochemical make-up of the product.^{78–83} These details are not readily available to the consumer (or post-harvest researcher); thus, choosing one product out of a multitude of available options constitutes a major challenge. Researchers apply a wide variety of criteria for choosing natural products for further study, including convenience,⁸⁴ consumer preference,²⁸ or a custom prepared extract formulation contracted by the research team.⁸⁵ As a result, the natural product interventions that are evaluated may or may not have a resemblance to the products that are currently on the market or used by the greatest numbers of consumers. While it is typically not possible to select a single product that represents all of the products used by consumers, it is worthwhile to capture a “snapshot” of the variability among the pool of products available, such that a desired product can be selected rationally. To conduct such a comparison, it is first necessary to obtain multiple products that represent those commercially available. For example, when the NaPDI Center prepared to conduct studies with green tea (*Camellia sinensis*) and goldenseal (*Hydrastis canadensis*), we purchased 34 and 35 commercially available samples of each botanical, respectively. We obtained fully characterized green tea samples from NIST, and several botanical reference standards of goldenseal from ChromaDex and laboratory collections.^{28, 29} These products were then all extracted in the same fashion, and their metabolite profiles were compared to those of the reference materials to select a representative product for *in vitro* and clinical studies. Here we describe our selection process in more detail, using green tea as an example. A similar approach has been employed by other laboratories for other studies of botanical natural products.^{30, 86, 87}

7.1 Selecting a formulation

Botanical natural products are available to the public in a variety of formulations. Commercially available formulations include tinctures (ethanolic or glycerol extractions of raw plant material), tea bags for making hot water extracts, raw plant material for human consumption, capsules, and tablets. A different constituent profile may be obtained by extracting in alcohol versus hot water and by using different plant material (for example, roots versus leaves) as starting material. Not all capsules are prepared in the same fashion. Some are made of raw plant material encapsulated in some coating (i.e. gelatin), some have a liquid extract directly encapsulated, whereas others are prepared by spray drying an extract on some solid support, such as cellulose. The formulation used can alter the chemical composition of the product. Spray-dried extract capsules usually are more concentrated than capsules prepared from raw plant material and may also be chemically simpler or modified by the extraction process. Because of these differences, we recommend that multiple representative formulations and, if applicable, multiple plant parts, of a given natural product be subjected to chemical evaluation. For example, for studies our group conducted with green tea, both spray-dried and raw plant material supplements and tea formulations in the 34 products were included in the initial evaluation. Metabolomics analysis demonstrated differences in profiles of the loose-leaf tea versus the encapsulated green tea supplements. The PCA loadings plot revealed several metabolites present in higher concentrations in the

green tea supplements than the leaf teas: myricetin, kaempferol, and quercetin aglycones, as well as theaflavin 3-O-(3-O-methyl) gallate and the dimer epicatechin (4 β →8)-epigallocatechin-3-O-gallate (Figure 4).²⁸

Another consideration when selecting natural product formulation is that the dissolution and absorption of the bioactive constituents, and by extension human systemic exposure, can differ depending on formulation.^{88, 89} These factors should be considered when designing *in vitro*, non-clinical *in vivo*, or clinical studies involving a botanical natural product.

Selection of product formulation should be guided by consumer usage. For example, we elected to conduct a clinical study with green tea as a beverage (hot tea) rather than a supplement based on popular use.¹⁶ Similarly, a goldenseal capsule containing dried botanical material was selected because this formulation is most commonly used as a botanical dietary supplement.

7.2 Standardized products

Some commercially available botanical natural products are sold with no information about chemical makeup, while others are “standardized” to (reportedly) contain a specific amount of a known constituent or constituents.²⁷ An advantage of standardization is that researchers, health care practitioners, and consumers are provided information that may be useful for selecting and adjusting dosage.²⁶ The standardization process may also provide consistency across products, with the caveat that methods used for standardization must be reproducible across multiple labs.^{90, 91} Standardization can involve multiple analytical methods, including qNMR,⁹² UV-VIS absorbance patterns,^{90, 93} and MS profiling.⁹⁴ Even for standardized products, label claims about content and identity of a natural product should be confirmed in-house or by a contract laboratory prior to conducting *in vitro*, non-clinical *in vivo*, and clinical studies.

A disadvantage of standardization is that the focus shifts from the quality of the entire mixture to a single or limited set of constituents. Standardized extracts may be chemically simpler than non-standardized counterparts, because the process of enhancing the concentration of constituents may involve extraction procedures that reduce extract complexity. Reduced extract complexity can be a disadvantage when the true “bioactive” constituent is not known, such that standardizing to one constituent may result in a lower dosage (or complete lack) of a constituent responsible for the presumed biological effect. How does a researcher decide whether to use a standardized product for an *in vitro*, non-clinical *in vivo*, or clinical study? One consideration is whether the most popular product(s) used by consumers are standardized. It is also advisable to compare metabolite profiles of standardized products to those of unstandardized counterparts to determine if components contributing to biological activity might have been lost during the standardization process (see Section 9.1).

7.3 Selection of extraction solvent

Once a series of representative products and authenticated standards has been selected, the first step in chemical analysis is to prepare extracts from these products. Typical methods of extracting botanical natural products for human consumption include the use of aqueous

ethanol or aqueous glycerin (to prepared tinctures), or hot water (to prepare a tea). Applying these same extraction procedures when preparing samples for chemical analysis may seem prudent. However, water, ethanol and glycerin extracts are difficult to manipulate in a laboratory setting. Drying botanical natural product extracts (removing solvents) for long term storage is ideal because extracts are most stable and least likely to degrade when solvent is not present.^{95, 96} Removing ethanol is a challenging process, removing water is more so, and removing glycerin is most arduous. Additionally, water extracts are unstable and contain polar constituents that confound chemical analysis and result in sticky, tar-like samples.

Unlike the aforementioned solvents, methanol is an outstanding solvent for preparing botanical natural product extracts prior to analysis. Methanol has an intermediate polarity that effectively solubilizes a wide range of structurally diverse small molecules, is inexpensive, and can be rapidly and efficiently removed under vacuum or a nitrogen stream. As such, many natural product extracts are prepared in methanol or similar organic solvents.^{77, 97, 98} A disadvantage of these solvents is that they are unsafe for human consumption, raising concern that the resulting extracts are not representative of consumer use. While characterizing various green tea products prior to clinical evaluation, we addressed this concern by comparing hot water extraction (an obvious choice for tea) to methanol extraction. Methanol and hot water extracts contained similar quantities of catechins²⁸ and showed qualitatively similar biological activity *in vitro*. Based on these results, methanol extraction was selected for metabolomics analysis of green tea products. Because methanol enabled rapid and reproducible extraction of samples, the extraction process did not confound comparisons of the chemical composition of the products themselves. A follow up study showed that accelerated solvent extraction improved extraction efficiency of green tea components in methanol.⁹⁹ An alternative to methanol extraction includes “comprehensive” extraction techniques, with a gradient of solvents producing a variety of crude extracts with differing polarities.¹⁰⁰ Ultimately, the ideal extraction solvent will depend on the type of material being extracted, the methods used to prepare preparations for consumer use, and the study question. Extraction solvent choice should be carefully considered and justified based both on experimental data and practical considerations. As the extracts prepared for research purposes often are not intended for human use, practical considerations of component solvation, ease of evaporation, and other physiochemical properties of the solvent may outweigh concerns about whether the extraction solvent is identical to that used to produce food grade products.

8. Quality Control and Comparison of Botanical Product Composition

Targeted chemical analysis of known markers and untargeted metabolomics to authenticate botanical natural product samples can be used to characterize and compare a series of candidate botanical natural products prior to selecting a product for *in vitro*, non-clinical *in vivo*, and/or clinical study. The advantage of untargeted chemical analysis for this purpose is that the analysis is relatively simple, can be highly sensitive and selective (leading to low limits of detection) and can be quantitative (when relevant reference standards are available; see Section 9.2). The disadvantage of targeted analysis for comparing samples is that the differences between samples may not be due to known “marker compounds.” For example,

in green tea studies conducted by the NaPDI Center, samples of tea leaves or powder were shown to differ chemically from green tea supplements when untargeted metabolomics was applied, as indicated by the distinct grouping in a PCA plot (Figure 3). However, when a PCA plot was generated using data from targeted analysis of 15 known green tea constituents, the differences among samples were less clear. As evidence of this lack of clear distinction, the points in the PCA plot in Figure 3B representing supplement and tea samples overlap, suggesting (incorrectly) that the chemical makeup of the different samples is indistinguishable.

8.1 Quantitative metrics for comparing similarity of natural product mixture

Untargeted metabolomics approaches comparing botanical natural products yield a tremendous quantity of data. Multivariate statistical techniques (e.g., PCA and partial least squares discriminate analysis (PLS-DA)) have become routine analytical methods used to analyze these data¹⁰¹ and serve as the backbone of metabolomic statistical analyses. PCA enables visualization of similarity and differences among samples by plotting them as points in 3D or 2D space (Figure 3).¹⁰² With appropriate validation, PCA serves as an effective approach to compare qualitative differences among botanical samples.^{103, 104} Other approaches, such as cosine calculations, measures based on vector distance, and measures based upon probabilistic difference, are used to generate *quantitative* metrics to compare chemometric or metabolomic spectral datasets.^{105–108} Recently, to facilitate selection of a green tea product for *in vitro* and clinical studies, we developed a new approach for quantitative comparison among samples in a metabolomics dataset, termed the composite score (CS).²⁸ The CS is based on the scores and loadings of the totality of a multivariate PCA model, rather than pairwise comparisons of components (as in traditional PCA scores plots). The scores and loadings are integrated into a single correlation matrix, which can be used to compare similarity between multiple samples (Figure 5). The CS ranges from –1.00 to 1.00. Values close to 1.00 indicate similarity among two samples (Figure 5b), whereas values further from 1.00 indicate more dissimilar samples (Figure 5c). The CS approach was used to evaluate which commercial sample was most similar to a given standard reference material.^{28, 29, 99} The PCA plot (Figure 3A) from the green tea study yielded distinct clusterings of green tea supplements versus loose leaf green tea products (teas and powders). However, resolution was lacking to discern which specific tea was most similar to the chosen standard, NIST T26; indeed, a different answer regarding which teas are most similar would be obtained depending on which two principal components were plotted against each other (PC1 versus PC2, PC2 versus PC3, etc.). CS analysis (Figure 5) was used to evaluate which products were most similar to the reference, as well as to select a commercial product (T21) for *in vitro* and clinical studies.²⁸

8.2 Adulteration of botanical natural products

Several reports have described suppliers of natural products selling products that are either unintentionally misidentified or intentionally adulterated (contaminated with botanical or other material different from that listed on the label).^{64, 87, 109} Adulteration can occur in a variety of ways, including spiking synthetic compounds into the plant material or adding or substituting a different (less expensive) species, representing cost-savings to a dishonest supplier.¹¹⁰ When studies are unknowingly conducted with adulterated or misidentified plant

material, results may be irreproducible in future studies or irrelevant to the botanical supposedly under evaluation. Thus, testing botanical natural products for adulteration is critical.

Several research groups have used targeted^{64, 87, 90, 111–113} or untargeted methods^{114–116} to identify adulterants in botanical natural products. For example, we studied 35 commercial products listed as “goldenseal”. PCA of untargeted metabolomics revealed a distinct grouping of goldenseal material, including both root/rhizome- and leaf-based samples (Figure 6), yet three samples (GS-07, GS-20, and GS-33) were distinctly separated from the other clusters.²⁹ Separation of these samples from the others in the PCA scores plot raised suspicion that they were misidentified or adulterated with other botanical products.

To tentatively identify the species present in the outlier materials, reference material for three non-goldenseal species that serve as common adulterants were included in the dataset: *Coptis chinensis* rhizome (GS-39) and root (GS-40), *Mahonia aquifolium* leaf (GS-41) and root (GS-42), and *Berberis vulgaris* root (GS-43).^{110, 117, 118} The resulting PCA scores plot (Figure 6B) showed close clustering of GS-33 with *C. chinensis*, GS-20 with *M. aquifolium*, and GS-07 in the middle, possibly as a blend of multiple species. Examining specific marker compounds for each of the four species (berberine, m/z 336.1229 [M]⁺; hydrastine, m/z 384.1440 [M+H]⁺; canadine, m/z 340.1545 [M+H]⁺; palmatine, m/z 352.1543 [M]⁺; coptisine m/z 320.0917 [M]⁺; and dihydrocoptisine, m/z 322.1074 [M]⁺) revealed specific concentration patterns (Figure 7) for the putative adulterants that supported the hypotheses generated by the metabolomic analysis.²⁹ This example illustrates the usefulness of both untargeted and targeted methods for identifying adulterated samples. PCA plots generated from untargeted metabolomics can be used to give a broad sweep analysis of the data and identify potential outliers, and follow up targeted analyses can help verify that adulteration has occurred and, if sufficient literature precedent exists, determine the identity of adulterants.²⁹

8.3 Consideration of contamination

A potential risk with botanical natural products is contamination by heavy metals,¹¹⁹ pesticides and herbicides,¹²⁰ residual solvents,¹²¹ fungal spores and mycotoxins,¹²² or microbes.¹²³ These concerns have received considerable attention, and recent reviews of the risks, policies, and analyses of contaminants in botanical natural products are available.^{27, 31, 112, 124} In 2010, the FDA instituted the requirement that botanical natural products must be prepared using good manufacturing practices (GMP) that provide guidance for safety, consistency, and reproducibility of these products.^{21, 125} GMP producers certify that their products are consistently produced and controlled according to quality standards. Whether the product is custom formulated²¹ or obtained from commercial sources, testing is required for contaminants, including “heavy metals, microbial limits, residual pesticides, adventitious toxins (e.g., aflatoxins), [and] endogenous toxins (e.g., pyrrolizidine alkaloids)” according to 21 CFR 111.70(b)(3). Product batches that do not meet the limits of contamination must be rejected.^{125, 126} FDA guidance also specifies that at least one test must be used to verify botanical ingredients using a validated method.¹¹² However, because the manufacturer is responsible for developing GMP procedures for each botanical natural

product, there is no standardized procedure across the industry. In addition, compliance enforcement is limited and does not require evaluation or verification of authenticity or screening for contaminants by a third party.¹²⁴ Given these limitations, testing botanical natural product study materials for contamination by pesticides, heavy metals, microorganisms, and residual solvents is advised. Such tests can be accomplished in house or through a contract laboratory. Screening is done in a targeted fashion, i.e. the analyst specifies a list of common or likely contaminants *a priori* and analyzes the sample to determine levels of these contaminants. The challenge with this approach is selecting which contaminants to test for among the almost unlimited number of potentials. Selection of contaminants can be guided by consultation with the literature or based upon knowledge regarding the environment in which the sample was grown and the processing it was subjected to post-harvest. Nonetheless, it is easy to imagine that some contaminants may be unexpected and, therefore, overlooked. Untargeted metabolomics may serve as a useful tool for identifying unexpected contaminants.^{29, 111, 115, 127} However, it should be noted that untargeted methods are likely to have higher limits of detection than targeted methods; thus, the former may fail to detect low-level contaminants. Furthermore, some contaminants are detectable only with specialized analytical methods. For example, GC-MS is appropriate for the analysis of certain pesticides and residual solvents, but does not respond to many non-volatile contaminants, and a specialized system such as an inductively coupled plasma mass spectrometer (ICP-MS) is needed to detect heavy metal contaminants. With these challenges in mind, the analyst should carefully consider what the most likely contaminants are in a given sample and apply multiple analytical techniques and a combination of both targeted and untargeted methods to maximize the likelihood of detecting as many contaminants as possible.

9. Identification, Structure Elucidation, and Quantitative Analysis of Bioactive Compounds in the Selected Study Material

Once a natural product has been selected for *in vitro*, non-clinical *in vivo*, or clinical studies, it should be thoroughly characterized, and concentrations and identities of key bioactive constituents (i.e., those responsible for the reported biological activity) should be determined. Such knowledge can be used to adjust dosage and conduct quality control. The challenge is that identities of bioactive constituents relevant to the activity being studied are often not known and are likely to differ depending on the biological activity evaluated.^{62, 128, 129} Assigning bioactive constituents is not trivial, as a given botanical extract may contain hundreds or thousands of constituents.¹³⁰ The collective activity of the extract may be due to the combined action of these constituents acting in an additive, synergistic, or antagonistic manner.^{131–135} Despite these challenges, focusing solely on one or two constituents in an extract is not prudent (discussed in Sections 6.3–6.5). For some of the well characterized botanicals, nearly complete characterization of major constituents can sometimes be possible. For example, in our work with green tea, 15 commercially available pure standards were obtained that represented major and minor peaks in a LC-MS chromatogram (Figure 2). The sections below highlight key points to consider when conducting qualitative and quantitative analysis of botanical extracts and assigning biologically active components. Based on practical experience, a summary of common

pitfalls during quantitative and qualitative analysis of botanical extracts is provided (Table 2).

9.1 Determining which constituents of a botanical natural product are biologically active

The gold standard approach for identifying bioactive mixture components is bioassay-guided fractionation, in which the extract is subjected to successive rounds of fractionation and purification (see *Section 9.2*), with each fraction prioritized for the next stage based on biological assay data. Such an approach has been used in countless studies, with perhaps the most well-known being discovery of the cancer chemotherapeutic agent taxol (aka paclitaxel) from the Pacific yew tree.^{136, 137} An alternate approach to bioassay-guided fractionation when targets are known and available in sufficient quantity is pulsed ultrafiltration mass spectrometry,^{138, 139} the advantages and disadvantages of which are detailed elsewhere.¹⁴⁰ In general, we find that bioassay-guided fractionation is the most effective strategy for identifying bioactive mixture components when the specific targets and mechanism of action of a given botanical natural product are not known, and when those activities can be modeled effectively *in vitro*. Bioassay-guided fractionation can be enhanced using statistical approaches to integrate biological assay data with chemical metabolite (metabolomics) profiles, a process often referred to as ‘biochemometrics’.^{77, 141}

9.1.1 Advantages and limitations of *in vitro* assays.—A critical element of bioassay-guided fractionation is biological evaluation. Biological activity of a given botanical extract constituent is typically assessed using an *in vitro* assay.^{21, 142, 143} Isolated compounds with promising activity may eventually be tested in non-clinical (animal) models and potentially in clinical studies. However, it is often infeasible to conduct such *in vivo* studies with the speed and scale necessary to facilitate bioassay-guided fractionation; thus, fractionation and isolation efforts are typically guided by *in vitro* evaluation.

In vitro assays are used to assess the potential for natural product constituents as lead compounds for a diverse array of illnesses, including cancer, bacterial infections, and diabetes.^{144–146} Cell-based assays can be used to guide future *in vivo* studies to test for safety and efficacy of natural products,¹⁴⁷ to give insight regarding potential interactions between natural products and conventional drugs,^{148–150} and to predict how food matrices could affect the pharmacokinetics of natural product constituents.¹⁵¹ For studies evaluating interactions between natural products and conventional drugs, mechanistic insight gained from *in vitro* studies is imperative to guide further evaluation via mathematical (e.g., physiologically-based pharmacokinetic) modeling or clinical studies.⁶ However, *in vitro* assay data may not always translate to *in vivo* activity, partially due to the inherent complexity of biological systems, and the effects of other biological factors (e.g., microbiota) that are not accounted for in most *in vitro* models.¹⁵² Additionally, many botanical natural products are used for purported health benefits that are non-specific and difficult to model *in vitro*, such as “improving immune health” or “enhancing adaptability to stress.” In light of these challenges, it is important to keep in mind that the extent to which any bioassay-guided fractionation experiment is useful is fundamentally limited by the availability of a relevant, robust and translatable biological assay.

9.1.2 Approaches for integrating *in vitro* data with chemical composition.—

Visual inspection of chromatographic data is often used to guide isolation efforts as part of bioassay-guided fractionation experiments. Limitations of this approach include a tendency to focus on constituents that are most abundant, most responsive to the detector, or most easily isolated. In addition, re-isolation of known active compounds has plagued traditional bioassay-guided fractionation experiments. This problem of re-isolation can be addressed using various “dereplication” approaches, which often rely on comparing experimentally measured spectroscopic data with databases for known compounds.^{158–161} When the active constituents of a botanical natural product are not known, several methods are available to correlate metabolite profiles with bioactivity data and guide isolation of constituents most likely to be active. Use of statistical approaches to integrate biological assay data with measurements of chemical composition is termed “biochemometrics”.¹⁴¹ Partial least-squares (PLS) modeling and associated variants (Partial least-squares discriminate analysis (PLS-DA) and orthogonal partial least-squares discriminant analysis (OPLS-DA)) have become major statistical methods for biochemometric analysis.^{162–164} PLS modeling has been adapted to identify individual (or several) metabolites that are predicted to be responsible for the bioactivity of a complex natural product mixture.^{77, 165, 166} For the NaPDI Center study of green tea, a biochemometric approach was used to predict which catechins were responsible for the *in vitro* inhibition of intestinal UDP-glucuronosyltransferases (UGTs). The selectivity ratio^{167–169} was used as a metric to demonstrate the extent to which a given mixture constituent was associated with biological activity.¹⁷⁰ Five catechins were identified as major constituents in a selected bioactive subfraction. Selectivity ratio analysis predicted (–)-epicatechin gallate (ECG) to be a major contributor to inhibition of UGTs (Figure 8). This prediction was confirmed with follow up *in vitro* studies using commercially available catechins and the clinically relevant intestinal UGT substrate raloxifene.¹⁷⁰

9.2 Obtaining “pure” standards

Efforts to identify and quantify constituents of botanical extracts often rely on the availability of purified standards. Several commercial suppliers provide pure samples of individual botanical compounds and include Sigma-Aldrich (St. Louis, MO, USA), ChromaDex (Irvine, CA, USA), NIST (Gaithersburg, MD, USA), and the United States Pharmacopeia (Rockville, MD, USA). The modest cost of these standards typically far outweighs the time and effort involved in obtaining standards via independent isolation methods. A standard comes with a “certificate of analysis” that reports identity and purity of the sample. Additional characterization (by LC-MS and/or NMR) is recommended to confirm accuracy of information included in the certificate of analysis.¹⁵⁵

When standards of the compounds of interest are not available commercially, isolating these compounds from the complex starting material is necessitated.¹⁷¹ A number of separation techniques can be used for this purpose, including solid phase extraction,¹⁷² counter current chromatography,^{173, 174} HPLC,¹⁷⁵ or thin layer chromatography (TLC).¹⁷⁶ The most common strategy for isolation of constituents from natural products is a combination of liquid-liquid partitioning and column chromatography approaches. A typical fractionation scheme involves several steps, as illustrated in Figure 9. The sample is extracted in

methanol, partitioned between chloroform and water or ethyl acetate/water, and subsequently subjected to an additional partitioning step between methanol:acetonitrile and hexanes.¹⁷⁷ These steps could be modified using acid/base conditions if the target metabolites are alkaloids. The resulting residue is dried under nitrogen and subjected to several stages of normal phase flash chromatography. Finally, pure compounds are isolated using reversed phase preparative HPLC. The solvents and stationary phases used for the separation can be varied to optimize separation of the compound of interest. Isolation of milligrams of pure material often takes weeks or months (or may be unsuccessful) depending on the difficulty of the separation and the skill of the analyst.¹⁷⁸ As such, isolation of all constituents in a given botanical extract is typically unfeasible, and it is helpful to guide isolation efforts with biological assay data, as described in *Section 9.1*.

9.3 Structure elucidation of constituents of botanical natural products

Identification of metabolites in botanical extracts has been reviewed previously,^{179, 180} and structure elucidation of natural product constituents is detailed in several textbooks.^{181, 182} Two approaches are commonly used: isolation and subsequent structure elucidation via NMR or characterization of mixture constituents without isolation via LC-MS. Conclusive assignments of structure are made via NMR and it is the preferred technique when sufficient material (typically 0.5 to 1 mg) has been isolated.¹⁸³ Structure elucidation via NMR is conducted in combination with other spectroscopic approaches, notably high resolution MS, which enables determination of molecular formula.¹⁸⁰ Accepted guidelines from peer-reviewed journals (e.g., *Journal of Natural Products*) recommend that assigned ¹H NMR and ¹³C NMR spectra be presented for review, as well as molecular formulae calculated by relying on accurate mass measurements with MS, UV absorptivity, and any additional stereoscopic information as needed.¹⁰ In the very near future, we anticipate that many journals will require the deposition of raw NMR data associated with the structure elucidation of organic molecules, akin to the way DNA data are deposited in GenBank; a recent review in *Natural Products Reports*, co-authored by over 70 scientists, advocates for this practice.¹⁸⁴

For comprehensive characterization of mixtures without isolation, LC-MS (or GC-MS for volatile samples), is the most effective approach. Challenges, pitfalls, and strategies to identify botanical natural product metabolites by mass spectrometry have been detailed previously.⁶² If high resolution mass spectrometry is used, accurate measurements of mass are possible to tentatively identify extract components based on calculations of molecular formula. Mass spectrometers with the ability to measure mass with sufficient accuracy include those with Orbitrap, Fourier transform inductively coupled resonance (FT-ICR) or quadrupole time-of-flight (Q-ToF) mass analyzers. The confidence with which structure can be assigned based on mass spectrometric data is increased when retention time and fragmentation patterns for the component of the unknown sample can be compared to those of pure, isolated standards.^{186–188} However, mass spectrometry is not a conclusive technique for assigning configuration of stereoisomers.

9.4 Quantitative analysis

The goal of quantitative analysis is to determine the absolute concentration of a known constituent (analyte) of a complex botanical extract. Effective quantitative analysis requires an isolated standard identical to the constituent to be quantified (see section 9.2, *Obtaining “Pure” Standards*). Dilutions of this standard are prepared and analyzed with the extract. A calibration curve of detector response (most commonly chromatographic peak area under the curve) vs. analyte concentration is plotted. The calibration curve is analyzed via regression over the linear range (range in which response increases linearly with increasing concentration). The response (peak area) for the analyte in the extract is substituted into the equation for the best fit line, and the concentration in the extract and the original plant sample is back calculated by adjusting for all necessary dilutions.¹⁸⁹ To improve the accuracy of quantitative analyses, methods using isotopically labeled standards (when available) can also be employed.^{190–192}

The most popular techniques for quantitative analysis of botanical natural products are GC or GC-MS (for volatile sample components) and LC-UV/VIS or LC-MS (for nonvolatile sample components). The relative merits of these techniques are detailed elsewhere.⁶² Briefly, LC-MS often (but not always) provides lower limits of detection compared to LC-UV/VIS. An additional advantage of LC-MS is that it can be used to obtain orthogonal data about analyte mass and fragmentation pattern that is helpful for deconvoluting chromatograms and confirming identification.¹⁵⁸ LC-UV/VIS, on the other hand, is typically less susceptible to matrix interference and for some applications demonstrates improved linearity, particularly at the high end of the calibration curve.⁹⁵ Instruments with UV/VIS detectors are also less expensive and available to a broader base of analysts. Having access to instruments equipped with both UV and MS detectors, such that identity of the quantified constituents can be confirmed using more than one data type, is ideal. When both UV and MS data are available, quantitative analysis can be accomplished with whichever technique provides the most linear data in the region of interest.

When conducting quantitative analysis, analyzing the extract (botanical natural product) at multiple dilutions is recommended. A typical analysis includes a series of 2-fold or 10-fold dilutions of the extract, with the highest concentration in the range of 1.0 mg/mL (expressed as dry weight of extract per volume of dilution solvent). It is often desirable to quantify more than one constituent, and the ideal dilution to achieve a concentration within the linear range of the calibration curve will vary from constituent to constituent. Thus, analyzing the extract at multiple dilutions is helpful. A series of 2-fold or 10-fold dilutions of the standard is also included in the same analysis with the extract to generate a calibration curve for quantitative analysis. The resulting calibration curve should span the maximum linear range, and several dilutions above and below this range should be included to enable the full range to be determined. AOAC INTERNATIONAL (The Association of Analytical Communities) recommends including 6–8 data points within the linear range.¹⁹³

One question that often arises when conducting quantitative analysis is the issue of whether (or how) a method has been validated. Validation is a process by which a given method is optimized for parameters related to accuracy, precision, limits of detection and quantitation, linear range and robustness.⁹⁰ Other elements of method evaluation in the context of

botanical natural product research include measurements of extraction efficiency and checking for potential matrix interference. Methods can be validated both within a single laboratory (intralaboratory validation)^{95, 194} or across multiple laboratories (interlaboratory validation).^{67, 90} The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) provides general guidelines for validation of analytical methods,¹⁹⁵ and more specific guidelines and definitions related to validation of constituents of botanicals have been drafted by a subcommittee of AOAC INTERNATIONAL.^{13, 193} When selecting a method for analysis of botanicals, we recommend reviewing multiple published methods (if available), with particular attention to those that have been validated. However, for practical reasons, including differences in available expertise, samples and instrumentation, it may be difficult to precisely replicate even a validated method for a new study. This is particularly true if the method has not already been subjected to interlaboratory validation. For this reason, it is typically useful to validate the selected method as part of the given project. One excellent strategy for checking accuracy is to obtain a botanical reference sample for which concentrations of the bioactives of interest are known, then compare the results obtained with the selected method to those reported on the certificate of analysis for the reference material. Such validation can be performed using reference standards from NIST (if available) for which marker compound concentrations have been reported. To avoid misleading results due to matrix interference, it is important to check the accuracy of the method *in the botanical matrix of interest*. For example, our methods for quantifying catechins as part of the NaPDI Center's green tea study were tested by analyzing a NIST green tea standard reference material, not isolated catechins.²⁸ If no commercial reference material of the botanical is available, alternate techniques for quantifying metabolites are available. One technique is to conduct a "spike recovery" study, in which a standard is spiked into the botanical matrix at known concentrations, and the concentration is back-calculated to evaluate accuracy.^{95, 196–198} Quantitative NMR and pre- or post-column derivatization¹⁹⁹ also can be used to quantify metabolites without known standards.²⁰⁰

We end this section with a word of caution, which is that it is easy to become so focused on method validation that one is distracted from the overall research question (which typically relates to safety and/or efficacy of a botanical). It is often true that variability in biological response is so significant that it is unnecessary (and even a waste of time) to pin down concentrations in a botanical extract to the highest level achievable by a trained analytical chemist. The extent to which validation is necessary should be carefully considered in the context of the research question being evaluated.

10. Conclusions

The process of selecting botanical natural products for *in vitro*, non-clinical *in vivo*, and clinical studies is complex, and the ideal approach must be tailored on a case by case basis. Nonetheless, as a result of extensive experience, and much trial and error, we have identified a few common themes, pitfalls, and questions that almost always arise. In the preceding sections, we sought to address these by presenting examples from work conducted as a part of the NaPDI Center in the context of current literature. While constructing the recommendations provided in this report, we have considered the need to balance the ideal,

i.e. what would be done in a situation where time, person power, and funding were unlimited, with the practical, i.e. what is possible subject to realistic constraints. We provide Table 3 as a summary of many of the recommendations from the text, and to highlight a few additional points for consideration. It is our expectation that consideration of the topics addressed herein will enable the design of rigorous and reproducible studies with complex botanical natural products, and prevent the wasted time, funds, and effort that goes into studies where product selection has not been carefully considered.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This project was supported by the National Institutes of Health via the National Center for Complementary and Integrative Health, specifically the Center of Excellence for Natural Product Drug Interaction Research (NaPDI Center, U54 AT008909). JSM was supported in part by the National Cancer Institute [R01 CA182963]. MFP was supported in part by the National Institute of General Medical Sciences [R01 GM077482]. We thank the following individuals for insightful comments and helpful discussions: Dr. D. Craig Hopp, NCCIH Program Officer; Dr. Amy Roe, Principal Toxicologist at Procter and Gamble; Dr. Huzefa Raja, research scientist at the University of North Carolina at Greensboro; Tyler Graf, research scientist at the University of North Carolina at Greensboro; and Dr. Danny D. Shen, Professor of Pharmaceutics at the University of Washington. We would like to thank Matt Bryant and Cindy Graf for their expertise and assistance in designing the workflow figures presented here. The high-resolution mass spectrometry data were collected at the Triad Mass Spectrometry Facility at the University of North Carolina at Greensboro (chem.uncg.edu/triadmslab/).

11. References

1. Wu CH, Wang CC, Tsai MT, Huang WT and Kennedy J, Evidence-based complementary and alternative medicine : eCAM, 2014, 2014, 872320. [PubMed: 25574184]
2. Job KM, Kiang TKL, Constance JE, Sherwin CMT and Enioutina EY, Expert Review of Clinical Pharmacology, 2016, 9, 1597–1609. [PubMed: 27644147]
3. Gufford BT, Barr JT, González-Pérez V, Layton ME, White JR, Oberlies NH and Paine MF, CPT: Pharmacometrics & Systems Pharmacology, 2015, 4, 701–710. [PubMed: 26904384]
4. Smith T, Kawa K, Eckl V, Morton C and Stredney R, Herbalgram, 2017, 115, 56–65.
5. Klein MA, Nahin RL, Messina MJ, Rader JI, Thompson LU, Badger TM, Dwyer JT, Kim YS, Pontzer CH, Starke-Reed PE and Weaver CM, Journal of Nutrition, 2010, 140, 1192S–1204S. [PubMed: 20392880]
6. Paine MF, Shen DD and McCune JS, Drug Metabolism and Disposition, 2018, dmd. 117.079962.
7. Dwyer J, Coates P and Smith M, Nutrients, 2018, 10, 41.
8. Center for Drug Evaluation and Research, Botanical Drug Development: Guidance for Industry, U.S. Department of Health and Human Services, 2015.
9. National Center for Complimentary and Integrative Health, Journal, 2016.
10. Journal of Natural Products, Preparation and Submission of Manuscripts, http://pubs.acs.org/paragonplus/submission/jnprdf/jnprdf_authguide.pdf, (accessed June 8, 2017).
11. Phytochemistry Letters, Guide for Authors, <https://www.elsevier.com/journals/phytochemistry-letters/1874-3900/guide-for-authors>, (accessed June 8, 2017).
12. Planta Medica, Guidelines for Authors, https://www.thieme.de/statics/dokumente/thieme/final/de/dokumente/zw_thieme-en/PlantaMedica-Authors-December-2016.pdf, (accessed June 8, 2017).
13. AOAC International, Journal of AOAC International, 2012, 95, 268–272. [PubMed: 22468370]
14. Kroll DJ and Oberlies NH, Focus on Alternative and Complementary Therapies, 2003, 8, 302–306.
15. Office of Dietary Supplements, Dietary Supplement Label Database, <https://www.dsld.nlm.nih.gov/dsld/index.jsp>, (accessed September 16, 2018).

16. Centers for Disease Control and Prevention, Journal, 2017.
17. Consumer Reports, Consumer Reports, 2003, 68, 40–43. [PubMed: 12599344]
18. J. Health, Green Tea Supplement Reviews, <http://www.benefitsofgreentea.info/2007/green-tea-supplement-reviews.php>, (accessed November 11, 2015).
19. Smith T, Kawa K, Eckl V and Johnson J, Herbalgram, 2016, 111, 67–73.
20. Amazon.com, Amazon Best Sellers, <http://www.amazon.com/Best-Sellers-Grocery-Gourmet-Food-Green-Tea-Beverages/zgbs/grocery/16318471>, (accessed November 11, 2015).
21. van Breemen RB, Journal of Medicinal Chemistry, 2015, 58.
22. National Park Service, Journal, 2009, 2009.
23. Stearn WT, Biological Journal of the Linnean Society, 1971, 3, 225–233.
24. Greve M, Lykke AM, Fagg CW, Gereau RE, Lewis GP, Marchant R, Marshall AR, Ndayishimiye J, Bogaert J and Svenning J-C, South African Journal of Botany, 2016, 105, 317–323.
25. Phytochemistry, Author Information Pack, <https://www.elsevier.com/journals/phytochemistry/0031-9422?generatepdf=true>, (accessed August 31, 2017).
26. Piersen CE, Booth NL, Sun Y, Liang W, Burdette JE, Breemen R. B. v., Geller SE, Gu C, Banuvar S, Shulman LP, Bolton JL and Farnsworth NR, Current Medicinal Chemistry, 2004, 11, 1361–1374. [PubMed: 15180571]
27. van Breemen RB, Fong HHS and Farnsworth NR, Chemical Research in Toxicology, 2007, 20, 577–582. [PubMed: 17362032]
28. Kellogg JJ, Graf TN, Paine MF, McCune JS, Kvalheim OM, Oberlies NH and Cech NB, Journal of Natural Products, 2017, 80, 1457–1466. [PubMed: 28453261]
29. Wallace ED, Oberlies NH, Cech NB and Kellogg JJ, Food and Chemical Toxicology, 2018, 120, 439–447. [PubMed: 30031041]
30. Harnly JM, Luthria D and Chen P, Journal of AOAC International, 2012, 95, 1579–1587. [PubMed: 23451372]
31. Posadzki P, Watson L and Ernst E, European Journal of Clinical Pharmacology, 2013, 69, 295–307. [PubMed: 22843016]
32. Kroll DJ, Shaw HS, Mathews JM and Oberlies NH, 2005.
33. Ding S, Dudley E, Plummer S, Tang J, Newton RP and Brenton AG, Phytochemistry, 2008, 69, 1555–1564. [PubMed: 18342344]
34. Chen P, Ozcan M and Harnly J, Analytical and Bioanalytical Chemistry, 2007, 389, 251–261. [PubMed: 17632706]
35. Li W, Sun Y, Fitzloff JF and Van Breemen RB, Chemical research in toxicology, 2002, 15, 1174–1178. [PubMed: 12230411]
36. Abourashed EA and Khan IA, Journal of pharmaceutical sciences, 2001, 90, 817–822. [PubMed: 11458331]
37. Henning SM, Fajardo-Lira C, Lee HW, Youssefian AA, Go VLW and Heber D, Nutrition and Cancer, 2003, 45, 226–235. [PubMed: 12881018]
38. Friedman M, Levin CE, Choi SH, Kozukue E and Kozukue N, Journal of food science, 2006, 71, C328–C337.
39. Soejarto DD, Journal of Ethnopharmacology, 1996, 51, 1–15. [PubMed: 9213606]
40. Soejarto DD, Fong HHS, Tan GT, Zhang HJ, Ma CY, Franzblau SG, Gyllenhaal C, Riley MC, Kadushin MR, Pezzuto JM, Xuan LT, Hiep NT, Hung NV, Vu BM, Loc PK, Dac LX, Binh LT, Chien NQ, Hai NV, Bich TQ, Cuong NM, Southavong B, Sydara K, Bouamanivong S, Ly HM, Thuy TV, Rose WC and Dietzman GR, Journal of Ethnopharmacology, 2005, 100, 15–22. [PubMed: 15993554]
41. Secretariat of the Convention on Biological Diversity, Handbook of the Convention on Biological Diversity, Earthscan Publications Ltd., London 2001.
42. Kursar TA, Caballero-George CC, Capson TL, Cubilla-Rios L, Gerwick WH, Gupta MP, Ibañez A, Linington RG, McPhail KL, Ortega-Barría E, Romero LI, Solis PN and Coley PD, BioScience, 2006, 56, 1005–1012.
43. N. Protocol, 2011.

44. International Institute for Environment and Development, Nagoya Protocol on Access to Genetic Resources and Benefit-Sharing, <https://biocultural.iied.org/nagoya-protocol-access-genetic-resources-and-benefit-sharing>, (accessed June 1, 2018).
45. Leal MC, Hilario A, Munro MHG, Blunt JW and Calado R, *Natural Product Reports*, 2016, 33, 747–750. [PubMed: 26892141]
46. Raja HA, Baker TR, Little JG and Oberlies NH, *Food Chemistry*, 2017, 214, 383–392. [PubMed: 27507489]
47. Upton R, Graff A, Jolliffe G, Langer R and Williamson R, *American Herbal Pharmacopoeia: Botanical Pharmacognosy - Microscopic Characterization of Botanical Medicines*, CRC Press, Boca Raton, FL, 2016.
48. Ivanova NV, Kuzmina ML, Braukmann TWA, Borisenko AV and Zakharov EV, *PLOS ONE*, 2016, 11, e0156426. [PubMed: 27227830]
49. Little DP, *Genome*, 2014, 57, 513–516. [PubMed: 25495290]
50. Coutinho Moraes DF, Still DW, Lum MR and Hirsch AM, *Planta Med*, 2015, 81, 687–695. [PubMed: 25856442]
51. New York State Office of the Attorney General, A.G. Schneiderman asks major retailers to halt sales of certain herbal supplements as DNA tests fail to detect plant materials listed on majority of products tested, <https://ag.ny.gov/press-release/ag-schneiderman-asks-major-retailers-halt-sales-certain-herbal-supplements-dna-tests>).
52. Parveen I, Gafner S, Techen N, Murch SJ and Khan IA, *Planta Med*, 2016, 82, 1225–1235. [PubMed: 27392246]
53. Hollingsworth PM, Li D-Z, van der Bank M and Twyford AD, *Phil. Trans. R. Soc. B*, 2016, 371, 20150338. [PubMed: 27481790]
54. Hollingsworth PM, Graham SW and Little DP, *PloS one*, 2011, 6, e19254. [PubMed: 21637336]
55. Clement WL and Donoghue MJ, *BMC evolutionary biology*, 2012, 12, 73. [PubMed: 22646220]
56. Fazekas AJ, Kesanakurti PR, Burgess KS, Percy DM, Graham SW, Barrett SC, Newmaster SG, Hajibabaei M and Husband BC, *Molecular Ecology Resources*, 2009, 9, 130–139. [PubMed: 21564972]
57. Pei N, Erickson DL, Chen B, Ge X, Mi X, Swenson NG, Zhang J-L, Jones FA, Huang C-L and Ye W, *Scientific reports*, 2015, 5, 15127. [PubMed: 26456472]
58. Fazekas AJ, Kuzmina ML, Newmaster SG and Hollingsworth PM, *Methods in molecular biology (Clifton, N.J.)*, 2012, 858, 223–252.
59. Blumenthal M and Busse WR, *The complete German Commission E monographs*, American Botanical Council, Austin, Texas, 1998.
60. U.S. Pharmacopeia, *Herbal Medicines Compendium*, <http://hmc.usp.org/>, (accessed September 19, 2017, 2017).
61. Tyler VE, *Herbs of choice: the therapeutic use of phytomedicinals*, Pharmaceutical Products Press (imprint of Haworth Press, Inc.), 1994.
62. Cech NB, *LCGC North America*, 2013, 31, 938–947.
63. Simmler C, Napolitano JG, McAlpine JB, Chen S-N and Pauli GF, *Current Opinion in Biotechnology*, 2014, 25, 51–59. [PubMed: 24484881]
64. Vaclavik L, Krynitsky AJ and Rader JI, *Analytical and Bioanalytical Chemistry*, 2014, 406, 6767–6790. [PubMed: 25270866]
65. Zhu J, Fan X, Cheng Y, Agarwal R, Moore CMV, Chen ST and Tong W, *PLOS ONE*, 2014, 9, e87462. [PubMed: 24498109]
66. Navarrete A, Avula B, Choi Y-W and Khan IA, *Journal of AOAC International*, 2006, 89, 8–15. [PubMed: 16512222]
67. Markus MA, Ferrier J, Luchsinger SM, Yuk J, Cuerrier A, Balick MJ, Hicks JM, Killday KB, Kirby CW, Berrue F, Kerr RG, Knagge K, Gödecke T, Ramirez BE, Lankin DC, Pauli GF, Burton I, Karakach TK, Arnason JT and Colson KL, *Planta Med*, 2014, 80, 732–739. [PubMed: 24963620]
68. Harnly J, Chen P, Sun J, Huang H, Colson KL, Yuk J, McCoy J-AH, Reynaud DTH, Harrington PB and Fletcher EJ, *Planta Med*, 2016, 82, 250–262. [PubMed: 26692457]

69. Avula B, Wang Y-H, Isaac G, Yuk J, Wrona M, Yu K and Khan IA, *Planta Med*, 2017, DOI: 10.1055/s-0043-109239.
70. Kumar D, *Critical Reviews in Analytical Chemistry*, 2016, 46, 400–412. [PubMed: 26575437]
71. Yuk J, McIntyre KL, Fischer C, Hicks J, Colson KL, Lui E, Brown D and Arnason JT, *Analytical and Bioanalytical Chemistry*, 2013, 405, 4499–4509. [PubMed: 23250379]
72. Gika HG, Theodoridis GA, Plumb RS and Wilson ID, *Journal of Pharmaceutical and Biomedical Analysis*, 2014, 87, 12–25. [PubMed: 23916607]
73. Jorge TF, Rodrigues JA, Caldana C, Schmidt R, van Dongen JT, Thomas-Oates J and António C, *Mass Spectrometry Reviews*, 2016, 35, 620–649. [PubMed: 25589422]
74. Markley JL, Brüschweiler R, Edison AS, Eghbalian HR, Powers R, Raftery D and Wishart DS, *Current Opinion in Biotechnology*, 2017, 43, 34–40. [PubMed: 27580257]
75. Sangster T, Major H, Plumb R, Wilson AJ and Wilson ID, *Analyst*, 2006, 131, 1075–1078. [PubMed: 17003852]
76. Dunn WB, Wilson ID, Nicholls AW and Broadhurst D, *Bioanalysis*, 2012, 4, 2249–2264. [PubMed: 23046267]
77. Kellogg JJ, Todd DA, Egan JM, Raja HA, Oberlies NH, Kvalheim OM and Cech NB, *Journal of Natural Products*, 2016, 79, 376–386. [PubMed: 26841051]
78. Lee J-E, Lee B-J, Chung J-O, Hwang J-A, Lee S-J, Lee C-H and Hong Y-S, *Journal of Agricultural and Food Chemistry*, 2010, 58, 10582–10589. [PubMed: 20828156]
79. Fujimura Y, Kurihara K, Ida M, Kosaka R, Miura D, Wariishi H, Maeda-Yamamoto M, Nesumi A, Saito T, Kanda T, Yamada K and Tachibana H, *PLoS One*, 2011, 6, e23426. [PubMed: 21853132]
80. Lee J-E, Lee B-J, Chung J-O, Shin H-J, Lee S-J, Lee C-H and Hong Y-S, *Food Res. Int*, 2011, 44, 597–604.
81. Lee J-E, Lee B-J, Hwang J-A, Ko K-S, Chung J-O, Kim E-H, Lee S-J and Hong Y-S, *Journal of Agricultural and Food Chemistry*, 2011, 59, 10579–10585. [PubMed: 21899366]
82. Lee L-S, Choi JH, Son N, Kim S-H, Park J-D, Jang D-J, Jeong Y and Kim H-J, *Journal of Agricultural and Food Chemistry*, 2013, 61, 332–338. [PubMed: 23256790]
83. Lee J-E, Lee B-J, Chung J-O, Kim HN, Kim E-H, Jung S, Lee H, Lee SJ and Hong Y-S, *Food Chemistry*, 2015, 174, 452–459. [PubMed: 25529705]
84. Gupta PK, Barone GW, Gurley B, Fifer EK and Hendrickson HP, *Drug Metabolism and Disposition*, 2015, DOI: 10.1124/dmd.114.059410.
85. Bogdanski P, Suliburska J, Szulinska M, Stepien M, Pupek-Musialik D and Jablecka A, *Nutrition research*, 2012, 32, 421–427. [PubMed: 22749178]
86. Bergeron C, Livesey JF, Awang DVC, Arnason JT, Rana J, Baum BR and Letchamo W, *Phytochemical Analysis*, 2000, 11, 207–215.
87. Wohlmuth H, Savage K, Dowell A and Mouatt P, *Phytomedicine*, 2014, 21, 912–918. [PubMed: 24566389]
88. McCloud TG, *Molecules*, 2010, 15, 4526–4563. [PubMed: 20657375]
89. Gurley BJ, *Clinical Pharmacology & Therapeutics*, 2011, 89, 915–919. [PubMed: 21544076]
90. Brown PN and Roman MC, *Journal of AOAC International*, 2008, 91, 694–701. [PubMed: 18727526]
91. Mudge EM, Betz JM and Brown PN, *Advances in Nutrition: An International Review Journal*, 2016, 7, 390–398.
92. Napolitano JG, Gödecke T, Rodríguez-Brasco MF, Jaki BU, Chen S-N, Lankin DC and Pauli GF, *Journal of Natural Products*, 2012, 75, 238–248. [PubMed: 22332915]
93. Luthria DL, Mukhopadhyay S, Robbins RJ, Finley JW, Banuelos GS and Harnly JM, *Journal of Agricultural and Food Chemistry*, 2008, 56, 5457–5462. [PubMed: 18572954]
94. Wang M, Zhao J, Avula B, Wang Y-H, Avonto C, Chittiboyina AG, Wylie PL, Parcher JF and Khan IA, *Journal of Agricultural and Food Chemistry*, 2014, 62, 12103–12111. [PubMed: 25383633]
95. Graf TN, Cech NB, Polyak SJ and Oberlies NH, *Journal of Pharmaceutical and Biomedical Analysis*, 2016, 126, 26–33. [PubMed: 27136284]

96. Coimbra M, Isacchi B, van Bloois L, Torano JS, Ket A, Wu X, Broere F, Metselaar JM, Rijcken CJ, Storm G, Bilia R and Schiffelers RM, *International journal of pharmaceutics*, 2011, 416, 433–442. [PubMed: 21291975]
97. Lapornik B, Prošek M and Golc Wondra A, *Journal of Food Engineering*, 2005, 71, 214–222.
98. Azmir J, Zaidul ISM, Rahman MM, Sharif KM, Mohamed A, Sahena F, Jahurul MHA, Ghafoor K, Norulaini NAN and Omar AKM, *Journal of Food Engineering*, 2013, 117, 426–436.
99. Kellogg JJ, Wallace ED, Graf TN, Oberlies NH and Cech NB, *Journal of Pharmaceutical and Biomedical Analysis*, 2017, 145, 604–610. [PubMed: 28787673]
100. Yuliana ND, Khatib A, Verpoorte R and Choi YH, *Analytical Chemistry*, 2011, 83, 6902–6906. [PubMed: 21800886]
101. Saccenti E, Hoefsloot HCJ, Smilde AK, Westerhuis JA and Hendriks MMWB, *Metabolomics*, 2014, 10, 361–374.
102. Abdi H and Williams LJ, *Wiley Interdisciplinary Reviews: Computational Statistics*, 2010, 2, 433–459.
103. LaBudde RA and Harnly J, *Journal of AOAC International*, 2012, 95, 273–285. [PubMed: 22468371]
104. López MI, Callao MP and Ruisánchez I, *Analytica Chimica Acta*, 2015, 891, 62–72. [PubMed: 26388364]
105. Sheen DA, Wocha WFC, Lippa KA and Bearden DW, *Chemometrics and Intelligent Laboratory Systems*, 2017, 162, 10–20. [PubMed: 28694553]
106. Grapov D, Wanichthanarak K and Fiehn O, *Bioinformatics*, 2015, 31, 2757–2760. [PubMed: 25847005]
107. Rider CV, Dinse GE, Umbach DM, Simmons JE and Hertzberg RC, in *Chemical Mixtures and Combined Chemical and Nonchemical Stressors*, Springer, 2018, pp. 235–270.
108. Rider CV, Walker NJ and Waidyanatha S, *Clinical Pharmacology & Therapeutics*, 2018.
109. Ernst E, *Journal of Internal Medicine*, 2002, 252, 107–113. [PubMed: 12190885]
110. Tims M, *Botanical Adulterants Bulletin*, 2016, 6, 1–6.
111. Geng P, Harnly JM, Sun J, Zhang M and Chen P, *Analytical and Bioanalytical Chemistry*, 2017, 409, 2591–2600. [PubMed: 28160032]
112. Sarma N, Giancaspro G and Venema J, *Drug testing and analysis*, 2016, 8, 418–423. [PubMed: 26857794]
113. Betz JM, Brown PN and Roman MC, *Fitoterapia*, 2011, 82, 44–52. [PubMed: 20884340]
114. Simmler C, Graham JG, Chen S-N and Pauli GF, *Fitoterapia*, 2018, **in press**.
115. Geng P, Harnly JM and Chen P, *Journal of Agricultural and Food Chemistry*, 2015, 63, 6189–6211. [PubMed: 26083013]
116. Mncwangi NP, Viljoen AM, Zhao J, Vermaak I, Chen W and Khan IA, *Phytochemistry*, 2014, 106, 104–115. [PubMed: 25041697]
117. Pengelly A, Bennett K, Spelman K and Tims M, *Appalachian Center for Ethnobotanical Studies*, 2012.
118. Weber HA, Zart MK, Hodges AE, Molloy HM, O'Brien BM, Moody LA, Clark AP, Harris RK, Overstreet JD and Smith CS, *Journal of agricultural and food chemistry*, 2003, 51, 7352–7358. [PubMed: 14640583]
119. Grippo AA, Hamilton B, Hannigan R and Gurley BJ, *American journal of health-system pharmacy*, 2006, 63.
120. Wong JW, Zhang K, Tech K, Hayward DG, Krynitsky AJ, Cassias I, Schenck FJ, Banerjee K, Dasgupta S and Brown D, *Journal of Agricultural and Food Chemistry*, 2010, 58, 5884–5896. [PubMed: 20225896]
121. Witschi C and Doelker E, *European Journal of Pharmaceutics and Biopharmaceutics*, 1997, 43, 215–242.
122. Veprikova Z, Zachariasova M, Dzuman Z, Zachariasova A, Fenclova M, Slavikova P, Vaclavikova M, Mastovska K, Hengst D and Hajslova J, *Journal of Agricultural and Food Chemistry*, 2015, 63, 6633–6643. [PubMed: 26168136]

123. Raman P, Patino LC and Nair MG, *Journal of Agricultural and Food Chemistry*, 2004, 52, 7822–7827. [PubMed: 15612762]
124. Brown AC, *Food and Chemical Toxicology*, 2017, 107, 449–471. [PubMed: 27818322]
125. Food and Drug Administration, *Journal*, 2010.
126. Danko R, *Natural Product Insider*, 2016, 21, 48.
127. Dowlatabadi R, Farshidfar F, Zare Z, Pirali M, Rabiei M, Khoshayand MR and Vogel HJ, *Metabolomics*, 2017, 13, 1–11. [PubMed: 27980501]
128. Alvarez-Zapata R, Sánchez-Medina A, Chan-Bacab M, García-Sosa K, Escalante-Erosa F, García-Rodríguez RV and Peña-Rodríguez LM, *Journal of Chromatography A*, 2015, 1422, 213–221. [PubMed: 26515384]
129. Simmler C, Kulakowski D, Lankin DC, McAlpine JB, Chen S-N and Pauli GF, *Advances in Nutrition*, 2016, 7, 179–189. [PubMed: 27180381]
130. Enke CG and Nagels LJ, *Analytical Chemistry*, 2011, 83, 2539–2546. [PubMed: 21366323]
131. Junio HA, Sy-Cordero AA, Etefagh KA, Burns JT, Micko KT, Graf TN, Richter SJ, Cannon RE, Oberlies NH and Cech NB, *Journal of Natural Products*, 2011, 74, 1621–1629. [PubMed: 21661731]
132. Chiari BG, Severi JA, Pauli-Credendio D, Abackerli P, Sylos C. M. d., Vilegas W, Correa MA and Isaac VLB, *International Journal of Pharmacy and Pharmaceutical Sciences*, 2012, 331–336.
133. Wagner H, *Fitoterapia*, 2011, 82, 34–37. [PubMed: 21075177]
134. Niemeyer K, Bell IR and Koithan M, *Journal of herbal medicine*, 2013, 3, 112–119. [PubMed: 24058898]
135. Burns J, Zhao L, Taylor EW and Spelman K, *Toxicology*, 2010, 278, 140–159. [PubMed: 19818374]
136. Oberlies NH and Kroll DJ, *Journal of natural products*, 2004, 67, 129–135. [PubMed: 14987046]
137. Wani MC, Taylor HL, Wall ME, Coggon P and McPhail AT, *Journal of the American Chemical Society*, 1971, 93, 2325–2327. [PubMed: 5553076]
138. van Breemen RB, Huang CR, Nikolic D, Woodbury CP, Zhao YZ and Venton DL, *Anal Chem*, 1997, 69, 2159–2164. [PubMed: 9183179]
139. Yang Z, Zhang Y, Sun L, Wang Y, Gao X and Cheng Y, *Anal Chim Acta*, 2012, 719, 87–95. [PubMed: 22340536]
140. Ciesla L and Moaddel R, *Natural Product Reports*, 2016, 33, 1131–1145. [PubMed: 27367973]
141. Martens H, Bruun SW, Adt I, Sockalingum GD and Kohler A, *Journal of Chemometrics*, 2006, 20, 402–417.
142. Balouiri M, Sadiki M and Ibsouda SK, *Journal of Pharmaceutical Analysis*, 2016, 6, 71–79. [PubMed: 29403965]
143. Brusotti G, Cesari I, Dentamaro A, Caccialanza G and Massolini G, *Journal of pharmaceutical and biomedical analysis*, 2014, 87, 218–228. [PubMed: 23591140]
144. Dietz BM, Hajirahimkhan A, Dunlap TL and Bolton JL, *Pharmacological reviews*, 2016, 68, 1026–1073. [PubMed: 27677719]
145. Harvey AL, Edrada-Ebel R and Quinn RJ, *Nature reviews. Drug discovery*, 2015, 14, 111. [PubMed: 25614221]
146. Newman DJ and Cragg GM, *J. Nat. Prod*, 2016, 79, 629–661. [PubMed: 26852623]
147. Chun-Guang W, Jun-Qing Y, Bei-Zhong L, Dan-Ting J, Chong W, Liang Z, Dan Z and Yan W, *European Journal of Pharmacology*, 2010, 627, 33–41. [PubMed: 19857484]
148. Gufford BT, Chen G, Vergara AG, Lazarus P, Oberlies NH and Paine MF, *Drug Metabolism and Disposition*, 2015, 43, 1353–1359. [PubMed: 26070840]
149. Li G, Huang K, Nikolic D and Van Breemen RB, *Drug Metabolism and Disposition*, 2015, dmd.115.065987.
150. Posadzki P, Watson L and Ernst E, *British journal of clinical pharmacology*, 2013, 75, 603–618. [PubMed: 22670731]
151. Van Loo-Bouwman CA, Naber THJ, Minekus M, van Breemen RB, Hulshof PJM and Schaafsma G, *Journal of Agricultural and Food Chemistry*, 2014, 62, 950–955. [PubMed: 24397305]

152. Possemiers S, Bolca S, Verstraete W and Heyerick A, *Fitoterapia*, 2011, 82, 53–66. [PubMed: 20655994]
153. Price JM, Little JG and Baker TR, *Planta Med*, 2014, 80, CL12.
154. Krivos KL, Regg BT, Price JM, McMillan DA and Baker TR, *Planta Med*, 2014, 80, PPL19.
155. Napolitano JG, Gödecke T, Lankin DC, Jaki BU, McAlpine JB, Chen S-N and Pauli GF, *Journal of pharmaceutical and biomedical analysis*, 2014, 93, 59–67. [PubMed: 23870106]
156. Broeckling CD, Afsar FA, Neumann S, Ben-Hur A and Prenni JE, *Analytical Chemistry*, 2014, 86, 6812–6817. [PubMed: 24927477]
157. Paguigan ND, El-Elimat T, Kao D, Raja HA, Pearce CJ and Oberlies NH, *The Journal of Antibiotics*, 2017, 70, 553–561. [PubMed: 28074050]
158. El-Elimat T, Figueroa M, Ehrmann BM, Cech NB, Pearce CJ and Oberlies NH, *Journal of Natural Products*, 2013, 76, 1709–1716. [PubMed: 23947912]
159. Wang M, Carver JJ, Phelan VV, Sanchez LM, Garg N, Peng Y, Nguyen DD, Watrous J, Kapon CA and Luzzatto-Knaan T, *Nature biotechnology*, 2016, 34, 828.
160. Hoffmann T, Krug D, Hüttel S and Müller R, *Analytical chemistry*, 2014, 86, 10780–10788. [PubMed: 25280058]
161. Gaudêncio SP and Pereira F, *Natural product reports*, 2015, 32, 779–810. [PubMed: 25850681]
162. Cox DG, Oh J, Keasling A, Colson KL and Hamann MT, *Biochimica et biophysica acta*, 2014, 1840, 3460–3474. [PubMed: 25151044]
163. Yuliana ND, Khatib A, Choi YH and Verpoorte R, *Phytotherapy Research*, 2011, 25, 157–169. [PubMed: 20658470]
164. Qiu F, Cai G, Jaki BU, Lankin DC, Franzblau SG and Pauli GF, *Journal of natural products*, 2013, 76, 413–419. [PubMed: 23356207]
165. Britton ER, Kellogg JJ, Kvalheim OM and Cech NB, *Journal of Natural Products*, 2018, 81, 484–493. [PubMed: 29091439]
166. Caesar LK, Kellogg JJ, Kvalheim OM, Cech RA and Cech NB, *Planta medica*, 2018.
167. Farrés M, Platikanov S, Tsakovski S and Tauler R, *Journal of Chemometrics*, 2015, 29, 528–536.
168. Rajalahti T, Arneberg R, Berven FS, Myhr K-M, Ulvik RJ and Kvalheim OM, *Chemometrics and Intelligent Laboratory Systems*, 2009, 95, 35–48.
169. Rajalahti T, Arneberg R, Kroksveen AC, Berle M, Myhr K-M and Kvalheim OM, *Analytical chemistry*, 2009, 81, 2581–2590. [PubMed: 19228047]
170. Tian D-D, Kellogg JJ, Okut N, Oberlies NH, Cech NB, Shen DD, McCune JS and Paine MF, *Drug Metabolism and Disposition*, 2018, **in press**.
171. Bucar F, Wube A and Schmid M, *Natural product reports*, 2013, 30, 525–545. [PubMed: 23396532]
172. Michalkiewicz A, Biesaga M and Pyrzynska K, *Journal of chromatography A*, 2008, 1187, 18–24. [PubMed: 18282581]
173. McAlpine JB, Friesen JB and Pauli GF, in *Natural Products Isolation*, eds. Sarker SD and Nahar L, Humana Press, Totowa, NJ, 2012, DOI: 10.1007/978-1-61779-624-1_9, pp. 221–254.
174. Kim N-C, Oberlies NH, Brine DR, Handy RW, Wani MC and Wall ME, *Journal of natural products*, 2001, 64, 251–253. [PubMed: 11430014]
175. Pfoze N, Myrboh B, Kumar Y and Rohman MR, *Journal of Medicinal Plants*, 2014, 2.
176. Agatonovic-Kustrin S, Morton DW and Yusof AP, *Modern Chemistry & Applications*, 2015.
177. El-Elimat T, Raja HA, Graf TN, Faeth SH, Cech NB and Oberlies NH, *Journal of natural products*, 2014, 77, 193–199. [PubMed: 24456525]
178. Graf TN, Wani MC, Agarwal R, Kroll DJ and Oberlies NH, *Planta medica*, 2007, 73, 1495–1501. [PubMed: 17948171]
179. Bross-Walch N, Kuhn T, Moskau D and Zerbe O, *Chemistry & Biodiversity*, 2005, 2, 147–177.
180. Bouslimani A, Sanchez LM, Garg N and Dorrestein PC, *Natural product reports*, 2014, 31, 718–729. [PubMed: 24801551]
181. Crews P, Rodriguez J and Jaspars M, *Organic Structure Analysis 2nd Ed.*, Oxford University Press, New York, NY, USA, 2009.

182. Linington RG, Williams PG and MacMillan JB, *Problems in Organic Structure Determination: A Practical Approach to NMR Spectroscopy*, CRC Press, Boca Raton, FL, USA, 2015.
183. Halabalaki M, Vougianniopoulou K, Mikros E and Skaltsounis AL, *Current opinion in biotechnology*, 2014, 25, 1–7. [PubMed: 24484874]
184. McAlpine JB, Chen S-N, Kutateladze A, MacMillan JB, Appendino G, Barison A, Beniddir MA, Biavatti MW, Bluml S, Boufridi A, Butler MS, Capon RJ, Choi YH, Coppage D, Crews P, Crimmins MT, Csete M, Dewapriya P, Egan JM, Garson MJ, Genta-Jouve G, Gerwick WH, Gross H, Harper MK, Hermanto P, Hook JM, Hunter L, Jeannerat D, Ji N-Y, Johnson TA, Kingston DGI, Koshino H, Lee H-W, Lewin G, Li J, Linington RG, Liu M, McPhail KL, Molinski TF, Moore BS, Nam J-W, Neupane RP, Niemitz M, Nuzillard J-M, Oberlies NH, Ocampos FMM, Pan G, Quinn RJ, Reddy DS, Renault J-H, Rivera-Chávez J, Robien W, Saunders CM, Schmidt TJ, Seger C, Shen B, Steinbeck C, Stuppner H, Sturm S, Tagliatalata-Scafati O, Tantillo DJ, Verpoorte R, Wang B-G, Williams CM, Williams PG, Wist J, Yue J-M, Zhang C, Xu Z, Simmler C, Lankin DC, Bisson J and Pauli GF, *Natural Product Reports*, 2018, **in press**.
185. Figueroa M, Jarmusch AK, Raja HA, El-Elimat T, Kavanaugh JS, Horswill AR, Cooks RG, Cech NB and Oberlies NH, *Journal of Natural Products*, 2014, 77, 1351–1358. [PubMed: 24911880]
186. Xu H, Niu H, He B, Cui C, Li Q and Bi K, *Molecules*, 2016, 21, 664.
187. Ekanayaka EP, Celiz MD and Jones AD, *Plant physiology*, 2015, 167, 1221–1232. [PubMed: 25659383]
188. Xie T, Liang Y, Hao H, Jiye A, Xie L, Gong P, Dai C, Liu L, Kang A and Zheng X, *Journal of Chromatography A*, 2012, 1227, 234–244. [PubMed: 22293283]
189. Reynertson KA and Mahmood K, *Botanicals: Methods and techniques for quality & authenticity*, CRC Press, Boca Raton, FL, USA, 2015.
190. Arrivault S, Guenther M, Fry SC, Fuenfgeld MM, Veyel D, Mettler-Altmann T, Stitt M and Lunn JE, *Analytical chemistry*, 2015, 87, 6896–6904. [PubMed: 26010726]
191. Chahrouh O, Cobice D and Malone J, *Journal of pharmaceutical and biomedical analysis*, 2015, 113, 2–20. [PubMed: 25956803]
192. Habler K and Rychlik M, *Analytical and bioanalytical chemistry*, 2016, 408, 307–317. [PubMed: 26514672]
193. AOAC International, *AOAC Official Methods of Analysis*, 2012, 1–32.
194. VanderMolen KM, Cech NB, Paine MF and Oberlies NH, *Phytochemical analysis : PCA*, 2013, 24, 10.1002/pca.2449.
195. ICH Harmonized Tripartite Guideline, Q2 (R1), 2005, 1.
196. Roman MC, Betz JM and Hildreth J, *Journal of AOAC International*, 2007, 90, 68–81. [PubMed: 17373438]
197. Jarmusch AK, Musso AM, Shymanovich T, Jarmusch SA, Weavil MJ, Lovin ME, Ehrmann BM, Saari S, Nichols DE and Faeth SH, *Journal of pharmaceutical and biomedical analysis*, 2016, 117, 11–17. [PubMed: 26340558]
198. VanderMolen KM, Ainslie GR, Paine MF and Oberlies NH, *Journal of pharmaceutical and biomedical analysis*, 2014, 98, 260–265. [PubMed: 24951959]
199. Lv Y, Yang X, Zhao Y, Ruan Y, Yang Y and Wang Z, *Food Chemistry*, 2009, 112, 742–746.
200. Liu Y, Chen S-N, McAlpine JB, Klein LL, Friesen JB, Lankin DC and Pauli GF, *Journal of Natural Products*, 2014, 77, 611–617. [PubMed: 24432981]
201. Booker A, Suter A, Krnjic A, Strassel B, Zloh M, Said M and Heinrich M, *Journal of Pharmacy and Pharmacology*, 2014, 66, 811–822. [PubMed: 24417505]
202. Todd DA, Gullede TV, Britton ER, Oberhofer M, Leyte-Lugo M, Moody AN, Shymanovich T, Grubbs LF, Juzumaite M, Graf TN, Oberlies NH, Faeth SH, Laster SM and Cech NB, *PLOS ONE*, 2015, 10, e0124276. [PubMed: 25933416]
203. Dai J, Gupte A, Gates L and Mumper R, *Food and chemical toxicology*, 2009, 47, 837–847. [PubMed: 19271318]

Box 1**Summary of Recommendations: Literature Review for Botanical Natural Products**

- Examine references to identify known constituents and biological targets.
- Investigate botanical usage information, including common species used, physiological portions harvested, and preparation.
- Utilize both peer-reviewed resources, government reports, and industry-based information.

Box 2**Summary of Recommendations: Obtaining Study Material**

- When harvesting cultivated or wild botanical specimens, collect additional material for identification and accession to an herbarium.
- Procure sufficient number of commercial products to be representative of the overall market population. In our studies we evaluated a minimum of 30 different products.
- For commercial samples that are dried, powdered, or pre-extracted, obtain additional authenticated reference material for identification.
- Either a commercial source or a contract laboratory can be used to provide botanical natural product study material. The advantage of a commercial source is that the product is representative of what is used by consumers, but care must be taken to ensure that the material is authentic, consistent, and stable.
- Obtain proper permissions before harvesting; collection of samples could be restricted due to local, state, federal, or international laws.

Box 3**Summary of Recommendations: Authentication of Natural Products**

- Authenticate botanical materials to ensure the validity of *in vitro*, non-clinical *in vivo*, and/or clinical studies.
- Select an authentication method that is suitable for the type of specimens being analyzed. For example, extracted botanical supplements cannot readily be identified using DNA barcoding techniques.
- Authentication using only a single technique may be insufficient to ensure accurate identification of the botanical material, and the application of orthogonal methods (i.e. targeted analysis of marker compounds plus untargeted metabolomics or DNA barcoding) is recommended.

Box 4**Summary of Recommendations: Selecting a Botanical Natural Product Sample**

- Acquire multiple representative examples of different preparations and/or plant parts.
- Use industry and consumer data to guide selection of products that are representative of the consumption patterns of the population of interest.
- Consider whether a standardized product is appropriate for a given application; there are advantages and disadvantages of standardization.
- Select an extraction solvent based on the research question, being mindful that there are many advantages to methanol extraction for stability, ease, and repeatability of laboratory work.

Box 5**Summary of Recommendations: Quality Control and Comparison of Botanical Material**

- PCA is an effective method to compare qualitative differences among botanical samples.
- The composite score (CS) represents a useful statistical approach for comparing the similarity or dissimilarity of multiple samples in large metabolomics datasets.
- Untargeted metabolomics combined with targeted methods can effectively discern potential adulterations in botanical natural products.
- Contamination testing is recommended prior to conducting studies involving botanical natural products.

Box 6**Summary of Recommendations: Identification, Structure Elucidation, and Quantitative Analysis of Bioactive Constituents**

- Knowledge of which constituents in the botanical extract are biologically active is needed to establish quality control procedures and determine dosage.
- Biological activity may be due to more than one constituent, and the active constituent profile may vary depending on which biological activity is tested.
- Because of the inherent complexity of botanical extracts, isolating all constituents and testing them individually for biological activity is infeasible.
- Bioassay-guided fractionation, which can be enhanced with biochemometric analyses, can be used to direct isolation efforts towards identifying compounds likely to be active, provided that a relevant and robust *in vitro* assay is available.
- Results of bioassay-guided fractionation experiments should be interpreted with consideration of the potential limitations of translatability to clinical studies.
- Quantitative analysis of active compounds can be accomplished using isotopically labeled standards or calibration curves prepared from standards of the compound of interest in the relevant matrix.

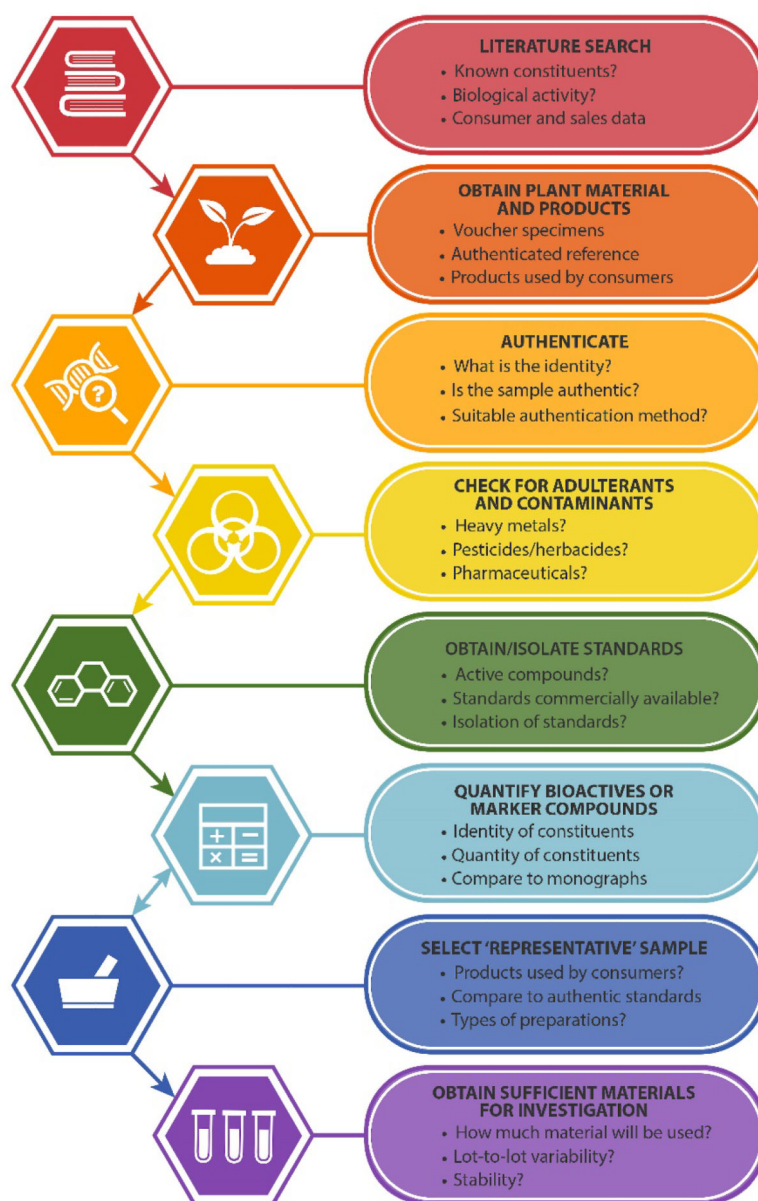


Figure 1. Recommended steps in the selection of botanical natural products for research purposes (*in vitro* and/or *in vivo* studies), and important questions for consideration with each step.

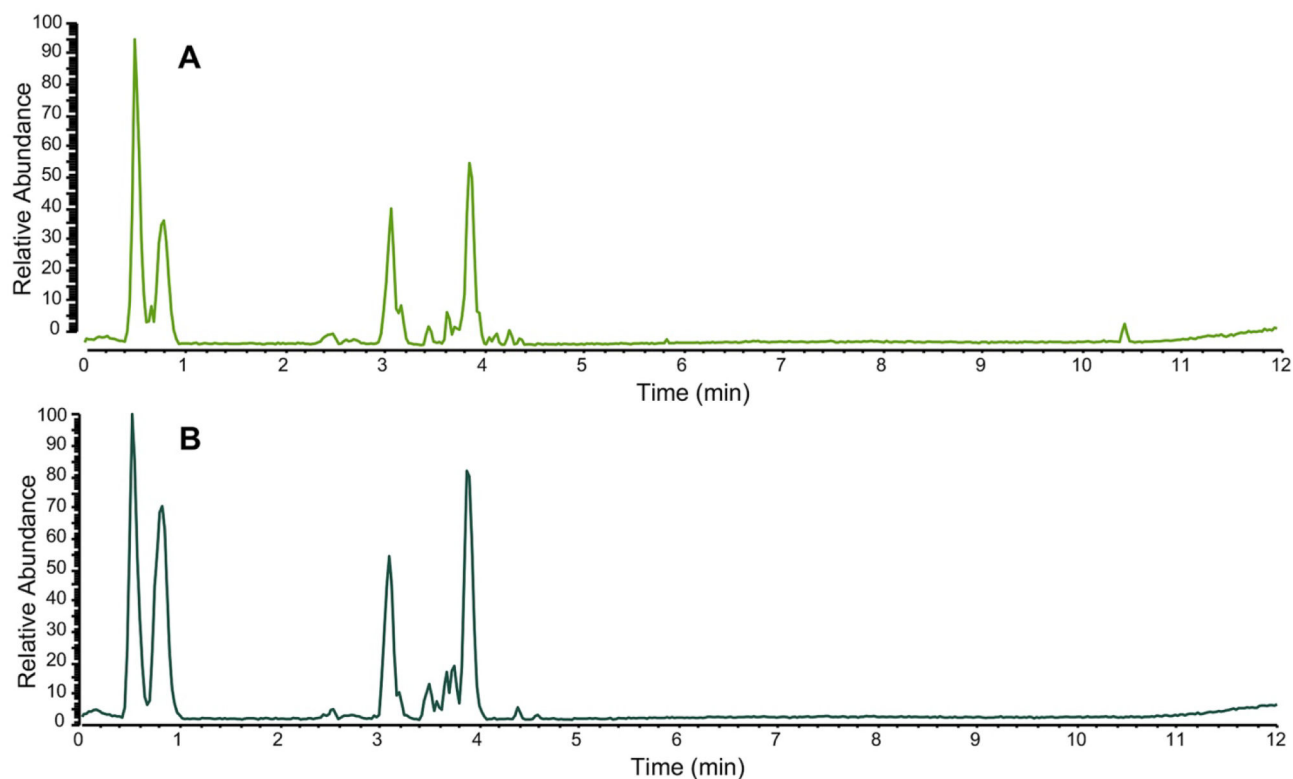


Figure 2. Representative liquid chromatography-mass spectrometry (LC-MS) fingerprinting of a green tea sample (A) compared against an authentic green tea (*Camellia sinensis*) standard from the National Institute of Standards and Technology (NIST 3254) (B). Both chromatograms were collected in the negative ion mode using ultrahigh performance liquid chromatography (UHPLC) coupled to electrospray ionization mass spectrometry; (reproduced with permission from ²⁸). Note that on balance, the majority of the peaks, both in retention time (x axis) and relative abundance (y axis), match between the two samples.

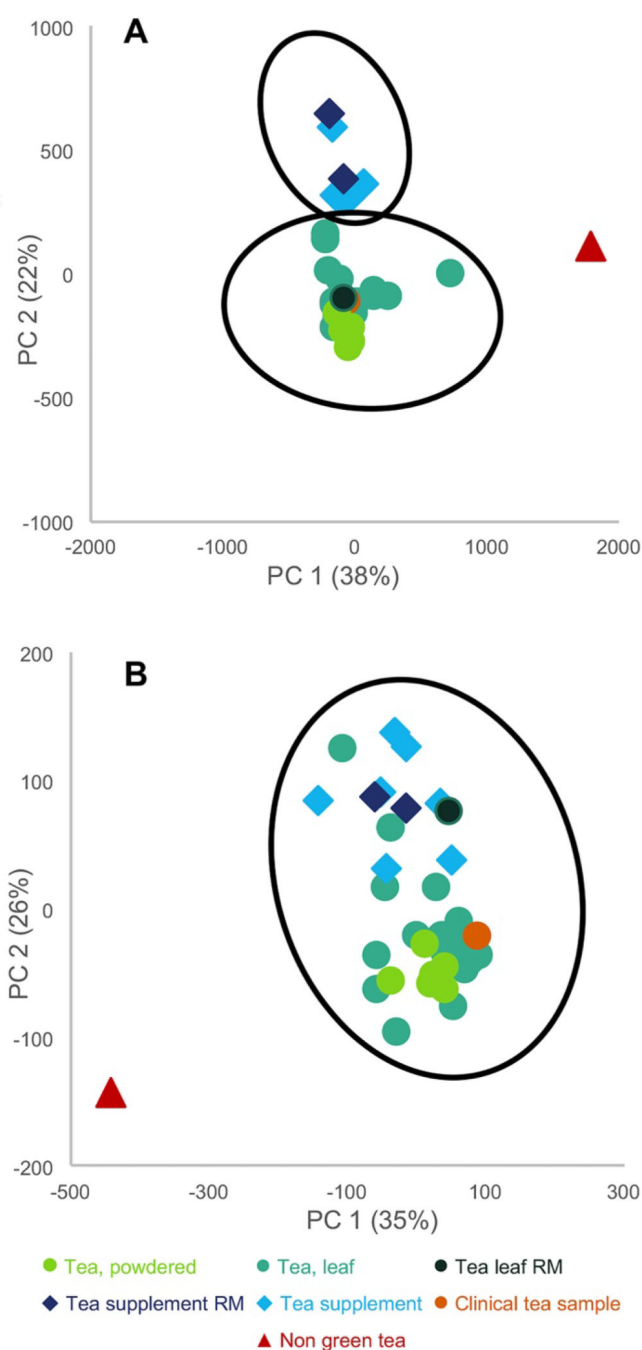


Figure 3.

An example of principal component analysis (PCA) scores plots for green tea samples. Data points represent averaged triplicate extractions and plots are drawn with Hotelling's 95% confidence ellipses. "RM" represents green tea standard reference material from the National Institute of Standards and Technology (NIST). A) Untargeted metabolomics analysis yielded clusters of distinct sample types (green tea supplements, green teas, and the negative control (indicated as "non-green tea")). B) Targeted mass spectrometry analysis, using 15 quantified standards, was employed to differentiate between green tea samples and

the negative control. However, distinction between green tea supplements and green teas was not as clearly defined, and the RM samples were clustered together, irrespective of their origin. Reproduced with permission from Kellogg et al.²⁸

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

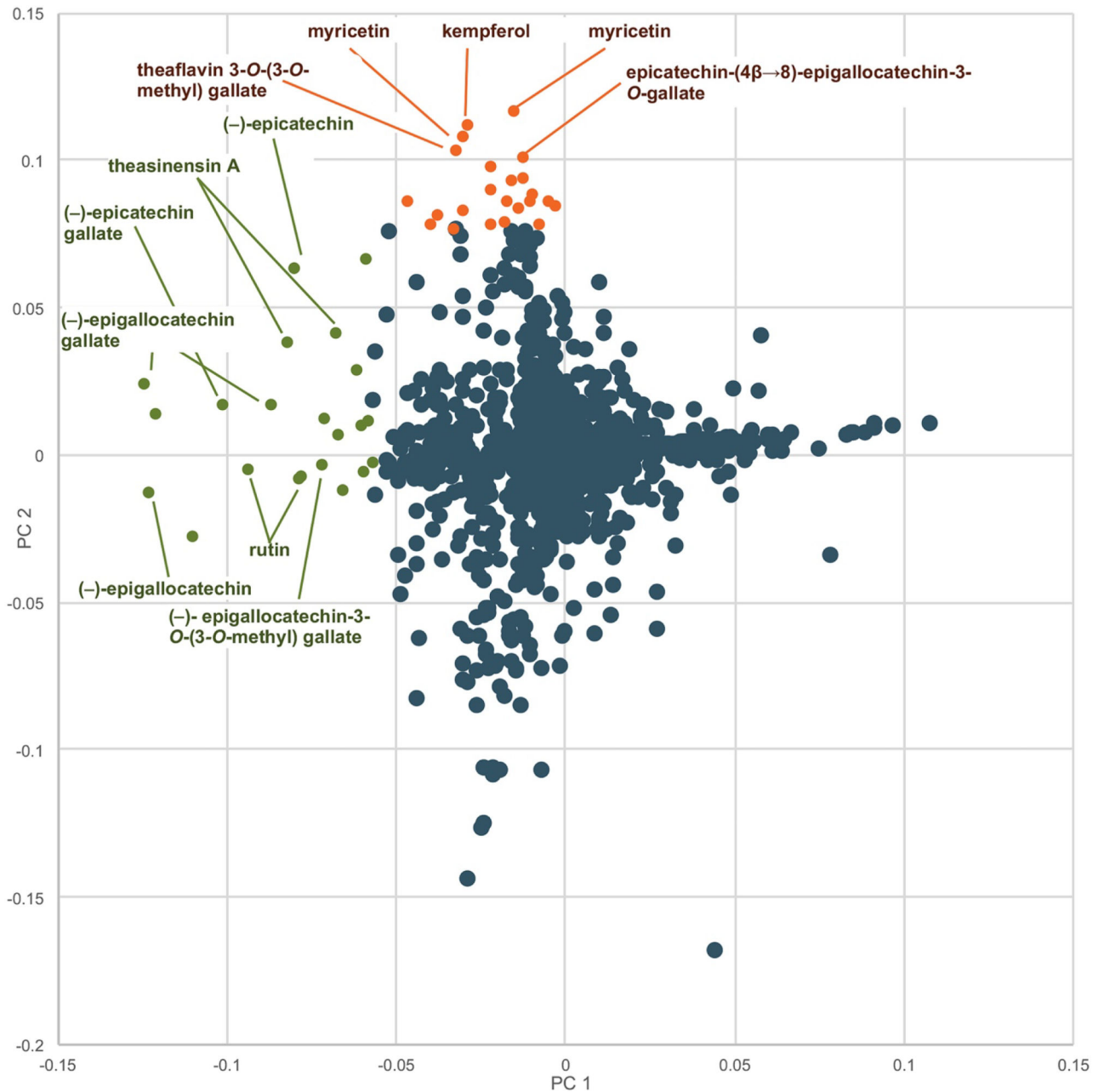


Figure 4.

Loadings plot to identify constituents responsible for differences among commercial preparations of green tea. Metabolites with more negative values along the x-axis (PC1, green labels) were present in higher concentrations in the green tea samples versus the non-green tea (negative) control. Labeled metabolites with greater positive values along the y-axis (PC2, brown labels) were more heavily represented in green tea supplement samples compared to green tea leaf and powder samples (reproduced with permission from Kellogg et al.²⁸)

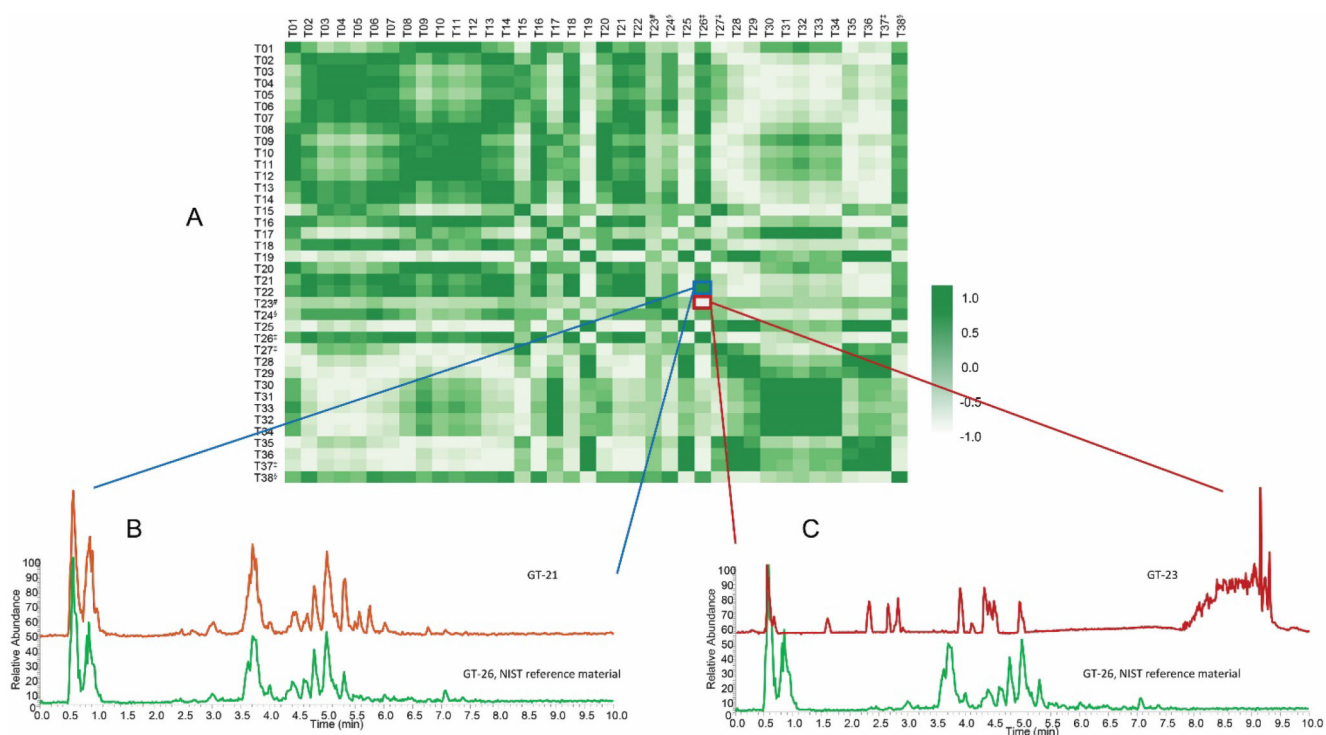


Figure 5.

Composite score (CS) analysis for green tea samples. (A) Heat map correlation matrix of all green tea samples (indicated by codes on the horizontal and vertical axes). CS was calculated from the reproduced correlation coefficient matrix comprised of a four-component PCA model. In the heat map, darker shades (values closer to 1.00) represent stronger correlation between samples, while lighter shades (values approaching -1.00) represent decreased correlation between samples. The correlations' relationship to the metabolome are represented in two examples. (B) High correlation between two samples ($CS = 0.98$, dark color) indicates a strong similarity of the metabolomic profiles. (C) Low correlation ($CS = -0.85$ light color) indicates two samples with diverging metabolomic profiles. #Negative control (T23); ‡NIST standard reference materials (T26, T27, and T37); §green teas with botanical additives (T24 and T38).

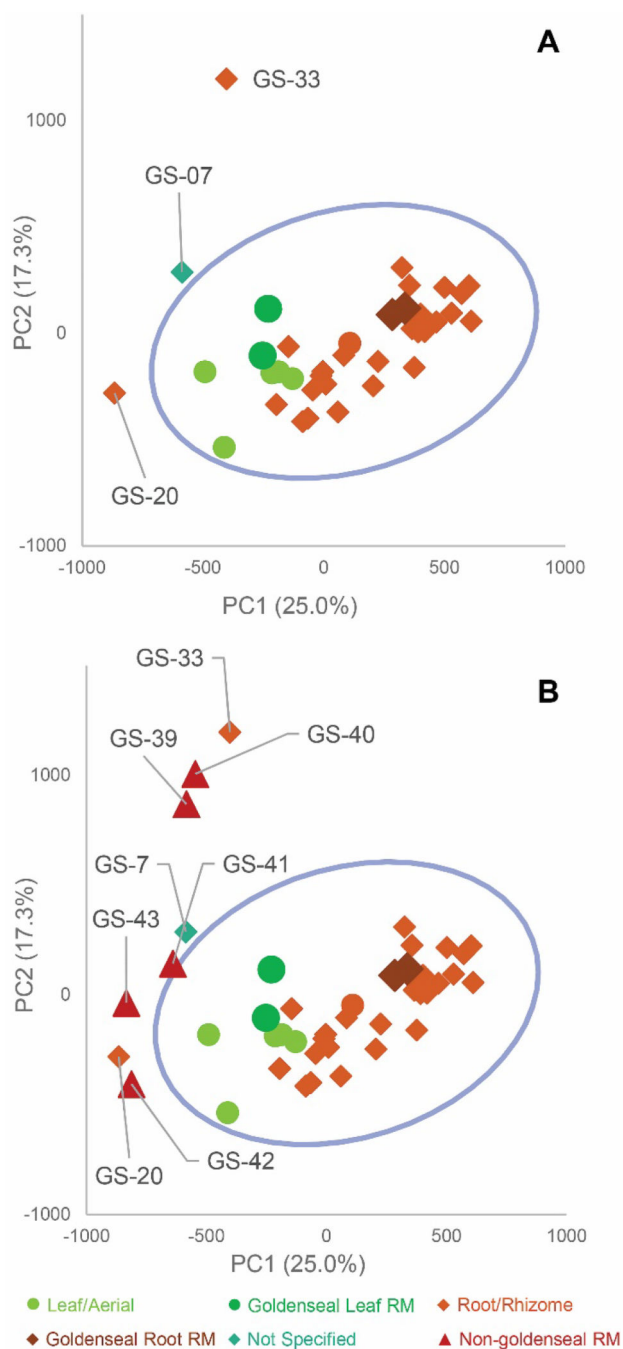


Figure 6. Principal component analysis (PCA) scores plots from untargeted mass spectrometry metabolomics analysis of commercial goldenseal product samples, plotted with Hotelling's 95% confidence interval (as shown by blue circle). **(A)** PC1 versus PC2 (25.0% and 17.3% explained variance, respectively) allowed for visualization of the samples (aerial portions and root/rhizome portions), with corresponding reference material (RM) located within the main cluster. The three samples labeled GS-07, GS-20, and GS-33 are located distinctly from other goldenseal samples and were considered outliers. **(B)** Follow-up metabolomics

profiling and PCA scores plot including non-goldenseal reference materials (*Coptis chinensis*, GS-39 and GS-40; *Mahonia aquifolium* GS-41 and GS-42; *Berberis vulgaris*, GS-43) highlighted the potential sources of adulteration in the three outlier samples. Reproduced with permission from Wallace et al. (2018).²⁹

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

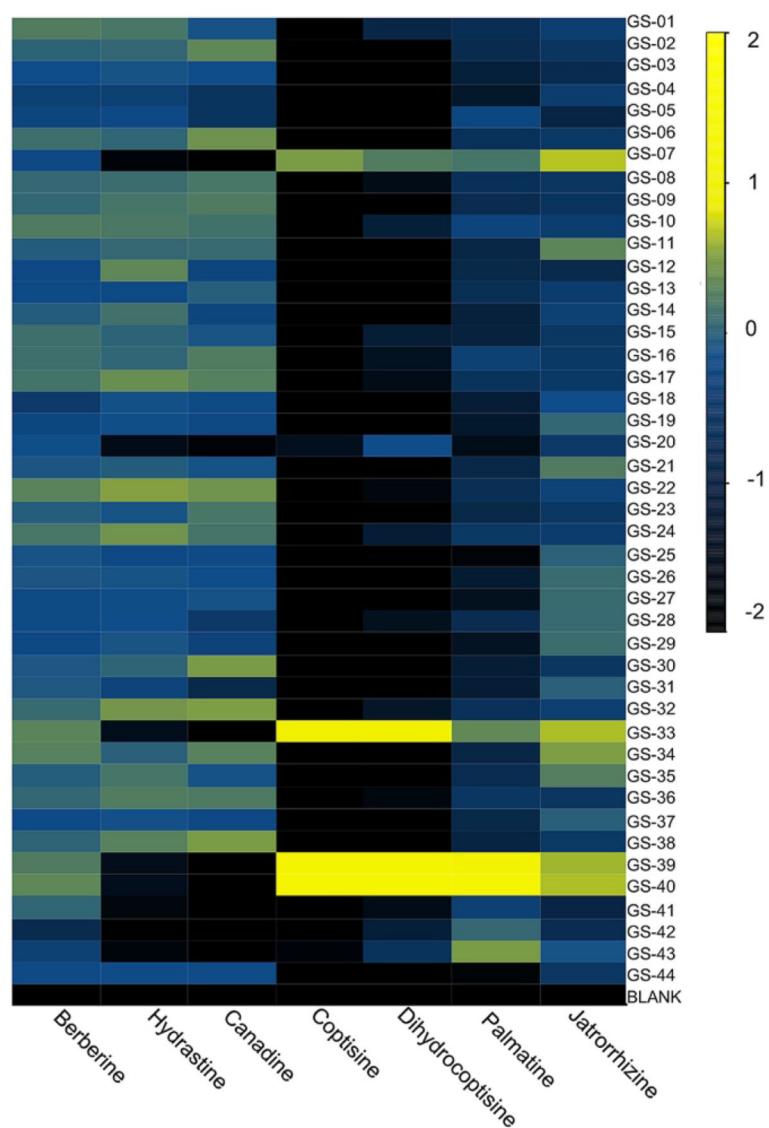


Figure 7. Heatmap of purported adulterants: six primary alkaloids present in *H. canadensis*, *C. chinensis*, *M. aquifolium*, and *B. vulgaris*. Heatmap was generated using the log₁₀ of the peak area for the relevant ion detected by high resolution LC-MS. Reproduced with permission from Wallace et al.²⁹

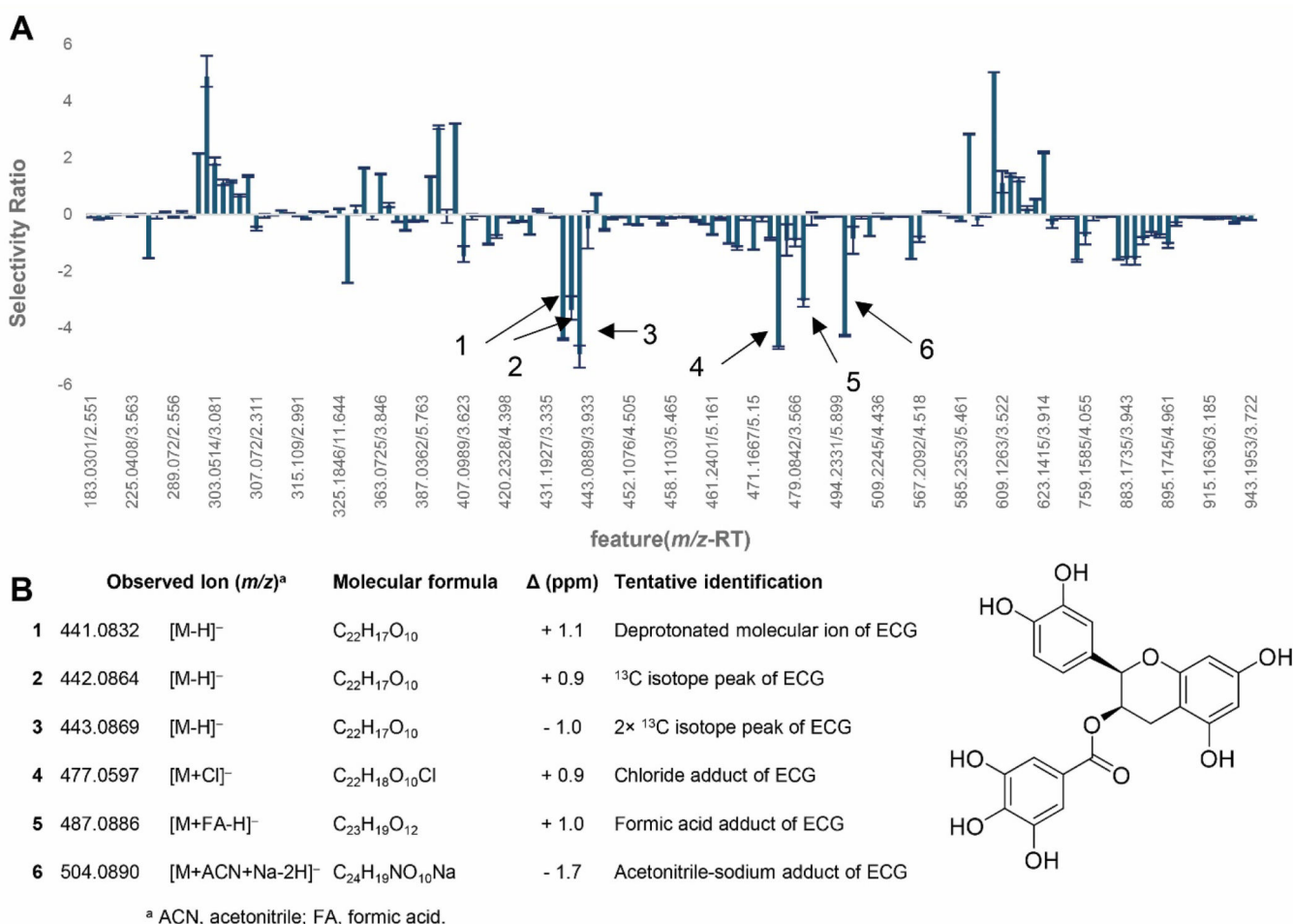


Figure 8. Biochemometric analysis of the inhibition of UDP-glucuronosyltransferase (UGT) activity by green tea constituents. Mass spectrometry-based metabolomics was correlated with bioactivity to generate a partial least squares (PLS) model for eight green tea subfractions. This model was used to determine the selectivity ratio for each feature detected in the green tea extract (**A**), the x-axis represents sample components detected as unique m/z -retention time (TR) pairs (features) in the extract. The more negative values on the y-axis represent more significant contribution to the observed UGT inhibitory activity. All six detected ions (spectral variables) represent isotope peaks and cluster ions of (-)-epicatechin gallate (ECG) (**B**) molecular ion, suggesting that ECG is the dominant bioactive ion in the green tea matrix. Adapted with permission from Tian et al. (2018).¹⁷⁰

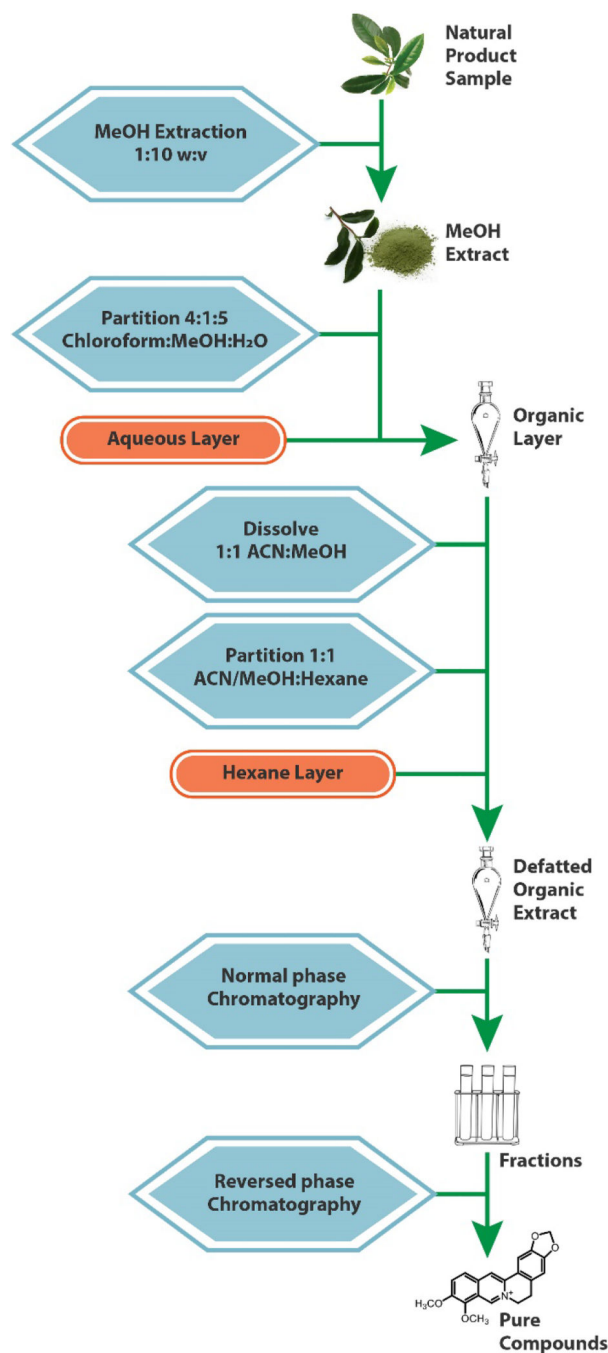


Figure 9. Standard extraction and fractionation scheme for a natural product. The sample material is macerated with an organic solvent, usually methanol, then subjected to repeated liquid-liquid partitions, yielding a final organic extract. This residue is fractionated with a normal-phase flash chromatography system, after which the fractions are further separated via a reverse-phase HPLC separation to yield purified compounds.^{77, 185}

Table 1.

Characteristics of the ideal botanical natural product for use in research studies

-
1. Representative of what is commonly used by consumers
 2. Authentic (species verified)
 3. Well-characterized
 - 3.1. Active constituents known
 - 3.2. Concentrations of active or marker metabolites have been characterized and comply with monograph
 4. Free of contamination and adulteration
 5. Sufficient material available to conduct *in vitro* and/or clinical studies
 6. Material consistent for duration of the trial
 - 6.1. Shelf life (i.e. stability)
 - 6.2. Batch-to-batch reproducibility
-

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Common pitfalls in quantitative and qualitative interpretation of data from botanical extracts

Table 2.

1.	<p><i>Assuming that relative peak intensities in a chromatogram represent relative concentrations in the sample.</i></p> <p>Do not assume that a large peak means a high concentration of analyte in an LC-MS or LC-UV chromatogram. Instead, recognize that they reflect the responsiveness of the analyte to the detector, i.e. its ionization efficiency (for MS) or its molar absorptivity (UV). One approach to circumvent this problem is to use an evaporative light scattering detector (ELSD) or charged aerosol detector (CAD)^{153, 154}. The response of these detectors closely reflects relative abundance of analytes in the sample, but they are less sensitive than MS or UV.</p>
2.	<p><i>Attempting to quantify analytes without reference standards.</i></p> <p>Because response of the detector (mass spectrometer or spectrophotometer) depends on structure, an authentic standard should be used to quantify the unknown of interest.</p>
3.	<p><i>Assuming that the absence of a constituent in the data means an absence of the compound in the sample.</i></p> <p>Failure to observe a compound in the data means only that it was not detectable by the technique used above a specific stated limit of detection. It does not mean that the analyte was absent from the sample. Including a positive control can help address this problem. If the analyte of interest is detectable in the positive control but not in the sample, the analyte likely is absent from the sample (above the limit of detection).</p>
4.	<p><i>Overstating the accuracy to which a concentration is known.</i></p> <p>Absolute quantitation (knowledge of the true amount of a given compound in a botanical sample) is difficult to achieve. Factors such as inefficient extraction, lack of purity in the standard, and drift in instrument response make knowledge of true absolute concentration extremely difficult to assess. Avoid chasing perfect and absolute quantitation when relative or approximate quantitation will be sufficient. On the other hand, knowing that most quantitative measures are not as absolute as the reports may imply, avoid the pitfall of comparing samples from different runs or different laboratories when the results may be different due to factors inherent in the analysis. Include appropriate replicates, references and controls in each experiment such that comparison among experiments is possible. Report the concentration of a given analyte with its associated uncertainty and to the correct significant figures.</p>
5.	<p><i>Assuming confirmation of identity with MS data.</i></p> <p>Even if appropriate standards are used to match retention time and fragmentation patterns of unknown compounds in a sample, mass spectrometry data does not confirm configuration of stereoisomers. NMR is needed for such confirmations. However, depending on the needs of the study, knowledge of configuration of stereoisomers may not be necessary, and MS data may be sufficient.</p>
6.	<p><i>Extrapolating results beyond the linear range of the calibration curve.</i></p> <p>Always dilute samples so that the analyte of interest falls within the linear range of the calibration curve. Failure to do so may result in serious underestimation of analyte concentration. Errors also result from attempting to quantify analytes present at too low of a concentration, i.e. below the limit of quantification for the analytical method.</p>
7.	<p><i>Failing to account for matrix interference.</i></p> <p>Matrix interference occurs when the matrix (everything <i>but</i> the analyte in the sample) alters the response of the analyte. This issue can be particularly pronounced with mass spectrometry as an analytical method. The best strategy for avoiding matrix interference is to dilute the sample as much as possible and subject it to chromatographic separation prior to analysis. A validation check that involves spiking the sample with a standard and comparing its response in matrix and solvent is necessary to check for matrix interference.⁹⁵</p>
8.	<p><i>Assuming a sample is “pure” based on LC-UV data.</i></p> <p>When interpreting LC-UV data, it is important to remember that contaminants will not be detectable unless they absorb light in the region used for the analysis. Quantitative NMR represents an alternate method to determine purity.¹⁵⁵</p>
9.	<p><i>Assuming that each peak in an LC-UV chromatogram represents a single compound.</i></p> <p>Multiple compounds in a complex sample may coelute in what appear as a single peak. Thus, one peak does not necessarily mean one compound. MS detection can be used to identify coeluting compounds if they differ in mass. Examination of UV spectra at multiple retention times across a single chromatographic peak can also aid in detecting coeluting compounds.</p>
10.	<p><i>Assuming that each peak in the mass spectrum represents a different ion.</i></p> <p>Clustering and in-source fragmentation in mass spectrometric analysis often lead to multiple masses that represent a single ion. Identifying the true “molecular ion” can be difficult. Software packages such as RamClust¹⁵⁶ and IntelliXtract¹⁵⁷ can help assign identities of clusters and fragments and group associated ions.</p>

Table 3.**Practical tips for selecting botanical natural products prior to conducting *in vitro*, non-clinical *in vivo*, and clinical studies****Screen many products**

Acquire and chemically evaluate as many different products as feasible. Most studies reviewed selected >20 products,^{28, 29, 86, 94, 201} although lower thresholds may be acceptable depending on limitations due to funding, logistics, and availability. Products selected for a given study or studies should be representative of consumer usage.

Confirm identity claims

Verify all claims on labels via an independent analysis (e.g., the researcher's laboratory or a contract laboratory). This is important both for commercial botanical products and reference standards.

Authenticate thoroughly

Authenticate using multiple orthogonal techniques and by comparing with trustworthy, vouchered samples. While this may seem obvious, avoid the wasted effort of conducting a research project starting with the wrong botanical.

Obtain vouchers

Collect a voucher specimen and submit it to a publicly accessible collection at an herbarium. When a commercial product is used and a voucher specimen is not accessible, compare the chemical profile of the botanical to be used in the study to that of a vouchered reference material.

Collect plant specimens using paper bags

Plant specimens should be stored in paper, rather than plastic, bags after collection. Plastic bags can contaminate the specimens with polymers that may mistakenly be attributed to the specimen. Plastic bags also prevent gas exchange, allowing moisture to accumulate rapidly, and accelerating fungal growth and potential mycotoxin contamination.

Reserve samples

When preparing an extract, always save a labelled and catalogued sample of the starting product (extract, supplement, or dried plant material). When purifying an extract for analysis (i.e. conducting preparative chromatography or liquid-liquid partitioning), save a small sample of the extract and of each successively purified fraction. This reserved aliquot can be re-analyzed as necessary for chemical composition and biological activity. It is better to err on the side of saving too much of these materials than to save too little (or none). Relative to the cost of the future studies, the price of starting with well more than is needed is modest.

Extract with methanol

Use methanol to prepare samples for chemical analysis and *in vitro* evaluation. Methanol is an effective extraction solvent because it facilitates reproducible production and stable storage of extracts. If necessary, conduct a parallel extraction with other solvents to ensure relevance to consumer usage. In the authors' experience, selection of extraction solvent *typically* does not change the profile of small molecule metabolites in a sample to a large extent, albeit relative amounts of these compounds may be altered. Methanol has a lower boiling point than ethanol or water, which facilitates its evaporation at lower temperatures, thus minimizing the chance of decomposition of the sample during evaporation. It is also our experience that the difficulty in working with and storing extracts prepared in other solvents introduces artefacts and unexpected complications into a research study, and some other solvents (such as DMSO) can contribute to the oxidation of extract constituents.⁸⁸

Include controls in analyses

Botanical reference standards aid in the detection of adulterated products and enable botanical authentication, whereas chemical standards aid in quantification and provide the opportunity to monitor the performance of the analytical method.

Purchase or collect more of a single lot than needed for *in vitro* and/or clinical studies

When using a commercial product, purchase double or triple the amount of product needed for the given study or studies. Where possible, purchase all of this material from a single lot to ensure consistency within and between studies.

Characterize fully (not just for marker compounds)

Obtain the most comprehensive profile of metabolites possible when characterizing an extract. Techniques that focus on a few metabolites (or purported "marker compounds") may produce misleading results and may unintentionally imply similarity among products that may possess different biological activities.

Consider the limitations of the analytical technique

A few common errors in interpreting data analysis of botanical extracts are listed in Table 2. Careful data analysis and inclusion of appropriate controls and replicates is important to avoid these errors.

Be cognizant of potential endotoxin contamination

Lipopolysaccharides (LPS, also referred to as endotoxins) are bacterial components that can profoundly influence *in vitro* assays, particularly those evaluating immunomodulatory activity. Because botanical extracts contain endophytic and epiphytic bacteria (and other microbes), it is virtually impossible to avoid contamination with LPS in extracts prepared from botanicals. Even ethanol extracts can contain sufficient LPS content to alter the results of biological assays²⁰².

Store samples cool, dark, and dry

Never store extracts in liquid form, even in the freezer. Freeze/thaw cycles, especially in solvents such as DMSO, can result in accelerated degradation. Storing samples dry and in dark and cool conditions is the best way to prevent degradation. Cooler temperatures, such as refrigeration or freezing, are indicated to prolong the activity and chemical stability of prepared extracts.²⁰³ However, the type and extent of degradation that occurs will differ depending on the natural product being evaluated. Thus, it is advisable to monitor for potential degradation products by collecting chromatographic profiles immediately after the sample is prepared, and then at regular intervals or prior to each experiment in which it will be evaluated for biological activity.