DNA Barcoding for the Identification of Botanicals in Herbal Medicine and Dietary Supplements: Strengths and Limitations

Authors   **Iffat Parveen**[1], **Stefan Gafner**[2], **Natascha Techen**[1], **Susan J. Murch**[3], **Ikhlas A. Khan**[1]

Affiliations   [1] National Center for Natural Products Research, Thad Cochran Research Center, School of Pharmacy, University of Mississippi, University, Mississippi, USA
[2] American Botanical Council, Austin, Texas, USA
[3] Department of Chemistry, University of British Columbia, Kelowna, British Columbia, Canada

Correspondence
Ikhlas A. Khan
TCRC 3012
National Center for Natural
Products Research
Thad Cochran Research Center
School of Pharmacy
University of Mississippi
University, Mississippi 38677
USA
Phone: + 16 62 91 57 8 21
Fax: + 16 62 91 57 9 89
ikhan@olemiss.edu

Correspondence
Dr. Stefan Gafner
American Botanical Council
PO Box 144345
Austin, Texas 78714
USA
Phone: + 1 51 29 54 74 87
Fax: + 1 51 29 26 23 45
stefan@herbalgram.org

Abstract
▼

In the past decades, the use of traditional medicine has increased globally, leading to a booming herbal medicine and dietary supplement industry. The increased popularity of herbal products has led to a rise in demand for botanical raw materials. Accurate identification of medicinal herbs is a legal requirement in most countries and prerequisite for delivering a quality product that meets consumer expectations. Traditional identification methods include botanical taxonomy, macroscopic and microscopic examination, and chemical methods. Advances in the identification of biological species using DNA-based techniques have led to the development of a DNA marker-based platform for authentication of plant materials. DNA barcoding, in particular, has been proposed as a means to identify herbal ingredients and to detect adulteration. However, general barcoding techniques using universal primers have been shown to provide mixed results with regard to data accuracy. Further technological advances such as mini-barcodes, digital polymerase chain reaction, and next generation sequencing provide additional tools for the authentication of herbs, and may be successful in identifying processed ingredients used in finished herbal products. This review gives an overview on the strengths and limitations of DNA barcoding techniques for botanical ingredient identification. Based on the available information, we do not recommend the use of universal primers for DNA barcoding of processed plant material as a sole means of species identification, but suggest an approach combining DNA-based methods using genus- or species-specific primers, chemical analysis, and microscopic and macroscopic methods for the successful authentication of botanical ingredients used in the herbal dietary supplement industry.

Supporting information available online at http://www.thieme-connect.de/products

Introduction
▼

DNA barcodes, a term introduced by Hebert et al. [1], describe short genomic regions from the nuclear and/or organelle genome used to distinguish animal, plant, fungal, and bacterial species. The use of these short genomic regions for biological species discrimination is called DNA barcoding. Applications include tracking the illegal trade of endangered species of both animals and plants (i.e., biopiracy) [2–5], forensic analysis [6, 7] identifying invasive species [8, 9], plant identification at any stage of the life cycle (juvenile or mature) [10], identifying complex food webs by studying species diversity in the gut contents of animals [11], analyzing herbivore's diet components [12], checking adulterations and substitutions in food products [13], and authentication of herbal medi-

cine and identification of botanical ingredient adulterants [14–17].

The herbal products industry is a multibillion-dollar industry and an important part of the world's economy. In a study funded by the Natural Products Foundation, estimates suggest that the total economic contribution of the dietary supplement industry to the U.S. economy is more than three times the annual consumer sales, or $61 billion dollars per year [18]. However, as the popularity of herbal dietary supplements has increased, so have reports of adulteration; this admixture, or substitution, of herbal products/supplements with materials of substandard quality is a growing concern since it may lead to decreased efficacy and the occurrence of serious adverse events. Price pressure, increased demand, limited availability of medicinal herbs, and greed

of unscrupulous suppliers are some of the reasons for the intentional substitution of botanical ingredients. Adulteration is often carried out using a closely related species having similar active ingredients (although in some cases, the use of closely related species is acceptable in pharmacopeial monographs) to fool the traditionally used authentication methods, but can occur with completely different lower-cost substitutes as well. Adulteration can also be accidental, for example, due to the complex nomenclature of medicinal herbs or the unintentional misidentification of plant species collected in the wild. Worldwide, a variety of different names to indicate a specific plant species is currently in use: pharmaceutical names, scientific binomials, older scientific names called synonyms, and common names (see review by Barnes et al. [19]). Confusion often arises when different plants have the same common names, which may lead to situations where the local gatherer collects the wrong material. For example, the name *fang ji* is used for two Chinese herbs – *han fang ji* (*Stephania tetranda*) and *guang fang ji* (*Aristolochia fangchi*). Both materials are used in traditional Chinese medicine (TCM) for the treatment of similar ailments, but *A. fangchi* roots contain aristolochic acids that are nephrotoxic and can cause urothelial carcinoma [20]. Notably, reports from Belgium in the 1990s detailed 128 cases of aristolochic acid nephropathy due to the ingestion of an herbal weight loss product where the roots of *S. tetranda* were replaced by *A. fangchi* roots [21,22]. Another example is the confusion of *Eleutherococcus gracilistylus* (the bark of which is known with herb traders as *wu jia pi*) with *Periploca sepium* (known as *bei wu jia pi*) [23]. Hence, the correct identification and authentication of botanicals is important for the safety and efficacy of herbal products. In the USA, it is mandatory for any dietary supplement manufacturing company to conduct at least one appropriate test to verify the components of the dietary material under the current Good Manufacturing Practice (cGMP) regulations of the US Food and Drug Administration.

Botanical taxonomy, macroscopic, organoleptic, microscopic, and chemical identification methods are typically employed for the authentication of herbal materials. Like all methodologies, they each have their advantages and disadvantages. Ideally, a plant is identified by an experienced botanist in its whole form in the environment where it grows. However, in the global herbal trade, most commodities are sold as cut, powdered, or extracted materials that may be difficult to differentiate. Macroscopic identification based on morphological characters requires an intact plant or plant parts for identification, and that the morphological characteristics available are unique for the plant species of interest. The technique requires an experienced taxonomist and can often be challenging because these morphology-based procedures are usually time consuming and may not always provide resolution to the species level. Macroscopic identification is not suitable for powdered and extracted herbal materials. Botanical microscopy is helpful for cut and powdered ingredients, and allows for the detection of inorganic materials such as sand and salts (which are sometimes added to increase the weight of a material). However, as with macroscopic identification, it may be difficult to determine the identity to the species level. Chemical identification methods are suitable for cut, powdered, and processed material, but they also require expertise as well as complex and often expensive instrumentation, and the phytochemical composition can be affected by geographic location, seasonal variations, storage conditions, and processing method [24]. Recently, DNA barcoding has been proposed as another possible method for identi-

fication and authentication of medicinal plants in herbal products [25,26].

The concept of the barcode is an analogy to the combination and spacing of black and white lines that can be found on the package of almost any commercial item. This "barcode" is a representation of the unique item. It can be scanned and compared to a database to identify the item and consequently its price. Barcodes of plants do not consist of black and white lines, but, instead, consist of a unique combination of nucleotides of roughly 300–1000 base pairs (bp) in length that are ideally specific to a plant of interest. These DNA barcodes are agreed-upon DNA sequences from either the nuclear or organelle (chloroplast, mitochondria) genome. In the strictest term, "DNA barcoding" refers to the technique where millions of copies of these diagnostic sequences are made using universal primer sets in a polymerase chain reaction (PCR), and which are subsequently identified by a sequencing method. The universal primer sets used for PCR were designed to bind to conserved flanking regions of the diagnostic/species identifying sequence found in most plants. The unique nucleotide sequence (DNA barcode) between the flanking regions can be determined and used to compare to voucher sequences stored in either in-house or openly accessible databases, such as GenBank, using the BLAST (Basic Local Alignment Search Tool) algorithm [27]. The open databases are collections of sequences submitted by individual researchers, labs, and institutions but are not curated, and the accuracy has not been verified or confirmed. Frequently, several different sequences are reported from different sources with multiple possible barcode regions of varying lengths. For example, **Table 1 S**, Supporting Information, depicts publicly available barcode sequences and their sources [28]. The data used to standardize the targeted barcodes comes from a range of different sources with different levels of certainty. Some of the barcodes come from wild species from unknown locations, without vouchers, the collector or the date of collection identified, the part of the plant may not be known, and the number of nucleotides sequenced in the barcode region may be different between samples of the same species (**Table 1 S**, Supporting Information). Given that a 99–100% match is used to confirm the species, a match of 98% or less gives only confidence to the genus level [29], the integrity and reliability of the reference materials are crucial to the accuracy of the technique. The best gene region for barcoding of land plants is a matter of debate, and molecular biologists have used a number of different single loci, or combinations of loci, to authenticate plant materials [30]. A recent development is the generation of "mini-DNA barcodes" (○ **Fig. 1**). They represent only a fraction of the agreed upon barcodes and are usually 100–200 bp in length. Amplification of mini-barcodes is done by primer sets that are less universal and more specific to either the genus of a plant or, if possible, unique to only one plant species of a genus.

Though DNA barcoding is independent of morphological characteristics and physical and seasonal variations, it has limitations as a stand-alone method for authentication of processed botanicals. Hence, the current review highlights the potentials and pitfalls of DNA barcoding with an emphasis on DNA barcoding methodology used for the identification of herbs and future directions for authentication of herbal ingredients using different techniques.
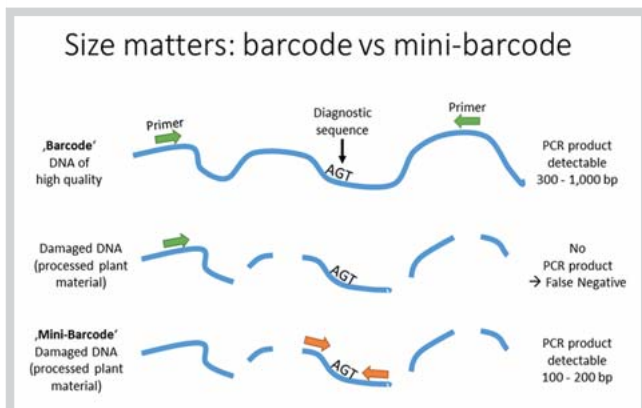
Fig. 1 Advantages of using mini-barcodes for amplification of DNA in processed plant materials. To amplify a full-length barcode (300–1000 bp in length), high-quality DNA as a template is required. When damaged/fragmented DNA serves as a template for PCR, a full-length barcode may not be amplified. Hence, the lack of a PCR product may lead to the wrong conclusion that the DNA template is absent. The use of primer combinations that result in a very small PCR product (100–200 bp), a mini-barcode, is a better approach to detect DNA that may be damaged/fragmented, e.g., found in processed plant materials. (Color figure available online only.)

DNA Barcoding for Identification of Herbal Materials
▼
DNA extraction methods

DNA barcoding can be performed on herbal material only when a minimum quantity and quality of DNA is present. A number of extraction methods and commercial kits are available to extract high-quality DNA from plants [31–39]. However, there is no universal method that can be applied to the isolation of DNA in herbal materials. DNA extraction from herbal materials should be performed shortly after plant collection to avoid DNA degradation due to DNA damaging storage conditions (UV, light, temperature, bacteria, fungi) and under good laboratory practice to avoid cross-contamination with other samples. DNA-based authentication methods work best with freshly collected whole plant material. However, the herbal materials used for dietary supplements are generally collected, dried, and stored for various periods of time before they are used for the preparation of the herbal product. Plant part, storage time, storage conditions, and processing methods affect the quality and quantity of DNA. In addition, plant metabolites such as polysaccharides, flavonoids, polyphenols, and terpene lactones may hinder DNA isolation. The polysaccharides and certain secondary metabolites reportedly coprecipitate with the DNA and prevent the complete dissolution of DNA, thus requiring modified protocols for the isolation of DNA from materials containing polysaccharides and problematic secondary metabolites [40–43]. The most widely used approaches to extract genomic DNA are the cetyl trimethyl ammonium bromide (CTAB) method [33] and commercial DNA extraction kits [37]. However, CTAB/kit methods are not successful in isolating DNA from plants or plant parts that contain high amounts of secondary metabolites. Roots, rhizomes, and tubers may contain particularly high levels of polysaccharides and polyphenols that must be removed using protocols with added high concentrations of CTAB, polyvinylpyrrolidone (PVP), and *β*-mercaptoethanol (*β*-Me) during the early stages of DNA extraction [41,44–46]. High-quality DNA may be obtained from leaves and flowers because of the low levels of interfering metabolites and fibers. Especially the

DNA obtained from fresh and young leaves and flowers can be used to prepare a crude extract, and this solution can be used in direct PCR to amplify the target DNA barcode [47]. However, some plant leaves like tomato, cotton, and tea contain high concentrations of polyphenols and tannins, which also hinder DNA isolation and PCR amplification, thus rendering them unsuitable for direct PCR. Therefore, modified DNA extraction methods were developed for the isolation of DNA from tissues or plants containing high amounts of phenolic compounds and tannins [45,48]. Similarly dried stems, roots, and fruits may not be suitable for direct PCR as they contain secondary metabolites that inhibit PCR amplification. While fresh tissues obtained right after the collection of medicinal herbs are preferred, most of the herbal materials are available for analysis either in dried or powdered form, or, after further processing, as capsules, tablets, or liquids. In modern phytomedicines, the plant DNA is often either removed or degraded during the manufacturing processes of herbal products. Hence, the DNA extracted from capsules, tablets, and liquid extracts appear often as a smear on an analytical agarose gel due to fragmentation, or no band appears due to complete removal of DNA (● **Fig. 2**). Wallace et al. [49] tested 95 natural health products using a standard DNA barcoding technique with multiple markers and primer sets. The authors were unable to retrieve DNA barcodes from 25% of the tested plant products of which 2 were plant roots and 14 were capsules, tablets, or caplets. Their results demonstrated that DNA amplification could be easily accomplished in botanical materials (33 in a total of 35 samples of teas and roots) as compared to pharmaceutical formulations (19 out of 33 samples of tablets, capsules and caplets). The difference in the results could be attributed to the fact that herbal product formulations often contain excipients (such as fillers, diluents, binders, glidants, lubricants, pigments, and stabilizers) that may affect the DNA extraction, or that the primer sets failed to amplify the targeted region. Costa et al. [50] evaluated the possible effect of four different pharmacological excipients on DNA isolation. All the tested excipients, talc, silica, iron oxide, and titanium dioxide, exhibited adsorbent properties that affected the extraction of DNA from natural products [50]. However, with slight modifications in the DNA extraction protocol, it is sometimes possible to obtain useful DNA as a template for PCR from a powder, capsules, or tablets [51]. In contrast to this, the processes by which tinctures and extracts are made, which can include extensive heat treatment, filtration, extractive distillation, or supercritical fluid extraction [52], thus degrading or removing the plant DNA completely, often make these materials unsuitable for DNA barcoding analysis [53].

Loci selected as DNA barcodes for herbal materials

The selection of a universal barcode region for identification of all land plants has proven to be quite challenging. Though the technique has successfully been used in the identification of animal species using cytochrome oxidase I (*COI*) from the mitochondrial genome as a universal barcode [1], barcoding plants is more difficult for many reasons. The slow evolutionary rate of the plant mitochondrial genome means that the mitochondrial gene regions, including the *COI* region, do not sufficiently distinguish plants species. Therefore, relatively fast-evolving plastid and nuclear genomes were proposed as alternative barcodes for plants [54–57]. The most common regions are *matK, rbcL,* ITS, ITS2, *psbA-trnH, atpF-atpH, ycf5, psbK-I, psbM, trnD, coxI, nad1, trnL-F, rpoB, rpoC1,* and *rps16* [54–57]. These regions have a relatively fast evolutionary rate compared to mitochondrial genes, can dis-
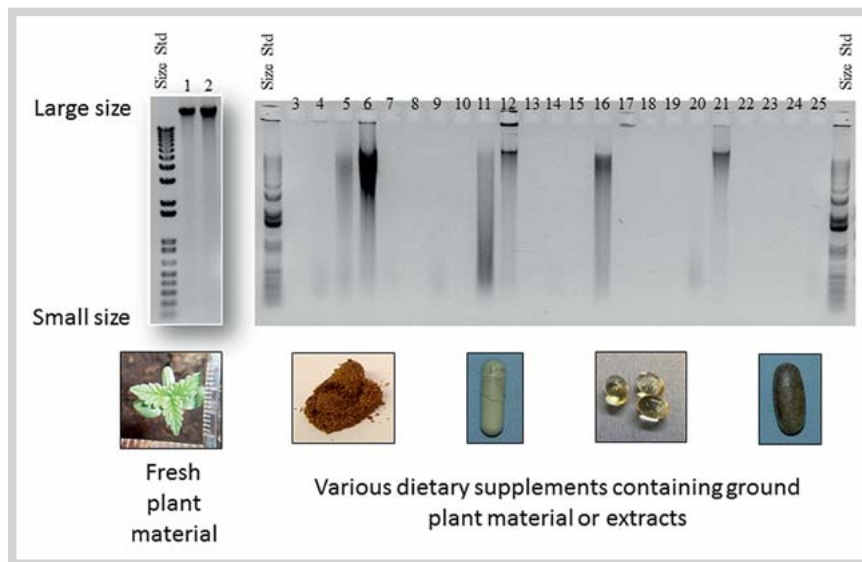
Fig. 2 Comparison of high-quality DNA (large size and usually received from fresh plant material) and genomic DNA from processed plant material, e.g., from dietary supplements, which is often damaged (appear as a smear on gel; Lanes 6, 11, 16), present in low quantities, or absent altogether and can often not be detected on an agarose gel. (Color figure available online only.)

tinguish the species based on differences in the genetic code, and have conserved regions flanking the ends of the DNA sequence for the binding of universal primers. None of the individual plant DNA barcodes described to date have both differentiating regions and universal primer regions. Hence, a multilocus plant barcode with combinations of two or three loci was recommended [55, 56]. The Consortium for the Barcode of Life (CBOL) Plant Working group [56] suggested *matK + rbcL* as the preferred plant barcode combination. The *matK* locus is difficult to amplify in some plant genera since the universal primer-binding site is not perfectly conserved and, therefore, may not be useful in a two-locus combination in some plants. Another suggestion was the *rbcL + trnH-psbA* combination, but due to the high variability of the *trnH-psbA* sequence, it was difficult to align, and thus the two-locus barcode approach was found to be problematic for some of the plants. To overcome this problem, a tiered approach was suggested by Newmaster et al. [54]. The method utilizes the easily amplifiable and alignable *rbcL* region as a scaffold on which data from highly variable non-coding regions such as ITS2 or the *trnH-psbA* region are employed for identification of plant species. Using this tiered approach, approximately 75–80% of the tested plant species can reportedly be barcoded [54,58].

Numerous studies have been published to identify medicinal plants using various suggested barcode loci. Techen et al. [30] reviewed the various barcode loci and methods used for the identification of medicinal plants. The most extensive study of DNA barcoding of medicinal plants was accomplished by Chen et al. [20]. Seven DNA regions, *psbA-trnH, matK, rbcL, rpoC1, ycf5*, ITS (consisting of both ITS1 and ITS2), and ITS2, were evaluated for identification of more than 6600 samples of medicinal plants (fresh leaves) and their closely related species [20]. Their data suggested that the nuclear ITS2 locus could identify 92.7% of the tested species (8557 medicinal plants and closely related samples belonging to 5905 species from 1010 diverse genera of 219 families in 7 phyla-angiosperms, gymnosperms, ferns, mosses, liverworts, algae, and fungi) and proposed ITS2 as the core barcode for medicinal plants [20]. Subsequently, ITS2 has been tested across a broad range of plant taxa with a large sample size and confirmed as an effective barcode for plants by the China Plant Barcode of Life (BOL) Group [59]. Chen and colleagues subsequently built a TCM barcode platform, called the Traditional Chi-

nese Medicine Database using a two-locus barcode system containing ITS2 and *psbA-trnH* sequences [47]. This database contains barcodes belonging to more than 23 000 medicinal plant species and known adulterants [47]. Additional investigations on different plant groups confirmed the effectiveness of ITS2 for the identification of medicinal plant species. For example, 24 medicinal plants from the Fabaceae family and their adulterants were identified using ITS2, with a success rate of 37–97% [60]. Sun and Chen [61] successfully used the ITS2 barcode to distinguish 19 cortex herbs listed in the Chinese Pharmacopoeia, and Pang et al. [62] also reported ITS2 as a suitable barcode for the identification of herbal materials. Zhu et al. [63] recommended ITS2 locus to discriminate between *Glehniae radix* and its common adulterants. The major advantage of ITS2 as a barcode for the identification of herbal supplements is its short length (200–230 bp on average). In most herbal products and dietary supplements, the DNA is highly degraded into pieces of less than 500 nucleotides in length due to various processing methods and, consequently, universal long barcodes ranging from 600 to 800 bp, on average, may not be amplified. Hence, short-length barcodes that can be easily retrieved from dried, powdered form, or sometimes even from extracts were recommended [51]. Despite this advantage, the ITS2 locus was not found to be suitable for global identification of plants because of the presence of multiple copies of ITS2 within one individual in all plant species. The multiple ITS2 copies, which are not always homogenized by concerted evolution, led to the incorrect identification of species due to their similarity with the copies of the more closely related species. Furthermore, the same problem might arise in hybrids due to the biparental inheritance of ITS2. Another disadvantage of ITS is the technical problems in amplification and sequencing that can arise due to the presence of DNA from other species (e.g., fungi, which coexist with plants as endophytes and/or mycorrhizal symbionts) [64,65]. Recently, Cheng et al. [66] reported that the most frequently used primer pairs (ITS1 + ITS4) for ITS were originally designed for fungi [67]. The less plant-specific primer sets led to low PCR and sequencing success rates in some of the plant groups, for instance, < 50% PCR success for algae [68,69] and ferns [20], 57.6% for gymnosperms [70], and 88.0% for angiosperms [70]. Therefore, Cheng et al. [66] designed universal

plant-specific ITS primers for both plant DNA barcoding and plant systematics.

Recently, the term "mini-barcodes" was introduced for short-length DNA markers used for the identification of botanical ingredients from processed herbal supplements [51, 71]. Mini-barcodes are short (< 200 bp) sequences of DNA from standardized *matK* and *rbcL* barcode regions, which have been used, e.g., to identify and authenticate herbal dietary supplements made from saw palmetto (*Serenoa repens*) fruit [71], *Ginkgo biloba* [51] leaf, and devil's claw (*Harpagophytum procumbens* and *Harpagophytum zeyheri*) root and rhizome [72] herbal dietary supplements available in the markets of North America. The advantages of mini-barcodes are the easy retrieval of DNA markers even from processed dietary materials due to their small amplicon length, and their ability to distinguish closely related species because of the genus/species specificity.

PCR amplification

PCR, first devised by Mullis [73], is a molecular method by which a single copy or a few copies of a piece of DNA is amplified and thousands to millions of copies of a particular DNA sequence are generated. The method employs a heat-stable DNA polymerase, the nucleotides, template DNA, and DNA oligonucleotides (also called DNA primers). The general barcoding technique uses universal primers for rapid identification of plant species [57, 58, 74, 75]. The universal primer sets recommended for barcoding of plant species are selected to amplify DNA from four genomic regions, namely ITS/ITS2 from the nuclear genome, and *matK, rbcL*, and *trnH-psbA* from the chloroplast genome. Inherent biases in the analysis can lead to false positives and false negatives [76]. Biases can occur when the sequence at one of the universal priming sites varies sufficiently to prevent efficient annealing, when different species have a different number of copies of the chosen region leading to over or underestimation unrelated to the actual composition of plant material, when the target DNA is degraded or fragmented, or from the manner of extraction and preparation of the DNA for sequencing [65, 76]. Soares et al. [77] and Costa et al. [78] demonstrated that the reliability of DNA barcode methods in complex plant samples varied depending on which of the commercial DNA extraction kits was used for the initial sample preparations. The efficiency of the PCR reaction itself can impose a bias on the barcoding results. Differences in the melting temperatures of the primers could lead to a reduced amplification rate and the affinity of universal primers to template DNA of all known and unknown organisms, and a balanced melting temperature of the primer pairs are two important criteria to produce robust amplification [79]. Further, the presence of inhibitory secondary metabolites and inactive ingredients in tablets and capsules can also reduce the efficiency of PCR amplification or lead to false negative results. The use of excipient materials made from wheat, rice, or soy is common in manufacturing processes of the herbal dietary supplement and pharmaceutical industries [50]. Small amounts of starch are often required in order to optimize the formulation and manufacture of pills, capsules, and tablets. Multiple nucleotide sequences may be obtained upon the sequencing of herbal supplements due to the presence of excipients or if the herbal product contains more than one plant species. Little [51] used digital PCR [80] for the verification of the presence of ginkgo DNA in herbal supplements that, due to the overpowering amount of excipient, only produced sequences from excipient materials or that produced sequencing chromatograms with multiple/overlaying signals indicating a mixed DNA sample.

In digital PCR, the samples are diluted (1:5–1:50000) in a suitable buffer to an extent that there is approximately one DNA template molecule per µL. The goal is to dilute the DNA to the extent that a few samples contain only molecules of low abundant DNA. The term "low abundant DNA" refers to DNA that originated either from the medicinal plant material present within a large amount of DNA from filler material (e.g., rice flour) or DNA derived from small amounts of adulterating plant material. By preparing several PCRs using one molecule/µL DNA solution as a template, chances are that DNA molecules of low abundance get amplified and, consequently, detected. The number of PCRs required depends on the expected frequency of the DNA to be detected. For a sample containing 10% of the low abundant DNA, 2 out of 20 PCRs may result in the amplification of low abundant DNA. The analysis showed 9 (24.3%) of the 37 herbal supplements required digital PCR to separate excipient DNA from possible ginkgo DNA. In the study, digital PCR produced amplicons of ginkgo, rice, and an unidentifiable species [51].

Sequencing methods

The conventional method used for generating DNA sequence data to obtain a barcode from PCR amplicons is Sanger's di-deoxy method of sequencing [81]. Sanger's sequencing technology is capable of generating sequencing reads of up to 1000 bases and has been the approach used for DNA sequencing in most of the DNA barcode analyses published to date. The inherent limitations of Sanger-based DNA sequencing are low throughput and the requirement for high concentrations of DNA (100–500 ng) to avoid biases and errors [82]. Moreover, the method provides two sequencing signal patterns, or electropherograms, for each sequence generated [83]. Hence, the Sanger sequencing method is suitable for herbal materials that contain only a single medicinal plant. If the herbal preparation (e.g., a dietary supplement) contains multiple plant species or excipients, co-amplification of barcode sequences from other material than the intended one can occur due to the nature of the universal primers during the PCR amplification step. This leads to the production of multiple/overlaying sequencing peaks and, consequently, a failure of sequencing because the correct DNA sequence of the barcode cannot be determined (○ **Fig. 3**). Moreover, multiple sequences may also create confusion in the identification of the "true" barcode and other sequences. Additionally, many plants have symbiotic associations with bacteria, algae, or fungi [64, 65]. The occurrence of multiple copies of fungal ITS barcodes creates difficulties in direct Sanger sequencing. Most of these situations can lead to ambiguity or false information when the Sanger sequencing method is employed in generating DNA barcodes, and may result in repeated or failed sequencing attempts.

The poor read quality can be improved by processes upstream to sequencing, for example cloning in a suitable bacterial or microbial host. The DNA fragments are ligated with a vector and cloned in bacteria. Several bacterial clones are then sequenced to identify the different DNA sequences. However, cloning introduces biases against extreme base composition (e.g., stretches with high guanine and cytosine contents), inverted repeats, and genes not accepted by the bacterial cloning host [84]. To overcome the limitations of Sanger-based sequencing for DNA barcoding of processed or mixed samples, a high-throughput sequencing method called next-generation sequencing (NGS) has been used [85]. The NGS technology allows parallel sequencing of multiple DNA fragments from various DNA templates in a single reaction [85]. It can generate up to one million DNA sequences that are
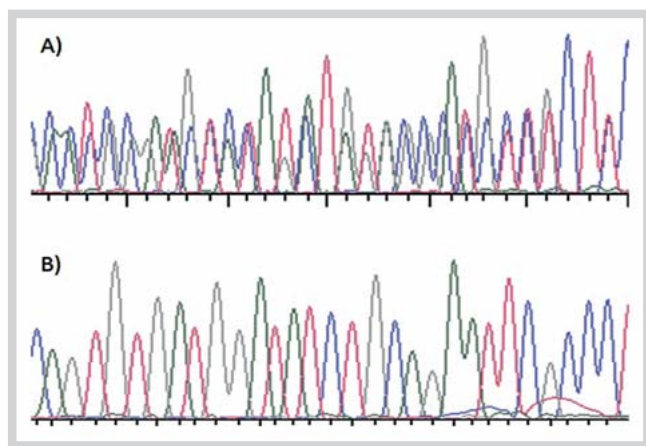
Fig. 3 Electropherograms showing sequencing signals obtained with Sanger sequencing. The overlaying peaks (**A**) make it difficult to determine the real sequence. It is an indication that the sequenced template consists of mixed DNA. Additional steps (digital PCR or cloning) are recommended to identify the DNA source(s). Single peaks (**B**) with a low background are desired to determine a sample's DNA sequence. (Color figure available online only.)

up to 700 bases in length in a single sequencing run, though the base length is highly variable depending on the NGS platform/ technology being used. The NGS platforms were originally developed to generate DNA sequence information from whole genomes or large environmental samples. For example, the whole chloroplast sequence of *Ceratophyllum demersum* was obtained by Moore et al. [86] using the 454 Life Sciences sequencing platform and complete plastomes of 37 *Pinus* species were assembled by Parks et al. [87] on a multiplex Illumina sequencing platform. The NGS method was also useful to verify the contents of multiple ingredient herbal products by parallel sequencing their barcodes [88]. This method prevents overlaying sequence peaks as found in Sanger sequencing and therefore facilitates the sorting of DNA barcodes, and, consequently, the identification of mixed plant material [65,89]. NGS proved to be superior to Sanger sequencing in a comparative analysis that included 15 commercial dietary supplements made from crude powdered material (7), extracts (7), or a mixture of both (1). Reproducible Sanger sequencing using the *rbcL* and ITS2 gene regions was achieved in four dietary supplements containing crude powdered material. None of the extracts provided sequences of the labeled ingredient using Sanger sequencing, but excipient DNA was detected in two supplements instead. The NGS method using the ITS2 locus yielded results in eight supplements, including three extracts. The use of the ITS2 gene region for the three valerian (*Valeriana officinalis*) root samples was unsuccessful, possibly due to the intraspecific variation of the plant, which is known to have variable gene size and ploidy levels depending on the population [65]. The NGS method is generally a less expensive method as compared to Sanger sequencing in terms of per base sequencing cost, however, the cost may increase if only a few samples are analyzed in a single run. Furthermore, there is an additional cost for bioinformatics due to the large amount of data obtained from NGS.

The NGS "meta-barcoding" method combines DNA barcoding and high-throughput DNA sequencing to mass analyze DNA barcodes from sediments or environmental, ancient/historical, or processed samples [90,91]. Coghlan et al. [92] used a meta-barcod-

ing technique for the detection of plant and animal DNA from highly processed TCM. A high-throughput NGS screen was used for 15 complex TCM samples. The NGS generated over 49 000 sequence reads and, according to the BLAST results, the analysis showed that the reads belonged to 68 plant families, including two genera containing possibly toxic species. Some of the TCM samples also contained traces of CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora)- listed animal and plant genera such as the Asiatic black bear (*Ursus thibetanus*), the Saiga antelope (*Saiga tatarica*), and Asian ginseng (*P. ginseng*) [92].

Accuracy of DNA Barcoding Techniques for Authentication of Botanicals in Herbal Products
▼

Most of the botanical ingredient DNA barcoding studies published to date focused on the identification of a universal marker or suitable barcode locus/loci for herbal raw material authentication. Only a small number of published papers discussed the use of DNA barcoding for the identification and authentication of botanicals from finished herbal products and dietary supplements. Srirama et al. [17] reported an investigation in which 25 *Phyllanthus* samples used as raw herbal drugs in the Indian market were assessed for their authenticity. Their analysis revealed that six different species of *Phyllanthus* were available on the market based on morphological studies. Seventy-six percent of the market samples contained *Phyllanthus amarus* as the predominant species (>95%) and the remaining 24% included five different species, namely *P. debilis, P. fraternus, P. urinaria, P. maderaspatensis*, and *P. kozhikodianus*. Species-specific DNA barcode signatures were developed for the tested *Phyllanthus* species using the chloroplast DNA region *psbA-trnH.* The trade sample identities were validated and confirmed by these species-specific DNA barcodes [17]. Ginseng samples (*Panax* spp.) were tested for their authenticity by both Zuo et al. [93] and Wallace et al. [49]. Zuo et al. [93] analyzed DNA from fresh leaves of 95 ginseng samples, representing all of the species in the genus *Panax*. The analysis showed that the combination of *psbA-trnH* and ITS was able to identify all of the species and clusters in the genus. Wallace et al. [49] tested 41 commercial ginseng samples (raw materials and finished products) in the North American market and found that the core barcodes *matK* and *rbcL* required additional data from ITS for successful species identification. Stoeckle et al. [94] analyzed commercial tea samples (*Camellia sinensis*) with 90% success identification rates using *rbcL* and *matK* barcode loci and reported 33% adulterations in herbal teas. Black cohosh (*Actaea racemosa*) samples were analyzed by Baker et al. [95], and with a mini-barcode approach using the *matK* locus they could identify 75% of the tested samples. They found that 25% of the tested black cohosh samples were adulterated. Little and Jeanson [71] used mini-barcodes from the *rbcL* and *matK* regions to authenticate saw palmetto (*S. repens*) herbal dietary supplements. The analysis of these tested supplements demonstrated that 85% contained saw palmetto and that 6% of the supplements contain related species (*Aceolorrhaphe wrightii* or an unidentified species) that cannot be legally sold as herbal dietary supplements in the United States [70]. Similarly, a mini-barcode assay of ginkgo (*G. biloba*) dietary supplements by Little [51] revealed that of the 40 supplements tested, 83.8% contained identifiable *G. biloba* DNA, and six supplements (16.2%) contained fillers without any detectable *G. biloba* DNA. Substitution of raw herbal materials in

local Indian markets was shown for *Sida cordifolia* [96] and in 50% of *Cassia fistula* and *Senna* spp. [97]. Palhares et al. [98] analyzed 257 dried or powdered samples from 8 medicinal plant species approved by the World Health Organization (WHO) for the production of herbal drugs sold in Brazilian markets. These included witch hazel (*Hamamelis virginiana*) leaves, chamomile (*Matricaria recutita*) flowers, espinheira santa (*Maytenus ilicifolia*) leaves, guaco (*Mikania glomerata*) leaves, Asian ginseng (*P. ginseng*) roots, passion flower (*Passiflora incarnata*) leaves, boldo (*Peumus boldus*) leaves, and valerian (*V. officinalis*) roots. The DNA barcoding analysis using *matK, rbcL*, and ITS2 regions confirmed species belonging to the correct genus in 42% of the samples. For the remainder, results suggested that the level of substitutions might be as high as 71% [98], although some of this is due to the misapplication of the common name, e.g., 100% of the samples presented as *P. ginseng* were actually from the genus *Pfaffia*, also called "Brazilian ginseng". Recently, Han et al. [47] investigated 1436 samples representing 295 medicinal species from 7 primary TCM markets in China. Their results indicated that of the 1260 samples, approximately 4.2% were identified as being adulterated and they suggested a regulatory platform based on DNA barcoding should be established for TCM market supervision. However, verification of the accuracy of the results using established compendial methods, e.g., those listed in the European Pharmacopoeia or the United States Pharmacopeia (USP), was not performed in any of these studies.

The first investigation into a larger set of diverse dietary supplements using DNA barcoding was published by Newmaster et al. in 2013 [99]. In the study, the authenticity of 44 (41 capsules, 2 powders, and 1 tablet) single ingredient herbal products from 12 companies was evaluated. The validity of the approach was evaluated by including a second set of samples grown from commercially available seeds in horticultural greenhouses, including all those plants listed on the herbal product labels and some closely related species. The authors of the study were able to recover DNA from 40 out of 44 samples. According to the results, 14 samples out of 40 were correctly labelled, 14 contained the correct species but included additional DNA from either another species or from an excipient, and 12 contained only DNA from other species or excipients. A number of concerns about the study were raised by DNA experts and members of the American Botanical Council [100], e.g., the results were not confirmed by orthogonal methods, such as chemical analysis, and the DNA barcoding method was not validated to the standard required of the industry. According to the authors, products contained solely crude powdered raw material; however, attempts to identify the contents using botanical microscopy were unsuccessful due to the lack of recognizable plant fragments (S. Newmaster personal communication, July 8, 2014). Also, although the products were manufactured and sold, the authors did not take into account the normal manufacturing process for commercial dietary supplement products. Another issue is the acceptable presence of other species. Monographs for herbal raw materials, such as those in the European Pharmacopoeia and the United States Pharmacopeia, allow a certain amount (e.g., 2% in USP) of foreign organic matter. DNA from foreign organic matter such as other plant species could be accidentally introduced at any stage of processing, for example, at the time of collection of herbal plant material, during storage, drying, grinding, at various stages of the product manufacturing process, or during the analysis in the quality control laboratory (○ Fig. 4). Therefore, it is normal to detect DNA from additional species in crude herbal raw materials.
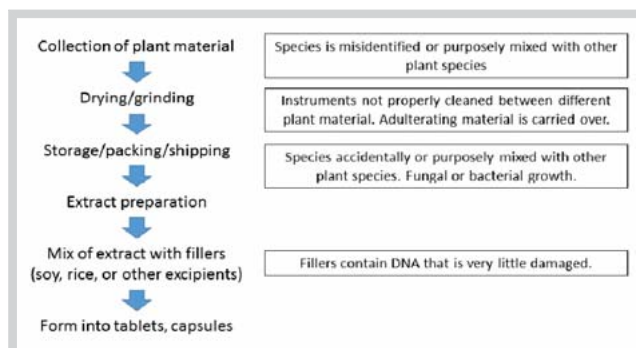

Fig. 4 Various steps in the processing of plant material during which exogenous DNA may be introduced into a sample. (Color figure available online only.)

The results and methodology published by Newmaster et al. [99] prompted the New York State Attorney General (NYAG) to launch his own investigation into the quality of herbal dietary supplements using DNA barcoding (details about the method have not been made publicly available) [101, 102]. The results from this investigation suggested that out of 24 commercial products [labeled to contain Echinacea (*Echinacea* spp.), garlic (*Allium sativum*), ginkgo (*G. biloba*), ginseng (*Panax* spp.), saw palmetto (*S. repens*), St. John's wort (*Hypericum perforatum*), or valerian (*V. officinalis*)] analyzed, only 5 contained DNA of the labeled species. The results led the NYAG to demand that the four retailers selling the supplements, GNC, Target, Walgreens, and Walmart, remove the products from their shelves [101]. The accuracy of the results was immediately questioned, mostly because the majority of the products were made from herbal extracts, where DNA was probably fragmented or degraded, and DNA barcoding (i.e., the amplification of the several hundred nucleotide long complete barcodes) has not been shown to provide useful results for these cases [103]. In addition, the investigation found DNA from *Oryza, Allium*, and *Dracaena* species in 19, 9, and 7 of the samples analyzed, respectively, strongly suggesting cross-contamination. Another puzzling result was the occurrence of saw palmetto DNA in one of the valerian samples, again raising the question about cross-contamination. And while the usefulness of full-length DNA barcoding in providing reliable information of what types of genera or species might be present in a sample containing crude or raw botanical materials is generally recognized, the full-length barcoding approach suffers from the misconception that DNA is always uniformly preserved and available [72, 104]. As mentioned before, DNA quality is affected by heat, freeze/thaw cycles, fungal or bacterial contamination, irradiation, and a number of chemicals. In addition, most of the DNA is typically removed during metabolite extraction, and any DNA that does remain in an herbal extract will usually be fragmented. Additional purification steps used in commercial extraction, such as column chromatography, may eliminate any remaining DNA altogether. Therefore, the relatively long genomic regions required for universal DNA barcoding are no longer present in most botanical extracts. According to Little [72], universal DNA barcoding rarely provides accurate species identification, since distinction among closely related species is difficult, often impossible, and that the diagnostic features examined may not be distinctive enough for a given species.

Method Validation

▼

A critical step in ensuring reliable results from any analytical method is the method validation. Guidelines for the validation of methods to identify botanical ingredients have been published, e.g., by AOAC International [105]. The guidelines were written with chemical methods in mind, so transferring these to DNA barcoding may not be straightforward. The method of choice has to be able to distinguish the species of interest from its adulterants, which may include different plant parts from the same species. It also needs to be able to determine at what level of contamination the method can detect an adulterant. This is done by preparing mixtures of the target species with known amounts of adulterant. These requirements can usually be fulfilled with raw herbal material, except that DNA barcoding will not be able to determine the plant part from which the article is derived from. However, method validations become complicated when finished products are evaluated. DNA barcoding methods are most often "validated" by using fresh or dried raw plant material in order to determine if the chosen DNA barcode region can be amplified [72, 98, 99] and distinguished from related species. The usefulness of such validation in finished products is questionable since processing methods may alter or eliminate the DNA, excipients and secondary metabolites may interfere with DNA extraction and amplification, and the presence of DNA from excipients and fillers may lead to erroneous results. Therefore, the development of guidelines to address the method validation requirements for DNA barcoding, in particular with regard to finished products, is much needed and one of the larger issues to be resolved.

Limitations and Future Challenges of DNA Barcoding for Authentication of Herbal Products

▼

The major limitations to DNA barcoding of herbal products are related to the quality of DNA, primer affinity, PCR amplification, and sequencing of amplicons. Plant DNA is a relatively stable molecule and can be easily extracted from fresh or dried plant material using simple DNA extraction methods. However, the manufacturing process of herbal products that involves extensive heat treatment, irradiation, distillation, filtration, UV light exposure, and/or supercritical fluid extraction leads to either complete removal of DNA or degradation of DNA into smaller fragments [52]. Hence, DNA barcoding is not feasible for processed herbal products such as extracts and tinctures in which the DNA is not present at all or highly degraded. Therefore, the stage at which DNA barcoding analysis should be performed is very important. For example, in the case of processed herbal materials, the molecular analysis is more successful if it is carried out at the initial stage during the collection of raw herbal materials from which the botanical preparation is manufactured. DNA barcoding is more feasible from dried or powdered raw material from which the extract or tincture is manufactured. One of the major advantages of DNA-based analysis is the ubiquitous presence of DNA in all parts of the plant and no effect of seasonal variations on the quality and quantity of DNA has been observed. However, in traditional systems of medicine, specific plant parts/tissues collected in a particular season when secondary metabolite production is highest have been prescribed to be used for therapeutic purposes [106]. DNA barcoding cannot differentiate the different tissues of the plant within the same species. For example, in the case of Asian ginseng (*P. ginseng*), the roots are often mixed or substituted with undeclared *P. ginseng* leaves. Roots and leaves both contain high levels of ginsenosides, but exhibit a different chemical profile. In such cases of substitution with the same species parts, DNA barcoding will not be able to detect the adulteration. Similarly, the approach will not detect if exhausted material is used, i.e., material where the putative active components have been removed previously via extraction. Likewise, if the prescribed plant is collected in the wrong season, the efficacy of the herbal material decreases and DNA barcoding will fail to detect the substitution with such low quality material. To overcome these limitations of DNA-based analysis, chemical profiling or other analytical chemistry methods should be adopted for authentication of herbal products. Thus, DNA barcoding should go hand-in-hand with chemical analysis and macroscopic and/or microscopic evaluation to tackle the adulteration problems prevailing in the herbal industry. Another limitation of DNA barcoding is the universal primer set affinity to the excipient DNA or DNA of another adulterating or substituted species. Generally, the excipients are added after the processing of the herbal material is completed and, hence, the filler's DNA remains intact. Consequently, DNA barcode primers are likely to preferentially amplify DNA from excipients, possibly yielding a false negative result for the herbal species that was intended to be detected. PCR bias can partly be overcome by NGS, digital PCR [51, 71], or by cloning specific PCR products into vectors and sequencing the amplicon. Moreover, it is preferable to design species/genus-specific primers for short barcode sequences (< 200 bp), so-called mini-barcodes, for successful amplification of species-specific barcodes, rather than using the longer universal barcodes. The shortcomings of Sanger-based sequencing methods can be overcome by NGS techniques for sequencing of herbal products containing multiple plant species or admixtures. DNA barcoding, like any other analytical method, has its own limitations, and hence, we view it as an additional identification tool for the authentication of herbal products in combination with other established methods.

One of the major challenges of DNA barcoding for authentication of herbal products is the lack of reference libraries and voucher specimens linked to reference DNA sequences in the GenBank database. The creation of a mini-barcode reference library or Herb-BOL (also suggested by Mishra et al. [107]), containing all of the authentic reference barcode sequences linked to the respective taxonomically validated herbarium vouchers, would be a useful tool to ensure access to reliable DNA barcodes. The use of a barcode reference library could provide a basis for using DNA technologies as a cGMP-compliant approach for the authentication of herbal products and dietary supplements in the future.

Conclusions

▼

The DNA sequencing technologies to identify medicinal plant species in herbal products and dietary supplements is a highly reliable and promising tool under specific conditions, such as the correct stage of analysis when the DNA could be detected, primer affinity for successful PCR amplification, and absence of contaminating DNA. The detection of adulteration of botanical ingredients could be improved if DNA barcoding is routinely and appropriately used for authentication of herbal materials. It is important to apply the most appropriate method to efficiently detect and identify the analyzed raw or processed material.

However, the inherent limitations of the DNA barcoding methods make it unsuitable as a stand-alone tool for identifying and authenticating the herbal plant species. Therefore, we advocate the addition of DNA barcoding to the other existing analytical methods for authentication of botanical ingredients in herbal medicines and dietary supplements.

Supporting information

Reference materials for the DNA barcoding of *H. perforatum* L. (St. John's wort) are available as Supporting Information.

Acknowledgements
▼

Conflict of Interest
▼

The authors declare no conflicts of interest.

References

1 *Hebert PDN, Cywinska A, Ball SL, de Waard JR.* Biological identifications through DNA barcodes. Proc R Soc Lond B 2003; 270: 313–321

2 *Eaton MJ, Meyers GL, Kolokotronis SO, Leslie MS, Martin AP, Amato G.* Barcoding bushmeat: molecular identification of Central African and South American harvested vertebrates. Cons Genet 2010; 11: 1389–1404

3 *Jeanson ML, Labat JN, Little DP.* DNA barcoding: a new tool for palm taxonomists? Ann Bot 2011; 108: 1445–1451

4 *Muellner AN, Schaefer H, Lahaye R.* Evaluation of candidate DNA barcoding loci for economically important timber species of the mahogany family (Meliaceae). Mol Ecol Resour 2011; 11: 450–460

5 *Yesson C, Bárcenas RT, Hernández HM, de la Luz Riz-Maqueda M, Prado A, Rodríguez VM, Hawkins JA.* DNA barcodes for Mexican Cactaceae, plants under pressure from wild collecting. Mol Ecol Resour 2011; 11: 775–783

6 *Coyle HM, Lee CL, Lin WY, Lee HC, Palmbach TM.* Forensic botany: using plant evidence to aid in forensic death investigation. Croat Med J 2005; 46: 606–612

7 *Ferri G, Alù M, Corradini B, Beduschi G.* Forensic botany: species identification of botanical trace evidence using a multigene barcoding approach. Int J Legal Med 2009; 123: 395–401

8 *Bleeker W, Klausmeyer S, Peintinger M, Dienst M.* DNA sequences identify invasive alien *Cardamine* at Lake Constance. Biol Conserv 2008; 141: 692–698

9 *Van de Wiel CCM, van der Schoot J, van Valkenburg JLCH, Duistermaat H, Smulders MJM.* DNA barcoding discriminates the noxious invasive plant species, floating pennywort (*Hydrocotyle ranunculoides* L.f.), from non-invasive relatives. Mol Ecol Resour 2009; 9: 1086–1091

10 *Gonzalez MA, Baraloto C, Engel J, Mori SA, Pétronelli P, Riéra B, Roger A, Thébaud C, Chave J.* Identification of amazonian trees with DNA barcodes. PLoS One 2009; 4: e7483

11 *Soininen EM, Valentini A, Coissac E, Miquel C, Gielly L, Brochmann C, Brysting AK, Sønstebø JH, Ims RA, Yoccoz NG, Taberlet P.* Analysing diet of small herbivores: the efficiency of DNA barcoding coupled with high-throughput pyrosequencing for deciphering the composition of complex plant mixtures. Front Zool 2009; 6: 16–24

12 *Valentini A, Miquel C, Nawaz MA, Bellemain EVA, Coissac E, Pompanon F, Gielly L, Cruaud C, Nascetti G, Wincker P, Swenson JE, Taberlet P.* New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the *trnL* approach. Mol Ecol Resour 2009; 9: 51–60

13 *Jaakola L, Suokas M, Häggman H.* Novel approaches based on DNA barcoding and high-resolution melting of amplicons for authenticity analyses of berry species. Food Chem 2010; 123: 494–500

14 *Yao H, Song JY, Ma XY, Liu C, Li Y, Xu HX, Han JP, Duan LS, Chen SL.* Identification of *Dendrobium* species by a candidate DNA barcode sequence: the chloroplast *psbA-trnH* intergenic region. Planta Med 2009; 75: 667–669

15 *Chen S, Yao H, Han J, Liu C, Song J, Shi L, Zhu Y, Ma X, Gao T, Pang X, Luo K, Li Y, Li X, Jia X, Lin Y, Leon C.* Validation of ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS One 2010; 5: e8613

16 *Asahina H, Shinozaki J, Masuda K, Morimitsu Y, Satake M.* Identification of medicinal *Dendrobium* species by phylogenetic analyses using *matK* and *rbcL* sequences. J Nat Med 2010; 64: 133–138

17 *Srirama R, Senthilkumar U, Sreejayan N, Ravikanth G, Gurumurthy BR, Shivanna MB, Sanjappa M, Ganeshaiah KN, Uma Shaanker R.* Assessing species admixtures in raw drug trade of *Phyllanthus*, a hepato-protective plant using molecular tools. J Ethnopharmacol 2010; 130: 208–215

18 *Natural Products Foundation.* New study shows dietary supplements industry contributes more than 60 billion to national economy. Available at http://www.naturalproductsinfo.org/index.php?src=news&srctype=detail&category=DSIB%20Releases&refno=181&view=DSI-B_Releases_Detail. Accessed April 18, 2016

19 *Barnes J.* Pharmacovigilance of herbal medicines. Drug Saf 2003; 26: 829–851

20 *Chen CH, Dickman KG, Moriya M, Zavadil J, Sidorenko VS, Edwards KL, Gnatenko DV, Wu L, Turesky RJ, Wu XR, Pu YS, Grollman AP.* Aristolochic acid-associated urothelial cancer in Taiwan. Proc Natl Acad Sci U S A 2012; 109: 8241–8246

21 *Debelle FD, Vanherweghem JL, Nortier JL.* Aristolochic acid nephropathy: a worldwide problem. Kidney Int 2008; 74: 158–169

22 *Vanherweghem JL, Tielemans C, Abramowicz D, Depierreux M, Vanhaelen-Fastre R, Vanhaelen M, Dratwa M, Richard C, Vandervelde D, Verbeelen D, Jadoul M.* Rapidly progressive interstitial renal fibrosis in young women: association with slimming regimen including Chinese herbs. Lancet 1993; 341: 387–391

23 *Hu SY.* The role of botany in Chinese medicinal material research: the case of eleuthero. In: Chang HM, Yeung HW, Tso W-W, Koo A, eds. Advances in Chinese medicinal materials research. Singapore: World Scientific Publishing; 1985: 28–33

24 *Upton R, Graff A, Jolliffe G, Länger R, Williamson E.* American Herbal Pharmacopoeia: Botanical Pharmacognosy – microscopic Characterization of botanical Medicines. Boca Raton: CRC Press; 2011

25 *Chen S, Pang X, Song J, Shi L, Yao H, Han J, Leon C.* A renaissance in herbal medicine identification: from morphology to DNA. Biotech Adv 2014; 32: 1237–1244

26 *De Boer HJ, Ichim MC, Newmaster SG.* DNA barcoding and pharmacovigilance of herbal medicines. Drug Saf 2015; 38: 611–620

27 *Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.* Basic local alignment search tool. J Mol Biol 1990; 215: 403–410

28 Barcode of Life Data System. Available at http://www.boldsystems.org/index.php/Public_SearchTerms. Accessed May 23, 2016

29 *Zemlak TS, Ward RD, Connell AD, Holmes BH, Hebert PDN.* DNA barcoding reveals overlooked marine fishes. Mol Ecol Resour 2009; 9: 237–242

30 *Techen N, Parveen I, Pan Z, Khan IA.* DNA barcoding of medicinal plant material for identification. Curr Opin Biotechnol 2014; 25: 103–110

31 *Murray M, Thompson WF.* Rapid isolation of high molecular weight plant DNA. Nucleic Acid Res 1980; 8: 4321–4325

32 *Dellaporta SL, Wood J, Hicks JB.* A plant DNA minipreparation: Version II. Plant Mol Biol Rep 1983; 1: 19–21

33 *Doyle JJ, Doyle JL.* A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 1987; 19: 11–15

34 *Lodhi MA, Ye GN, Weeden NF, Reisch BI.* A simple and efficient method DNA extraction from grapevine cultivars and *Vitis* species. Plant Mol Biol Rep 1994; 12: 6–13

35 *Aljanabi SM, Martinez I.* Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acids Res 1997; 25: 4692–4693

36 *Alaey M, Naderi R, Vezvaei A, Khalighi A, Salami A.* Comparing study between four different methods of genomic DNA extraction from *Cyclamen persicum* Mill. Int J Agric Biol 2005; 7: 882–884

37 *Akkurt M.* Comparison between modified DNA extraction protocols and commercial isolation kits in grapevine (*Vitis vinifera* L.). Genet Mol Res 2012; 11: 2343–2351

38 *Xin Z, Chen J.* A high throughput DNA extraction method with high yield and quality. Plant Methods 2012; 8: 26–32

39 *Sika KC, Kefela T, Adoukonou-Sagbadja H, Ahoton L, Saidou A, Baba-Moussa L, Baptiste LJ, Korconi SO, Gachomo EW.* A simple and efficient

genomic DNA extraction protocol for large scale genetic analyses of plant biological systems. Plant Gene 2015; 2: 43–45

40 *Porebski S, Bailey LG, Baum BR.* Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Mol Biol Report 1997; 15: 8–15

41 *Barnwell P, Blanchard AN, Bryant JA, Smirnoff N, Weir AF.* Isolation of DNA from the highly mucilaginous succulent plant *Sedum telephium.* Plant Mol Biol Rep 1998; 16: 133–138

42 *Khanuja SPS, Shasany AK, Darokar MP, Kumar S.* Rapid isolation of DNA from dry and fresh samples of plants producing large amounts of secondary metabolites and essential oils. Plant Mol Biol Rep 1999; 17: 1–7

43 *Sahu SK, Thangaraj M, Kathiresan K.* DNA extraction protocol for plants with high levels of secondary metabolites and polysaccharides without using liquid nitrogen and phenol. ISRN Mol Biol 2012; 2012: 1–6

44 *Fang G, Hammar S, Grumet R.* A quick and inexpensive method for removing polysaccharides from plant genomic DNA. Biotechniques 1992; 13: 52–54

45 *Peterson DG, Boehm KS, Stack SM.* Isolation of milligram quantities of nuclear DNA from tomato (*Lycopersicon esculentum*), a plant containing high levels of polyphenolic compounds. Plant Mol Biol Rep 1997; 15: 148–153

46 *Cavallari MM, Siqueria MVBM, Val TM, Pavanelli JC, Monteiro M, Grando C, Pinheiro JB, Zucchi MI, Gimenes MA.* A modified acidic approach for DNA extraction from plant species containing high levels of secondary metabolites. Gen Mol Res 2012; 13: 6497–6502

47 *Han J, Pang X, Liao B, Yao H, Song J, Chen S.* An authenticity survey of herbal medicines from markets in China using DNA barcoding. Sci Rep 2016; 6: 18723

48 *Katterman FRH, Shattuck VI.* An effective method of DNA isolation from the mature leaves of *Gossypium* species that contain large amounts of phenolic terpenoids and tannins. Prep Biochem 1983; 13: 347–359

49 *Wallace L, Boilard SMAL, Eagle SHC, Spall JL, Shokralla S, Hajibabaei M.* DNA barcodes for everyday life: routine authentication of natural health products. Food Res Int 2012; 49: 446–452

50 *Costa J, Amaral JS, Fernandes TJR, Batista A, Oliveira MB, Mafra I.* DNA extraction from plant food supplements: Influence of different pharmaceutical excipients. Mol Cell Probes 2015; 29: 473–478

51 *Little DP.* Authentication of *Ginkgo biloba* herbal dietary supplements using DNA barcoding. Genome 2014; 57: 513–516

52 *Novak J, Grausgruber-Gröger S, Lukas B.* DNA-based authentication of plant extracts. Food Res Int 2007; 40: 388–392

53 *Harnly J, Chen P, Sun J, Huang H, Colson KL, Yuk J, McCoy JA, Reynaud DT, Harrington PB, Fletcher EJ.* Comparison of flow injection MS, NMR and DNA sequencing: methods for identification and authentication of black cohosh (*Actaea racemosa*). Planta Med 2016; 82: 250–262

54 *Newmaster SG, Fazekas AJ, Ragupathy S.* DNA barcoding in land plants: evaluation of *rbcL* in a multigene tiered approach. Botany 2006; 84: 335–341

55 *Kress WJ, Erickson DL.* A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH–psbA* spacer region. PLoS One 2007; 2: e508

56 *CBOL Plant Working Group.* A DNA barcode for land plants. Proc Natl Acad Sci U S A 2009; 106: 12794–12797

57 *Hollingsworth PM, Graham SW, Little DP.* Choosing and using a plant DNA barcode. PLoS One 2011; 6: e19254

58 *Burgess KS, Fazekas AJ, Kesanakurti PR, Graham SW, Husband BC, Newmaster SG, Percy DM, Hajibabaei M, Barrett SCH.* Discriminating plant species in a local temperate flora using the *rbcL + matK* DNA barcode. Methods Ecol Evol 2011; 2: 333–340

59 *China Plant BOL Group.* Comparative analysis of a large dataset indicates that ITS should be incorporated into the core barcode for seed plants. Proc Natl Acad Sci U S A 2011; 108: 19641–19646

60 *Gao T, Yao H, Song J, Liu C, Zhu Y, Ma X, Pand X, Xu H, Chen S.* Identification of medicinal plants in the family Fabaceae using a potential DNA barcode ITS2. J Ethnopharmacol 2010; 130: 116–121

61 *Sun ZY, Chen SL.* Identification of cortex herbs using the DNA barcode nrITS2. J Nat Med 2013; 67: 296–302

62 *Pang XH, Shi LC, Song JY, Chen XC, Chen SL.* Use of the potential DNA barcode ITS2 to identify herbal materials. J Nat Med 2013; 67: 571–575

63 *Zhu X, Zhang Y, Liu X, Hou D, Gao T.* Authentication of commercial processed *Glehniae Radix* (Beishashen) by DNA barcodes. Chin Med 2015; 10: 35–43

64 *Rodriguez R, White J Jr., Arnold A, Redman R.* Fungal endophytes: diversity and functional roles. New Phyto 2009; 182: 314–330

65 *Ivanova NV, Kuzmina ML, Braukmann TWA, Borisenko AV, Zakharov EV.* Authentication of herbal supplements using next-generation sequencing. PLoS One 2016; 11: e0156426

66 *Cheng T, Xu C, Lei L, Li C, Zhang Y, Zhou S.* Barcoding the kingdom Plantae: new PCR primers for ITS regions of plants with improved universality and specificity. Mol Eco Res 2016; 16: 138–149

67 *White TJ, Bruns T, Lee S, Taylor J.* Amplification and direct Sequencing of fungal ribosomal RNA Genes for Phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, editors. PCR Protocols: a Guide to Methods and Applications. San Diego: Academic Press; 1990: 315–322

68 *Tai V, Lindstrom SC, Saunders GW.* Phylogeny of the Dumontiaceae (Gigartinales, Rhodophyta) and associated families based on SSU rDNA and internal transcribed spacer sequence data. J Phycology 2001; 37: 184–196

69 *Saunders GW, Kucera H.* An evaluation of *rbcL, tufA*, UPA, LSU and ITS as DNA barcode markers for the marine green macroalgae. Cryptogam Algol 2010; 31: 487–528

70 *Li DZ, Gao LM, Li HT, Wang H, Ge XJ, Liu JQ, Chen ZD, Zhou SL, Chen SL, Yang JB, Fu CX, Zeng CX, Yan HF, Zhu YJ, Sun YS, Chen SY, Zhao L, Wang K, Yang T, Duan GW.* Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. Proc Natl Acad Sci U S A 2011; 108: 19641–19646

71 *Little DP, Jeanson ML.* DNA barcode authentication of saw palmetto herbal dietary supplements. Sci Rep 2013; 3: 3518

72 *Little DP.* Confirming species identity of herbal dietary supplements, an example from devil's claw. Adulteration and Fraud in Food Ingredients and Dietary Supplements Workshop, Rockville, MD; 2015

73 *Mullis K.* The unusual origin of the polymerase chain reaction. Sci Am 1990; 262: 56–61, 64–65

74 *Kress WJ, Wudrack KJ, Zimmer EA, Weigt LA, Janzen DH.* Use of DNA barcodes to identify flowering plants. Proc Natl Acad Sci U S A 2005; 102: 8369–8374

75 *Lahaye R, van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, Maurin O, Duthoit S, Barraclough TG, Savolainen V.* DNA barcoding the floras of biodiversity hotspots. Proc Natl Acad Sci U S A 2005; 105: 2923–2928

76 *Pawluczyk M, Weiss J, Links MG, Aranguren ME, Wilkinson MD, Egea-Cortines M.* Quantitative evaluation of bias in PCR amplification and next-generation sequencing derived from metabarcoding samples. Anal Bioanal Chem 2015; 407: 1841–1848

77 *Soares S, Amaral JS, Oliveira MB, Mafra I.* Improving DNA isolation from honey for the botanical origin identification. Food Control 2015; 48: 130–136

78 *Costa J, Melo VS, Santos CG, Oliveira MB, Mafra I.* Tracing tree nut allergens in chocolate: A comparison of DNA extraction protocols. Food Chem 2015; 187: 469–476

79 *Green SJ, Venkataraman R, Naqib A.* Deconstructing the polymerase chain reaction: understanding and correcting bias associated with primer degeneracies and primer-template mismatches. PLoS One 2015; 10: e0128122

80 *Vogelstein B, Kinzler KW.* Digital PCR. Proc Natl Acad Sci U S A 1999; 96: 9236–9241

81 *Sanger F, Nicklen S, Coulson AR.* DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 1977; 74: 5463–5467

82 *Polz MF, Cavanaugh CM.* Bias in template-to-product ratios in multi template PCR. Appl Environ Microbiol 1998; 64: 3724–3730

83 *Shokralla S, Gibson JF, Nikbakht H, Janzen DH, Hallwachs W, Hajibabaei M.* Next-generation DNA barcoding: using next-generation sequencing to enhance and accelerate DNA barcode capture from single specimens. Mol Ecol Resour 2014; 14: 892–901

84 *Aird D, Ross MG, Chen WS, Danielsson M, Fennell T, Russ C, Jaffe DB, Nusbaum C, Gnirke A.* Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. Genome Biology 2011; 12: R18

85 *Kircher M, Kelso J.* High-throughput DNA sequencing-concepts and limitations. Bioessays 2010; 32: 524–536

86 *Moore MJ, Bell CD, Soltis PS, Soltis DE.* Using plastid genome scale data to resolve enigmatic relationships among basal angiosperms. Proc Natl Acad Sci U S A 2007; 104: 19363–19368

87 *Parks M, Cronn R, Liston A.* Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. BMC Biol 2009; 7: 84

88 *Sucher NJ, Hennell JR, Carles MC.* DNA fingerprinting, DNA barcoding, and next generation sequencing technology in plants. Methods Mol Biol 2012; 862: 13–22

89 *Sarwat M, Yamdagni MM.* DNA barcoding, microarrays and next generation sequencing: recent tools for genetic diversity estimation and authentication of medicinal plants. Crit Rev Biotechnol 2016; 36: 191–203

90 *Shokralla S, Spall JL, Gibson JF, Hajibabaei M.* Next-generation sequencing technologies for environmental DNA research. Mol Ecol 2012; 21: 1794–1805

91 *Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E.* Towards next generation biodiversity assessment using DNA metabarcoding. Mol Ecol 2012; 21: 2045–2050

92 *Coghlan M, Haile J, Houston J, Murray D, White N, Moolhuijzen P, Bellgard MI, Bunce M.* Deep sequencing of plant and animal DNA contained within traditional Chinese medicines reveals legality issues and health safety concerns. PLoS Genet 2012; 8: e1002657

93 *Zuo Y, Chen Z, Kondo K, Funamoto T, Wen J, Zhou S.* DNA barcoding of Panax species. Planta Med 2011; 77: 182–187

94 *Stoeckle MY, Gamble CC, Kirpekar R, Young G, Ahmed S, Little DP.* Commercial teas highlight plant DNA barcode identification successes and obstacles. Sci Rep 2011; 1: 1–7

95 *Baker DA, Stevenson DW, Little DP.* DNA barcode identification of black cohosh herbal dietary supplements. J AOAC Int 2012; 95: 1023–1034

96 *Vassou SL, Kusuma G, Madasamy P.* DNA barcoding for species identification from dried and powdered plant parts: a case study with authentication of the raw drug market samples of *Sida cordifolia*. Gene 2015; 559: 86–93

97 *Seethapathy GS, Ganesh D, Kumar JUS, Senthilkumar U, Newmaster SG, Ragupathy S, Uma Shaanker R, Ravikanth G.* Assessing product adulteration in natural health products for laxative yielding plants, *Cassia, Senna,* and *Chamaecrista,* in Southern India using DNA barcoding. Int J Legal Med 2014; 129: 693–700

98 *Palhares RM, Drummond MG, Alves BS, Brasil F, Cosenza GP, Brandao MG, Oliveira G.* Medicinal plants recommended by the World Health Organization: DNA barcode identification associated with chemical analyses guarantees their quality. PLoS One 2015; 10: e0127866

99 *Newmaster SG, Grguric M, Shanmughanandhan D, Ramalingam S, Ragupathy S.* DNA barcoding detects contamination and substitution in North American herbal products. BMC Med 2013; 11: 222–234

100 *Gafner S, Blumenthal M, Reynaud DH, Foster S, Techen N.* ABC Review and Critique of the Research Article "DNA Barcoding Detects Contamination and Substitution in North American Herbal Products" by Newmaster et al. Available at: http://cms.herbalgram.org/heg/volume10/11November/DNAbarcodingReviewandCritique.html?t=1383684796. Accessed June 30, 2016

101 *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. Press release, February 3, 2015. Available at http://www.ag.ny.gov/press-release/ag-schneiderman-asks-major-retailers-halt-sales-certain-herbal-supplements-dna-tests. Accessed January 21, 2016

102 *O'Connor A.* New York attorney general targets supplements at major retailers. New York Times. February 3, 2015. Available at http://well.blogs.nytimes.com/2015/02/03/new-york-attorney-general-targets-supplements-at-major-retailers/. Accessed June 13, 2015

103 *Harbaugh Reynaud DT, Mischler BD, Neal-Kababick J, Brown PN.* The capabilities and limitations of DNA barcoding of dietary supplements. March 2015. Available at http://www.crnusa.org/NYAG/The-Capabilities-and-Limitations-of-DNA-Testing-FINAL-3–10-2015.pdf. Accessed March 16, 2015

104 *Coutinho Moraes DF, Still DW, Lum MR, Hirsch AM.* DNA-based authentication of botanicals and plant-derived dietary supplements: where have we been, and where are we going? Planta Med 2015; 81: 687–695

105 *AOAC International.* Appendix K: Guidelines for dietary supplements and botanicals. Available at: http://www.eoma.aoac.org/app_k.pdf. Accessed June 30, 2016

106 *Shukla AK, Mall M, Rai SK, Singh S, Nair P, Parashar G, Shasany AK, Singh SC, Joshi VK, Khanuja SPS.* A transcriptomic approach for exploring the molecular basis for dosha-balancing property-based classification of plants in Ayurveda. Mol Biol Rep 2013; 40: 3255–3262

107 *Mishra P, Kumar A, Nagireddy A, Mani DN, Shukla AK, Tiwari R, Sundaresan V.* DNA barcoding: an efficient tool to overcome authentication challenges in the herbal market. Plant Biotech J 2016; 14: 8–21