DNA Metabarcoding for Quality Control of Basil, Oregano, and Paprika

Ancuța Cristina Raclariu-Manolică[1,2]*, Jarl Andreas Anmarkrud[1], Marcin Kierczak[3], Nima Rafati[3], Birgitte Lisbeth Graae Thorbek[1], Audun Schrøder-Nielsen[1] and Hugo J. de Boer[1]

[1] Natural History Museum, University of Oslo, Oslo, Norway, [2] Stejarul Research Centre for Biological Sciences, National Institute of Research and Development for Biological Sciences, Piatra Neamt, Romania, [3] Researgene, Uppsala, Sweden

Herbs and spices are some of the most vulnerable products in terms of fraud and adulteration in the food sector. Although standard analytical methods are accurate for quality control of specific lead or marker compounds, they cannot accurately assess the entire species composition of many marketed products. Complementary analytical approaches are thus often used for comprehensive screening of herbs and spices. In this study we evaluate DNA metabarcoding for the identification and authentication of 62 products, containing basil, oregano, and paprika collected from different retailers and importers in Norway. Our results show varying degrees of discrepancy between the constituent species and those listed on the product labels, despite high product authenticity. We suggest the false positives result from the sensitivity of DNA metabarcoding and filtering thresholds should be integrated into protocols to reduce false positives. Our results highlight how integrating DNA metabarcoding into the toolbox of analytical methods for quality control of fresh and/or processed plant-based food can improve product quality.

Keywords: authentication, DNA metabarcoding, food control, food safety, high-throughput sequencing, molecular identification

INTRODUCTION

Herbs and spices have been used for food and beverage flavoring since the beginning of human history. Herbs generally refer to the leafy parts of a plant, and spices are derived from any other plant part, including the root, stem, bulb, bark or seeds. Herbs and spices may derive from the same plant and they are usually rich sources of phytochemicals. Some herbs and spices may be categorized as functional foods or nutraceuticals with health properties beyond basic nutrition, such as a reduction of the risks related to cardiovascular disease, diabetes, obesity, cancer, and Alzheimer's disease (Tapsell et al., 2006; Viuda-Martos et al., 2011).

Globalization and the search for curative properties are key drivers behind the notable growth in both the value and volume of trade in herbs and spices (Amcham and Trade USA, 2015). The global market of these commodities is projected to grow in the near future to US$6.5 billion, with Asia and Europe the largest consumer markets worldwide (CBI, 2018).

Herbs and spices products are however some of the most vulnerable segments within the food sector (Galvin-King et al., 2018; van Ruth et al., 2018). Apart from their popularity and long

historical use, there are increasing concerns over the product quality and safety (Galvin-King et al., 2018). There are several recent cases of adulteration involving the deliberate inclusion of substances whose presence is not legally declared and leaving out expensive ingredients (Galvin-King et al., 2018). Additionally food safety concerns can arise from issues related to chemical, microbial and physical hazards. Another major issue is the adulteration of herbs and spices. This is driven by the increasing level of consumption which exceeds the supply capacity as well as price increases, especially when complex supply chains are involved (Black et al., 2016; Galvin-King et al., 2018).

Herbs and spices as high-price commodities are among some of the most reported products involved in food fraud (Silvis et al., 2017; Galvin-King et al., 2018). The effects of a food fraud scandal can be detrimental for a business operator and result in varying degrees of prejudice and public distrust toward the industry sector in which it occurs (Aung and Chang, 2014). Some herbs and spices are especially well documented to be susceptible to substitution with unauthorized synthetic flavor and colors, including the use of the Sudan Red G dye, which is illegal and comes with genotoxic and carcinogenic effects, to color chili, curry and curcuma (Cornet et al., 2006; Galvin-King et al., 2020).

Generally a fast, reliable and comprehensive analytical approach is required to detect the authenticity and integrity of food items. Also, efficient traceability tools are necessary to minimize the production and distribution of unsafe or poor quality products in order to protect the consumer's health and confidence (Aung and Chang, 2014). However, many important issues influence the quality of the herbs and spices. These need to be carefully considered when deciding on the analytical method for quality control. Herbs and spices are often marketed as complex mixtures containing one or more species that have been through many processing steps. The resulting phytochemical diversity across these products usually complicates further investigations, posing serious challenges to the accuracy of the quality control. Thus, modern research in food science requires innovative approaches using advanced analytical tools for comprehensive screening and analysis to detect and prevent fraud in the industry (Galvin-King et al., 2018).

With DNA metabarcoding, or multispecies identification with extracellular and/or total DNA extracted from complex samples containing DNA of different origins (Taberlet et al., 2012) followed by high throughput sequencing (HTS), it is possible to analyze species composition in mixtures of DNA in a range of products. This includes environmental samples (Sogin et al., 2006; Epp et al., 2012; Nielsen and Wall, 2013) or marketed herbal medicines (Coghlan et al., 2015; Ivanova et al., 2016; Raclariu et al., 2017a,b, 2018b; Anthoons et al., 2020). DNA metabarcoding is a highly sensitive method, and *a priori* and *a posteriori* data can inform and aid interpretation of metabarcoding results. *A priori* and *a posteriori* data can help to classify positive identifications that are relevant to consumer safety, as well as contaminants that are not likely to be present in significant quantities. It will also help to interpret false negative detections and make assumptions on the possible cause of the overlooked species.

In this study we develop and test DNA metabarcoding for identifying plant species in marketed herbs and spices. We targeted two herbs, oregano *(Origanum vulgare* L., Lamiaceae), and basil *(Ocimum basilicum* L., Lamiaceae) and one spice, paprika *(Capsicum annuum* L., Solanaceae), purchased in Norway. These target plant species were selected due to their popularity and leading consumer preference. Despite the fact that Europe is a large producer of native herbs, such as basil and oregano, these are still imported in large quantities as European production fails to meet the growing market demand (CBI, 2018).

Our hypothesis was that DNA metabarcoding can detect plant ingredients in herbs and spices-based marketed products, but that blind use of these molecular identification methods can result in an underestimation of product quality. Specific research objectives were to (i) evaluate DNA metabarcoding for detecting substitution in marketed herbs and spices; (ii) assess the use of *a priori* and *a posteriori* data to enhance interpretation of DNA metabarcoding data; (iii) contribute to developing a complementary tool to secure the food supply chain against accidental substitution and/or deliberate adulteration.

MATERIALS AND METHODS

Samples Collection and Preparation
62 herb and spice samples were examined that included 23 samples of oregano, 21 of paprika, 17 of basil, and one mix of oregano and basil. Information on the samples including label information, but not the producer/importer name, lot number, expiration date, or any other information that could lead to the identification of that specific product can be found in **Supplementary File 1**: Information about the investigated herbs and spices products.

DNA Extraction
A 200 mg (dry weight) of each sample was transferred to a 2 ml microcentrifuge tube with two 3 mm tungsten-carbide beads (Qiagen) and ground to a fine powder on a Retsch mill MM 301 (Retsch GmbH, Haan Germany) (2X cycles for 45 s at 25 Hz). Total DNA was extracted from the homogenized contents using a modified CTAB extraction method as described by Doyle and Doyle (1990), and adapted by Raclariu et al. (2017b). The final elution volume was 100 μl.

DNA Libraries Preparation and Sequencing
The amplicon library and high-throughput sequencing were performed as previously described (Raclariu et al., 2017b) with slight modifications. All amplicon libraries were prepared in three technical replicates on 96-well PCR amplification plates. On each plate we also included negative controls (also in triplicate), including extraction blanks (extraction negative controls, ENC) and PCR controls (PCR negative controls, PNC). The amplicon libraries were prepared using fusion

primers for the nuclear ribosomal target sequences, internal transcribed spacer nrITS2. Ion Torrent fusion primers were based on 5.8I2 and 26SE as described by Sun et al. (1994), ITS2_5.8I_F 5′-GCCTGGGCGTCACGC and_26SE_R 5′-CCCGGTTCGCTCGCCGTTAC. The primer match/mismatch was checked before the wet lab work for the target species. The forward primer was labeled with a unique 10 bp multiplex identifier (MID) tag and the reverse primer with a uniform truncated P1 (trP1) tag. Thermal cycling was carried out in 25 μl final reaction volumes, and each reaction contained 1X Q5 hot start high fidelity mastermix (New England Biolabs Inc., United Kingdom), 0.5 μM of each primer (Biolegio, Netherlands), and 1 μl of template DNA. The following thermocycling protocol was used: 30 s of initial denaturation at 98°C, followed by 30 cycles of denaturation at 98°C for 10 s, annealing at 71°C for 30 s, and elongation at 72°C for 30 s, followed by a final elongation step at 72°C for 2 min. Agarose gel electrophoresis was used to inspect amplification of all PCR reactions. Based on the relative intensity of the PCR fragments on the gels, we merged uniform amounts of each amplicon in a separate microcentrifuge tube using a Biomek 4000 Laboratory Automation Workstation (Beckman Coulter, United States) liquid handling system. Each normalized library was size selected using BluePippin (Sage Science, United States) and BDF1510 cassettes, targeting the desired fragment length within the range of 300-550 bp. The library was then purified with 0.8X AMpure XP beads (Beckman Coulter, United States) following the manufacturer's instructions. The total concentration of the purified pooled amplicon library stock was measured on both Qubit (Invitrogen, United States) with DNA High Sensitivity Kit (Thermo Fischer Scientific, United States) and on Fragment Analyzer™ (Advanced Analytical Technologies, Inc., United States) using the DNF-488 High Sensitivity Genomic DNA Analysis Kit, in order to identify the optimum concentration range for the template preparation. An Ion Chef (Life Technologies (LT), Thermo Fisher Scientific, United States) was used to prepare pooled Ion AmpliSeq libraries (LT) for emulsion PCR and to load the sequencing chips. The sequencing was performed on Ion 318 v2 chips and sequenced on an Ion Torrent PGM (LT). Sequencing data was analyzed and demultiplexed into FASTQ files per sample using Torrent Suite version 5.0.4 (LT).

Bioinformatics Data Analysis

MultiQC (Ewels et al., 2016) was used to perform quality control of the raw FASTQ files delivered by the sequencing facility. Using the Flexbar tool the reads were demultiplexed first by library barcode and then by the nrITS2 primer sequences used for amplification (Dodt et al., 2012; Roehr et al., 2017). The following selection criteria were used to select the remaining reads: minimal read length of 200 bp, base-quality threshold set to 22, and alignment mode LEFT leaving the right side of the sequence after primer removal. We set the barcode error rate to 2 mismatches per 10 nucleotides. Cutadapt was used to trim the reads to a maximum of 360 bp and merge the reads that were amplified by both the forward and reverse primers into a single file (Martin, 2011). Dada2 was used for the following analysis (Callahan

et al., 2016, 2017) taking the amplicon sequence variants (ASV) approach as described in Callahan et al. (2017). The 1.6.0 version of the dada2 R package and R 3.4.3 was used (Kite-eating Tree). In the initial step of the ASV pipeline, we estimated read quality using 5 randomly (seed 77) selected samples (products). After estimating the read quality, we trimmed the 15 left most nucleotides and removed any truncated reads if the base quality dropped to 2. Reads beyond 320bp were discarded. The truncated reads were used to train for base-call error pattern in our dataset. The learned patterns were then used to correct the trimmed and dereplicated reads. Following this, sample inference was performed using minimal Hamming distance within a single cluster set to 3. We set the minimal number of reads to form a cluster to 5. From the resulting samples a set S of the most abundant sequences was created. Every sequence was checked for whether it was a chimera of any two sequences from S, and all chimeric sequences were removed. The species identification was performed by aligning (BLAST) all inferred amplicon sequence variants to the ITS2db (Ankenbrand et al., 2015) database (ver. V, access: June 03, 2018) of ITS2 sequences for green plants (Viridiplantae), using 97% sequence identity threshold. For each ASV, we report the top 4 BLAST hits.

RESULTS

The success rate in identifying ASVs varied between the samples (see **Supplementary File 3**: Species identification report and BLAST queries). 8 out of 62 samples (KR18-09, KR18-16, KR18-33, KR18-46, KR18-52, KR18-53, KR18-54 and KR18-60) did not pass our quality and trimming quality criteria, and they yielded no taxonomic identifications after the analysis. The extraction negative controls (ENC) and the PCR negative controls (PNC) did not yield any taxonomic identification.

Paprika

13 out of 21 samples with paprika listed as the only ingredient listed on the label yielded taxonomic identifications (See **Figure 1**, **Table 1**, and **Supplementary File 4**: DNA metabarcoding based identification of the products). The low success rate of DNA metabarcoding for paprika identification could be the result of DNA degradation due to excessive heating in the drying and/or smoking process of the peppers. The eight non-identified samples were excluded from further analyses. Of the identified 13 total samples, ten included and three lacked paprika (KR18-07; KR18-23; and KR18-58). However, the samples lacking paprika did include DNA from wind-pollinated plant species (KR18-07; KR18-23) or DNA from wind-spread plant species (KR18-58). Five samples yielded identifications of one to several wind-pollinated species. Three samples yielded identifications for one to several wind-spread species. Three samples yielded identifications for other plant species than those listed on the product label, or reasonably expected through contamination from wind-pollinated or wind-spread species. The detection of *Zea mays* in KR18-03 suggested the presence of corn flour, a common gluten-free filler found in several DNA barcoding studies of herbal medicines.

FIGURE 1 | Product fidelity rates per target species using DNA metabarcoding based identification.
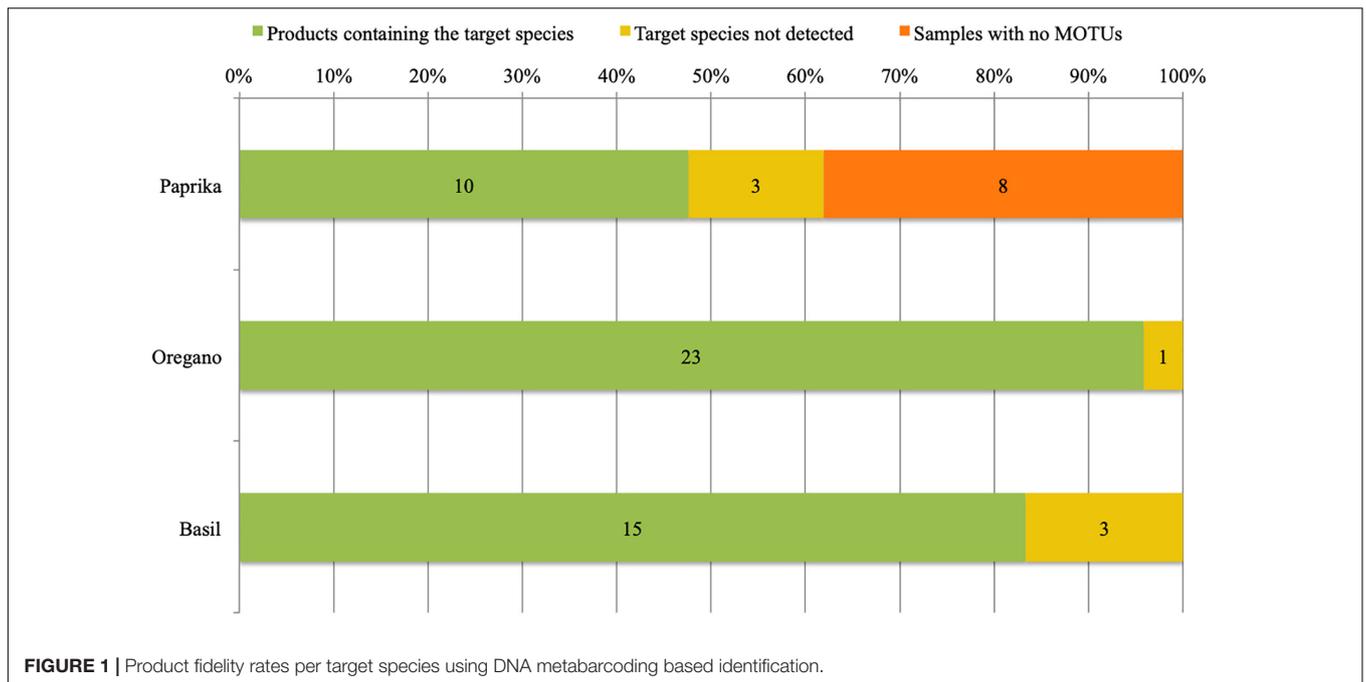
Oregano

All 23 samples of oregano, and the single sample containing a mixture of oregano and basil, could be identified based upon the obtained sequencing results (See **Figure 1**, **Table 2**, and **Supplementary File 4**: DNA metabarcoding based identification of the products). In only a single sample (KR18-01) the target *Origanum vulgare* was not detected. However, the identification of *Thymus* sp. in this product does raise concerns about its quality. 10 samples yielded identifications for one to several wind-pollinated species. 15 samples yielded identification for one to several wind-spread species. 19 samples yielded identifications for other plant species than those listed on the product label, or reasonably expected through contamination from wind-pollinated or wind-spread species. Notable identifications include *Thymus* sp. (KR18-01), *Mentha longifolia* (KR18-44), *Thymus* sp. and *Polygonum* sp. (KR18-47), and *Veronica* sp. (KR18-62). The identification of unexpected species in these samples suggests contamination in the production chain, and likely during harvesting (KR18-44; KR18-47; KR18-62), or processing (*Thymus* sp. in KR18-01 and KR18-47).

Basil

All 17 samples of basil, and the one sample containing a mixture of oregano and basil, yielded DNA based identifications (See **Figure 1**, **Table 3**, and **Supplementary File 4**. DNA metabarcoding results for product identification). Three samples (KR18-32, KR18-51 and KR18-59) did not yield DNA identifications for *Ocimum basilicum*. Again, the identification of other plant species raised concerns about the quality of this product. 14 samples yielded identifications of one to several wind-pollinated species. One sample yielded identification of the wind-spread species. All 17 samples yielded identifications

of other plant species than those listed on the product label, or reasonably expected through contamination from wind-pollinated or wind-spread species. Notable identifications include *Origanum vulgare* (KR18-25; KR18-30; KR18-51, KR18-59), *Physalis* sp. (KR18-25), *Ipomoea* sp. (KR18-43; KR18-49).

DISCUSSION

Overall, DNA metabarcoding detected considerable inconsistencies between the identified species and those listed on the product labels. Our results indicate that only five products (four paprika and one oregano) come with correct label information, and contained precisely the one listed plant-based ingredient. For 47 samples (10 paprika, 22 oregano, 14 basil, and one oregano-basil), the target plant species as well as additional other plant species were detected. For seven samples (three paprika, one oregano, and three basil) the target plant species was not identified. Eight samples (all labeled as containing paprika) did not yield useful information for the taxonomic identification after applying the quality filtering criteria. In summary 54 out of 62 samples contained other species than the ones listed on the product.

The discrepancies between the species detected using DNA metabarcoding and those listed on the product labels require a careful consideration of possible contamination. In this study we used *a priori* information from the product labels to establish a hypothesis of the target species for each product, as well as a putative origin of the raw plant material and cultivation conditions. The origin and cultivation conditions can help us narrow down identifications from the sequencing data. *A posteriori* data in this sense are the putative ASV identifications, i.e., the species detected in the products. Information on the

geographic range, growth habit or cultivation status of the detected species can aid in evaluating the results. For instance, in this study, we detect numerous wind-pollinated and wind-spread species, and their presence could be the result of (insignificant) trace contamination. However, when interpreting these results, we should also bear in mind that DNA metabarcoding is a highly sensitive detection method, and even trace amounts of pollen or plant dust left in production equipment can be detected. It is therefore important that herbal product authentication using DNA metabarcoding focuses on the presence of a target species. However, this does require a case-based analysis that considers the experimental steps (e.g., sample preparation, library preparation, and HTS) and post-bioinformatics analysis that may yield false positive or false negative results (Robasky et al., 2014; Ficetola et al., 2015). In this study, technical replicates were used to limit the effects of false positive and false negative results (Robasky et al., 2014; Ficetola et al., 2015). In addition, efforts were made to overcome the amount of sequencing errors associated with the Ion Torrent sequencing platform (Loman et al., 2012). Such sequencing errors may lead to the formation of false ASVs. Hence, we included three technical replicates per sample and applied strict filtering and trimming thresholds for base call length and quality. Speranskaya et al. (2018) performed a comparative study of the two most widely used sequencing platforms, Illumina and Ion Torrent, to assess the composition of herbal teas. The two methods were found to be qualitatively and quantitatively consistent, with a certain level of variation between runs on the same platform which were more likely due to the stochastic dynamics of PCR, or other reactions during the library preparation. The large number of non-listed plant species detected is therefore likely a result of contamination and amplification bias, i.e., PCR chimeras, sequencing errors, or false-positive taxonomic identifications due to error-prone barcode sequences reference databases (Ficetola et al., 2010; Robasky et al., 2014; Pawluczyk et al., 2015). Additionally metabarcoding does not always resolve closely related plant species. This can lead to incorrect identifications if the database used does not contain all possible species. For example, in this study it is very unlikely that the products labeled as containing basil contain *Ocimum americanum,* even though this is actually the result of our data analysis. The presence of this closely related species in our results is more likely due the fact that *Ocimum basilicum* cannot

TABLE 1 | DNA metabarcoding based authentication of the products labeled as containing paprika.

Sample ID	Paprika	Unlabeled ingredients		
	# of replicates detected	Plant species	# of replicates detected	ASV identifier
KR18-03	2	*Zea mays*	3	PTU_ 034
KR18-04	3	–	–	–
KR18-07	0	*Phragmites australis*	2	PTU_305
		Phragmites japonicus	2	PTU_305
KR18-13	2	*Molinia caerulea*	2	PTU_354
		Phragmites australis	2	PTU_354
		Phragmites japonicus	2	PTU_354
KR18-17	2	–	–	–
KR18-19	2	–	–	–
KR18-23	0	*Amaranthus blitoides*	3	PTU_466
		Amaranthus capensis	3	PTU_466
		Amaranthus tuberculatus	3	PTU_466
KR18-26	2	*Echinochloa colona*	2	PTU_072
		Echinochloa crus-galli	2	PTU_072
KR18-28	2	*Galinsoga parviflora*	3	PTU_209
		Galinsoga quadriradiata	3	PTU_209
		Sabazia sarmentosa	3	PTU_209
KR18-34	2	*Convolvulus arvensis*	3	PTU_010
		Galinsoga parviflora	3	PTU_209
		Galinsoga quadriradiata	3	PTU_209
		Sabazia sarmentosa	3	PTU_209
		Solanum lycopersicum	3	PTU_115
		Solanum pennellii	3	PTU_115
KR18-38	2	*Origanum vulgare*	3	PTU_036
KR18-45	3	–	–	–
KR18-58	0	*Galinsoga parviflora*	2	PTU_209
		Galinsoga quadriradiata	2	PTU_209
		Sabazia sarmentosa	2	PTU_209

TABLE 2 | DNA metabarcoding based authentication of the products labeled as containing oregano.

Sample ID	Oregano	Unlabeled ingredients		
	# of replicates detected	Plant species	# of replicates detected	ASV identifier
KR18-01	0	*Thymus altaicus*	3	PTU_265
		Thymus marschallianus	3	PTU_265
		Thymus mongolicus	3	PTU_265
		Thymus pulegioides	3	PTU_265
KR18-05	3	*Arnebia euchroma*	3	PTU_159
		Tragopogon porrifolius	3	PTU_159
		Tragopogon pratensis	3	PTU_159
		Convolvulus arvensis	3	PTU_166
		Iris ruthenica	3	PTU_432
		Lactuca saligna	3	PTU_432
		Lactuca sativa	3	PTU_432
		Silene vulgaris	3	PTU_496
		Vaccaria hispanica	3	PTU_496
		Linaria genistifolia	2	PTU_552
		Linaria michauxii	2	PTU_552
		Linaria salangensis	2	PTU_552
		Linaria vulgaris	2	PTU_552
		Silene dichotoma	2	PTU_284
		Silene fabaria	2	PTU_284
KR18-06	3	*Crepis foetida*	3	PTU_155
		Linaria genistifolia	3	PTU_252
		Linaria japonica	3	PTU_252
		Linaria michauxii	3	PTU_252
		Linaria vulgaris	3	PTU_252
		Convolvulus arvensis	2	PTU_257
		Geranium rotundifolium	2	PTU_616
		Linaria salangensis	2	PTU_361
KR18-10	3	*Convolvulus arvensis*	2	PTU_362
		Iris ruthenica	2	PTU_062
		Lactuca saligna	2	PTU_062
		Lactuca sativa	2	PTU_062
KR18-12	3	*Arnebia euchroma*	3	PTU_188
		Convolvulus arvensis	3	PTU_198
		Crepis foetida	3	PTU_155
		Tragopogon porrifolius	3	PTU_188
		Tragopogon pratensis	3	PTU_188
KR18-14	3	*Olea europaea*	3	PTU_293
KR18-15	3	–	–	–
KR18-21	3	*Convolvulus arvensis*	3	PTU_098
		Conyza bonariensis	3	PTU_110
		Erigeron rosulatus	3	PTU_110
		Echinochloa colona	3	PTU_123
		Echinochloa crus-galli	3	PTU_123
		Echinochloa oryzicola	3	PTU_123
		Polygonum arenastrum	3	PTU_160
KR18-24	3	*Convolvulus arvensis*	2	PTU_218
KR18-27	3	*Convolvulus arvensis*	3	PTU_267
		Arnebia euchroma	2	PTU_270
		Phoebe puwenensis	2	PTU_101
		Tragopogon dubius	2	PTU_210
		Tragopogon pratensis	2	PTU_210

(Continued)

TABLE 2 | Continued

Sample ID	Oregano	Unlabeled ingredients		
	# of replicates detected	Plant species	# of replicates detected	ASV identifier
KR18-29	3	*Bidens andicola*	3	PTU_228
		Bidens aurea	3	PTU_228
		Bidens alba	2	PTU_220
		Bidens cronquistii	2	PTU_220
		Bidens pilosa	2	PTU_220
KR18-31	2	*Convolvulus arvensis*	3	PTU_317
KR18-35`	3	*Convolvulus arvensis*	3	PTU_328
		Echinochloa colona	3	PTU_049
		Echinochloa crus-galli	3	PTU_049
		Echinochloa oryzicola	3	PTU_049
KR18-36	0	*Bidens andicola*	3	PTU_087
		Bidens aurea	3	PTU_087
		Bromus erectus	2	PTU_095
		Bromus sterilis	2	PTU_095
		Bromus tectorum	2	PTU_095
		Taraxacum amplum	2	PTU_322
		Taraxacum officinale	2	PTU_322
		Taraxacum sp. CF-2016	2	PTU_322
KR18-41	2	*Cirsium acaule*	3	PTU_538
		Cirsium arvense	3	PTU_538
		Cirsium dissectum	3	PTU_538
		Cirsium tuberosum	3	PTU_538
		Convolvulus arvensis	3	PTU_112
		Melica ciliata	2	PTU_286
		Melica nutans	2	PTU_286
		Melica picta	2	PTU_286
KR18-42	3	*Knautia degenii*	3	PTU_349
		Knautia integrifolia	3	PTU_349
KR18-44	3	*Amaranthus blitoides*	3	PTU_149
		Amaranthus capensis	3	PTU_149
		Amaranthus tuberculatus	3	PTU_149
		Amaranthus powellii	3	PTU_117
		Amaranthus retroflexus	3	PTU_117
		Convolvulus arvensis	3	PTU_407
		Rhaponticum repens	3	PTU_048
		Ziziphora clinopodioides	2	PTU_048
		Arnebia euchroma	2	PTU_130
		Chenopodium album	2	PTU_192
		Iris ruthenica	2	PTU_062
		Lactuca saligna	2	PTU_062
		Lactuca sativa	2	PTU_062
		Mentha longifolia	2	PTU_235
		Tragopogon lainzii	2	PTU_280
		Tragopogon porrifolius	2	PTU_291
		Tragopogon pratensis	2	PTU_291
KR18-47	3	*Convolvulus arvensis*	3	PTU_457
		Polygonum arenastrum	3	PTU_170
		Polygonum aviculare	3	PTU_170
		Thymus altaicus	3	PTU_265
		Thymus marschallianus	3	PTU_265

(Continued)

TABLE 2 | Continued

Sample ID	Oregano	Unlabeled ingredients		
	# of replicates detected	Plant species	# of replicates detected	ASV identifier
		Thymus mongolicus	3	PTU_265
		Thymus pulegioides	3	PTU_265
		Polygonum rurivagum	2	PTU_431
KR18-48	3	*Dactylis glomerata*	2	PTU_334
		Petrorhagia nanteuilii	2	PTU_499
KR18-50	3	*Bromus erectus*	3	PTU_365
		Bromus sterilis	3	PTU_365
		Bromus tectorum	3	PTU_365
		Senecio vulgaris	3	PTU_194
		Erodium cicutarium	2	PTU_510
		Erodium lebelii	2	PTU_510
		Erodium moschatum	2	PTU_510
KR18-55	3	*Filago pyramidata*	3	PTU_352
		Gnaphalium supinum	3	PTU_352
		Logfia gallica	3	PTU_352
		Logfia minima	3	PTU_352
KR18-57	3	*Amaranthus blitoides*	3	PTU_149
		Amaranthus capensis	3	PTU_149
		Amaranthus tuberculatus	3	PTU_149
		Xanthium sibiricum	3	PTU_223
		Xanthium strumarium	3	PTU_223
		Arnebia euchroma	2	PTU_231
		Convolvulus arvensis	2	PTU_461
		Linaria genistifolia	2	PTU_330
		Linaria michauxii	2	PTU_330
		Linaria salangensis	2	PTU_330
		Linaria vulgaris	2	PTU_330
		Tragopogon porrifolius	2	PTU_501
		Tragopogon pratensis	2	PTU_501
KR18-61	3	*Alopecurus myosuroides*	3	PTU_282
		Poa trivialis	3	PTU_282
		Arbutus canariensis	3	PTU_292
		Arnebia euchroma	3	PTU_319
		Tragopogon porrifolius	3	PTU_319
		Tragopogon pratensis	3	PTU_319
		Convolvulus arvensis	3	PTU_481
		Iris ruthenica	3	PTU_062
		Lactuca saligna	3	PTU_062
		Lactuca sativa	3	PTU_062
		Scorzonera cana	3	PTU_393
		Scorzonera laciniata	3	PTU_393
		Salvia spinosa	2	PTU_358
KR18-62	3	*Bromus erectus*	3	PTU_392
		Bromus sterilis	3	PTU_392
		Bromus tectorum	3	PTU_392
		Veronica agrestis	3	PTU_667
		Veronica persica	3	PTU_667
		Bidens andicola	2	PTU_232
		Bidens aurea	2	PTU_232

TABLE 3 | DNA metabarcoding based authentication of the products labeled as containing basil.

Sample ID	Basil	Unlabeled ingredients		
	# of replicates detected	Plant species	# of replicates detected	ASV identifier
KR18-02	3	*Amaranthus hybridus*	3	PTU_083
		Amaranthus palmeri	3	PTU_083
		Amaranthus retroflexus	3	PTU_083
		Damrongia orientalis	3	PTU_253
KR18-08	2	*Amaranthus viridis*	3	PTU_023
		Convolvulus arvensis	3	PTU_595
		Echinochloa colona	3	PTU_409
		Echinochloa frumentacea	3	PTU_409
		Hymenachne grumosa	3	PTU_047
		Portulaca oleracea	3	PTU_051
		Dactyloctenium aegyptium	2	PTU_059
		Damrongia orientalis	2	PTU_059
KR18-11	3	*Amaranthus viridis*	3	PTU_147
		Damrongia orientalis	3	PTU_263
		Amaranthus hybridus	2	PTU_029
KR18-18	3	*Cynodon dactylon*	3	PTU_035
		Dactyloctenium aegyptium	3	PTU_187
		Damrongia orientalis	3	PTU_344
KR18-20	3	*Amaranthus viridis*	3	PTU_147
		Damrongia orientalis	3	PTU_015
		Amaranthus hybridus	2	PTU_027
KR18-22	3	*Damrongia orientalis*	3	PTU_447
KR18-25	0	*Amaranthus viridis*	3	PTU_023
		Damrongia orientalis	3	PTU_500
		Echinochloa colona	3	PTU_409
		Echinochloa frumentacea	3	PTU_409
		Hymenachne grumosa	3	PTU_061
		Origanum vulgare	3	PTU_057
		Portulaca oleracea	3	PTU_051
		Physalis angulata	2	PTU_276
		Physalis pubescens	2	PTU_276
KR18-30	3	*Amaranthus hybridus*	3	PTU_207
		Convolvulus arvensis	3	PTU_668
		Damrongia orientalis	3	PTU_662
		Echinochloa colona	3	PTU_409
		Echinochloa frumentacea	3	PTU_409
		Hymenachne grumosa	3	PTU_090
		Amaranthus palmeri	2	PTU_083
		Amaranthus retroflexus	2	PTU_083
		Amaranthus viridis	2	PTU_023
		Origanum vulgare	2	PTU_006
KR18-32	0	*Convolvulus arvensis*	3	PTU_668
		Hymenachne grumosa	3	PTU_249
		Amaranthus viridis	2	PTU_023
		Damrongia orientalis	2	PTU_447
KR18-35`	2	*Convolvulus arvensis*	3	PTU_457
		Echinochloa colona	3	PTU_123
		Echinochloa crus-galli	3	PTU_123
		Echinochloa oryzicola	3	PTU_123
KR18-37	2	*Convolvulus arvensis*	3	PTU_131
		Dactyloctenium aegyptium	2	PTU_187

(Continued)

TABLE 3 | Continued

Sample ID	Basil	Unlabeled ingredients		
	# of replicates detected	Plant species	# of replicates detected	ASV identifier
		Damrongia orientalis	2	PTU_253
KR18-39	3	*Damrongia orientalis*	3	PTU_263
KR18-40	3	*Damrongia orientalis*	3	PTU_500
KR18-43	3	*Amaranthus hybridus*	3	PTU_146
		Amaranthus viridis	3	PTU_146
		Convolvulus arvensis	3	PTU_166
		Dactyloctenium aegyptium	3	PTU_187
		Damrongia orientalis	3	PTU_253
		Hymenachne grumosa	3	PTU_249
		Ipomoea batatas	2	PTU_165
		Ipomoea sp. GO-2017	2	PTU_165
		Ipomoea triloba	2	PTU_165
KR18-49	3	*Cenchrus americanus*	3	PTU_096
		Cynodon dactylon	3	PTU_119
		Dactyloctenium aegyptium	3	PTU_187
		Echinochloa colona	3	PTU_409
		Echinochloa frumentacea	3	PTU_409
		Ipomoea batatas	3	PTU_165
		Ipomoea sp. GO-2017	3	PTU_165
		Ipomoea triloba	3	PTU_165
		Portulaca oleracea	3	PTU_051
		Damrongia orientalis	2	PTU_500
KR18-51	0	*Amaranthus hybridus*	3	PTU_146
		Amaranthus viridis	3	PTU_146
		Convolvulus arvensis	3	PTU_595
		Corchorus olitorius	3	PTU_628
		Dinebra panicea subsp. brachiata	3	PTU_251
		Dinebra panicea subsp. mucronata	3	PTU_251
		Dinebra retroflexa	3	PTU_251
		Echinochloa colona	3	PTU_409
		Echinochloa frumentacea	3	PTU_409
		Hymenachne grumosa	3	PTU_090
		Leptochloa filiformis	3	PTU_251
		Origanum vulgare	2	PTU_254
KR18-56	3	*Damrongia orientalis*	3	PTU_662
KR18-59	0	*Amaranthus hybridus*	3	PTU_146
		Amaranthus viridis	3	PTU_146
		Convolvulus arvensis	3	PTU_010
		Hymenachne grumosa	3	PTU_249
		Origanum vulgare	3	PTU_339

be discriminated from *Ocimum americanum* using the genetic marker internal transcribed spacer 2 (ITS2) due to the low level of interspecific genetic diversity between these species. On the other hand, overlooked species can also be explained by false negative detections, which can be the case if the plant DNA is too heavily degraded or even removed during the production process (Novak et al., 2007; Cheng et al., 2014), there is a poor primer-template match (Piñol et al., 2015), amplification stochasticity due to low DNA concentration (Giguet-covex et al., 2014), or incomplete barcode sequences in the reference databases. In this regard, an

interesting result of this study is that we are confident in our negatives for basil and oregano, but not those of paprika. For instance, we can visually observe that the processing of basil and oregano should not degrade its DNA, and thus we expect to detect basil and oregano DNA if the species are present. Consequently, the absence of basil or oregano DNA in a product is a strong indicator that they are truly absent. For paprika powder on the other hand, it is very hard to determine with certainty that these products include paprika. When we cannot detect paprika DNA, and consider all the processing steps that might have

degraded or removed the DNA, it is difficult to ascertain whether paprika is truly absent or whether its DNA is degraded beyond the point that amplification of the ITS2 marker is possible. To reduce the uncertainty of the results, we used three technical replicates per sample, in addition to strict filtering and trimming thresholds for base call length and quality, and strict clustering criteria for ASV formation. The use of replicates has been reported to considerably reduce the risk of missing any present taxa at the expense of substantially increasing sequencing costs and time (Ficetola et al., 2015). Moreover, PCR-free approaches have been proposed to overcome limitations associated with amplification and to improve barcoding in taxonomically difficult plant groups in which traditional barcodes do not always provide sufficient taxonomic resolution (Kress and Erickson, 2007; Fazekas et al., 2009). For instance, "extended barcodes" that use low-coverage shotgun sequencing of genomic DNA (Straub et al., 2012; Coissac et al., 2016) and "super-barcodes" that use target capture sequencing (Mamanova et al., 2010; Li et al., 2015) are two PCR-free techniques that can be used for sample and product identification and authentication (Manzanilla et al., 2018; Haiminen, 2019). However, insufficient testing and the unavailability of extensive plastome or mitochondrial genome reference databases currently limit the applicability of these methods for the authentication of complex multi-ingredient plant-based products.

Any authentication strategy for quality control needs to pass analytical validation before it can be used in quality monitoring programs in a regulatory context and/or in supply chain management systems by the industry sector. This is particularly important since the quality, reliability, and consistency of analytical results are used to judge the quality and authenticity of a product. DNA-based methods for molecular diagnosis are some of the most promising prospective standards for quality confidence, though they require further validation (Bruno et al., 2019; Newmaster et al., 2019). Validating DNA metabarcoding is however challenging. Firstly, DNA metabarcoding does not provide information on the active metabolites in the raw plant material or the resulting preparations. This narrows its applicability to the identification of target plant species and confirmation of presence – but not absence – of other species, possible contaminants, and adulterants. Thus, DNA metabarcoding is recommended as a method to complement traditional analytical methods rather than be used alone (Raclariu et al., 2018a). Additionally, DNA metabarcoding cannot quantify relative species abundance since there are a number of confounding factors that can affect read numbers (Staats et al., 2016). Nevertheless, DNA metabarcoding can be used to analyze a diverse range of samples that are very often multi-ingredient and highly processed. It is a powerful method for the non-targeted identification of all taxa simultaneously in a product at any processing or production stage. This represents a key advantage over traditional DNA barcoding that is limited to the identification of single ingredients in raw materials (Raclariu et al., 2018a). However, proper analytical validation of the method is challenging since a product's various extractions and processing steps can lead to loss, degradation, or mixing of DNA. While some qPCR and DNA barcoding methods are validated and

standardized for quality control in commercial applications and regulatory contexts (Sgamma et al., 2017; Newmaster et al., 2019), no DNA metabarcoding protocols are yet established in these areas. Nevertheless, DNA metabarcoding addresses a number of the limitations of the currently used analytical methods for quality control, and we expect that validation studies will contribute to making it applicable in quality control systems.

CONCLUSION

Biological identification and authentication approaches based on DNA metabarcoding can be successfully used for the authentication of herbs and spices-based products, for post-marketing control and to provide insight into the total species diversity in processed, multi-ingredient products. The use of DNA metabarcoding in combination with appropriate traditional chemical methods can considerably increase the reliability of the quality control. However, proper analytical validation of DNA metabarcoding is necessary before it can be implemented for molecular diagnostics, both in quality monitoring programs in a regulatory context, and in supply chain management systems by the industry sector.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the ZENODO repository, accession number 4730123.

AUTHOR CONTRIBUTIONS

AR-M, JA, and HB conceived the experiment. AR-M carried out the molecular lab work together with JA, BT and AS-N and wrote the manuscript with HB. MK and NR carried out the bioinformatics analysis. All authors have read and approved the final version of the manuscript.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.665618/full#supplementary-material

Supplementary File 1 | Information about the investigated herbs and spices products.

Supplementary File 2 | DNA gel electrophoresis images.

Supplementary File 3 | Species identification report and BLAST queries.

Supplementary File 4 | DNA metabarcoding results for product identification.

REFERENCES

Amcham, and Trade USA (2015). *Spices and Herbs Market in USA: Market Analysis & Consumer Trends.* Trade USA: American ellenic Chamber of Commerce.

Ankenbrand, M. J., Keller, A., Wolf, M., Schultz, J., and Förster, F. (2015). ITS2 database V: twice as much. *Mol. Biol. Evol.* 32, 3030–3032. doi: 10.1093/molbev/msv174

Anthoons, B., Karamichali, I., Schrøder-Nielsen, A., Drouzas, A. D., de Boer, H., and Madesis, P. (2020). Metabarcoding reveals low fidelity and presence of toxic species in short chain-of-commercialization of herbal products. *J. Food Composition Anal.* 97:103767. doi: 10.1016/j.jfca.2020.103767

Aung, M. M., and Chang, Y. S. (2014). Traceability in a food supply chain: safety and quality perspectives. *Food Control* 39, 172–184. doi: 10.1016/j.foodcont.2013.11.007

Black, C., Haughey, S. A., Chevallier, O. P., Galvin-King, P., and Elliott, C. T. (2016). A comprehensive strategy to detect the fraudulent adulteration of herbs: the oregano approach. *Food Chem.* 210, 551–557. doi: 10.1016/j.foodchem.2016.05.004

Bruno, A., Sandionigi, A., Agostinetto, G., Bernabovi, L., Frigerio, J., Casiraghi, M., et al. (2019). Food tracking perspective: DNA metabarcoding to identify plant composition in complex and processed food products. *Genes* 10:248. doi: 10.3390/genes10030248

Callahan, B. J., McMurdie, P. J., and Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* 11, 2639–2643. doi: 10.1038/ismej.2017.119

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: high resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. doi: 10.1038/nmeth.3869

CBI (2018). *Which Trends Offer Opportunities on the European Spices and Herbs Market?* | *CBI - Centre for the Promotion of Imports from Developing Countries.* New Delhi: CBI.

Cheng, X., Su, X., Chen, X., Zhao, H., Bo, C., Xu, J., et al. (2014). Biological ingredient analysis of traditional Chinese medicine preparation based on high-throughput sequencing: the story for Liuwei Dihuang Wan. *Sci. Rep.* 4:5147. doi: 10.1038/srep05147

Coghlan, M. L., Maker, G., Crighton, E., Haile, J., Murray, D. C., White, N. E., et al. (2015). Combined DNA, toxicological and heavy metal analyses provides an auditing toolkit to improve pharmacovigilance of traditional Chinese medicine (TCM). *Sci. Rep.* 5:17457. doi: 10.1038/srep17475

Coissac, E., Hollingsworth, P. M., Lavergne, S., and Taberlet, P. (2016). From barcodes to genomes: extending the concept of DNA barcoding. *Mol. Ecol.* 25, 1423–1428. doi: 10.1111/mec.13549

Cornet, V., Govaert, Y., Moens, G., Van Loco, J., and Degroodt, J.-M. (2006). Development of a fast analytical method for the determination of sudan dyes in chili- and curry-containing foodstuffs by high-performance liquid chromatography-photodiode array detection. *J. Agric. Food Chem.* 54, 639–644. doi: 10.1021/jf0517391

Dodt, M., Roehr, J. T., Ahmed, R., and Dieterich, C. (2012). FLEXBAR—Flexible barcode and adapter processing for next-generation sequencing platforms. *Biology* 1, 895–905. doi: 10.3390/biology1030895

Doyle, J. J. and Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. *Focus* 12, 13–15.

Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., et al. (2012). New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems. *Mol. Ecol.* 21, 1821–1833. doi: 10.1111/j.1365-294X.2012.05537.x

Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, 3047–3048. doi: 10.1093/bioinformatics/btw354

Fazekas, A. J., Kesanakurti, P. R., Burgess, K. S., Percy, D. M., Graham, S. W., Barrett, S. C. H., et al. (2009). Are plant species inherently harder to discriminate than animal species using DNA barcoding markers? *Mol. Ecol. Resources* 9, 130–139. doi: 10.1111/j.1755-0998.2009.02652.x

Ficetola, G. F., Coissac, E., Zundel, S., Riaz, T., Shehzad, W., Bessière, J., et al. (2010). An in silico approach for the evaluation of DNA barcodes. *BMC Genom.* 11:434. doi: 10.1186/1471-2164-11-434

Ficetola, G. F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., et al. (2015). Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Mol. Ecol. Resources* 15, 543–556. doi: 10.1111/1755-0998.12338

Galvin-King, P., Haughey, S. A., and Elliott, C. T. (2018). Herb and spice fraud; the drivers, challenges and detection. *Food Control* 88, 85–97. doi: 10.1016/j.foodcont.2017.12.031

Galvin-King, P., Haughey, S. A., and Elliott, C. T. (2020). The detection of substitution adulteration of paprika with spent paprika by the application of molecular spectroscopy tools. *Foods* 9:944. doi: 10.3390/foods9070944

Giguet-covex, C., Pansu, J., Arnaud, F., Rey, P., Griggo, C., Gielly, L., et al. (2014). Long livestock farming history and human landscape shaping revealed by lake sediment DNA. *Nat. Commun. London* 5:3211. doi: 10.1038/ncomms4211

Haiminen, N. (2019). Food authentication from shotgun sequencing reads with an application on high protein powders. *NPJ Sci. Food* 3:24.

Ivanova, N. V., Kuzmina, M. L., Braukmann, T. W. A., Borisenko, A. V., and Zakharov, E. V. (2016). Authentication of herbal supplements using next-generation sequencing. *PLoS One* 11:e0156426. doi: 10.1371/journal.pone.0156426

Kress, W. J., and Erickson, D. L. (2007). A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. *PLoS One* 2:e508. doi: 10.1371/journal.pone.0000508

Li, X., Yang, Y., Henry, R. J., Rossetto, M., Wang, Y., and Chen, S. (2015). Plant DNA barcoding: from gene to genome. *Biol. Rev. Camb. Philos. Soc.* 90, 157–166. doi: 10.1111/brv.12104

Loman, N. J., Misra, R. V., Dallman, T. J., Constantinidou, C., Gharbia, S. E., Wain, J., et al. (2012). Performance comparison of benchtop high-throughput sequencing platforms. *Nat. Biotechnol.* 30, 434–439. doi: 10.1038/nbt.2198

Mamanova, L., Coffey, A. J., Scott, C. E., Kozarewa, I., Turner, E. H., Kumar, A., et al. (2010). Target-enrichment strategies for next-generation sequencing. *Nat. Methods* 7, 111–118. doi: 10.1038/nmeth.1419

Manzanilla, V., Kool, A., Nguyen Nhat, L., Nong Van, H., Le Thi, Thu, H., et al. (2018). Phylogenomics and barcoding of Panax: toward the identification of ginseng species. *BMC Evol. Biol.* 18:44. doi: 10.1186/s12862-018-1160-y

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17, 10–12. doi: 10.14806/ej.17.1.200

Newmaster, S. G., Shanmughanandhan, D., Kesanakurti, P., Shehata, H., Faller, A., Noce, I. D., et al. (2019). Recommendations for validation of real-time PCR methods for molecular diagnostic identification of botanicals. *J. AOAC Int.* 102, 1767–1773. doi: 10.5740/jaoacint.18-0321

Nielsen, U. N., and Wall, D. H. (2013). The future of soil invertebrate communities in polar regions: different climate change responses in the Arctic and Antarctic? *Ecol. Lett.* 16, 409–419. doi: 10.1111/ele.12058

Novak, J., Grausgruber-Gröger, S., and Lukas, B. (2007). DNA-based authentication of plant extracts. *Food Res. Int.* 40, 388–392. doi: 10.1016/j.foodres.2006.10.015

Pawluczyk, M., Weiss, J., Links, M. G., Egaña Aranguren, M., Wilkinson, M. D., and Egea-Cortines, M. (2015). Quantitative evaluation of bias in PCR amplification and next-generation sequencing derived from metabarcoding samples. *Anal. Bioanal. Chem.* 407, 1841–1848. doi: 10.1007/s00216-014-8435-y

Piñol, J., Mir, G., Gomez-Polo, P., and Agustí, N. (2015). Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Mol. Ecol. Resources* 15, 819–830. doi: 10.1111/1755-0998.12355

Raclariu, A. C., Heinrich, M., Ichim, M. C., and de Boer, H. (2018a). Benefits and limitations of DNA barcoding and metabarcoding in herbal product authentication. *Phytochem. Anal.* 29, 123–128. doi: 10.1002/pca.2732

Raclariu, A. C., Țebrencu, C. E., Ichim, M. C., Ciuperca, O. T., Brysting, A. K., and de Boer, H. (2018b). What's in the box? authentication of Echinacea herbal products using DNA metabarcoding and HPTLC. *Phytomedicine* 44, 32–38. doi: 10.1016/j.phymed.2018.03.058

Raclariu, A. C., Mocan, A., Popa, M. O., Vlase, L., Ichim, M. C., Crisan, G., et al. (2017a). Veronica officinalis product authentication using DNA metabarcoding and HPLC-MS reveals widespread adulteration with Veronica chamaedrys. *Front. Pharmacol.* 8:378. doi: 10.3389/fphar.2017.00378

Raclariu, A. C., Paltinean, R., Vlase, L., Labarre, A., Manzanilla, V., Ichim, M. C., et al. (2017b). Comparative authentication of Hypericum perforatum herbal products using DNA metabarcoding, TLC and HPLC-MS. *Sci. Rep.* 7:1291. doi: 10.1038/s41598-017-01389-w

Robasky, K., Lewis, N. E., and Church, G. M. (2014). The role of replicates for error mitigation in next-generation sequencing. *Nat. Rev. Genet.* 15, 56–62. doi: 10.1038/nrg3655

Roehr, J. T., Dieterich, C., and Reinert, K. (2017). Flexbar 3.0 – SIMD and multicore parallelization. *Bioinformatics* 33, 2941–2942. doi: 10.1093/bioinformatics/btx330

Sgamma, T., Lockie-Williams, C., Kreuzer, M., Williams, S., Scheyhing, U., Koch, E., et al. (2017). DNA barcoding for industrial quality assurance. *Planta Med.* 83, 1117–1129. doi: 10.1055/s-0043-113448

Silvis, I. C. J., van Ruth, S. M., van der Fels-Klerx, H. J., and Luning, P. A. (2017). Assessment of food fraud vulnerability in the spices chain: an explorative study. *Food Control* 81, 80–87. doi: 10.1016/j.foodcont.2017.05.019

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., et al. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere.". *Proc. Natl. Acad. Sci. U S A.* 103, 12115–12120. doi: 10.1073/pnas.0605127103

Speranskaya, A. S., Khafizov, K., Ayginin, A. A., Krinitsina, A. A., Omelchenko, D. O., Nilova, M. V., et al. (2018). Comparative analysis of Illumina and Ion Torrent high-throughput sequencing platforms for identification of plant components in herbal teas. *Food Control* 93, 315–324. doi: 10.1016/j.foodcont.2018.04.040

Staats, M., Arulandhu, A. J., Gravendeel, B., Holst-Jensen, A., Scholtens, I., Peelen, T., et al. (2016). Advances in DNA metabarcoding for food and wildlife forensic species identification. *Anal. Bioanal. Chem.* 408, 4615–4630. doi: 10.1007/s00216-016-9595-8

Straub, S. C. K., Parks, M., Weitemier, K., Fishbein, M., Cronn, R. C., and Liston, A. (2012). Navigating the tip of the genomic iceberg: next-generation sequencing for plant systematics. *Am. J. Botany* 99, 349–364. doi: 10.3732/ajb.1100335

Sun, Y., Skinner, D. Z., Liang, G. H. and Hulbert, S. H. (1994). Phylogenetic analysis of *Sorghum* and related taxa using internal transcribed spacers of nuclear ribosomal DNA. *Theor. Appl. Genet.* 89, 26–32 doi: 10.1007/BF00226978

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., and Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Mol. Ecol.* 21, 2045–2050. doi: 10.1111/j.1365-294X.2012.05470.x

Tapsell, L. C., Hemphill, I., Cobiac, L., Sullivan, D. R., Fenech, M., Patch, C. S., et al. (2006). Health benefits of herbs and spices: the past, the present, the future. *Med. J. Aust.* 185, S1–S24. doi: 10.5694/j.1326-5377.2006.tb00548.x

van Ruth, S. M., Luning, P. A., Silvis, I. C. J., Yang, Y., and Huisman, W. (2018). Differences in fraud vulnerability in various food supply chains and their tiers. *Food Control* 84, 375–381. doi: 10.1016/j.foodcont.2017.08.020

Viuda-Martos, M., Ruiz-Navajas, Y., Fernández-López, J., and Pérez-Alvarez, J. A. (2011). Spices as functional foods. *Crit. Rev. Food Sci. Nutr.* 51, 13–28. doi: 10.1080/10408390903044271