

Accepted Manuscript

Title: Metabarcoding reveals low fidelity and presence of toxic species in short chain-of-commercialization of herbal products

Authors: Bastien Anthoons, Ioanna Karamichali, Audun Schrøder-Nielsen, Andreas D. Drouzas, Hugo de Boer, Panagiotis Madesis

To appear in: Journal of Food Composition and Analysis

Reference: YJFCA 103767

DOI: <https://doi.org/10.1016/j.jfca.2020.103767>

Received Date: 12 August 2020

Revised Date: 15 November 2020

Accepted Date: 12 December 2020

Metabarcoding reveals low fidelity and presence of toxic species in short chain-of-commercialization of herbal products

Bastien Anthoos^a, Ioanna Karamichali^a, Audun Schrøder-Nielsen^b, Andreas D. Drouzas^c, Hugo de Boer^b, Panagiotis Madesis^{a,d*}

^a Institute of Applied Biosciences, CERTH, 6th km Charilaou-Thermis Road, Thermi, 57001, Thessaloniki, Greece

^b Natural History Museum, University of Oslo, Postboks 1172, Blindern, 0318 Oslo, Norway

^c Lab. of Systematic Botany and Phytogeography, School of Biology, P.O. Box: 104, Aristotle University of Thessaloniki, 54124, Thessaloniki, Greece

^d Laboratory of Molecular Biology of Plants, School of Agricultural Sciences, University of Thessaly, Thessaly 38446 Greece

*padesis@certh.gr

Highlights

- DNA metabarcoding is a promising tool for authentication of complex herbal products
- Detection of toxic species raises concerns about products' reliability and safety
- Bar-HRM is an efficient verification tool for metabarcoding analysis
- Stricter quality control is necessary to combat adulteration and fraudulent practices

Abstract

Herbal medicinal products gain increasing popularity. The growing demand for herbal medicine along with a lack of regulation render herbal products subject to intentional adulteration. The substitution of costly ingredients with unlabelled plant-based fillers of inferior quality has been widely reported. Such fraudulent practices erode consumer trust, but can also pose serious health risks. In this work, 71 herbal medicinal products were randomly purchased from Greek markets and analysed using ITS2 metabarcoding for species identification. The aim was to investigate possible adulterations and assess the efficacy of metabarcoding in plant-based product authentication. Of the 131 detected species in our analyses, 87 were not listed on the product labels. This indicates a high level of adulteration and/or contamination during processing and distribution. Two toxic species, *Chelidonium majus* and *Nicotiana tabacum*, were also detected as major ingredients of two herbal mixtures for medicinal purposes. Furthermore, the detection of wheat in eight samples raises concerns for people with gluten intolerance. This study stresses the need for stricter quality control of herbal products. In addition, to overcome the limitations of metabarcoding and augment the approach we used Bar-HRM for the first time as a

verification tool. The combination of metabarcoding with species-specific Bar-HRM analysis enhances the reliability of the results.

Keywords

Adulteration, authentication, Bar-HRM, DNA metabarcoding, herbal products, ITS2, toxic plants

1. Introduction

Plants have always been an important food source while also being used as spices and traditional medicines, whose wide use has been documented through the ages. Spices are used to enhance palatability in foods and beverages. In addition, spices and herbs can be used in medical and cosmetic applications, as well as for the production of food additives and preservatives (Peter, 2012). Greece is characterized by high plant diversity (6620 taxa) and endemism (1459 taxa), due to its climate, geomorphology, geological history and the interplay of biotic and abiotic factors (Dimopoulos *et al.* 2013; Dimopoulos *et al.* 2016; Solomou *et al.* 2016). In particular, medicinal and aromatic plants are an important part of the Greek flora, comprising a large number of taxa, mainly of the families: Apiaceae, Asteraceae, Lamiaceae, Lauraceae, Myrtaceae, Pinaceae (Kokkini *et al.* 1988). Small-scale cultivation and wild harvesting of herbs and spices for household use is a cultural characteristic of Greece, since the use of herbs in every-day life is common since antiquity. Thus, herbs and spices are commonly found and sold in public markets (usually raw dried material), pharmacies, natural products shops, as well as supermarkets. Given the suitable soil and climate conditions it is not uncommon that most of the raw materials used to supply those markets are derived from native species cultivated in Greece or, in some cases, imported from abroad (Hanlidou *et al.* 2004; Stefanou *et al.* 2015).

During the past decade, traditional medicinal plant-based products have gained popularity in local healthcare as a source of perceived healthier alternatives to commercial pharmaceuticals. Furthermore, in many developing countries, a large number of people depend on complementary, alternative healthcare (CAM) and medical practices, including medicinal herbs (WHO, 2013; Osathanunkul *et al.* 2016). The popularity of plant based products is determined by their health attributes, their high nutritional value, cultural factors and the perception that these products are natural and thus with no or minimum side effects (Lynch and Berry, 2007; Ipsos MORI, 2008). In the United States alone, people spend over US\$ 7.4 billion a year on herbal food supplements (Smith *et al.* 2018; Nutrition Business Journal, 2016). Complementary to the food supplement industry, the global herb and spice trade was estimated to be worth US\$ 2.97 billion with the EU market accounting for a value of € 1.8 billion (Marieschi *et al.* 2009). In the UK, supermarket sales in 2014 reached £ 173 million for dried herbs and spices, and £107 million for fresh herbs (spices not included) (Nielsen, 2015). However, the growing demand for qualitative herbal medicine renders them subject to intentional and/or unintentional adulteration (Lupien, 2005).

Herbal products are usually complex mixtures of different plant materials, and are manufactured using various processing steps. Adulteration and substitution of these products has become a widespread problem in the herbal and food industry, owing to the lack of standardized methods for quality assessment, lucrative profits and a highly competitive market. The substitution of costly ingredients with

unlabelled plant fillers or plant materials of inferior quality has been reported in numerous studies (Ouarghidi *et al.* 2012; Zhang *et al.* 2012; Raclariu *et al.* 2017b; Seethapathy *et al.* 2019). This can lead to a loss in consumer trust and can even pose potential serious health risks (Seethapathy *et al.* 2019; Newmaster *et al.* 2013). Likewise, substitution of toxic plant species for edible species has occasionally been reported (Garzo *et al.* 2002; Johanns *et al.* 2002). Finally, the recent trend for spice fraud appears to be economically driven to gain greater profit margins (Osman *et al.* 2019).

Current quality and authentication assessments rely heavily on morphology and analytical phytochemistry-based methodologies. Herbal material and products are routinely assessed using standard morphology, microscopy and chemical fingerprinting, such as high-performance liquid chromatography (HPLC), and gas chromatography (GC) (de Boer *et al.* 2015; Parveen *et al.* 2016). These methods are quick and cost-effective for primary qualitative analysis of raw material and derived herbal products. However, they are usually not suitable for the identification of targeted and non-targeted plant species in complex mixtures, such as herbal products (de Boer *et al.* 2015). DNA barcoding is a validated molecular identification method that is commonly used in the authentication of plant-based products by providing species-level resolution (Newmaster *et al.* 2013). The technique is industrially used today for quality assessment of a wide range of herbal products (Sgamma *et al.* 2017). High-throughput sequencing based amplicon metabarcoding enables simultaneous multi-taxa identification from samples containing DNA from different origins (Taberlet *et al.*, 2012). The technique can provide further insights into species composition of complex mixtures, such as processed herbal products. Several studies have utilized this approach in authenticating medicinal plants and derived herbal products (Coghlan *et al.* 2012; Raclariu *et al.* 2017a; Raclariu *et al.* 2017b; Speranskaya *et al.* 2018; Lo and Shaw, 2019; Seethapathy *et al.* 2019). Raclariu *et al.* (2017) showed that DNA metabarcoding was able to authenticate *Hypericum perforatum* in complex herbal products and to detect incongruences between constituent species and the labelled species. Cheng *et al.* (2014) found that their selection of 27 Chinese herbal medicinal products contained contaminants (Cheng *et al.* 2014). Similarly, the analyses of Seethapathy *et al.* (2019) revealed that only two out of 12 single ingredient Ayurveda based herbal products contained only one species as labelled, and eight out of 27 multiple ingredient herbal products contained none of the species listed on the labelling (Seethapathy *et al.* 2019).

Although metabarcoding proved an effective method for species composition assessment and authentication, verification of the results is recommended. Barcoding and metabarcoding may provide false positives based on any amplifiable DNA and false negatives due to degraded DNA or post-harvest processing (de Boer *et al.* 2015, Raclariu *et al.* 2018). An effective DNA-based identification approach combines DNA barcoding with High Resolution Melting (HRM), called Bar-HRM (Reed & Wittwer, 2004; Jaakola *et al.*, 2010). High Resolution Melting is a post-PCR method for monitoring DNA dissociation kinetics ("melting"), and is a powerful technique to detected even point mutations and thus small DNA variations between closely related species/taxa (Reed & Wittwer, 2004; Lagiotis *et al.* 2020). The melting profile and peak of a sample, plotted as fluorescence measurements over time, are characteristic for this sample, which allows comparison and discrimination among samples (Osathanunkul *et al.* 2016). Bar-HRM has been effectively used in a number of applications including authentication of food and herbal products (e.g. Ganopoulos *et al.* 2012; Osathanunkul *et al.* 2017). The technique has been effectively used for quality control of vegetables and for differentiation between edible and poisonous plants (Thongkhao *et al.* 2020). Lagiotis *et al.* (2020) demonstrated that the HRM technology is a strong molecular approach in

tea authentication, capable of detecting very low adulterations in DNA admixtures. In this paper, we used Bar-HRM to verify the amplicon metabarcoding results as to enhance the confidence of these results.

We report here a novel authentication protocol applied on a selection of herbal products sold in Greek markets. The composition and fidelity of the herbal mixtures was assessed via amplicon metabarcoding (AMB) methodology and verified by Bar-HRM, aiming to investigate product fidelity in short supply chains. Our hypothesis is that fidelity is low even in short chains of commercialization. To test this hypothesis, we pose the following research questions: 1) Can we use AMB to detect species composition; 2) What read thresholds and replicate thresholds should be set to filter out noise or trace contamination; 3) Can we confirm AMB identifications using Bar-HRM.

2. Material and methods

2.1 Sample collection

In this study, we used 71 herbal products purchased randomly from pharmacies and local markets, presumably of Greek origin; 2 products were sold as capsules, 27 as powders, 2 plant-based extracts, and 40 dried plant parts. Labels were missing for 13 products (samples 19 – 31, Supplementary Table S1). Based on the label information, 5 were single plant ingredient products, while 66 contained 2 to 13 plant ingredients. All of the products contained a total of 131 plant species belonging to 115 genera and 52 families according to the products' labelling. An overview of the products, 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 Table S1.

2.2 DNA Extraction, Amplicon Generation, and High Throughput Sequencing

About 1 g of each herbal product was ground with liquid nitrogen and 150 mg were used for DNA extraction using the CTAB method (Doyle and Doyle, 1987). Extracted DNA was diluted in 100 µL and quantified using a Qubit 2.0 Fluorometer and Qubit dsDNA HS Assay Kit (Invitrogen, United States).

Amplicon libraries were prepared for all samples in triplicates. Three amplification blanks were also included per PCR plate (two 96-well plates). For each replicate the nuclear ribosomal target sequence of the Internal Transcribed Spacer (nrITS2) was amplified. The fusion primers included the annealing motif from the Sun et al. (1994) plant-specific primer pair for nrITS2 5.8l2 and 26SE. The forward primers included the Ion Torrent A adapter, a 10 bp multiplex identifier tag following the IonXpress setup for Ion Torrent (Thermo Fisher Scientific, United States). The reverse primer included the truncated P1 (trP1) tags in addition to the annealing motif. Expected amplicon sizes were 300–350 bp.

Polymerase chain reactions were carried out using DNA extracted from the herbal products in final reaction volumes of 25 µL including 0.5 µL of template DNA solution (ranging from 0.5 to 2 ng µL⁻¹), 1X Q5 reaction buffer (New England Biolabs Inc., United Kingdom), 0.6 mM of each primer (Biolegio B.V., Netherlands), 200 nM dNTPs, 5 U Q5 High-Fidelity DNA Polymerase (New England Biolabs Inc., United Kingdom) and 1X Q5 High GC enhancer. The PCR cycling protocol consisted of initial denaturation at 98 °C for 30 s, followed by 35 cycles of denaturation at 98 °C for 10 s, annealing at 71 °C for 30 s and extension at 72 °C for 30 s. All amplicons were visualized using agarose gel electrophoresis, and concentration of each amplicon was calculated based on fluorescence intensities using the Image Lab v.6.0.0 software (Bio-

Rad Laboratories, USA). Short DNA sequences, <150 bp, showing up on the gel were not included in concentration estimation, as these likely were primer dimers and would be removed prior to sequencing. All amplicons were then normalized (40 ng/sample) using a Biomek 4000 Laboratory Automation Workstation (Beckman Coulter, USA) liquid handling system. For amplicons with concentrations lower than 2 ng/μl, the entire amplicon was added to the pool (e.g. non-positive amplifications and all PCR- and extraction negatives). The normalized library pool was cleaned using 0.8X Ampure XP beads (Beckman Coulter, USA) before size selection was performed on a BluePippin (Sage Science, USA) using a 2% agarose cassette, V1 internal marker and 300-550 bp target range. Molarities of the size selected pools were estimated using a Fragment Analyzer DNF-488 kit (Advanced Analytical Technologies, Inc., USA). Template preparation and chip loading was performed on an Ion Chef (Thermo Fisher Scientific (TF), USA) using 50 pM input per Ion 530 chip (TF) before the libraries were sequenced on an Ion GeneStudio S5 system (TF) using 850 flows.

2.3 Bioinformatic analysis

The reads of the ITS2 region from both chips analysed were demultiplexed per chip cluster and per sample replicate, quality controlled and further analysed using the local Galaxy (<http://galaxy.naturalis.nl/>) instance Naturalis Galaxy Production v.19.01 (Ubuntu 18.04 LTS) provided by Naturalis Biodiversity Center of Leiden. All tools were executed with the default parameters, unless otherwise stated. The pipeline used included the tools 'Sequence analyzer' (V: 1.0), 'Sequence trimmer' (V: 1.0, length: 360-240, minimum mean Phred quality: 24 and the removal of the 5'end GAT tail for the CHIP2 sequences), 'Split on primer' (V: 1.1), 'Make OTU table' (V: 1.0.0, clustering: DADA2) and 'Identify reads with Blastn and find taxonomy' (V: 1.0.0, query coverage percentage cutoff: 85, identity percentage cutoff: 96, maximum 4 blast hits per sequence). We used the GenBank ITS database incorporated in the Galaxy pipeline.

The clustering similarity level thresholds (being >97%, >99% and 100%) are an important factor that affects the final number and size of the assigned MOTUs (Molecular Operational Taxonomic Units) (Raclariu *et al.* 2017a). In the present study, we used a 97% clustering threshold. Despite the proposed 99% threshold by Yang and Rannala (2017), in our case there was no significant difference in yield of MOTUs between 97% and 99% thresholds. Moreover, we tested the clustering thresholds (>97%, >99% and 100%) and the 97% threshold was found to correspond best to the expected species. Also, to limit the impact of sequencing errors which could lead to false MOTUs, we used only the clusters that contained a minimum of 100 reads. Further noise was reduced by accepting MOTUs only when found in two out of three replicates per sample. Those factors in addition to the applied filtering and trimming for length and quality criteria, increase the confidence of the results.

Further analyses and comparison between samples were performed using tailored Python scripts (V: 3.6.8, Ubuntu 18.04.2 LTS). Detected species with MOTUs count lower than 100, or presence in only one of the three replicates were filtered out in order to minimize false positive signals. Each unique MOTU was assigned to a single species that based on the Blastn analysis scored the highest in different values with order of significance bitscore>evalue>coverage>identity-percentage. The final list of the assigned species per sample constituted the detected species dataset. In parallel, taxonomy information of the labelled organisms of each sample were extracted from NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), constituting the expected species (i.e. species on the product label) dataset. It was confirmed that nrITS sequences of all the 131 plant species labelled in the analysed herbal products were available within the NCBI/GenBank database (<http://www.ncbi.nlm.nih.gov>). The accepted binomial names of the plant

species used as ingredients were validated using Dimopoulos *et al.* (2013) except from the species that are not found in Greece for which "Plants of the World Online. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet; <http://www.plantsoftheworldonline.org/> Retrieved 20/09/2019" (POWO, 2019) was used. The two datasets were compared at species level or if not applicable at genus level to further classify the expected and detected datasets into the categories "expected-detected", "expected-not-detected" and "not-expected-detected". Additionally, the main number of reads from the three replicates of each detected organism, as well as the corresponding percentage were calculated. Finally, the observed adulterations and fidelity were estimated for each sample.

All datasets were presented using Matplotlib (Hunter, 2007) for creating the dot plot and the Venn diagrams. Furthermore, the web tools ClustVis (in BETA, <https://biit.cs.ut.ee/clustvis/>) and RAWGraphics (<https://rawgraphics.io/>) were used for creating the heatmap and the circle packing representations, respectively.

2.4 PCR amplification and HRM analysis for species confirmation

New primers were designed based on the ITS2 sequences from the species used, available at the GenBank database, using the software Prime3 Output designing tool (<http://frodo.wi.mit.edu/prime3>) (Supplementary Table S2). The nucleotide sequences were submitted to a basic local alignment search tool BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Primer specificity was assessed using the Primer-BLAST tool, which reveals homologies in relation to sequences available in the GenBank database.

The ITS2 locus from three species was amplified (*Chelidonium majus*: private collection I. Tsiripidis (Assoc. Prof., School of Biology, Aristotle University of Thessaloniki), *Nicotiana tabacum*: commercial sample Xanthi81, *Triticum aestivum*: GRC 182/05). Cycling and HRM analyses of reference samples and herbal products (#19, #26, #29, #57, #64) were performed on a Rotor-Gene 6000 real-time 5-Plex HRM PCR Thermocycler (Corbett Research, Sydney, Australia), using the Rotor-Gene Q software 2.0.2 (Corbett Life Science, Cambridge, UK). PCR reactions were carried out in final reaction volumes of 20 μ l consisting of 2 ng of template DNA, 0.2 mM of each dNTPs, 1.5 mM MgCl₂, 0.5 μ M of forward and reverse primer, 2 μ l of 10x *Taq* DNA polymerase buffer, 1.5 mM Syto[®] 9, a third-generation green fluorescent nucleic acid stain (Invitrogen, Eugene, OR, USA) and 1 U Kapa *Taq* DNA polymerase (Kapa Biosystems, MA, USA). The accumulation of the amplified product during PCR was monitored by using Syto[®] 9. The universal ITS2 region was amplified as follows: an initial denaturation step of 4 min at 94 °C, followed by 40 cycles, each one including 30 s at 94 °C for denaturation, 30 s at 68 °C for annealing and 30 s at 72 °C for elongation. A 7 min step at 72 °C was programmed as a final extension. Amplification products were analysed by electrophoresis on a 1.5% agarose gel stained with ethidium bromide. Gels and images were analysed using UV Minibis Pro (DNR Bio-Imaging Systems, Jerusalem, Israel) to quantify signal intensity.

HRM was performed as follows: pre-melt at the first appropriate temperature for 90 s, and melt at a ramp of 10 °C in an appropriate temperature range at 0.1 °C increments every 2 s. The fluorescence data was acquired at the end of each annealing step during the PCR cycles. In order to further increase the reproducibility and reliability of the HRM curve analysis (by obtaining similar amplified quantities of final PCR products before melting), finer adjustments by dilutions were made to the genomic DNA templates.

The Rotor-Gene Q software was used to test the presence of the different species. The negative derivative of fluorescence (F) over temperature (T) (dF/dT) curve, the normalized raw curve depicting the decreasing fluorescence vs increasing temperature, and the difference curves were used. Furthermore, a

two-step procedure was followed according to Ganopoulos *et al.* (2012) to assess similarity of unknown HRM curves with a known one. The average HRM genotype confidence percentages (GCP's) were tabulated for the replicates and compared to decide about HRM curve similarities.

3. Results

Metabarcoding libraries were prepared for 61 herbal samples with DNA concentrations ranging from 0.6–60 ng μl^{-1} . We decided not to include 10 samples with concentrations of less than 0.6 ng μl^{-1} or not measurable at all (#1-2, #6-7, #9, #12, #23, #34, #45 and #52). These products consisted of 6 powders, 1 extract and 2 capsules and 1 mixture of dried plant parts. PCR amplification for the nrITS2 region was performed and amplicons were generated for all replicates. The sequencing resulted in 11 804 714 reads, which were further reduced to 11 465 601 after demultiplexing and trimming. DADA2 clustering of the reads resulted in 11 644 MOTUs (Supplementary Table S3). 20 products out of the 61 (32%) yielded no MOTUs in any of the replicates that fulfilled the quality criteria in the bioinformatics analysis, as described in 2.3, and they were excluded from the analysis (#8, #11, #13-14, #17, #20-22, #25, #34-37, #39-40, #47-48, #51, #65 and #71). The products that yielded MOTUs were represented by 15 powders and 27 mixtures of dried plant parts. Finally, the filtering criteria and the set DNA concentration threshold reduced the total amount of samples from 71 to 39. All of the analysed products contained between one to eight species listed on the label along with other species not on the product's label. Ingredient fidelity in each sample was defined as the ratio of the detected species to the total number of species on the labelling (Supplementary Table S4). Table 1 shows the top ten products with the highest fidelity and their associated level of adulteration, whereas Table 2 shows the top ten products with highest adulteration and their fidelity.

The number of species detected per sample through this analysis ranged from 1 to 31. In order to reduce the experimental biases, three replicates per sample and three negative controls per 96-well PCR plate were included. The extraction blanks as well as the PCR negative controls yielded no MOTUs with nrITS2 primers. After applying the quality selection criteria, where a species was present within the product only if it was detected in at least two of the three replicates, one additional product (herbal mixture #68) failed to yield the same MOTU in any of the replicates and was discarded. A total of 186 different plant species were identified from the retained MOTUs using BLAST. The total number of expected and detected species identified through metabarcoding are visualized with two- and three- way Venn diagrams (Figure 1). Figure 1 shows the detection of species through metabarcoding in two-way (Figure 1A) and three-way (Figure 1B) Venn diagrams. As shown on Figure 1A, the number of expected species detected in the herbal products is quite low (44 out of 152 expected species), which indicates a high level of adulteration. The three-way Venn diagram gives a more detailed representation of the number of individual species that were found between the different samples (Figure 1B). Fourteen species were classified as both expected-detected and not-expected-detected through the total number of samples. An example of a species belonging to this subgroup is *Urtica dioica* L., a species that is both very common and commonly used for medicinal purposes. The rest of the subgroups similarly represent the number of species classified as both “expected-not-detected” and “expected-detected” (8), or as both “not-expected-detected” and “expected-not-detected” (4), or belonging to all categories (3), (Figure 1B).

The overall ingredient fidelity of the successfully analysed samples is visualized with a 3D dot plot (Figure 2). Each symbol in the figure represents a single sample, different symbols represent different types of products. The fidelity is defined as the percentage of the number of the detected species out of the total number of species on the product's label (expected species per sample, y-axis). The sequence reads identified via metabarcoding and BLAST were classified as belonging to either the not-expected-detected or expected-detected category and the percentage per sample is represented on the x-axis for the not-expected-detected samples, and the z-axis for the expected-detected species, respectively. Most of the samples have a fidelity ranging from 40 to 70%. Not only most of the expected species were not detected, but also many unexpected species were found in many samples. In addition, no significant difference in fidelity was observed between the two product types (powder and dried plant parts).

Remarkably, two toxic species were identified in the examined herbal products. *Chelidonium majus* L. (greater celandine) was detected in both samples 57 and 64, having a relative abundance of 0.25% and 85.02%, respectively. In addition to *Chelidonium majus*, *Nicotiana tabacum* L. (tobacco) was detected through the metabarcoding analysis as a major ingredient of sample 29 with a relative abundance of 13.9%. Other species detected in the latter sample are *Capsicum annuum* L. (69.44%) and *Secale cereale* L. (16.66%), (Supplementary Figure 1). The species detected through the metabarcoding analysis are shown as the relative abundance of normalized read numbers in Figure 3. The species are listed per sample, per product type and category (expected-detected, not expected-detected and expected-not detected).

To confirm the presence of the detected toxic species and undetected labelled species by metabarcoding, we used barcoding coupled with High Resolution Melting (HRM) analysis (Figures 5 to 7). The results of the HRM analysis confirmed the presence of *Nicotiana tabacum* in the spice mixture sample 29 (Figure 4) and *Chelidonium majus* in the herbal mixture sample 64 (Figure 5). The GCP of *Chelidonium majus* in herbal mixture 64 is 95.62%, while the negative control resulted in a 0% correspondence. *Nicotiana tabacum* was detected with 98.35% certainty in mixture 29. Furthermore, the not-expected species *Triticum aestivum* L. was detected by HRM in both samples 19 and 26 (spice mixtures) with GCP values of 86.10% and 78.31%, respectively (Figure 6).

The following species were found in over 20% of the products (detected but not expected): *Convolvulus arvensis* L. (48%), *Urtica dioica* L. (41%), *Ocimum basilicum* L. (22%), *Rosmarinus officinalis* L. (22%), and *Sideritis scardica* Griseb. (22%) (Figure 3 and Supplementary Tables S5, S6). Eleven species are present in more than 10% of the samples and are listed in the Supplementary Table S6.

4. Discussion

The highly competitive herbal market, the high demand for herbal products and the associated drive for greater profit margins have rendered herbal medicinal mixtures and spices subject to intentional adulteration. Thus, there is an urgent need for standardized methods to combat these fraudulent practices. When dealing with plant powders and herbal mixtures it is not always possible to identify the labelled species based on traditional methods. In several cases, morphological identification becomes impossible or time-consuming. Phytochemical approaches such as chemical profiling are frequently being used for authentication of botanicals. However, such methodologies require the use of a large number of authenticated reference samples and are less suitable when evaluating complex mixtures (Grazina *et al.*

2020). The advances on DNA-based methods resulted in simple, reliable, cost-effective, and high-throughput tools for the accurate authentication of herbal products. Amplicon metabarcoding has demonstrated a high potential to authenticate complex herbal products (Coghlan *et al.* 2012). In this study, we used ITS2 as the barcode of choice as it has been proven to be accurate for the authentication of herbal products and the identification of medicinal plants in particular (Chen *et al.* 2010; Raclariu *et al.* 2017a; Raclariu *et al.* 2017b; Seethapathy *et al.* 2019). The aim of authentication studies using metabarcoding is checking the presence of labelled ingredients and possible contaminants introduced in different steps of the supply chain of medicinal plants, i.e. cultivation, packaging, transport, storage, manufacturing process and commercialization of the herbal products (Ivanova *et al.* 2016, Raclariu *et al.* 2017b). Among various NGS technologies, Illumina and Ion Torrent sequencing remain the two most widely used for authentication of herbal products. The reason for this is the short read-length of both techniques which fit the average DNA fragment length in the processed herbal products with degraded DNA (Lo and Shaw, 2019). Ion Torrent has some advantages over Illumina in terms of cost effectiveness and sample preparation time (Li *et al.* 2018). Bar-HRM has been applied in the identification and authentication of medicinal plants, including the detection of adulterants in herbal products (Ganopoulos *et al.* 2012, Osathanunkul *et al.* 2017). The method provides straightforward results, without requiring post-PCR analysis, and it proved adequate for processed samples (Grazina *et al.* 2020). Furthermore, HRM may be used for the quantification of adulterants in food products. This can be done by making serial dilutions of the DNA sample mixed with adulterant DNA, corresponding to different percentage levels of adulterant content, and comparing it to a reference curve with known concentrations (Lagiotis *et al.* 2020). The scope of this study was the authentication of herbal products in a short-chain-of-commercialization using metabarcoding technologies, as well as enhancing the accuracy of the analysis by Bar-HRM verification.

Considerable incongruences were found between the detected species and those listed on the label of the herbal products (expected species). In the present study, we identified 131 species in the 39 samples that passed the quality criteria. The overall fidelity of the products was estimated to be 49% and most of the species occurrences are ingredients of herbal products for food and medicinal purposes. *Convolvulus arvensis* L. is one of the most abundant not-expected components found in 48% of the products. Many species of this genus are invasive weeds that climb and attach to other plants. The presence of this species in many herbal products is likely because of insufficient quality control during harvesting (Speranskaya *et al.* 2018). Another occurrence worth noticing is *Urtica dioica*, which was found in 41% of the samples. This species co-exist in the category not-expected-detected and the category expected-detected. The overlap between both categories includes plant species that are both expected in some samples but not expected (not mentioned on the labelling) in others. *Urtica dioica* was mentioned on the labelling of many products because of its medicinal properties but was also found in samples where it was not mentioned on the labelling. *Urtica dioica* is a very common “weed” and found in great numbers in and around agricultural fields. It is very likely that this species ended up accidentally in a package during harvesting or served as a filler to multiply the product by adding a cheap “harmless” ingredient. Even though all products were presumably of Greek origin, some species found were very improbable based on their geographical distribution or unlikely use. Some of the aforementioned species include: *Taraxacum alaskanum* Rydb., *Betula occidentalis* Hook., *Betula kenaica* W.H. Evans, *Achillea alpina* L., *Crataegus laevigata* (Poir.) DC., *Symphotrichum boreale* (Torr. & A.Gray) Á.Löve & D.Löve and *Polygonum boreale* Lange (Small). Even though all the steps and procedures were carefully carried out and optimised, the presence of those species in the herbal products could be due to methodology artifacts, like: (1) PCR chimeras (PCR

amplicons generated erroneously from more than one DNA template), (2) misidentification due to sequence homology, (3) database sequence misannotations, and (4) sample contaminations (such as pollen) which confirms concerns about the hypersensitivity of the metabarcoding methodology (de Boer *et al.*, 2017). We only sequenced DNA extracts with a minimum concentration of 0.6 ng/μL, thus the species should be indeed present and not misidentified due to hypersensitivity of the metabarcoding method. High throughput sequencing is very sensitive and avoiding contamination is a primary concern. It is essential that DNA extraction, library preparation and PCR reactions are carried out according to the guidelines of “Laboratory management methods of clinical gene amplification test in medical institutions”. In addition, strict experimental zoning is necessary to eliminate contamination (Li *et al.* 2018).

It is worth stressing that in this study we identified two toxic species: *Nicotiana tabacum* (tobacco) and *Chelidonium majus* (greater celandine). Both of the species were detected through metabarcoding and their presence confirmed by Bar-HRM analysis. For single ingredient products, a cut-off genotype confidence percentage (GCP) of 90% is used to assign a genotype to a barcode region in HRM analysis (Ganopoulos *et al.* 2012). We opted for a cut-off value of GCP of 80% because we amplified species-specifically in multi-species mixtures instead of single-species samples. Besides, the medicinal plants present in the herbal products contain secondary metabolites that could inhibit amplification and samples were also dried which is known to hinder DNA isolation. The results of the HRM analysis confirmed the presence of *Nicotiana tabacum* in the spice mixture sample 29 (Figure 4) and *Chelidonium majus* in the herbal mixture sample 64 (Figure 5). Tobacco is a plant with proven toxicity when ingested. The plant contains high amounts of toxic alkaloids that can cause abdominal pain, hypertension, tachycardia, and tremors in an early phase and hypotension, bradycardia, and dyspnea, finally leading to coma and respiratory failure in the second phase (Schep *et al.* 2009). Similarly, cases of hepatotoxicity have been reported following the ingestion of *Chelidonium majus* (Pantano *et al.* 2017).

The detection of both of the above species in herbal mixtures and the many incongruences found, raises concerns about the reliability and safety of the products sold in the herbal markets, i.e. short supply chains. We hereby confirm our hypothesis that the products’ fidelity is low even in short-chains-of-commercialization. Adulteration has proven to be a problem in long supply chains (Raclariu *et al.* 2017a; Raclariu *et al.* 2017b, Raclariu *et al.* 2018b). Long and complex supply chains tend to be more prone to food fraud, such as adulteration, than shorter ones (Galvin-King, 2018). Herein, we report low fidelity and the presence of toxic species in herbal products coming from a short-chain-of-commercialization. As such, we can conclude that long and short chains of commercialization are equally sensitive to fraudulent practices including misrepresentation, adulteration and substitution.

The complementary Bar-HRM analysis was used to corroborate the metabarcoding method and verify the presence of the toxic species detected, thus improving the fidelity of the metabarcoding analysis. Although DNA barcoding and metabarcoding have been proven reliable methods in authenticating a wide range of herbal products, both have not yet been implemented in controlling their quality (Ivanova *et al.* 2016, Raclariu *et al.* 2017a). Several studies have used DNA metabarcoding as a method to investigate the discrepancies between the expected and detected plant species based on label claims of marketed herbal products (Cheng *et al.* 2014; Raclariu *et al.* 2017b; Seethapathy *et al.* 2019). According to the literature, there are some possible biases associated with molecular identification, such as extraction of DNA, DNA amplification, the sequencing process and the bioinformatics analysis (Ivanova *et al.* 2016; Raclariu *et al.* 2017a; Raclariu *et al.* 2017b). In order to minimize or avoid possible PCR biases, we carefully performed and optimised the different steps and procedures to ensure accurate amplification. In addition, we

increased the reliability of the results by adopting strict quality checks during the bioinformatics analysis and by verifying some important findings with Bar-HRM.

Total species composition within complex multi-ingredient and processed mixtures can be assessed qualitatively with DNA metabarcoding. One of the main challenges in authentication of herbal mixtures is the type of the raw material used and the heavy fragmentation of DNA in the extraction substrate. Traditional medicinal or food products may contain only very low amounts of DNA or contain ingredients that have been subjected to various treatments during the production process (Staats *et al.* 2016). Therefore, we used Bar-HRM with species specific primers to verify the validity of the metabarcoding results. As it is shown herein, metabarcoding is a suitable method for identification and authentication procedures. However, verification of the AMB results is necessary given the limitations of the method in detecting traces of contaminants.

5. Conclusions

The use of DNA metabarcoding for the authentication of herbal products from local markets is able to assess species composition of complex products and provides insights into possible fraudulent practices in the supply chain. We observed a clear conflict between the listed species and the detected species of herbal products on the Greek market. The percentage of adulteration suggests irregularities even in short supply chains. Substitution could be accidental, for example contamination during harvesting, contaminated bags and equipment or at any step of the manufacturing process. Nevertheless, fraud or adulteration could also be the case. However, we should focus on the detected species rather than on the failure of detecting species, as the supply chain includes many steps that might lead to degradation or loss of DNA beyond detectable limits, e.g., alcoholic extraction, and drying of material. The present study reveals a need for stricter quality control of herbal products, especially because of the presence of toxic species. Low fidelity was found to be an issue, also in short-chains-of-commercialization. DNA metabarcoding was shown in this study to be a promising tool for quality control and the authentication of complex herbal products. The detection of toxic species was further confirmed by Bar-HRM, which was proven here to be an efficient tool in verifying the results of the metabarcoding analysis. Nevertheless, for the authentication of complex samples, like herbal mixtures, metabarcoding remains the method of choice. In order to obtain a complete quality assessment, high throughput sequencing methods should be combined with Bar-HRM using species-specific primers as a verification tool. However, standardization of protocols is necessary before DNA metabarcoding can be implemented as a routine analytical tool and approved by competent authorities for use in regulatory frameworks. Last but not least, the availability of a rich and reliable DNA reference database is a prerequisite for successful implementation.

6. Acknowledgments

This work was supported by H2020-MSCA-ITN-ETN Plant.ID which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 765000.

7. References

- 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. (2010). Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species, *PLoS One*, 7, 5(1). <https://doi.org/10.1371/journal.pone.0008613>
- Cheng, X., Su, X., Chen, X., Zhao, H., Bo, C., Xu, J., Bai, H., Ning, K. (2014). Biological ingredient analysis of traditional Chinese medicine preparation based on high-throughput sequencing: the story for Liuwei Dihuang Wan, *Scientific Reports*, 4, 1–12. <https://doi.org/10.1038/srep05147>
- Coghlan, M. L., Haile, J., Houston, J., Murray, D. C., White, N. E., Moolhuijzen, P., Bellgard, M.I., Bunce, M. (2012). Deep sequencing of plant and animal DNA contained within traditional Chinese medicines reveals legality issues and health safety concerns, *PLoS Genetics*, 8(4). <https://doi.org/10.1371/journal.pgen.1002657>
- de Boer, H. J., Ichim, M. C., & Newmaster, S. G. (2015). DNA Barcoding and Pharmacovigilance of Herbal Medicines, *Drug Safety*, 38(7), 611–620. <https://doi.org/10.1007/s40264-015-0306-8>
- de Boer, H. J., Ghorbani, A., Manzanilla, V., Raclariu, A. C., Kreziou, A., Ounjai, S., Osathanunkul, M., Gravendeel, B. (2017). DNA metabarcoding of orchid-derived products reveals widespread illegal orchid trade. *Proceedings of the Royal Society B: Biological Sciences*, 284(1863). <https://doi.org/10.1098/rspb.2017.1182>
- Dimopoulos, P., Raus, T., Bergmeier, E., Constantinidis, T., Iatrou, G., Kokkini, S., Strid, A., Tzanoudakis, D. (2013). Vascular Plants of Greece: An Annotated Checklist. Vol 31. Berlin: Botanic Garden and Botanical Museum Berlin-Dahlem; Athens: Hellenic Botanical Society.
- Dimopoulos, P., Raus, T., Bergmeier, E., Constantinidis, T., Iatrou, G., Kokkini, S., Strid, A., Tzanoudakis, D. (2016). Vascular Plants of Greece: An Annotated Checklist. Supplement. *Willdenowia* 46 (3): 301–47. <https://doi.org/10.3372/wi.46.46303>
- Galvin-King, P., Haughey, S.A., Elliott, C.T. (2018). Herb and Spice Fraud; the Drivers, Challenges and Detection. *Food Control*, 88, 85–97. <https://doi.org/10.1016/j.foodcont.2017.12.031>.
- Ganopoulos, I., Madesis, P., Darzentas, N., Argiriou, A., & Tsiftaris, A. (2012). Barcode High Resolution Melting (Bar-HRM) analysis for detection and quantification of PDO “fava Santorinis” (*Lathyrus clymenum*) adulterants. *Food Chemistry*, 133(2), 505–512. <https://doi.org/10.1016/j.foodchem.2012.01.015>
- Garzo, C.F., Gómez P.P., Barrasa, A.B., Martínez, R.A., Ramírez, R.F., Ramón, F.R. (2002). Cases of Neurological Symptoms Associated with Star Anise Consumption Used as a Carminative. *Anales Espanoles de Pediatria*, 57(4), 290–94.
- Grazina, L., Amaral, J.S., Mafra, I. (2020) Botanical origin authentication of dietary supplements by DNA-based approaches. *Comprehensive Reviews in Food Science and Food Safety*, 19(3). <https://doi.org/10.1111/1541-4337.12551>
- Hanlidou, E., Karousou, R., Kleftoyanni, V., Kokkini, S. (2004). The Herbal Market of Thessaloniki (N Greece) and Its Relation to the Ethnobotanical Tradition. *Journal of Ethnopharmacology*, 91(2–3), 281–99. <https://doi.org/10.1016/j.jep.2004.01.007>

- Hunter, J.D., (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, 9(3), 90-95, 2007
- Ipsos MORI, (2008). Public Perceptions of Herbal Medicines General Public Qualitative & Quantitative Research, (November), 1–102.
- Ivanova, N. V., Kuzmina, M. L., Braukmann, T. W. A., Borisenko, A. V., Zakharov, E. V. (2016). Authentication of herbal supplements using next-generation sequencing. *PLoS ONE*, 11(5), 1–24. <https://doi.org/10.1371/journal.pone.0156426>
- Johanns, E.S, Sijben, A.E., Peters, P.W. (2002). An epidemic of Epileptic Seizures after Consumption of Herbal Tea. *Nederlands Tijdschrift voor Geneeskunde*, 146(17), 813-16
- Kokkini, S., Karagiannakidou, V., Hanlidou, E., Vokou, D. (1988). Geographical and Altitudinal Distribution of the Lamiaceae in Greece. *Phyton-Annales Rei Botanicae*, 28(2), 215–228.
- Lagiotis, G., Stavridou, E., Bosmali, I., Osathanunkul, M., Haider, N., Madesis, P. (2020). Detection and quantification of cashew in commercial tea products using High Resolution Melting (HRM) analysis. *Journal of Food Science*, 85(6), 1629-34. <https://doi.org/10.1111/1750-3841.15138>
- Li, Q., Sun, Y., Guo, H., Sang, F., Ma, H., Peng, H., Zheng, N., Xu, L. (2018). Quality control of the traditional Chinese medicine Ruyi Jinhuang powder based on high-throughput sequencing and real-time PCR. *Scientific Reports*, 8 (1), 8261. doi: 10.1038/s41598-018-26520-3
- Lo, Y.T., Shaw, P.C (2019). Application of next-generation sequencing for the identification of herbal products. *Biotechnology Advances*, 37. <https://doi.org/10.1016/j.biotechadv.2019.107450>
- Lupien, J. R. (2005). Food quality and safety: Traceability and labeling. *Critical Reviews in Food Science and Nutrition*, 45(2), 119–123. <https://doi.org/10.1080/10408690490911774>
- Lynch, N., Berry, D. (2007). Differences in perceived risks and benefits of herbal, over-the-counter conventional, and prescribed conventional, medicines, and the implications of this for the safe and effective use of herbal products. *Complementary Therapies in Medicine*, 15(2), 84–91. <https://doi.org/10.1016/j.ctim.2006.06.007>
- Marieschi, M., Torelli, A., Poli, F., Sacchetti, G., Bruni, R. (2009). RAPD-Based method for the quality control of mediterranean oregano and its contribution to pharmacognostic techniques. *Journal of Agricultural and Food Chemistry*, 57(5), 1835–1840. <https://doi.org/10.1021/jf8032649>
- Newmaster, S.G, Grguric M., Shanmughanandhan D., Ramalingam S., Ragupathy, S. (2013). DNA Barcoding detects contamination and substitution in North American herbal products. *BMC Medicine*, 11(1), 222. <https://doi.org/10.1016/j.sajb.2017.01.162>
- Nielsen (2015). <http://www.nielsen.com/uk/en.html/> Accessed 20 October 2019.
- Nutrition Business Journal (2016), Supplement Business Report 2016, Penton, New York, NY.
- Osathanunkul, M., Suwannapoom, C., Osathanunkul, K., Madesis, P., de Boer, H. (2016). Evaluation of DNA barcoding coupled high resolution melting for discrimination of closely related species in phytopharmaceuticals. *Phytomedicine*, 23(2), 156–165. <https://doi.org/10.1016/j.phymed.2015.11.018>

- Osathanunkul, M., Ounjai, S., Osathanunkul, R., Madesis, P. (2017). Evaluation of a DNA-based method for spice/herb authentication, so you do not have to worry about what is in your curry, buon appetito! *PLoS ONE*, 12(10). <https://doi.org/10.1371/journal.pone.0186283>
- Osman, A. G., Raman, V., Haider, S., Ali, Z., Chittiboyina, A. G., Khan, I. A. (2019). Overview of analytical tools for the identification of adulterants in commonly traded herbs and spices. *Journal of AOAC International*, 102(2), 376–385. <https://doi.org/10.5740/jaoacint.18-0389>
- Ouarghidi, A., Powell, B., Martin, G. J., de Boer, H., Abbad, A. (2012). Species Substitution in Medicinal Roots and Possible Implications for Toxicity of Herbal Remedies in Morocco. *Economic Botany*, 66(4), 370–382. <https://doi.org/10.1007/s12231-012-9215-2>
- Pantano, F., Mannocchi, G., Marinelli, E., Gentili, S., Graziano, S., Busardò, F. P., Di Luca, N. M. (2017). Hepatotoxicity induced by greater celandine (*Chelidonium majus* L.): A review of the literature. *European Review for Medical and Pharmacological Sciences*, 21(1), 46–52.
- Parveen, I., Gafner, S., Tehen, N., Murch, S. J., Khan, I. A. (2016). DNA barcoding for the identification of botanicals in herbal medicine and dietary supplements: Strengths and limitations. *Planta Medica*, 82(14), 1225–1235. <https://doi.org/10.1055/s-0042-111208>
- Peter, K. V. (2012). *Handbook of herbs and spices*. (2nd ed.). Elsevier, (Chapter 1)
- POWO (2019). Plants of the World Online. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet; <http://www.plantsoftheworldonline.org/> Retrieved 20 September 2019.
- Raclariu, A.C., Mocan, A., Popa, M.O., Vlase, L., Ichim, M.C., Crisan, G., Brysting, A.K., de Boer, H. (2017a). *Veronica officinalis* product authentication using DNA metabarcoding and HPLC-MS reveals widespread adulteration with *Veronica chamaedrys*. *Frontiers in Pharmacology*, 8, 1–13. <https://doi.org/10.3389/fphar.2017.00378>
- Raclariu, A.C., Paltinean, R., Vlase, L., Labarre, A., Manzanilla, V., Ichim, M.C., Crisan, G., Brysting, A.K., de Boer, H. (2017b). Comparative authentication of *Hypericum perforatum* herbal products using DNA metabarcoding, TLC and HPLC-MS. *Scientific Reports*, 7(1), 8–10. <https://doi.org/10.1038/s41598-017-01389-w>
- Raclariu, A.C., Heinrich, M., Ichim, M.C., de Boer, H. (2018a). Benefits and Limitations of DNA Barcoding and Metabarcoding in Herbal Product Authentication. *Phytochemical Analysis*, 29(2), 123–128. <https://doi.org/10.1002/pca.2732>
- Raclariu, A.C., Tebrencu, C.E., Ichim, M.C., Ciupercă, O.T., Brysting, A.K., de Boer, H. (2018b). What's in the Box? Authentication of *Echinacea* Herbal Products Using DNA Metabarcoding and HPTLC. *Phytomedicine*, 44, 32–38. <https://doi.org/10.1016/j.phymed.2018.03.058>
- Reed, G.H., Wittwer, C.T. (2004). Sensitivity and specificity of single-nucleotide polymorphism scanning by high-resolution melting analysis. *Clinical Chemistry*, 50(10), 1748–1754. <https://doi.org/10.1373/clinchem.2003.029751>
- Schep, L.J., Slaughter, R.J., Beasley, D.M.G. (2009). Nicotinic plant poisoning. *Clinical Toxicology*, 47(8), 771–781. <https://doi.org/10.1080/15563650903252186>

- Seethapathy, G.S., Raclariu, A.C., Anmarkrud, J.A., Wangenstein, H., de Boer, H. (2019). DNA Metabarcoding Authentication of Ayurvedic Herbal Products on the European Market Raises Concerns of Quality and Fidelity. *Frontiers in Plant Science*, 10, 1–11. <https://doi.org/10.3389/fpls.2019.00068>
- Sgamma, T., Lockie-Williams, C., Kreuzer, M., Williams, S., Scheyhing, U., Koch, E., Slater, A., Howard, C. (2017). DNA Barcoding for Industrial Quality Assurance. *Planta Medica*, 83(14–15), 1117–1129. <https://doi.org/10.1055/s-0043-113448>
- Smith, T., Kawa, K., Eckl, V., Morton, C., Stredney, R. (2018). Herbal supplement sales in US increase 8.5% in 2017, topping \$8 billion. *HerbalGram*, 119, 62–71.
- Solomou, A.D., Martinos, K., Skoufogianni, E., Danalatos, N.G. (2016). Medicinal and Aromatic Plants Diversity in Greece and Their Future Prospects: A Review. *Agricultural Science*, 4(1), 9–20. <https://doi.org/10.12735/as.v4i1p09>
- Speranskaya, A.S., Khafizov, K., Ayginin, A.A., Krinitsina, A.A., Omelchenko D.O., Nilova, M.V., Severova, E.E., Samokhina, E.N., Shipulin, G.A., Logacheva, M.D. (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. <https://doi.org/10.1016/j.foodcont.2018.04.040>
- Staats, M., Arulandhu, A.J., Gravendeel, B., Holst-Jensen, A., Scholtens, I., Peelen, T., Prins, T.W., Kok, E. (2016). Advances in DNA metabarcoding for food and wildlife forensic species identification. *Analytical and Bioanalytical Chemistry*, 408(17), 4615–4630. <https://doi.org/10.1007/s00216-016-9595-8>
- Stefanou, P., Baloutas, D., Katsinikas, D., Abraham, E.M., Kyriazopoulos, A.P., Parissi, Z.M., Arabatzis, G. (2014). Cultivation and production of aromatic plants in Greece: present situation, possibilities and prospects. In *Proceedings of the 8th Panhellenic Rangeland Congress, Thessaloniki* (pp. 1–3).
- Sun, Y., Skinner, D. Z., Liang, G. H., 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., 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
- Thongkhao, K., Tungphatthong, C., Phadungcharoen, T., Sukrong, S. (2020). The use of plant DNA barcoding coupled with HRM analysis to differentiate edible vegetables from poisonous plants for food safety. *Food Control*, 109. <https://doi.org/10.1016/j.foodcont.2019.106896>
- World Health Organisation (WHO) (2013). *WHO traditional medicine strategy 2014–2023*. 2013. Geneva: World Health Organization.
- Yang, Z., Rannala, B. (2017). Bayesian species identification under the multispecies coalescent provides significant improvements to DNA barcoding analyses. *Molecular Ecology*, 26(11), 3028–3036. <https://doi.org/10.1111/mec.14093>
- Zhang, J., Wider, B., Shang, H., Li, X., Ernst, E. (2012). Quality of herbal medicines: Challenges and solutions. *Complementary Therapies in Medicine*, 20(1–2), 100–106. <https://doi.org/10.1016/j.ctim.2011.09.004>

8. Figures

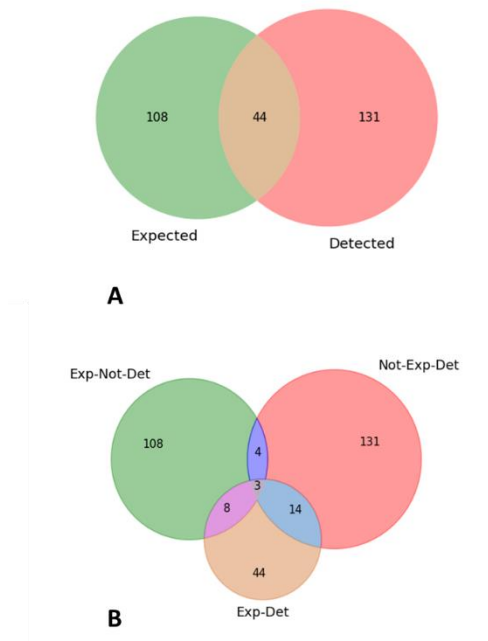


Figure 1. Venn diagrams representing the detection of individual species via Metabarcoding in all the samples. **(A)** Expected-Detected two ways Venn diagram. Comparison of the expected as labelled on the products analysed organisms (green group) and the detected via Metabarcoding individual species (red group). Only 44 expected species were detected (brown subgroup). **(B)** Expected-Detected three ways Venn diagram. More detailed representation of the number of species that were found in different groups between different samples. Green and red groups represent single species classified as expected-not-detected and not-expected-detected, respectively. The brown group represents the expected-detected species. The blue-shaded subgroups represent the number of species classified in different groups between different samples. For example, the dark blue subgroup includes species classified as both expected-not detected and not-expected-detected species through the total number of samples. The rest of the subgroups similarly represent the number of species classified as both expected-not-detected and expected-detected (purple subgroup), or as both not-expected-detected and expected-detected (light blue subgroup), or as expected-not-detected, not-expected-detected and expected-detected (grey subgroup).

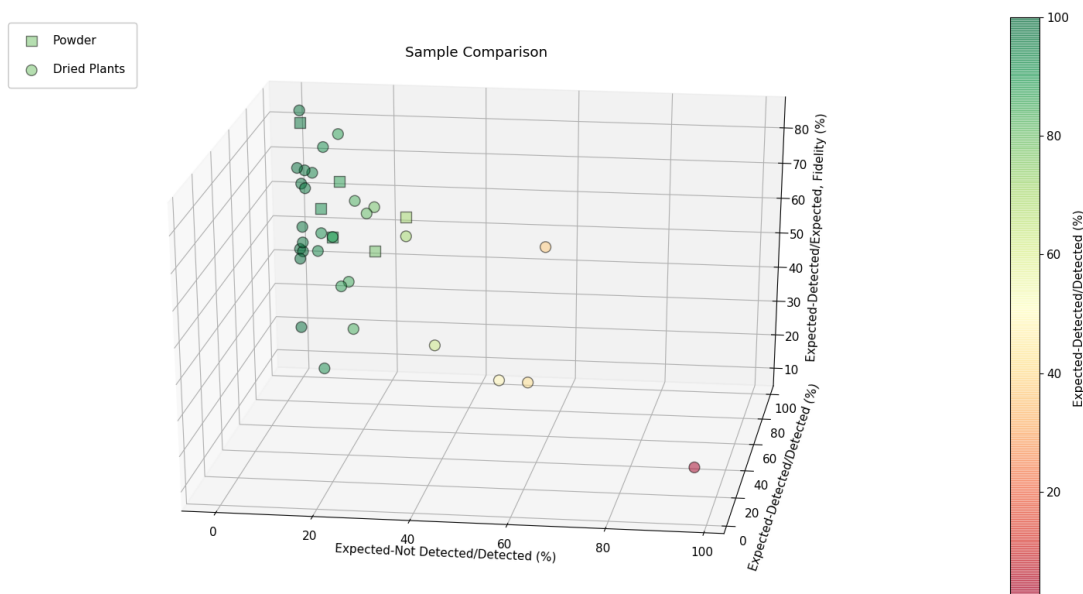


Figure 2. 3D dot plot representing the ingredient fidelity of the successfully analysed samples. Each symbol represents a single sample, different symbols represent different types of products (circle: dried plant parts, square: powder). The fidelity (y axes) is represented as the percentage of the number of the expected-detected species out of the total expected per sample species, as labelled on the products analysed. In parallel, the sequence reads identified via Metabarcoding and analysed by Blast were classified as belonging to either not-expected-detected species or expected-detected species and their percentage per sample is here presented on the x, z axes for the not-expected-detected and the expected-detected species, respectively. The percentage of the expected-detected species out of the total detected per sample species is also presented with the green to red colour code which follows the colour code of the rest of the analyses (green: detection of expected species, red: detection of not expected species).

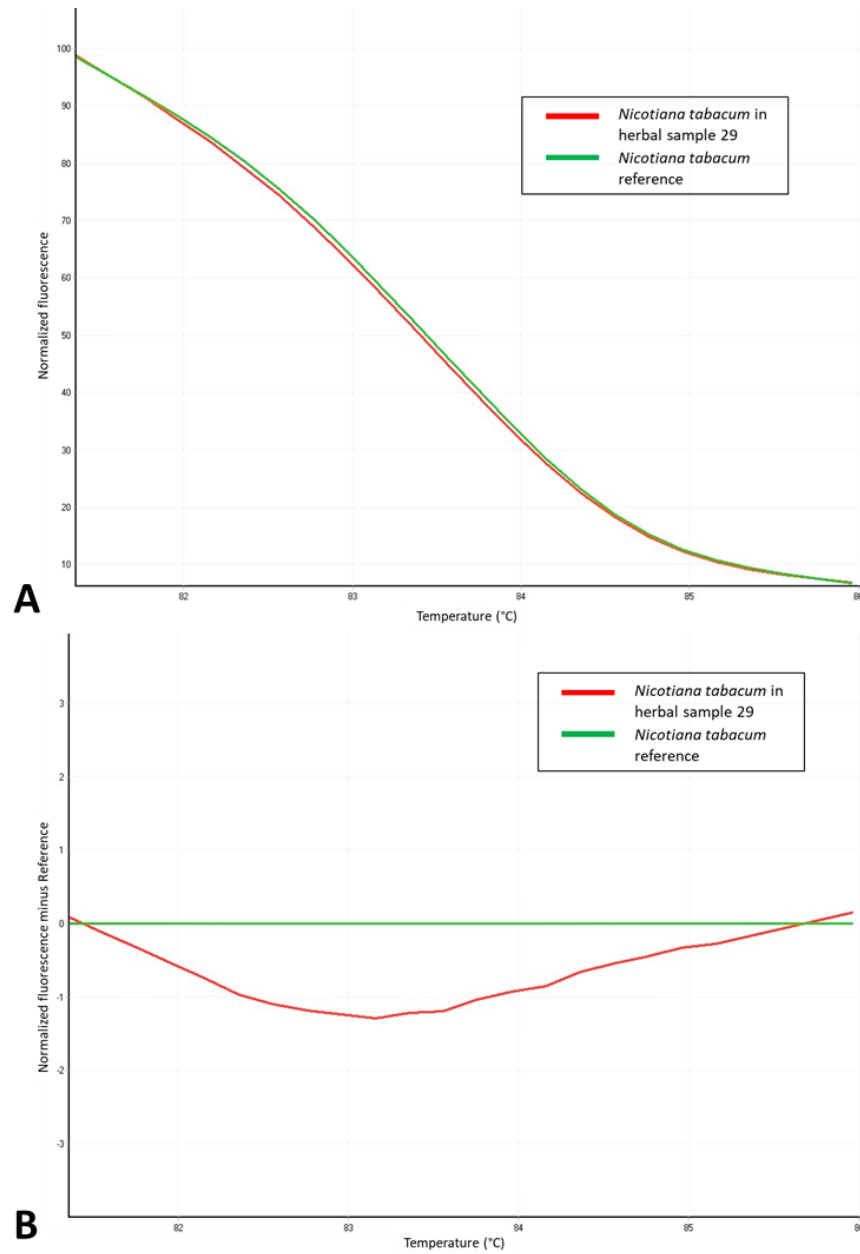


Figure 4. Melting profiles of *Nicotiana tabacum* in spice mixture (sample 29) (A) Normalized fluorescence plot of barcoding coupled high resolution melting (Bar-HRM) (B) Reference-corrected normalized fluorescence plot of barcoding coupled high resolution melting (Bar-HRM).

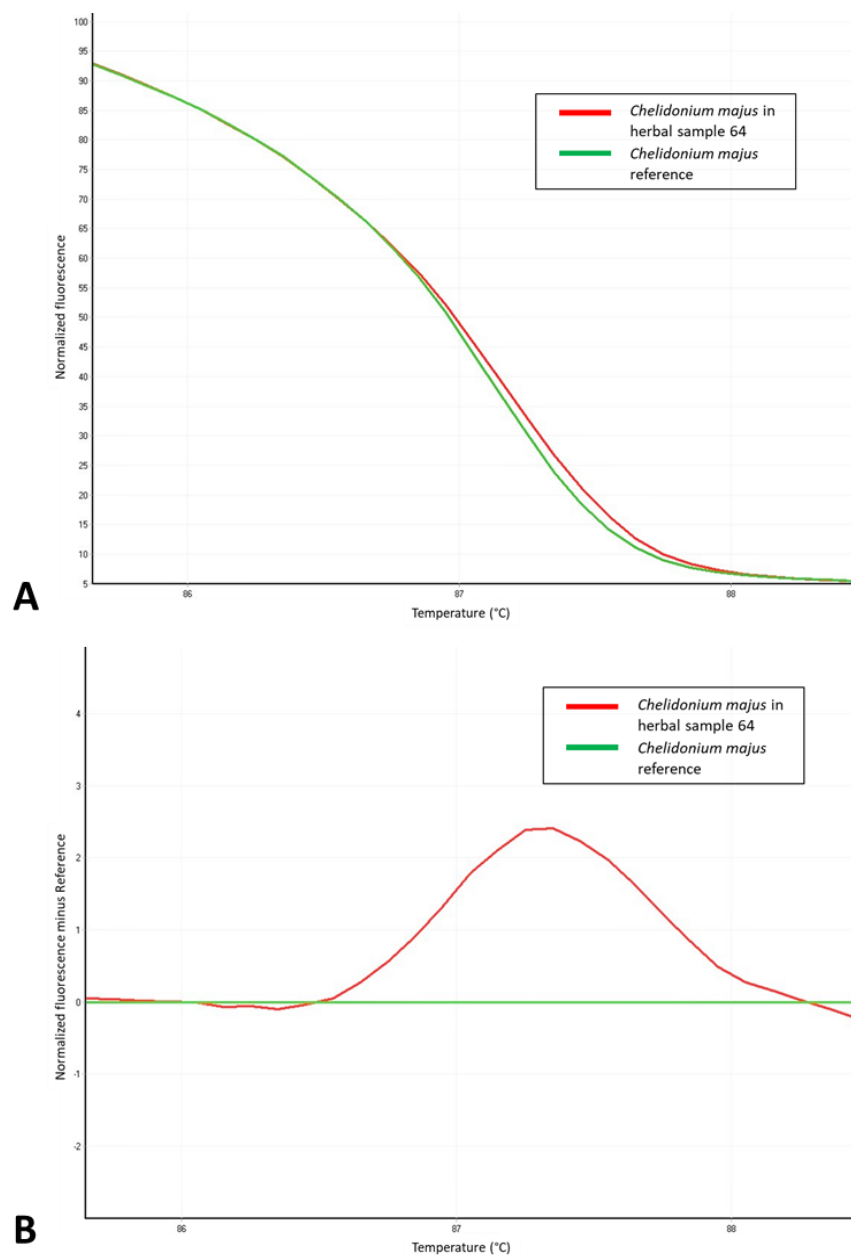


Figure 5. Melting profiles of *Chelidonium majus* in herbal mixture (sample 64) (A) Normalized fluorescence plot of barcoding coupled high resolution melting (Bar-HRM) (B) Reference-corrected normalized fluorescence plot of barcoding coupled high resolution melting (Bar-HRM).

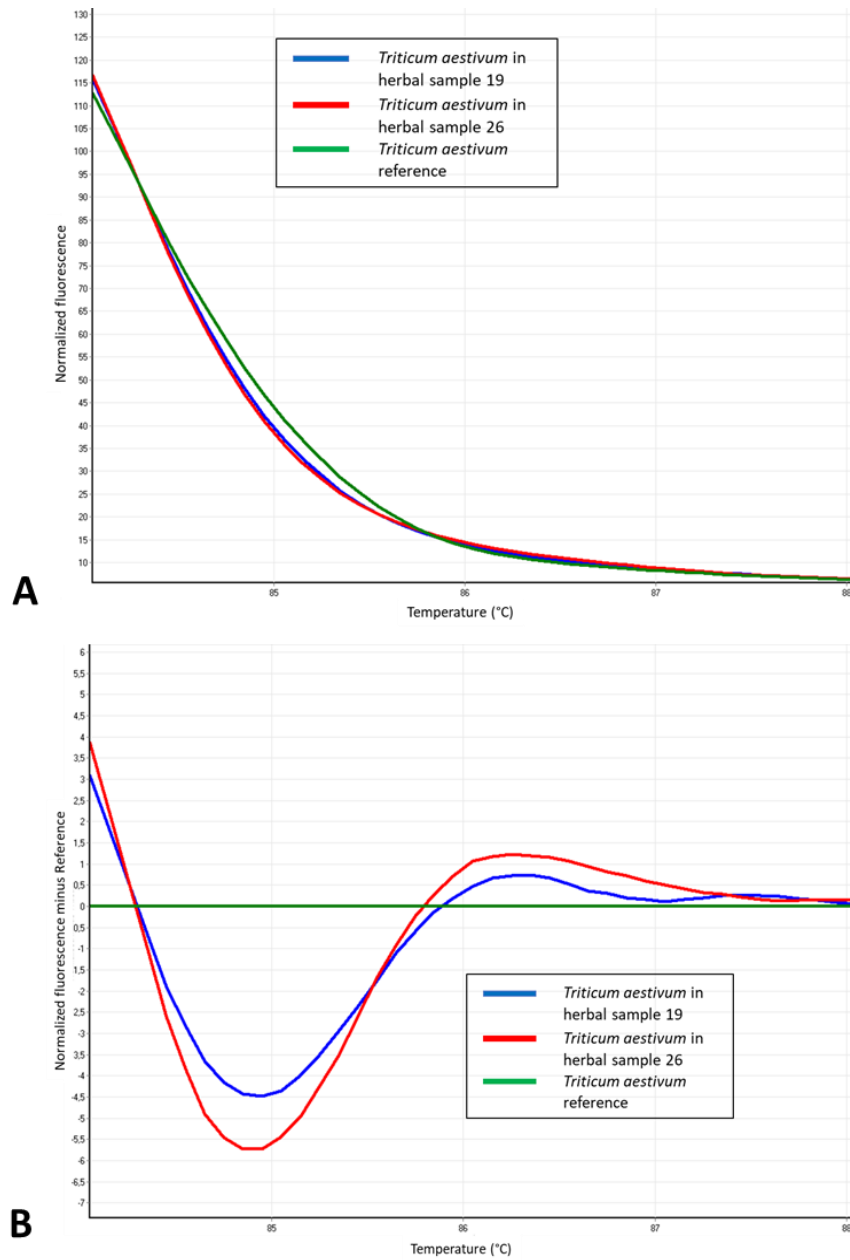


Figure 6. Melting profiles of *Triticum aestivum* in spice mixtures (sample 19 and 26) (A) Normalized fluorescence plot of barcoding coupled high resolution melting (Bar-HRM) (B) Reference-corrected normalized fluorescence plot of barcoding coupled high resolution melting (Bar-HRM).

Tables

Table 1. Top ten products with the highest fidelity and their level of adulteration.

| Herbal product code | Product type | Number of species on label | Species detected through AMB | Detected species from label | Fidelity | Detected species not on product label | Adulteration |
|---------------------|-------------------|----------------------------|------------------------------|-----------------------------|----------|---------------------------------------|--------------|
| 62 | Dried plant parts | 6 | 6 | 5 | 83.3 | 1 | 16.7 |
| 55 | Dried plant parts | 8 | 12 | 6 | 75.0 | 6 | 50.0 |
| 16 | Powder | 6 | 15 | 4 | 66.7 | 11 | 73.3 |
| 41 | Dried plant parts | 8 | 7 | 5 | 62.5 | 2 | 28.6 |
| 53 | Dried plant parts | 8 | 7 | 5 | 62.5 | 2 | 28.6 |
| 4 | Powder | 8 | 26 | 5 | 62.5 | 21 | 80.8 |
| 58 | Dried plant parts | 13 | 9 | 8 | 61.5 | 1 | 11.1 |
| 5 | Powder | 7 | 21 | 4 | 57.1 | 17 | 81.0 |
| 49 | Dried plant parts | 11 | 6 | 5 | 45.5 | 1 | 16.7 |
| 54 | Dried plant parts | 13 | 11 | 5 | 38.5 | 6 | 54.5 |

Table 2. Top ten products with the highest adulteration and their level of fidelity.

| Herbal product code | Product type | Number of species on label | Species detected through AMB | Detected species from label | Fidelity | Detected species not on product label | Adulteration |
|---------------------|-------------------|----------------------------|------------------------------|-----------------------------|----------|---------------------------------------|--------------|
| 56 | Dried plant parts | 8 | 21 | 2 | 25.0 | 19 | 90.5 |
| 15 | Dried plant parts | 11 | 25 | 3 | 27.3 | 22 | 88.0 |
| 3 | Powder | 6 | 22 | 3 | 50.0 | 19 | 86.4 |
| 38 | Dried plant parts | 7 | 12 | 2 | 28.6 | 10 | 83.3 |
| 5 | Powder | 7 | 21 | 4 | 57.1 | 17 | 81.0 |
| 4 | Powder | 8 | 26 | 5 | 62.5 | 21 | 80.8 |
| 57 | Dried plant parts | 6 | 15 | 3 | 50.0 | 12 | 80.0 |
| 16 | Powder | 6 | 15 | 4 | 66.7 | 11 | 73.3 |
| 18 | Powder | 5 | 13 | 4 | 80.0 | 9 | 69.2 |
| 54 | Dried plant parts | 13 | 11 | 5 | 38.5 | 6 | 54.5 |