DNA Barcoding for Industrial Quality Assurance

Authors

Tiziana Sgamma[1], Claire Lockie-Williams[2], Marco Kreuzer[2,4], Sarah Williams[1], Ulrike Scheyhing[3], Egon Koch[3], Adrian Slater[1], Caroline Howard[1,2]

Affiliations

1 Biomolecular Technology Group, School of Allied Health Sciences, De Montfort University, Leicester, UK
2 BP-NIBSC Herbal Laboratory, National Institute for Biological Standards and Controls, Potters Bar, UK
3 Dr. Willmar Schwabe Pharmaceuticals, Karlsruhe, Germany
4 Section of Ecology and Evolutionary Biology, School of Biological Sciences, University of Reading, Reading, UK

Key words

authentication, *Rhodiola rosea*, *Phyllanthus amarus*, *Ocimum tenuiflorum*, next-generation sequencing, metabarcoding, Crassulaceae, Lamiaceae

received March 10, 2017
revised May 16, 2017
accepted May 27, 2017

Bibliography

DOI https://doi.org/10.1055/s-0043-113448
Published online June 29, 2017 | Planta Med 2017; 83: 1117–1129 © Georg Thieme Verlag KG Stuttgart · New York | ISSN 0032-0943

Correspondence

Professor Adrian Slater
Biomolecular Technology Group, School of Allied Health Sciences, Faculty of Health and Life Sciences, Hawthorn Building
The Gateway, Leicester LE1 9BH, UK
Phone: + 44 11 62 57 78 43, Fax: + 44 11 62 01 39 25
ads@dmu.ac.uk

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

ABSTRACT

DNA barcoding methods originally developed for the identification of plant specimens have been applied to the authentication of herbal drug materials for industrial quality assurance. These methods are intended to be complementary to current morphological and chemical methods of identification. The adoption of these methods by industry will be accelerated by the introduction of DNA-based identification techniques into regulatory standards and monographs. The introduction of DNA methods into the British Pharmacopoeia is described, along with a reference standard for use as a positive control for DNA extraction and polymerase chain reaction (PCR). A general troubleshooting chart is provided to guide the user through the problems that may be encountered during this process. Nevertheless, the nature of the plant materials and the demands of industrial quality control procedures mean that conventional DNA barcoding is not the method of choice for industrial quality control. The design of DNA barcode-targeted quantitative PCR and high resolution melt curve tests is one strategy for developing rapid, robust, and reliable protocols for high-throughput screening of raw materials. The development of authentication tests for wild-harvested *Rhodiola rosea* L. is used as a case study to exemplify these relatively simple tests. By way of contrast, the application of next-generation sequencing to create a complete profile of all the biological entities in a mixed herbal drug is described and its potential for industrial quality assurance discussed.

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

Introduction

There has been growing interest in the use of DNA-based methods for the identification of medicinal plants and authentication of herbal products. Several recent reviews have provided a comprehensive overview of the background literature to DNA barcoding [1, 2], DNA technologies [3], authentication tests [4, 5], and the detection of adulteration in commercial products [1, 3, 5]. This review focuses on the practical implications of incorporating DNA tests into the quality-assurance procedures of the herbals industry. A major aim is to give practical guidance on the implementation of DNA tests in terms of infrastructure and procedures, data analysis, and troubleshooting, by way of selected case studies. It also provides confirmation that, despite the recent controversy around the misapplication of DNA authentication methods to processed herbal products [6, 7], DNA-based technology is a valuable addition to the toolkit of industrial quality assurance.

Quality Assurance of Herbal Medicinal Products

The medicinal status and regulation of herbal medicines vary considerably in different parts of the world [8–10]. In some countries

ABBREVIATIONS

BP	British Pharmacopoeia
BPNARM	British Pharmacopoeia Nucleic Acid Reference Material
dPCR	digital PCR
EFSA	European Food Standards Authority
Ph.Eur	European Pharmacopoeia
HPTLC	high performance thin layer chromatography
HRM	high resolution melt curve analysis
ITS	internal transcribed spacer
MIQE	minimum information for publication of quantitative real-time PCR experiments
NGS	next-generation sequencing
nrITS	nuclear ribosomal internal transcribed spacer
qPCR	quantitative real-time PCR
SNP	single nucleotide polymorphism
THM	traditional herbal medicine

such as the United States, all herbal products are treated as food supplements [11, 12], while elsewhere they are considered to be medicines or are treated as a distinct category of medicinal product [8, 13]. It is also still the case that many countries lack a regulatory framework to control the quality and safety of herbal products [8].

This review focuses on the regulation of herbal medicines in the European Union (EU). Legislation in the EU recognizes three categories of herbal product.

Well-established herbal medicinal products (HMPs) are regulated by Directive 2001/83/EC and subsequent amendments. This requires that all medicinal products (including HMPs) can be marketed in the EU only when they have obtained a marketing authorization. Authorization is dependent on the specification of pharmaceutical quality, safety, and efficacy information, along with the medicinal product name, pharmaceutical form, indication, dosage, and risk information [14].

The Traditional Herbal Medicines Directive 2004/24/EC amendment to Directive 2001/83/EC aimed to simplify the procedure for THMs by allowing for registration of THMs shown to have been used traditionally for 30 years including at least 15 years within the EU [8, 15]. It also requires that safety data are provided, though this can be literature-based, and that the producer can guarantee the quality of its product with reference to good manufacturing process [10, 15, 16] and the World Health Organization (WHO) Good Agricultural and Collection Practices for Medicinal Plants, which spell out requirements for species identification, collection practices, and cultivation of medicinal plant species [17]. The claims on the labels of medicinal plant products also fall under this legislation and will become standardized for therapeutic claims and safety information [14, 15].

Thus, HMPs can be sold in the EU either with a marketing authorization or with THM registration. Marketing authorizations require, among other measures, that the product be proved to be efficacious. This can either be shown in new clinical trials for safety and efficacy or via a "bibliographic application" in which the active compound of the medicinal plant in question can be shown to have been well established in the EU for 10 years [14, 15]. If the therapeutic efficacy of a herbal preparation is not fully proved, a THMP (traditional herbal medicinal product) registration may be obtained. This requires that the product proves not to be harmful in the specified conditions of use and the pharmacological effects or efficacy are plausible on the basis of long-standing use and experience [14].

A complex boundary exists between herbal medicines and food supplements (botanicals), which are regulated by the EFSA under the Food Supplements Directive (2002/46/EC) [12, 14]. A herbal medicine is defined as a medicinal product intended for treating or preventing disease. In contrast, a food supplement, which is a concentrated source of a substance(s) with a nutritional or physiological effect and is sold in dose form, is described as a product designed to supplement the normal diet. The main distinction between these two definitions is that only medicinal products can claim to be used for the treatment and prevention of disease. However, it is possible that the same substance could be sold under either scheme, depending on whether the producer wishes to make a health claim for the product. These definitions refer only to EU regulations and differ in other jurisdictions, presenting a difficult challenge for harmonization of herbal product quality and safety regulation [11, 12, 18].

One of the key requirements of herbal medicine market authorization or THM registration is that the quality standards are defined. For well-established medicines, the quality standards are detailed in monographs of the Ph.Eur under the *aegis* of the European Directorate for the Quality of Medicines and Healthcare. Monographs for registered THMPs are gradually being incorporated into the Ph.Eur. The Committee on Herbal Medicinal Products of the European Medicines Agency produces community herbal monographs for THMPs that contain safety data but not quality standards [11].

Herbal products described in Ph.Eur monographs must be prepared in accordance with the published quality standards. These prescribe benchmarks for identification and authentication of plant material upstream of manufacturing and processing and standards for purity. These include tests for loss on drying, water and total ash content, and pesticide, heavy metal, and microbial contamination. Specific toxins and radioactive contamination testing may be stipulated in certain circumstances. A test for foreign matter is also required. The material is typically sampled and sorted by eye for the presence of foreign matter, which is weighed and given as a percentage of the total. The acceptable foreign matter threshold is usually set at 2% (i.e., 98% purity), but different levels may be authorized for individual products. The authentication of the plant material is generally based on microscopic and macroscopic botanical identification and on simple chemical assays such as HPTLC (high performance thin layer chromatography) and HPLC (high performance liquid chromatography) [9, 19].

With regard to herbal food supplements, the EFSA has published guidance for the safety assessment of botanicals and derived preparations that are intended for use in food supplements [20]. The recommendations for identification and product specification are to follow the Ph.Eur standards where possible. A com-

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

pendium of botanicals that have been reported to contain substances that may be of health concern when used in food or food supplements has also been published by the EFSA and is subject to regular updates [21].

Introduction of DNA Testing into the BP

Within the EU, national bodies are responsible for the implementation of community HMPs regulations. For example, in the United Kingdom, the Medicines and Healthcare Products Regulatory Agency (MHRA) administers the THMP registration scheme and is also responsible for production of the BP. The monographs of the BP are harmonized with the Ph.Eur, but additional monographs that are not found in the Ph.Eur may be included in the BP.

In 2016 the BP published a new appendix method: "Deoxyribonucleic acid (DNA) based identification techniques for herbal drugs" [22]. The method sets out the basic requirements for molecular identification that are applicable to any investigation. General procedures are given for DNA extraction, polymerase chain reaction (PCR), gel electrophoresis, and DNA sequencing and analysis. The second section, now listed in a supplementary chapter [23], details the steps that are particular to the monograph species. For each subject, information is given on DNA purification requirements, PCR primer use, and most importantly the reference sequence.

The initial subject for this identification method was holy basil, defined in the monograph as *Ocimum tenuiflorum* L., Lamiaceae. The first point to be considered was sampling. BP monographs are elaborated by analyzing material that is traded in the United Kingdom. This is based on the practice for allopathic pharmaceutical drugs that have a specific chemical formula that can be directly measured and industrial production processes. Herbal drugs are less uniform and are ideally assessed in their "raw" form, making sample size and variety all the more important. The sampling methods and results used in this study are described in more detail in the Supporting Information.

The next step in the process was barcode region selection. Initially, five regions were amplified, sequenced, and assessed for useful sequence variation–variation that distinguished the commercial samples from different *Ocimum* species. The sequence analysis of the plastid *trnH-psbA* region revealed variation that was suitable in terms of both number and size of nucleotide polymorphisms, and this result was also informed by published literature [24].

Five distinct regions of sequence variation were highlighted as "key bases" for identification using a reference sequence (**Fig. 1S**, Supporting Information). Along the length of the ~ 500 bp region, the first three regions are between 200 and 250 bp (**Fig. 2S**, Supporting Information), and the final two were between 300 and 370 bp (**Fig. 3S**, Supporting Information). The key bases were selected with care to ensure that repeat regions would not cause false positive matches by "anchoring" the regions with non-variable bases at either end.

The result of this project is a thoroughly tested DNA extraction, purification, and amplification method for holy basil herbal drug material. Further, a reference sequence is published that can be used to identify *O. tenuiflorum* and distinguish it from closely related species known to contaminate holy basil herbal material. The method and structure that was designed and implemented has formed a model for high quality molecular characterization that can be used with different species. Fundamentally, the data generated from the DNA-based identification work were used in conjunction with results from macroscopic and microscopic analysis, foreign matter testing, loss on drying, acid insoluble ash, and HPTLC to ensure that only information from genuine and good quality samples be used to produce the benchmarks in the BP.

Introduction of DNA Tests into Quality Control Protocols

The appearance of DNA-based test quality standards indicates that industrial quality control laboratories hitherto focused on analytical chemistry may need to adapt to the introduction of molecular technologies. The next section considers some of the practical issues involved in developing a DNA testing facility within an industrial setting.

When establishing a DNA barcoding facility in existing laboratory space, there are a few simple rules:

1. Routine molecular biology can co-exist with other analytical techniques–there is no necessity to physically separate DNA work from analytical chemistry work, for example.
2. It is, however, essential to separate some DNA techniques from others. Putting aside a single room for all DNA work could be a recipe for disaster.

The central paradox is that the extraordinary power of PCR (the massive amplification of a small number of template molecules) is also its "Achilles heel." In a routine DNA testing lab, the same few barcode amplicons will be generated on an enormous scale. Careless tube opening or pipetting of a PCR after thermocycling can generate an aerosol of potential contaminant amplicons. Contamination of a new PCR with just a single amplicon from a previous reaction could then generate a cycle of false positive results.

There are several strategies to mitigate DNA contamination of new PCRs, based on the experience of forensic and clinical diagnostic PCR laboratories:

- Separation of "pre-PCR" and "post-PCR" operations in physical space and/or time is a key strategy [25]. Thus, a series of small clean rooms or compartments are more suitable than a single open laboratory for the core infrastructure.
- Establishment of a workflow that ensures that samples and procedures can only move downstream from plant material → DNA extraction → DNA quantitation → pre-PCR setup → PCR → post-PCR analysis → nested PCR. The most critical backflow to avoid is from post-PCR to pre-PCR (e.g., by transfer of amplicons from the "dirty" post-PCR area back into the "clean" pre-PCR area) [25, 26].
- Where possible, pre-PCR reagents should be dispensed in a separate clean area free from any potential template DNA (plant material, genomic DNA, and PCR products). It is good practice to aliquot reagents such as primers and PCR reagents in this area.

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

- Designation of dedicated lab coats, gloves, pipettes, and plasticware for the pre-PCR areas is an important element in the strategy, enforced by training and clear standard operating procedures. Most critically, these should spell out the practical issues around movement of personnel and materials in and out of the clean area.
- Procedures for regular decontamination of clean DNA areas should be implemented, by cleaning with 10% bleach, commercial DNA-destroying cleaners, and/or UV irradiation.

The positive side is that equipping a routine molecular biology laboratory is relatively inexpensive compared to analytical chemistry. A simple PCR facility comprising a conventional PCR machine, gel electrophoresis tank plus power supply, and a gel imager could cost less than €10 000. A real-time PCR machine, especially with HRM capability, would be more expensive (€10 000–25 000). Other specialized equipment to consider includes a dedicated spectrophotometer or fluorimeter for DNA quantitation, a mechanical homogenizer for tissue disruption, and a PCR clean hood with air filtration and/or UV irradiation, but these are all relatively inexpensive items.

A more important consideration is the personnel implications of introducing molecular techniques. While the core techniques of DNA barcoding are relatively simple to learn and perform by competent technical staff, there is a need for more specialist expertise for troubleshooting poor or unexpected results and for the bioinformatic analysis of results.

Troubleshooting DNA Barcoding Methods

The BP test is based on the standard procedure for DNA barcoding. The limitations of DNA barcoding for routine industrial quality testing will be addressed later, but the procedure has particular value for the identification of reference specimens and DNA sequences. ▶ **Fig. 1** shows a typical workflow for DNA barcoding. The process starts from DNA extraction of the plant sample, through methods for quantitation, PCR and gel electrophoresis, and DNA sequencing. A number of decision points are included for quality control and troubleshooting. These cover most eventualities and should facilitate the development of routine, reliable protocols. Further guidance on the generic downstream processes in ▶ **Fig. 1** is widely available [27–30]. The upstream steps that are specific to typical raw materials for herbal medicines will be briefly considered.

Plant sample collection for DNA barcoding has been described in detail for taxonomic specimen identification [28, 31–33]. However, herbal drug material is usually not of this type. Wild harvested source materials are typically mature tissues, often roots or rhizomes, that may have been dried slowly under poor conditions. In consequence, DNA extraction faces three main challenges: low yield, contamination with PCR inhibitory molecules, and DNA degradation [27, 28, 32]. Low yield is best solved by optimizing tissue disruption, increasing the proportion of extraction reagent volumes to tissue mass, and extending the tissue solubilization time. The effect of contamination by PCR inhibitory storage carbohydrates and polyphenolic secondary products can be mitigated by diluting the DNA template before PCR or clean-up of the genomic DNA by alcohol precipitation. Phenolics can also be removed by adding polyvinylpolypyrrolidone at an early stage in the extraction [34]. The issue of DNA degradation is more problematic. Upstream solutions would involve less-damaging procedures for drying and processing of the source material. Downstream solutions may employ the type of DNA repair and amplification procedures developed for forensic profiling of degraded DNA or target shorter "mini-barcode" regions (see next section) [35–38].

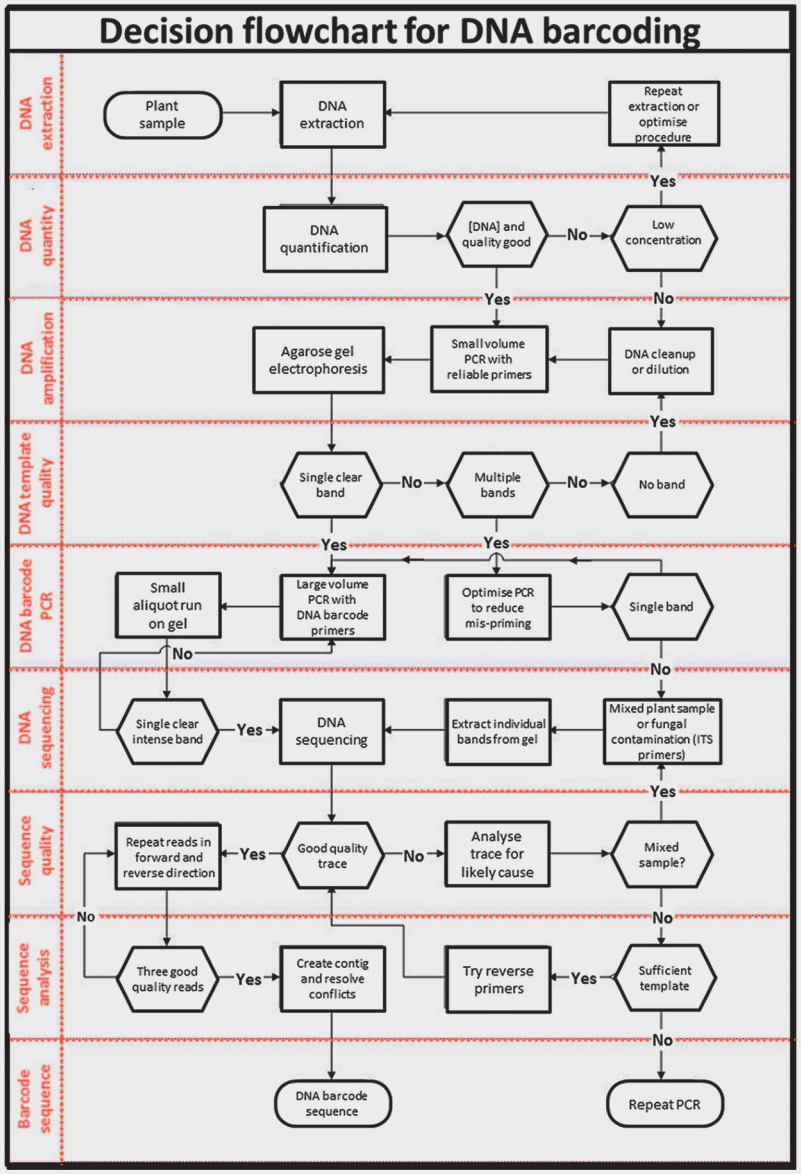
If several different types of material are going to be tested, it is advisable to adopt a general-purpose extraction method. This will typically be either a commercial kit based on ion exchange mini-columns or a variant of the CTAB (cetyltrimethylammonium bromide) method [39]. A number of comparisons of DNA extractions have been carried out, which indicate some variability between commonly used methods and their suitability for different tissue types in terms of quantity and quality of DNA extracted [40–42]. However, a reliable standard procedure is usually adequate to obtain sufficient DNA for subsequent PCR amplification and analysis.

DNA quantitation is an important indicator of the success of an extraction. Spectrophotometric measurement of UV absorbance at 260 nm can be used to directly quantify a double-stranded DNA sample. Readings at 280 nm can indicate residual phenol carryover from the extraction protocol or contamination of the extract with phenolic secondary metabolites from the plant material. Absorbance at 230 nm may indicate carbohydrate contamination, residual guanidinium salts (often used in column-based kits), or carryover of glycogen used for DNA precipitation [27, 30]. However, carbohydrates also absorb at 260 nm and can interfere with the quantitation of DNA. In contrast, fluorometric methods make use of DNA intercalating dyes, so they are less susceptible to interference by impurities. There are a number of commercial kits available for this purpose, some of which come with a simple dedicated fluorimeter with direct calculation of DNA concentration. The DNA is mixed with the reagents and the fluorescence compared to reference standards. Fluorescent readings are more specific to double-stranded DNA, so they are more reliable, particularly at low concentrations, but do not detect the presence of contaminants.

While quantitation gives some indication of the success of an extraction procedure, the ultimate test is whether the DNA is amplifiable by PCR. It is therefore common practice to perform a routine DNA barcode PCR after extraction to determine whether the DNA is either contaminated or degraded. The aim at this stage is to obtain a clear band, not necessarily to determine the sequence of the barcode.

PCR protocols for plant DNA barcoding have been well described, with specific information about the choice of barcode regions and individual primer pairs [27–30]. There are modifications to routine PCR protocols that address the particular problems of secondary metabolite PCR inhibitors from plant materials [43, 44] and the effects of DNA degradation [35, 36, 45, 46]. In all stages of the DNA barcoding process, it is good practice to include a positive control to ensure the quality of the procedure and troubleshoot problems. One example is described in the next section.

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.



▶ **Fig. 1** DNA barcoding flowchart. The chart indicates the steps involved in a normal DNA barcoding workflow, with quality control decision points to ensure positive outcomes and troubleshoot problematic steps in the process.

The BPNARM

The intention of BP Appendix XI V [22] is to provide generic procedures for molecular identification of herbal drugs, while the supplementary chapter "DNA barcoding as a tool for botanical identification of herbal drugs" [23] provides the information that differs between target samples, or species. To illustrate this, a worked example was published in the appendix for holy basil [22]. This directs the user to the generic protocols where appropriate (e.g., DNA extraction) and details the specific information for holy basil (e.g., DNA purification, the reference sequence). As this was the first protocol to be published, information was given as to the primer sequence and PCR cycling parameters required for the *trnH-psbA* region.

The appendix method recommends that the efficiency of DNA extraction methods be confirmed and that amplification protocols are controlled by the use of a known DNA sample. For this purpose, the *trnH-psbA* BPNARM was developed, which enables the user to confirm the suitability of his or her systems to conduct the analyses [47]. It is designed for use in two different ways:

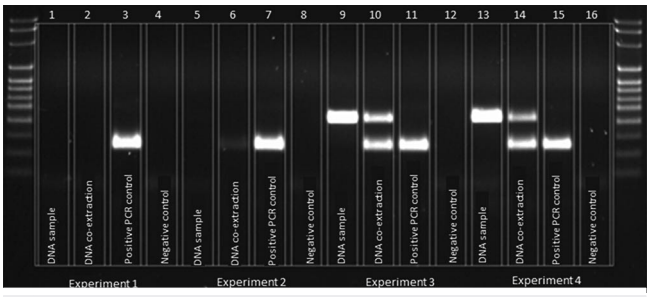
1. The BPNARM is mixed with the plant material prior to DNA extraction, and the product of this method is then a "co-extraction" of both plant and BPNARM DNA. The recovery of the BPNARM DNA confirms that the extraction process has been conducted effectively, and the efficient amplification of this

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

DNA indicates that PCR inhibitors are not preventing amplification.

2. The BPNARM is used as a positive control for the PCR, demonstrating the suitability of the PCR reagents, experimental setup, and amplicon detection method.

The BPNARM was developed alongside the holy basil method, and ▶ **Fig. 2** shows four examples of its application in troubleshooting: unsuccessful DNA extraction (Experiment 1), PCR inhibitors in the DNA samples (Experiment 2), demonstration of removal of PCR inhibitors by DNA purification (Experiment 3), and reduction of the effect of PCR inhibitors by dilution of the DNA template (Experiment 4). A detailed description of these experiments can be found in the Supporting Information. This example demonstrates the value of using a known and reliable control substance, without which the troubleshooting process can be much lengthier and technically demanding.



▶ **Fig. 2** The use of a standardized positive control for DNA extraction and PCR troubleshooting. Experiment 1: Results from an unsuccessful DNA extraction. Experiment 2: Effect of inhibitors present in the DNA samples. Experiment 3: Effect of removal of inhibitors from DNA samples, demonstrating that all inhibitors had now been removed from the DNA samples. Experiment 4: Effect of dilution of purified DNA. In each experiment the samples are as follows: DNA sample, DNA extracted from plant material; DNA co-extraction, DNA extracted from plant material mixed with the *trnH-psbA* BPNARM; Positive PCR control, PCR using the *trnH-psbA* BPNARM as template; Negative control, PCR with molecular grade water added in place of a DNA template. (Experimental details in Supporting Information.)

The Design of Simple DNA Tests for Industrial Quality Control

The scheme for DNA barcoding outlined in ▶ **Fig. 1** is ideal for the identification of individual plant specimens. However, as indicated by the many recommended troubleshooting steps, Sanger sequencing of a DNA barcode region is not suited to routine, robust, high-throughput screening of dried and processed mixtures of sub-optimal plant tissue types containing degraded DNA. The authors have developed a strategy for industrial quality control that involves the design of simple, reliable PCR tests that target individual differences in DNA barcode sequences [38, 48, 49] and only use DNA barcode sequencing as a confirmatory rather than routine assay for species authentication. These types of specific PCR assay targeting a short DNA barcode region have been developed by a number of authors [3–5, 43]. The following section describes the development of such tests in an industrial setting, using *Rhodiola rosea* L., Crassulaceae, as an example.

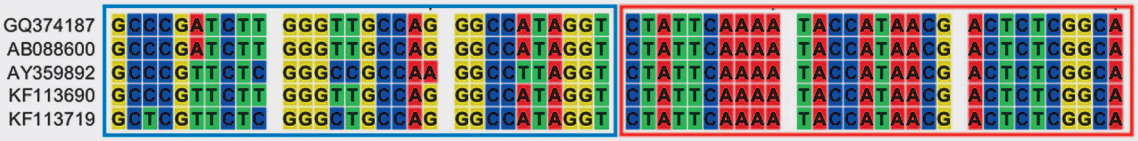
R. rosea root extract is a common medicinal plant in many countries used to stimulate the nervous system, decrease depression, enhance work performance, eliminate fatigue, and prevent high-altitude sickness thanks to its constituents such as salidroside (rhodioloside), rosavins, and p-tyrosol [50]. It grows in cold regions of the world, typically at high altitude on rocky outcrops, and is wild harvested in regions of Central Asia such as the Altai Mountains and the Qinghai-Tibetan Plateau [51–53]. A number of other *Rhodiola* species grow in similar arctic/alpine habitats, creating the potential for mistaken identity and nomenclature confusion as the roots are harvested and traded [52, 54].

Within an industrial quality control laboratory there are a number of tests used to confirm plant identity. The first stage is botanical morphology, both macroscopic and microscopic [8, 17, 18]. Further testing employs a range of analytical chemistry methods. TLC is one of the standard methods of pharmacopeia monographs. The technique is relatively simple and straightforward, and precision and standardization has been improved by the development of HPTLC methods. (**Fig. 4S** (Supporting Information) shows an example of HPTLC analysis of *R. rosea* samples.) One problem is that chemical markers used for identification have been found in similar species from the same genus, including *R. rosea* and related species [52]. Studies of commercially available products have also found that the species attribution is incorrect even though the target marker is detected [9]. Many herbal medicines are manufactured to contain a specified amount of certain chemical markers, as shown on the label. To ensure that a sample has the correct amount of marker, HPLC can be used to quantify the marker. The concentration of a certain marker can also be linked to the potential efficacy of a herbal medicine, as with rosavin, rosarin, and rosin in *R. rosea* products [52] (**Fig. 5S**, Supporting Information), and threshold concentrations are built into the quality control protocols.

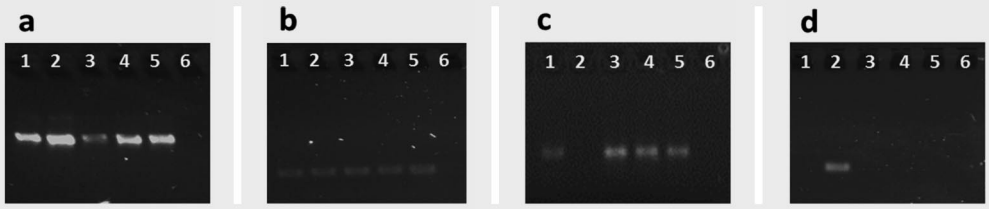
At the start of this study there were very few *Rhodiola* DNA barcode sequences available in the databases. Therefore, a reference collection of samples of 10 different *Rhodiola* species was created. Genomic DNA was isolated and PCR amplification and DNA sequencing of four barcode loci (*rbcL, trnH-psbA, matK,* and nrITS [nuclear ribosomal internal transcribed spacer]) according to the scheme in ▶ **Fig. 1** was successful for all the samples. Comparison of the four barcodes indicated that the ITS (internal transcribed spacer) region would be the most suitable target for discrimination between *R. rosea* and other *Rhodiola* species in terms of the extent, nature, and distribution of sequence variation through the barcode regions. The initial PCR assays were designed to these sequences along with a small number of accessions from GenBank.

More recently, a large number of *Rhodiola* barcode sequences have been deposited in the database, and a total of 438 ITS sequences from 36 *Rhodiola* species are now available (**Table 2S**, Supporting Information). These sequences were used to (a) perform a retrospective confirmation of the reference collection specimens by constructing a phylogenetic tree from a multiple alignment of database and reference sequences and (b) check that the species-

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.



▶ **Fig. 3** Fragment of a multiple alignment of the ITS region from a selection of *Rhodiola* species (GQ374187 *R. rosea*; AB088600 *Rhodiola ishidae*; AY359892 *R. crenulata*; KF113690 *Rhodiola coccinea*; KF113719 *Rhodiola sacra*). The highlighted sections show variable (left) and conserved (right) regions where species-specific and generic primers could be designed, respectively.



▶ **Fig. 4** Agarose gel electrophoresis of PCRs using (**a**) ITS primers, (**b**) generic *Rhodiola* primers, (**c**) non-*rosea* primers, and (**d**) *R. rosea*-specific primers (experimental details in Supporting Information). Gel lanes: 1. *Rhodiola pachyclados*; 2. *R. rosea*; 3. *Rhodiola heterodonta*; 4. *Rhodiola saxifragoides*; 5. *R. crenulata*; 6. Negative (no template) control. The sizes of the bands were confirmed by DNA sequencing.

specific primers matched all instances of the target species and would not cross-react with any non-target species. This iterative confirmation of DNA test protocols is an important consequence of the accumulation of DNA sequences in the databases.

In this example, PCR primers were designed to small (80–120 bp) variable, species-discriminatory regions of the barcodes (▶ **Fig. 3**). These are less vulnerable to DNA degradation and are ideal for qPCR analysis. The primary target was *R. rosea*, but primers for individual potential adulterant species and exclusive non-*rosea* primers were also designed. Generic primers were also designed to amplify any template DNA from the *Rhodiola* genus (▶ **Fig. 3**). As an illustration, ▶ **Fig. 4 d** shows the results of conventional PCR tests with a pair of *R. rosea*-specific primers, giving a positive reaction only with the target template. A non-*rosea* primer pair amplified the other *Rhodiola* species but not *R. rosea* (▶ **Fig. 4 c**). The ITS (▶ **Fig. 4 a**) and generic primers (▶ **Fig. 4 b**) amplified all samples.

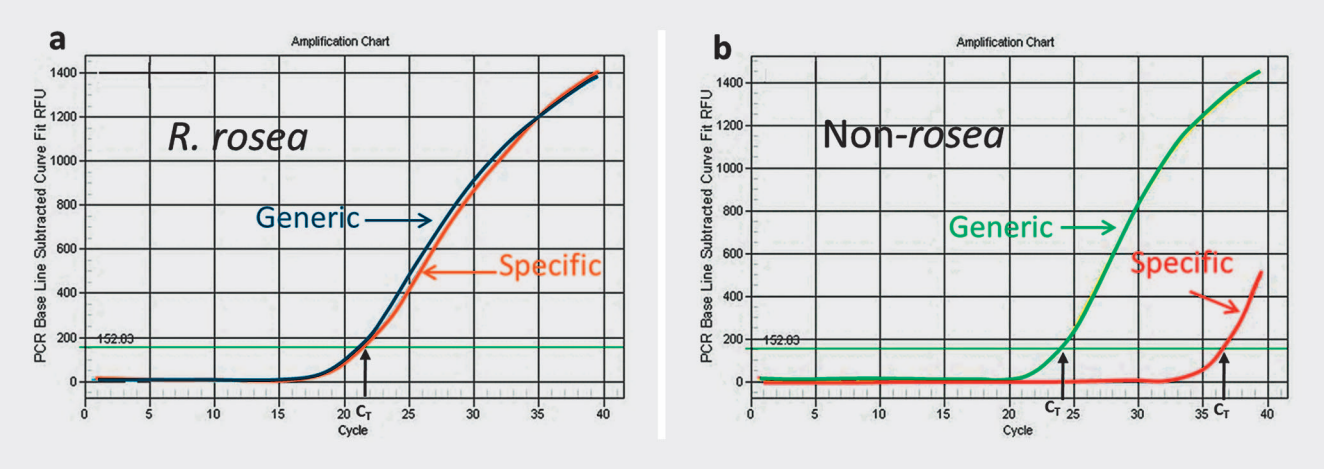
The process from DNA extraction to gel imaging (▶ **Fig. 1**), could be carried out in a basic molecular biology facility within one working day. However, qPCR is actually quicker and simpler to run than conventional PCR, since it does not require gel electrophoresis (with associated problems of reproducibility and gel image interpretation). The primer sets illustrated in ▶ **Fig. 4** were tested using qPCR analysis after optimizing the thermocycling setting and primers concentrations. In this example, the amplification curves obtained using generic and *rosea*-specific primers are very similar when amplifying a *R. rosea* template (▶ **Fig. 5 a**). The near-identical C_t value between the two assay indicates a positive identification of the template as *R. rosea*, whereas the marked difference in C_t value (around 12 cycles) between the generic and specific primers using a panel of non-target species is the template (▶ **Fig. 5 b**) is a clear negative result (see Supporting Information for a detailed explanation).

There are a number of strategies to improve the specificity of this type of PCR, including the design of amplification refractory mutation system primers with deliberately destabilized mismatch bases [43, 55–57] and the use of locked nucleic acid primers to enhance specific primer binding [58, 59].

The importance of DNA method validation has been stressed by various authors [6, 7, 30, 60]. In the *Rhodiola* study, around 40 *R. rosea* commercial samples were tested in parallel for rosavin content and by qPCR assay for *R. rosea* identification, with a small number of non-*rosea* samples as controls (see Supporting Information for details). The conclusions from this validation study were as follows:

1. The qPCR assay is rapid, robust, and reliable. In theory, the 40 samples could be processed and analyzed within 1 d. Correct identification results were obtained for all samples tested and were more reliable than the chemical tests.
2. This type of test can be applied to samples where Sanger-based DNA barcoding would fail due to DNA degradation and mixed samples (see ▶ **Figs. 1** and **2**).
3. DNA tests are complementary to chemical assays–in this example, samples lacking the rosavin chemical markers would still be rejected, even if shown genetically to be the correct species. Conversely, a sample containing the correct chemical markers would be rejected if shown to be a substituted or adulterated species.

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.



▶ **Fig. 5** Quantitative real-time PCR assay to distinguish (**a**) *R. rosea* and (**b**) adulterants. The assay is based on the difference in C_t (threshold cycle) between generic and specific primers. Similar C_t values indicate a positive identification (**a**), while a marked rightward shift in the specific primer C_t value indicates a negative result (**b**). (Experimental details in Supporting Information.)

The qPCR approach can also be used to quantify the amount of target DNA in a mixture, but this has limitations unless the precise nature of the adulterant is also known. It is much easier to distinguish between 0% and 2% contamination by a known adulterant, for example, than it is to determine that only 98% of a sample is the required target species.

Another approach that has been more successful in this regard is HRM [61, 62] (more recently termed Bar-HRM when applied to species identification [63, 64]). This technique can potentially discriminate between sequences containing a single base pair change and can detect the presence of both sequences in an admixture. Using *Rhodiola* as an illustration again, HRM primers were designed to conserved regions of the ITS sequences on either side of a short region of inter-specific variation, targeting species-specific SNPs (**Fig. 6S**, Supporting Information). A total of three different sets of primers were designed and tested against a reference panel of 12 *Rhodiola* species. All three sets of primers could discriminate *R. rosea* samples from *R. crenulata*, one of the most common contaminants of *R. rosea* (**Fig. 7S**, Supporting Information). When tested against all 12 *Rhodiola* species, the primers were not uniquely specific for *R. rosea* but could distinguish a number of variant types (**Fig. 8S**, Supporting Information).

dPCR offers the prospect of absolute quantitation of the number of starting template molecules in the PCR by partition of each molecule into a separate compartment (on a chip or in an emulsion droplet) [65–67] and has been used for species authentication [4, 68]. Because quantitation of the level of contamination by adulterants has proved to be a challenge for qPCR-based assays, the aim is to determine whether the dPCR system can provide an authentication test capable of detecting contamination down to the required < 2% foreign matter contamination threshold. An example of the use of species-specific qPCR primers in a dPCR system is shown in the Supporting Information (**Fig. 9S**, Supporting Information).

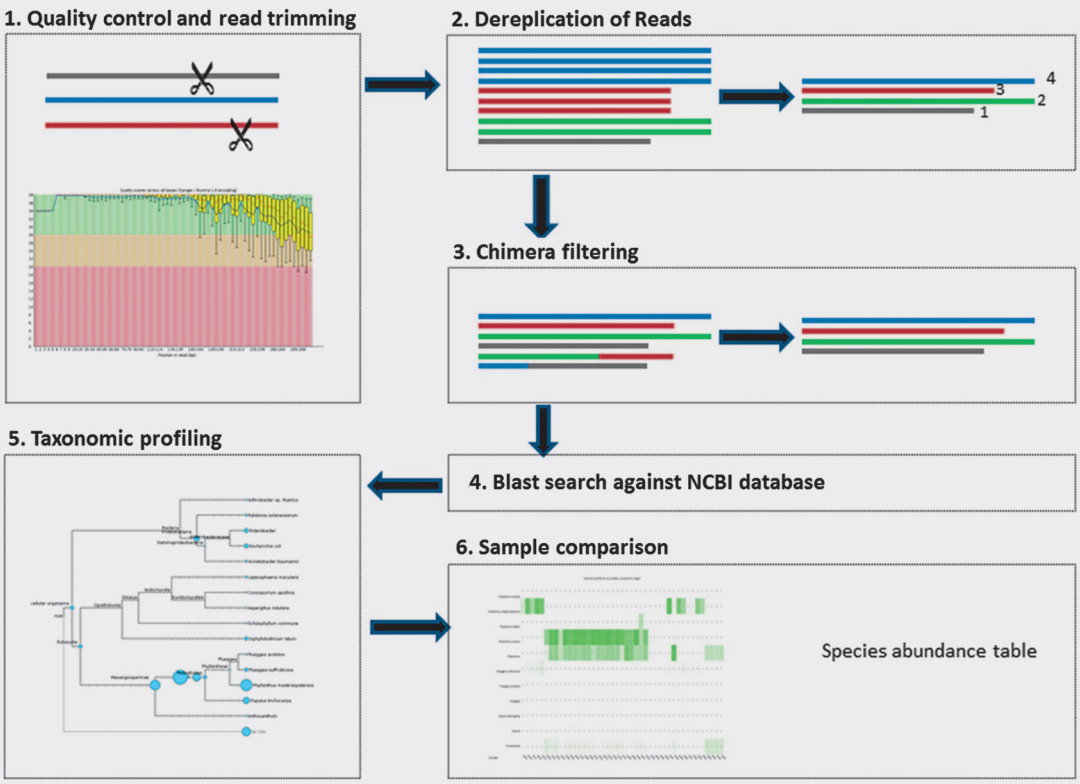
Prospects for NGS in Industrial Quality Control

In recent years, DNA sequencing technology has made dramatic steps forward. The conventional Sanger method, where relatively short, targeted sequences are produced, is being replaced by so called NGS technologies, deploying high-throughput massively parallel methods. NGS is based on a sequencing-by-synthesis approach along millions of single-stranded DNA templates, where each newly incorporated base produces a signal that is translated into a sequence. A major difference to Sanger sequencing is that every DNA fragment that is sequenced will result in an individual DNA output sequence (termed "read"), whereas in Sanger sequencing, the base call signal derives from numerous template molecules.

Next-generation amplicon sequencing is the process of enriching a small DNA fragment of interest through PCR and simultaneously sequencing each of the resulting fragments. The major advantage over Sanger sequencing is that the sequence diversity of fragments from the PCR is maintained in the reads. This technology has revolutionized the field of DNA metabarcoding, where mixed samples (e.g., environmental water samples [69]) are sequenced and analyzed for species diversity by comparing the reads to a reference database [70]. The main concept of DNA barcoding is maintained by using common DNA barcodes that are flanked by universal primers. This concept has been successfully applied in pharmacovigilance and market studies of HMPs [1, 2, 5, 71–73]. Here, an example of the application of these methods to commercial herbal products is described, using NGS and bioinformatics analysis to assess the composition of medicinal *Phyllanthus amarus* Schumach. & Thonn., Phyllanthaceae, samples. The workflow of the analysis is summarized in ▶ **Fig. 6**.

A library preparation method was designed that would enable four barcode regions to be sequenced for each of the DNA samples from ten commercial products in one run and the resulting

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.



▶ **Fig. 6** Overview of the bioinformatics pipeline for NGS data analysis.

data to be sorted. The workflow for amplicon sequencing needs to incorporate platform-specific particularities. On Illumina platforms, the workflow includes two separate PCRs. In the first step, the *trnH-psbA*, ITS2, *trnL-F*, and *rbcL* regions were amplified separately using standard amplification protocols, and in the second step, a PCR with platform-specific adapters was applied to each sample. The samples were then pooled together and sequenced on an Illumina MiSeq instrument (Supporting Information).

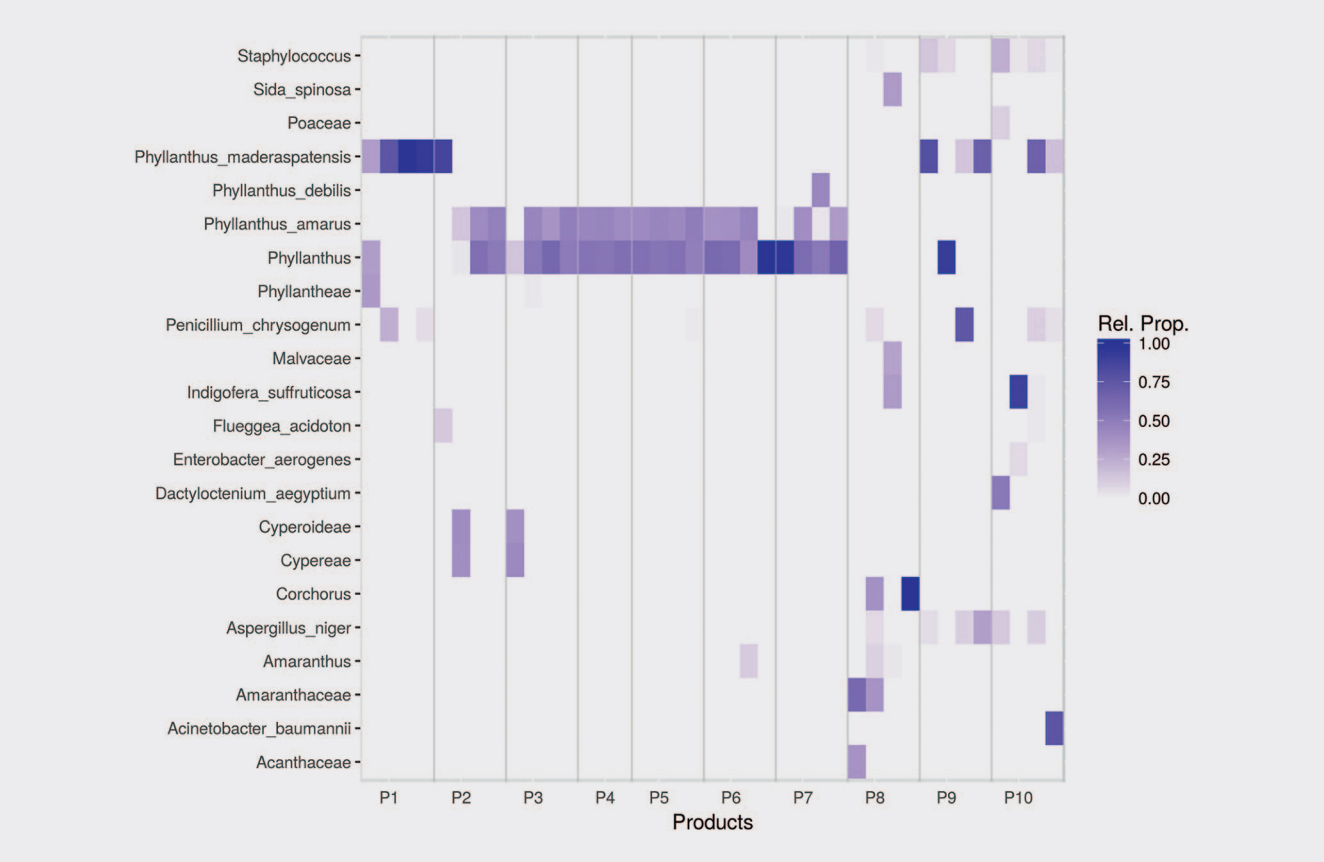
A major challenge of establishing standardized NGS protocols is to choose an appropriate and replicable bioinformatics pipeline. The bioinformatics pipeline applied in this study encompasses trimming of low quality reads, removal of optical duplicates and chimeric sequences, and clustering of sequences that share 99% sequence similarity (for details see Supporting Information). The pipeline should ideally include a standardized procedure for storing data files and logs along with the analysis.

As an example, the output from the NGS analysis of samples sold as *P. amarus* is shown in ▶ **Fig. 7**. The abundance of sequence data produced and analyzed lends itself to representation as a "heat map." The more intense the blue color, the greater the number of sequences matching to the species named in that row (see Supporting Information for a more detailed explanation). Our analysis suggests that many of the *Phyllanthus* samples are not pure and potentially intentionally substituted. Other studies draw similar conclusions when applying NGS amplicon sequencing to other HMPs [73, 46]. However, a comprehensive benchmarking study that identifies acceptable levels of contamination and that investigates the limits of the method is lacking.

The accuracy of the analysis depends the quality of the input library and the sequencing. Similar to Sanger sequencing, successful primer binding and amplification is crucial. The major source of errors in amplicon sequencing derives from PCR artifacts and sequencing errors that may in some cases be difficult to detect in downstream analyses [74]. It is therefore highly recommended to use high-fidelity polymerase enzymes with a minimal number of PCR cycles and to include quality control steps in the analysis pipeline.

The reference database used to find the best match for DNA sequences is of fundamental importance, as is the case for all sequencing methods. This is of particular concern for NGS methods, as the number of reads generated makes meticulous analysis of all the sequences involved impossible. To some degree, the number of sequences matched will dilute out any anomalous results, and a threshold of the number of matches required to assign any significance to a result can be incorporated. However, a fundamentally biased reference database (e.g., caused by a large study of one species within a genus) can make false positive identification an issue. Thus, a well-maintained and comprehensive database containing a range of sister species and potential adulterant and substitute species is necessary. It should also be noted that detection to species level often depends on the discriminatory power of the barcode. The concept of the barcoding gap describes the distance between intra- and inter-specific genetic variation. The assump-

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.



▶ **Fig. 7** Heat map showing the number of sequences matching to different species in the NCBI database for each of 10 products. The intensity of the color represents the number of matches, shown on the scale. Where a sequence matches to more than one species in the genus, the genus is listed as the category. Four DNA samples were analyzed for each of Products P1–P10; each of these is shown independently (experimental details in Supporting Information).

tion is that a barcoding gap exists, when the intra-specific genetic distance of the barcode is smaller than the inter-specific distance [75]. A careful evaluation of the barcoding gap minimizes the impact of false positive and false negative identifications.

The choice of barcode locus is also constrained by the length of the region. The read length derived from next-generation sequencers is smaller than the sequences from Sanger sequencing, and thus, the discriminatory power is potentially smaller. Whereas Sanger methods can produce reads of > 1000 bp, NGS methods can sequence up to 300 bp from either end of an amplicon. This produces 600 bp of sequence data once the reads have been paired, including primers and binding regions (~ 550 bp of sample sequence data). This means that longer barcode regions are problematic to analyze (e.g., *matK* and the full nrITS). Use of the shorter ITS2 region instead of the nrITS can circumvent this limitation for this region, but there is currently no universal shorter *matK* region.

NGS is a major forward step for molecular methods and allows an extremely high number of sequence reads for any sample. However, the amount of data generated by this method also requires that it is analyzed using specialized bioinformatics programs and by highly skilled individuals. The lack of skilled bioinformaticians has been highlighted and is being addressed by universities and researchers, but it remains a limitation of NGS methods. Another issue is the access to NGS equipment. Although the cost per read for this method is actually much lower than for Sanger methods, the price of equipment and consumables is not easily accessible for those starting out in molecular methods. The use of external companies to conduct the final few steps in the procedure, and also some of the data analysis, is widespread and many companies offer this service.

Conclusion

Despite reservations about the application of DNA testing to herbal medicine authentication, industry is adopting DNA quality control tests in response to two drivers: (1) the increased confidence in the identity of the herbal material and compliance with regulatory requirements (as exemplified by the BP) that the techniques can bring and (2) the potential cost saving of utilizing inexpensive high-throughput tests for authentication of plant raw materials (as exemplified by the *R. rosea* tests). Although phytopharmaceutical quality control infrastructure and expertise has historically been largely based around analytical chemistry, the introduction of molecular biological testing is relatively straightforward and is gradually being taken up by the industry. While con-

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

ventional DNA barcoding by Sanger sequencing is not ideal for rapid, robust quality testing, it does provide considerable genetic information upon which to base simpler DNA-based tests for industrial quality control.

NGS is a powerful tool for the molecular analysis of herbals; it provides a depth to the analysis and a direct identification of the target species. The area where NGS is unrivalled is in identifying unknown or unexpected species that are present as adulterants or contaminants to the "headline" species. Sanger sequencing can mask the presence of additional species, as the results generated are based on calling each base from a large number of amplicons simultaneously. NGS methods allow each contributing species to be identified individually, within one sample. These results can also be semi-quantitative, providing some measure of the proportion of each species in a sample (although biases in the PCR process prevent this from being absolute).

While this method shows the promise to answer all the possible requirements of industry, the barriers in terms of cost and expertise required currently place it firmly in the "future prospects" category, until the methods become more accessible and cost effective.

Supporting information

Five examples of the authors' unpublished works have been used to illustrate this review. These are as follows.

Example 1: The BP DNA-based identification method for *O. tenuiflorum*; Example 2: The BPNARM; Example 3: A specific qPCR assay for *R. rosea*; Example 4: A HRM curve assay for *Rhodiola* species; Example 5: NGS assay of *Phyllanthus* samples. The experimental details (Materials and Methods, Results) for each example are included as Supporting Information.

There are two supporting tables.

Table 1S: Commercial "holy basil" samples used for development of the BP DNA-based identification method for *O. tenuiflorum*; **Table 2S**: *Rhodiola* ITS sequences used in this study.

There are ten supporting figures.

Fig. 1S: Multiple alignment of *trnH-psbA* region sequences from holy basil commercial samples; **Fig. 2S**: Bases 190–260 of a multiple alignment of *trnH-psbA* region sequences from holy basil commercial samples; **Fig. 3S**: Bases 300–370 of a multiple alignment of *trnH-psbA* region sequences from holy basil commercial samples; **Fig. 4S**: HPTLC of *Rhodiola* samples and reference compounds; **Fig. 5S**: HPLC trace of an extract from *R. rosea*, showing the presence of the rosavine marker compounds rosarin, rosavin, and rosin; **Fig. 6S**: Two *Rhodiola* ITS regions selected for HRM curve analysis primers design; **Fig. 7S**: HRM curve analysis results using the HRM1, HRM2, and HRM3 primer pairs with *R. rosea* and *R. crenulata* templates; **Fig. 8S**: HRM curve analysis results using the HRM1, HRM2, and HRM3 primer pairs with 11 different *Rhodiola* templates; **Fig. 9S**: Clarity dPCR results with species-specific primers; **Fig. 10S**: Representation of the constructs produced for NGS barcoding.

Acknowledgements

VITANGO was a Marie Curie Actions Industry-Academia Partnerships and Pathways collaboration between De Montfort University, the United Kingdom, and Dr. Willmar Schwabe Pharmaceuticals, Germany. The project was funded by the EU under the Seventh Framework Programme (FP7/2007–2013) under grant agreement PIAP-GA-2011-286328 VITANGO.

The authors thank the following colleagues for their help and advice: Ms. Simone Bacher, Dr. Dipanwita Banik, Dr. Caroline Cayrou, Professor Julie Hawkins, Dr. Sabrina Kraft, Dr. Eva Masiero, Dr. Edward Mee, Ms. Attiyyah Omerjee, and Dr. Karl Schötz. This paper is dedicated to Professor Dr. Max Wichtl in recognition of his outstanding contribution to pharmacognosy research.

Marco Kreuzer was supported by MEDPLANT, a Marie Curie Actions Initial Training Network (ITN), funded by the European Union under the Seventh Framework Program.

Conflict of Interest

Caroline Howard and Claire Lockie Williams are employed by the UK Medicines and Healthcare products Regulatory Agency (MHRA). The BPNARM control product described in this paper is a commercial product marketed by the MHRA. Adrian Slater is a member of the Expert Panel DNA: Identification Techniques, an advisory board of the British Pharmacopoeia. The other authors declare no conflicts of interest.

References

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

[2] De Boer H, Ichim M, Newmaster S. DNA barcoding and pharmacovigilance of herbal medicines. Drug Saf 2015; 38: 611–620

[3] Ganie S, Upadhyay P, Das S, Sharma M. Authentication of medicinal plants by DNA markers. Plant Gene 2015; 4: 83–99

[4] Parveen I, Gafner S, Techen N, Murch SJ, Khan IA. DNA barcoding for the identification of botanicals in herbal medicine and dietary supplements: strengths and limitations. Planta Med 2016; 82: 1225–1235

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

[6] Fabricant D, Hilmas C. NPA White Paper: DNA barcoding for botanical authentication. National Products Association 2015. Available at https://www.npainfo.org/App_Themes/NPA/docs/regulatory Legislative/White%20Paper/NPAWhitePaper_DNABarcoding.pdf. Accessed February 28, 2017

[7] Reynaud D, Mishler BD, Neal-Kababick J. The capabilities and limitations of DNA barcoding of botanical dietary supplements. Authentechnologies 2015. Available at http://www.authentechnologies.com/wp-content/uploads/2015/04/Reynaud_DNA_Barcoding_White_Paper.pdf. Accessed February 28, 2017

[8] Sahoo N, Manchikanti P, Dey S. Herbal drugs: standards and regulation. Fitoterapia 2010; 81: 462–471

[9] Palhares RM, Gonçalves Drummond M, Dos Santos Alves Figueiredo Brasil B, Pereira Cosenza G, das Graças Lins Brandão M, Oliveira G. Medicinal plants recommended by the world health organization: DNA barcode identification associated with chemical analyses guarantees their quality. PLoS One 2015; 10: e0127866

[10] Kamil M. Assuring production of quality raw botanical materials – cultivation, collection and post-harvest practices. Planta Med 2012; 78: PC20

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

[11] Van Breemen RB. Development of safe and effective botanical dietary supplements. J Med Chem 2015; 58: 8360–8372

[12] Jordan SA, Cunningham DG, Marles RJ. Assessment of herbal medicinal products: challenges, and opportunities to increase the knowledge base for safety assessment. Toxicol Appl Pharmacol 2010; 243: 198–216

[13] Heinrich M. Quality and safety of herbal medical products: regulation and the need for quality assurance along the value chains. Br J Clin Pharmacol 2015; 80: 62–66

[14] Vlietinck A, Pieters L, Apers S. Legal requirements for the quality of herbal substances and herbal preparations for the manufacturing of herbal medicinal products in the European Union. Planta Med 2009; 75: 683–688

[15] Da-Costa-Rocha I, Edwards S. Quality and safety of herbal products: Part 1 – new legislation and production. Pharm J 2012; 288: 685–687

[16] Barnes J. Pharmacovigilance of herbal medicines. Drug Saf 2003; 26: 829–851

[17] WHO guidelines for assessing quality of herbal medicines with reference to contaminants and residues. World Health Organization 2007. Available at http://apps.who.int/medicinedocs/en/d/Js14878e/. Accessed February 28, 2017

[18] Van Breemen R, Fong H, Farnsworth N. The role of quality assurance and standardization in the safety of botanical dietary supplements. Chem Res Toxicol 2007; 20: 577–582

[19] Ali A, Gyulai G, Hidvégi N, Kerti B, Hemaid F, Pandey A, Lee J. The changing epitome of species identification – DNA barcoding. Saudi J Biol Sci 2014; 21: 204–231

[20] Speijers G, Bottex B, Dusemund B, Lugasi A, Tóth J, Amberg-Müller J, Galli CL, Silano V, Rietjens IMCM. Safety assessment of botanicals and botanical preparations used as ingredients in food supplements: testing an European Food Safety Authority-tiered approach. Mol Nutr Food Res 2010; 54: 175–185

[21] Pilegaard K. Compendium of botanicals reported to contain naturally occurring substances of possible concern for human health when used in food and food supplements. EFSA J 2012; 10: 2663

[22] British Pharmacopoeia Commission. Deoxyribonucleic acid (DNA) based identification techniques for herbal drugs. British Pharmacopoeia Appendix XI V. London: TSO; 2017

[23] British Pharmacopoeia Commission. DNA barcoding as a tool for botanical identification of herbal drugs. British Pharmacopoeia Supplementary Chapter SC VII D, London: TSO; 2017

[24] Christina V, Annamalai A. Nucleotide based validation of *Ocimum* species by evaluating three candidate barcodes of the chloroplast region. Mol Ecol Resour 2014; 14: 60–68

[25] Mifflin TE. Setting up a PCR Laboratory. Cold Spring, Harbor: Cold Spring Harbor Laboratory Press; 2003: 5–14

[26] Forensic Science Regulator. The control and avoidance of contamination in laboratory activities involving DNA evidence recovery and analysis. London: UK Government Publications; 2015

[27] Cowan RS, Fay MF, Carles MC. Challenges in the DNA Barcoding of Plant Material. In: Sucher NJ, Hennell JR, eds. Plant DNA Fingerprinting and Barcoding: Methods and Protocols, Methods in Molecular Biology. Vol. 862. Berlin: Springer; 2012: 23–33

[28] Fazekas AJ, Kuzmina ML, Newmaster SG, Hollingsworth PM. DNA Barcoding Methods for Land Plants. In: Kress WJ, Erickson DL, eds. DNA Barcodes: Methods and Protocols, Methods in Molecular Biology. Vol. 858. Berlin: Springer; 2012: 223–252

[29] De Vere N, Rich T, Trinder S, Long C. DNA Barcoding for Plants. In: Batley J, ed. Plant Genotyping: Methods and Protocols. Vol. 1245. Berlin: Springer; 2015: 101–118

[30] Reynaud DT. The DNA Toolkit: a practical User's Guide to genetic Methods of botanical Authentication. In: Reynertson K, Mahmood K, eds. Botanicals: Methods and Techniques for Quality & Authenticity. Boca Raton: CRC Press; 2015: 43–68

[31] Gemeinholzer B, Rey I, Weising K. Short Manual for Collection and Documentation of Plant Tissue Samples for DNA Analysis. Berlin: Freie Universität Berlin; 2010

[32] Hollingsworth P, Graham S, Little D. Choosing and using a plant DNA barcode. PLoS One 2011; 6: e19254

[33] Gemeinholzer B, Rey I, Weising K, Grundmann M, Muellner AN, Zetzsche H, Droege G, Seberg O, Petersen G, Rawson D. Organizing specimen and tissue preservation in the field for subsequent molecular analyses. In: Eymann J, Degreef J, Hauser C, Monje JC, Samyn Y, Van den Spiegel D, eds. Manual on Field Recording Techniques and Protocols for all Taxa Biodiversity Inventories and Monitoring. Brussels, Belgium: ABC Taxa; 2010: 129–157

[34] Varma A, Padh H, Shrivastava N. Plant genomic DNA isolation: an art or a science? Biotechnol J 2007; 2: 386–392

[35] Staats M, Cuenca A, Richardson JE. DNA damage in plant herbarium tissue. PLoS One 2011; 6: e28448

[36] McCord B, Opel K, Funes M, Zoppis S. An investigation of the effect of DNA degradation and inhibition on PCR amplification of single source and mixed forensic samples. US Department of Justice 2011. Available at https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=258707. Accessed February 28, 2017

[37] Hajibabei M, Smith A, Janzen D, Rodriguez J, Whitfield J, Hebert P. A minimalist barcode can identify a specimen whose DNA is degraded. Mol Ecol Notes 2006; 6: 959–964

[38] Kazi T, Hussain N, Bremner P, Slater A, Howard C. The application of a DNA-based identification technique to over-the-counter herbal medicines. Fitoterapia 2013; 87: 27–30

[39] Doyle JJ. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin 1987; 19: 11–15

[40] Pirttilä A, Hirsikorpi M, Kämäräinen T, Jaakola L, Hohtola A. DNA isolation methods for medicinal and aromatic plants. Plant Mol Biol Report 2001; 19: 273

[41] Llongueras J, Nair S, Salas-Leiva D, Schwarzbach A. Comparing DNA extraction methods for analysis of botanical materials found in anti-diabetic supplements. Mol Biotechnol 2013; 53: 249–256

[42] Schmiderer C, Novack J. Effect of different DNA extraction methods and DNA dilutions on the amplification success in the PCR of different medicinal and aromatic plants. J Herbs Spices Med Plants 2013; 18: 65–72

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

[44] Schori M, Appel M, Kitko A, Showalter AM. Engineered DNA polymerase improves PCR results for plastid DNA. Appl Plant Sci 2013; 1: 120519

[45] Särkinen T, Staats M, Richardson J, Cowan R, Bakker F. How to open the treasure chest? Optimising DNA extraction from herbarium specimens. PLoS One 2012; 7: e43808

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

[47] British Pharmacopoeia Commission. British Pharmacopoeia Nucleic Acid Reference Material Information Leaflet. British Pharmacopoeia 2016. Available at https://www.pharmacopoeia.com/Catalogue/Preview?uri=%2Fcontent%2Ffile%2Fproducts%2Fleaflets%2FBPNARM-Leaflet_Cat-1185_BPCRS3837_6.pdf. Accessed February 28, 2017

[48] Howard C, Bremner P, Fowler M, Isodo B, Scott N, Slater A. Molecular identification of *Hypericum perforatum* by PCR amplification of the ITS and 5.8S rDNA region. Planta Med 2009; 75: 864–869

[49] Howard C, Socratous E, Williams S, Graham E, Fowler MR, Scott NW, Bremner PD, Slater A. PlantID – DNA-based identification of multiple medicinal plants in complex mixtures. Chin Med 2012; 7: 18

[50] Panossian A, Wikman G, Sarris J. Rosenroot (*Rhodiola rosea*): traditional use, chemical composition, pharmacology and clinical efficacy. Phytomedicine 2010; 17: 481–493

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

[51] Zhang JQ, Meng SY, Allen G, Wen J, Rao GY. Rapid radiation and dispersal out of the Qinghai-Tibetan plateau of an alpine plant lineage *Rhodiola* (Crassulaceae). Mol Phylogenet Evol 2014; 77: 147–158

[52] Booker A, Jalil B, Frommenwiler D, Reich E, Zhai L, Kulic Z, Heinrich M. The authenticity and quality of *Rhodiola rosea* products. Phytomedicine 2016; 23: 754–762

[53] Xia T, Chen S, Chen S, Zhang D, Zhang D, Gao Q, Ge X. ISSR analysis of genetic diversity of the Qinghai-Tibet plateau endemic *Rhodiola chrysanthemifolia* (Crassulaceae). Biochem Syst Ecol 2007; 35: 209–214

[54] Xin T, Li X, Yao H, Lin Y, Ma X, Cheng R, Song J, Ni L, Fan C, Chen S. Survey of commercial *Rhodiola* products revealed species diversity and potential safety issues. Sci Rep 2015; 5: 8337

[55] Wang H, Sun H, Kwon WS, Jin H, Yang DC. A PCR-based SNP marker for specific authentication of Korean ginseng (panax ginseng) cultivar "Chunpoong". Mol Biol Rep 2010; 37: 1053–1057

[56] Heubl G. New aspects of DNA-based authentication of Chinese medicinal plants by molecular biological techniques. Planta Med 2010; 76: 1963–1974

[57] Sasaki Y, Fushimi H, Cao H, Cai SQ, Komatsu K. Sequence analysis of Chinese and Japanese Curcuma drugs on the 18S rRNA gene and trnK gene and the application of amplification-refractory mutation system analysis for their authentication. Biol Pharm Bull 2002; 25: 1593–1599

[58] Ikenaga M, Sakai M. Application of locked nucleic acid (LNA) oligonucleotide-PCR clamping technique to selectively PCR amplify the SSU rRNA genes of bacteria in investigating the plant-associated community structures. Microbes Environ 2014; 29: 286–295

[59] Handy S, Timme R, Jacob S, Deeds J. Development of a locked nucleic acid real-time polymerase chain reaction assay for the detection of *Pinus armandii* in mixed species pine nut samples associated with dysgeusia. J Agric Food Chem 2013; 61: 1060–1066

[60] Pawar R, Handy SM, Grundel E, Cheng R, Shyong N. Assessment of the authenticity of herbal dietary supplements: Comparison of chemical and DNA barcoding methods. Planta Med 2016; 82: OA17

[61] Ruzicka J, Schmiderer C, Novak J. DNA-based identification of *Valeriana officinalis* s.l: a multiplexed amplification refractory mutation system (MARMS) and high resolution melting curve analysis (HRMA). Plant Mol Biol Report 2016; 34: 909–922

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

[63] Sun W, Li J, Xiong C, Zhao B, Chen S. The potential power of Bar-HRM technology in herbal medicine identification. Front Plant Sci 2016; 7: 367

[64] Osathanunkul M, Madesis P, de Boer H. Bar-HRM for authentication of plant-based medicines: evaluation of three medicinal products derived from Acanthaceae species. PLoS One 2015; 10: e0128476

[65] Pohl G, Shih IM. Principle and applications of digital PCR. Expert Rev Mol Diagn 2004; 4: 41–47

[66] Huggett JF, Cowen S, Foy CA. Considerations for digital PCR as an accurate molecular diagnostic tool. Clin Chem 2014; 61: 79–88

[67] Corbisier P, Pinheiro L, Mazoua S, Kortekaas AM, Chung P, Gerganova T, Roebben G, Emons H, Emslie K. DNA copy number concentration measured by digital and droplet digital quantitative PCR using certified reference materials. Anal Bioanal Chem 2015; 407: 1831–1840

[68] Floren C, Wiedemann I, Brenig B, Schütz E, Beck J. Species identification and quantification in meat and meat products using droplet digital PCR (ddPCR). Food Chem 2015; 173: 1054–1058

[69] Valentini A, Taberlet P, Miaud C, Civade R. Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. Mol Ecol 2016; 25: 929–942

[70] Taberlet P, Coissac E, Pompanon F. Towards next-generation biodiversity assessment using DNA metabarcoding. Mol Ecol 2012; 21: 2045–2050

[71] Coghlan ML, Maker G, Crighton E, Haile J. Combined DNA, toxicological and heavy metal analyses provides an auditing toolkit to improve pharmacovigilance of traditional Chinese medicine (TCM). Sci Rep 2015; 5: 17475

[72] Cheng X, Su X, Chen X, Zhao H, Bo C, Xu J, Bai H. Biological ingredient analysis of traditional Chinese medicine preparation based on high-throughput sequencing: the story for Liuwei Dihuang Wan. Sci Rep 2014; 4: 5147

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

[74] Amore RD, Ijaz UZ, Schirmer M, Kenny J, Gregory R, Darby AC, Quince C, Hall N. A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling. BMC Genomics 2016; 17: 55

[75] Meyer CP, Paulay G. DNA barcoding: error rates based on comprehensive sampling. PLoS Biol 2005; 3: e422

Please note: this article was changed according to the following erratum: The following Acknowledgements have been added: Marco Kreuzer was supported by MEDPLANT, a Marie Curie Actions Initial Training Network (ITN), funded by the European Union under the Seventh Framework Program.

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.