RESOURCE ARTICLE

Using target capture to address conservation challenges: Population-level tracking of a globally-traded herbal medicine

Vincent Manzanilla[1] | Irene Teixidor-Toneu[1] | Gary J. Martin[2] | Peter M. Hollingsworth[3] | Hugo J. de Boer[1] | Anneleen Kool[1]

[1]Natural History Museum, University of Oslo, Oslo, Norway

[2]Global Diversity Foundation, Canterbury, UK

[3]Royal Botanic Garden Edinburgh, Edinburgh, UK

Correspondence
Vincent Manzanilla, Natural History Museum, University of Oslo, Sars gate 1, 0562 Oslo, Norway.
Email: vincent.manzanilla@gmail.com

Abstract

The promotion of responsible and sustainable trade in biological resources is widely proposed as one solution to mitigate current high levels of global biodiversity loss. Various molecular identification methods have been proposed as appropriate tools for monitoring global supply chains of commercialized animals and plants. Here, we demonstrate the efficacy of target capture genomic barcoding in identifying and establishing the geographic origin of samples traded as *Anacyclus pyrethrum*, a medicinal plant assessed as globally vulnerable in the IUCN Red List of Threatened Species. Samples collected from national and international supply chains were identified through target capture sequencing of 443 low-copy nuclear makers and compared to results derived from genome skimming of plastome and DNA barcoding of standard plastid regions and ITS. Both target capture and genome skimming provided approximately 3.4 million reads per sample, but target capture largely outperformed standard plant barcodes and entire plastid genome sequences. We were able to discern the geographical origin of *Anacyclus* samples collected in Moroccan, Indian and Sri Lankan markets, differentiating between plant materials originally harvested from diverse populations in Algeria and Morocco. Dropping costs of analysing samples enables the potential of target capture to routinely identify commercialized plant species and determine their geographic origin. It promises to play an important role in monitoring and regulation of plant species in trade, supporting biodiversity conservation efforts, and in ensuring that plant products are unadulterated, contributing to consumer protection.

KEYWORDS
Anacyclus, DNA barcoding, genomic barcoding, international trade, supply chain, target capture

1 | INTRODUCTION

Human exploitation of biological resources is a major challenge for biodiversity conservation and sustainable development. Global trade and consumer demand for natural products provide increasing threats to species (Lenzen et al., 2012), and at the same time create markets in which regulation and authentication are extremely difficult (Newmaster et al., 2013; WHO, 2004). Many traded plant

products from wild populations are over-harvested. Increasing scarcity results in higher prices, and incentivises adulteration, substitution and poaching (Hamilton, 2004; Schippmann et al., 2002; Veldman et al., 2017). In recent decades, multilateral environmental agreements including the Convention on Biological Diversity (CBD) and the Convention on International Trade in Endangered Species (CITES) have addressed the trade of threatened species. In parallel, the World Health Organisation has developed guidelines for safety monitoring of herbal medicines in pharmacovigilance systems (WHO, 2004) and the Food and Agriculture Organisation, through the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA), has set up a multilateral system to promote sustainable agriculture (Esquinas-Alcazar, 2004). However, implementing these regulations and guidelines is hampered by difficulties identifying plant products in trade. Multiple, complex and interacting supply chains can coexist for a single plant product (Booker et al., 2012). Traded plant products are often not identifiable to species by their morphology or chemistry, as they may be dried, powdered, processed, or commercialised in mixtures with other ingredients. The design, implementation and enforcement of successful conservation actions as well as assessments of product authenticity and quality often require the identification of the geographic origin of species in trade. This is difficult as the development of efficient methods to identify and trace traded products are still in their infancy.

DNA barcoding has quickly gained popularity since Hebert et al. (2003) first advocated the use of short and variable DNA sequences, amplified using universal primers, for species identification and discovery of new taxa. DNA barcoding is highly effective for species-level identification in animals using a portion of the mitochondrial marker cytochrome cxidase 1 (COI) (Hebert et al., 2016; Hebert, Ratnasingham, et al., 2016). In plants, standard DNA barcoding involves one to four plastid DNA regions (*rbcL*, *matK*, *trnH-psbA*, *trnL*), sometimes in combination with internal transcribed spacers of nuclear ribosomal DNA (nrDNA, ITS) (CBOL et al., 2009; Kress, 2017). Although these markers are very informative in many cases, no single marker or combination of markers routinely provide complete species-level resolution, especially in species-rich groups, let alone population-level assignment. The development of high throughput sequencing (HTS) with new reagents and platforms expands the application of DNA barcoding in plants in a cost-effective fashion, in part because it removes the need of targeting short universal barcodes (Hollingsworth et al., 2016; Lemmon & Lemmon, 2013). Two major approaches, shallow pass shotgun sequencing and target capture sequencing, have been proposed for increasing the coverage and resolution of plant DNA barcoding.

Shallow pass shotgun sequencing (commonly referred to genome skimming) is used to recover organellar genomes and nuclear ribosomal DNA sequences, increasing the amount of data per sample and leading to some increases in resolution (Manzanilla et al., 2018; Parks et al., 2009). Although workflows and bioinformatic pipelines are increasingly refined for this approach, current cost constraints mean that most genome skimming barcoding projects only have sufficient sequencing depth to generate comparative data for multicopy

regions such as plastid genomes and ribosomal DNA. These regions represent a limited number of independent loci, ultimately constraining resolving power (Soltis & Soltis, 2009; Wood et al., 2009).

Target capture sequencing offers the potential to overcome these deficiencies by efficiently targeting hundreds of low-copy nuclear markers, providing access to a much greater number of independent data points per unit of sequencing effort (Degnan & Rosenberg, 2009). Similar to genome skimming, target capture is successful in sequencing samples with poor DNA integrity (Brewer et al., 2019; Forrest et al., 2019) and allows sequencing hundreds of samples at the same time (Mamanova et al., 2010). It can also be designed to recover standard DNA barcodes in the same assay (Schmickl et al., 2016). Although target capture has been advocated as a powerful tool for molecular identification of plants (Pillon et al., 2013), its usefulness for determining and tracing the origin of plants in trade remained untested.

We evaluated the power of target-capture DNA barcoding by investigating the traceability of plant products reputedly derived from a vulnerable medicinal plant species, *Anacyclus pyrethrum* (L.) Lag. In addition to having a well-established international trade chain (Rankou et al., 2015), this species presents classic challenges for plant molecular identification such as recent radiation, frequent hybridization (Humphries, 1979) and large genome size (Garcia et al., 2013).

The genus *Anacyclus* (Asteraceae) comprises 12 species of annual and perennial weedy herbs with partly overlapping geographic ranges around the Mediterranean basin (Humphries, 1979; Rosato et al., 2017). Some species are abundant and have wide geographical ranges (for example, *A. clavatus* (Desf.) Pers. and *A. radiatus* Loisel.), whereas others are rare and have restricted ranges (for example, *A. maroccanus* (Ball) Ball and *A. pyrethrum* (L.) Lag.).

A. pyrethrum has a long history of use in traditional Arabic and Islamic, Ayurvedic and European medicine (Adams et al., 2011; De Vos, 2010; Pittle, 2005). In the 13th century, Ibn al-Baytār wrote that the plant was known across the world and traded from the Maghreb to all other areas (Leclerc, 1877). Its popularity as a medicinal plant stems from the many attested and putative pharmacological activities of its roots (Manouze et al., 2017). Traded from the Maghreb to India (Ved & Goraya, 2007) and Nepal (Tiwari et al., 2004), it is known to be over-harvested and is increasingly difficult to find in local markets in Morocco (Ouarghidi et al., 2012, 2013).

There are two accepted varieties, *A. pyrethrum* var. *pyrethrum* and var. *depressus* (Ball) Maire, both endemic to Algeria, Morocco, and southern Spain (Humphries, 1979; Rosato et al., 2017). In Morocco, *A. pyrethrum* var. *pyrethrum* is considered more potent and is up to 10 times more expensive than var. *depressus* (Ouarghidi et al., 2012). Both varieties are harvested from the wild and used extensively for the treatment of pain and inflammatory disorders across Morocco (Ouarghidi et al., 2012, 2013) and Algeria (Benarba, 2016; Ouelbani et al., 2016), as well as the Middle East (Pittle, 2005) and the Indian subcontinent (Tiwari et al., 2004). Although collectors are proficient in differentiating the two varieties, material traded as *A. pyrethrum* is adulterated and misidentified along the chain of commercialisation (de Boer et al., 2014; Ouarghidi et al., 2012, 2013).

We applied target-capture genomic barcoding to distinguish *Anacyclus* species and geographical races in root samples in the national and international supply chains. We compared this novel approach with plastid genome and nrDNA ITS barcodes obtained from genome skimming data.

2 | MATERIALS AND METHODS

2.1 | Sample collection

We made a reference collection of 72 accessions, consisting of 67 of *Anacyclus* and five of closely related genera, including 56 *Anacyclus* accessions and three of closely related genera we collected in Morocco and Spain during field work. We selected a total of 11 samples of *Anacyclus* and two of closely related genera from herbarium voucher specimens of species occurring elsewhere in the Mediterranean. Voucher specimen identifications, collection numbers and locality data are listed in Table S1 and the collection locations are mapped in Figure S1. We purchased 110 samples, each consisting of approximately 50 g of roots from collectors, herbalists, middle-men, traditional healers, wholesalers, and export companies in Morocco and India (Table S2).

2.2 | Trade information

We obtained samples of commercialized species by asking collectors and traders in Morocco for tiguendizt and iguendez, two vernacular names used for both *A. pyrethrum* varieties (Ouarghidi et al., 2012). We asked traders in India for akarkara, a common name used there (Ved & Goraya, 2007). When acquiring the samples, we conducted semi-structured interviews about the roots' trade with 39 informants, enquiring where the plant material was sourced, to whom it was sold and in what quantities, for what price, and if some types of roots were of higher quality than others. We followed the International Society of Ethnobiology Code of Ethics ("ISE Code of Ethics Online," 2018), ensuring free, prior and informed consent, full disclosure and respect for the confidentiality of informants, during all interviews.

2.3 | Extraction and library preparation

We extracted DNA of reference and traded vouchers from approximately 40 mg of dry leaf or root material using the DNeasy Plant Mini Kit (Qiagen). When DNA integrity was preserved, between 0.2–1.0 µg was sheared to 500 bp fragments using a Covaris S220 sonicator (Woburn) (Tables S1–S2). Libraries for samples with degraded DNA were purified with 1x AMPure Beads (Agencourt) to preserve short fragments. We prepared dual indexed libraries using the Meyer and Kircher protocol (Meyer & Kircher, 2010) for genome skimming and target capture (BioProject PRJNA631886).

2.4 | Target capture

Using the genome assembly of *A. radiatus* Loisel. subsp. *radiatus*, we designed 872 low-copy nuclear markers and associated RNA probes by following the Hyb-Seq pipeline (Weitemier et al., 2014: Supporting Information). For target capture enrichment, we prepared twelve equimolar pools with ten to 24 samples and an average 300 ng of input DNA per pool. The RNA probes were hybridized for 16 h before target baiting, and 14 PCR cycles were carried out after enrichment following the MyBaits v3 manual. The enriched libraries and genome skimming libraries were sequenced on two Illumina HiSeq 3000 lanes (150 bp paired-end). Sequencing runs have been deposited under a single NCBI BioProject accession PRJNA631886.

2.5 | Data processing

We retrieved four data sets from the genome skimming and target capture sequencing methods: (1) standard barcode markers (*matK*, *trnH-psbA*, the *trnL* intron and *rbcL*), (2) ITS, (3) complete plastid genomes (from shotgun genome skimming), and (4) hundreds of nuclear markers (from target capture) (Figure S2). The sequencing runs were trimmed and quality filtered using Trimmomatic (Bolger et al., 2014). Low-copy nuclear markers and their alleles were retrieved for each sample. First, the reads were mapped against the selected low-copy nuclear markers (Supporting Information) using BWA v0.7.5a-r40 (Li & Durbin, 2009). Duplicate reads were removed using Picard v2.10.4 (Wysoker et al., 2015). Alleles were phased for each marker and individual using SAMtools v1.3.1 (Li et al., 2009). The last step of the pipeline combined the retrieved alleles into single gene matrices. We recovered plastome and ITS sequences by pooling shotgun and target enrichment sequencing data. Plastid genomes were built using MITOBim v1.8 (Hahn et al., 2013). ITS sequences were recovered using BWA by mapping the reads to the reference ITS of *Anacyclus pyrethrum* (KY397478) for *Anacyclus* species and traded samples, to the reference ITS of *Achillea pyrenaica* Sibth. ex Godr. (AY603247) for *Otanthus* and *Achillea*, and to the reference ITS of *Matricaria aurea* (Loefl.) Sch.Bip. (KT954177) for *Matricaria* samples. During the mapping and the assembly steps, we retained only sequences with a minimum coverage of 10x.

2.6 | Phylogenomics

We aligned the recovered matrices (nuclear markers, ITS and plastomes) with MAFFT v7.471 (Katoh & Standley, 2013), refined with MUSCLE (-refine) (Edgar, 2004) and filtered with Gblocks v0.91b (Talavera & Castresana, 2007). Phylogenies were inferred using RAxML v8.0.26 (Stamatakis, 2006), with 1,000 bootstrap replicates under the GTRGAMMA model. For the low-copy nuclear markers, the species tree was inferred from the individual nuclear markers trees under the multispecies coalescence (MSC) framework with ASTRAL-III v5.5.9 (Zhang et al., 2017). We used the multialleles option in ASTRAL-III for reconciliation of the independent evolutionary

histories of the alleles. The molecular identification of traded roots was assessed from the MSC tree and posterior probabilities (PP) greater than 0.95. For the ITS phylogenetic reconstruction, we used additional Genbank references (Table S3).

We identified traded roots based on morphological characters (described in Supporting Information), and we used these to triangulate molecular identifications. Samples were categorised according to their position in the supply chain and geographical origin.

3 | RESULTS

We constructed a reference database of DNA sequences from fresh and herbarium specimens, consisting of 83 individuals of 10 *Anacyclus* species, and five individuals representing outgroup species (Figure S1). We used this reference database to assess the identity and geographic origins of 110 root samples acquired from traded materials by comparing the results from the four different data sets (Figure 1). We show that the target capture approach is the most powerful method to identify plant species in trade and discover their geographic origin.

3.1 | Data recovery for genome skimming and target capture data

After quality control filtering, an average 2.8 million reads (0.84 GB/sample) were obtained per sample from shotgun sequencing and

2.99 million reads (0.90 GB/sample) per sample for target capture (Figure S3, Table S4). The target capture yielded an average coverage of 303x for the 443 nuclear markers, whereas the unenriched genome skimming yielded an average coverage of 12x for the 443 nuclear markers, 20x for the plastome data and 131x for the ribosomal data (Table S5). The loci coverage was calculated with bedtools coverage (v 2.29.2). Samples below an average of 50x coverage in the nuclear data set show a higher missing data rate. Samples having over 7% missing data were excluded as stated in the Supporting Information (Figure 2, samples in blue). Adulterated samples from other genera have a coverage close to zero (Figure 2, samples in orange, yellow and green). To obtain 100x coverage for the nuclear regions using a genome skimming approach, with an average genome size of the targeted species of 11.72 Gb (Garnatje et al., 2011) and a duplication level of the genome skimming libraries of 9% (Supporting Information), it would require 14 HiSeq 3000/4000 lanes (Illumina). From the genome skimming data we assembled ITS and the standard barcodes markers, as well as the plastome (Figure S2). Out of the 110 trade samples, we succeeded in assembling ITS for 102 samples (93%), the standard barcode regions for 51 to 61 samples (46% to 55%), and plastomes for 49 samples (44%) (Figure 1). The resulting aligned matrices for each of the data sets were 633 bp for ITS (including 5.8S), 4,408 bp for the standard barcoding regions, 110,003 bp for the plastome and 289,236 bp for the 443 nuclear markers recovered from the target capture approach (Table S1). The standard barcoding regions included the full coding regions of *matK* 1,523 bp and *rbcL* 1,438 bp, as well as
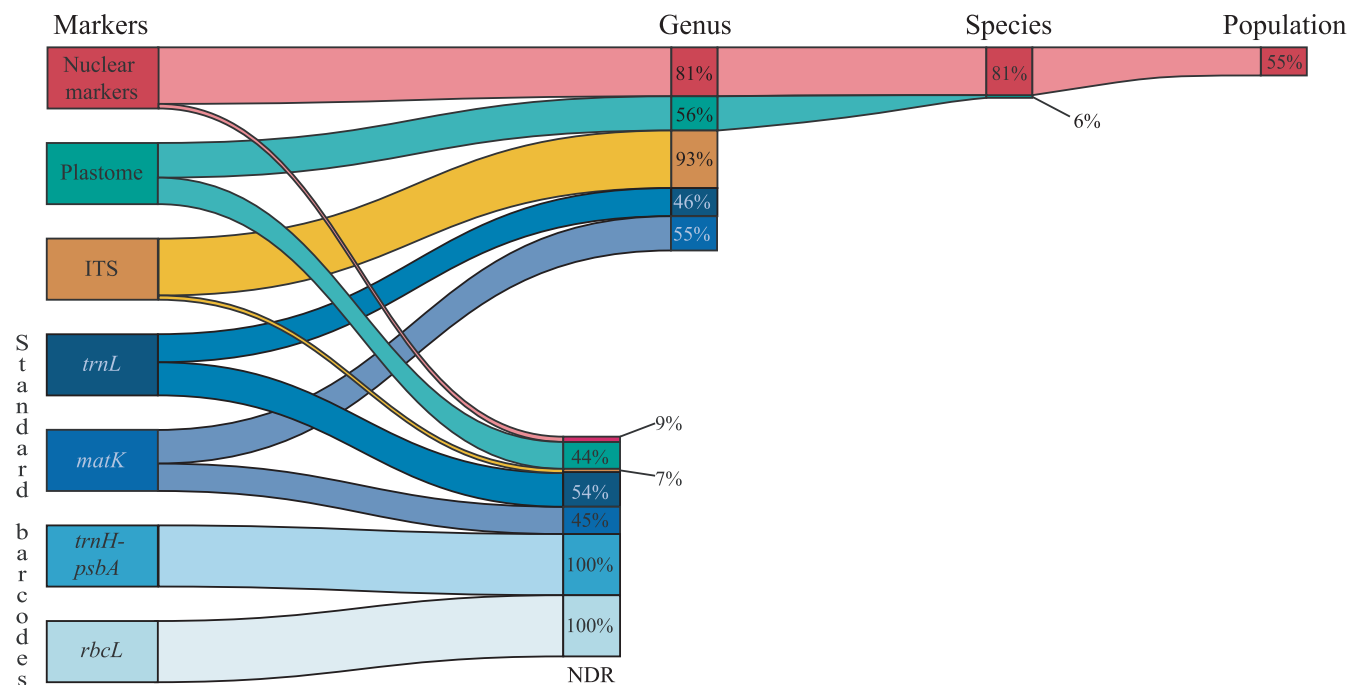

FIGURE 1 Sequencing recovery and identification success for the traded samples for each data set. The figure shows the percentage of samples for which useful marker sequences were successfully retrieved for molecular identification for standard barcodes, ITS, plastomes and nuclear markers. No data recovered (NDR) is used to indicate samples for which no sequence data was recovered or where no identification at genus level or below could be made. For the samples that produced useable data, the proportion of samples that resulted in identification at the genus, species and population levels is given
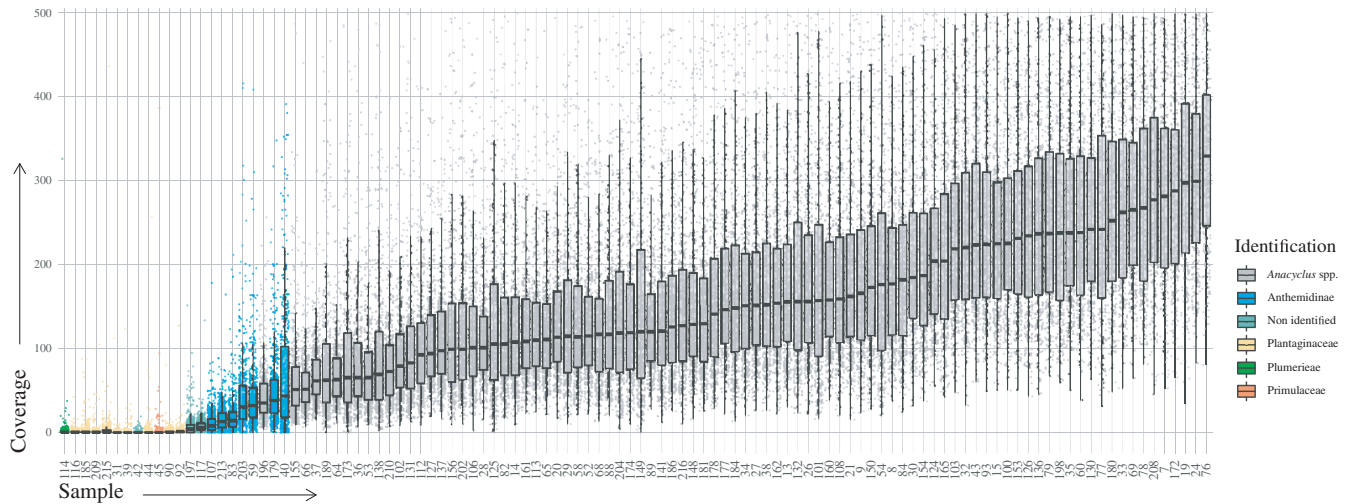
FIGURE 2 Box-and-whiskers plots showing retrieved coverage of 443 targeted markers for each sample. Identified and nonidentified samples are colour coded

trnH-psbA 500 bp and trnL 947 bp. The bioinformatic workflow for data analyses is described in Figure S2.

3.2 | Comparative levels of species discrimation using different approaches

The ITS, plastome, and standard barcodes phylogenies highlight the complex evolutionary history of *Anacyclus*. The ITS phylogeny lacks resolution in general (Figures S4–S5). The outgroups *Tanacetum*, *Matricaria*, *Achillea*, *Otanthus* and *Tripleurospermum* have well-supported bootstrap values, but within the genus *Anacyclus*, only *A. atlanticus* Litard. & Maire, *A. maroccanus* Ball and *A. radiatus* are highly supported. The plastome phylogeny shows very good support at genus level for the *Anacyclus* node, and at species level for the outgroups. The lack of variation in the plastid genome within the genus *Anacyclus* results in little phylogenetic support with no species-specific clusters recovered (Figures S6–S7). The standard barcode regions, *matK*, *rbcL*, *trnH-psbA* and *trnL* (Figures S8–S11) displayed low levels of resolution at the species level, even using the full coding regions of *matK* and *rbcL* (e.g., rather than the 800–900 bp of *matK* and 654 bp of *rbcL* typically recovered using standard barcoding primers (Alsos et al., 2020; Hollingsworth et al., 2011).

The 443 nuclear markers recovered by target capture led to a well-resolved phylogeny and high levels of species discrimination: all the genera in the Matricariinae tribe and all interspecific relationships are well-supported, with most nodes showing posterior probabilities (PP) of 1 (Figure 3 and Figure S12). Within *Anacyclus*, all species, sub-species and varieties are well supported. PP are lower for *A. monanthos* (L.) Thell. (PP =0.75). The complex of hybrid species composed of *A. clavatus* (Desf.) Pers., *A. homogamos* (Maire) Humphries, and *A.* x *valentinus* L. is polyphyletic and shows signs of hybridization and incomplete lineage sorting. Intraspecific nodes have PP varying between 0.27 and 1, mostly depending on species population structure.

3.3 | Assessment of *Anacyclus* trade

Interviews with 39 harvesters, middlemen, retailers, and wholesalers in various Moroccan cities indicate that the national and international trade of *Anacyclus pyrethrum* follow two separate supply chains (Supporting Information). Retailer herbalists in Moroccan cities are supplied by middlemen who acquire the plant material from local harvesters from rural communities. These retailers typically hold between a few hundred grams to one kilogram of the plant material in their shops. In contrast, wholesalers who export the plant internationally hire professional harvesters that travel across the geographical range of the species to collect plant material. Harvested roots are brought directly from the wild to the export companies in Rabat, Casablanca and Tangier, from where they enter the international market, including supply of material to India. According to informants from export companies, between 3–10 tons of the plant product can be stocked at a time.

Our examination of material in trade involved screening a total of 66 samples each containing an average of 25 g of dry roots. Initial morphological examination identified obvious non-*Anacyclus* adulterants in 39/66 batches. The adulterants were present in a proportion from 3% to 100% with an average of 42%. The non-*Anacyclus* adulterants were found at high frequency in collections from traditional healers and herbalists, less so from collectors, wholesalers and export companies (Figures 4–5, Table S6).

We selected 110 individual roots for DNA analysis from the 66 root batches. Of these 99 had a morphology consistent with *Anacyclus*, and 11 were classed as similar to *Anacyclus* but likely to be non-*Anacyclus* based on their morphology. We recovered partial plastome assemblies of these 11 non-*Anacyclus* roots and using sequence queries against GenBank, we obtained identifications for nine *Plantago* spp., one *Primula* spp. and one *Plumeria* spp. (Figure S13, Table S2).

Of the 99 *Anacyclus* roots, 10 had no identifiable DNA sequences via any of our methods, with all of the remaining samples identified
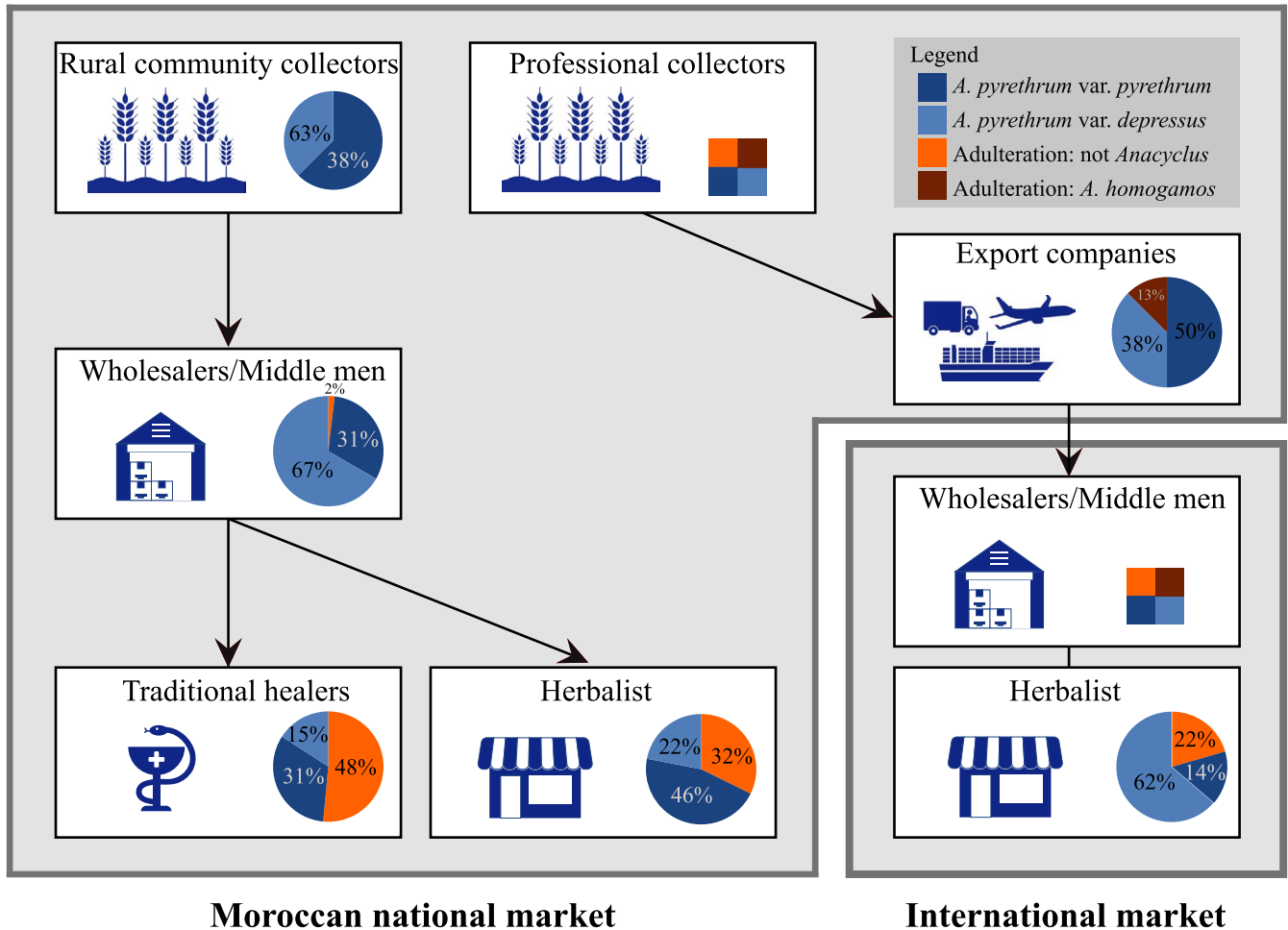
FIGURE 3 Multispecies coalescent phylogenetic tree of the nuclear loci data set. The reference data set includes taxon labels and herbarium accession numbers and the traded samples are numbered. Supported clades with an associated geographic origin are indicated by coloured bars

FIGURE 4 National and international supply chains of *A. pyrethrum*. Pie charts represent the proportion of *A. pyrethrum* (light and dark blue represent var. *depressus* and var. *pyrethrum* respectively) and adulterated samples (orange and brown for *A. homogamos* and other adulterants) by each stakeholder. We were unable to obtain samples from wholesalers/middlemen in India or professional collectors in Morocco (indicated by square boxes)
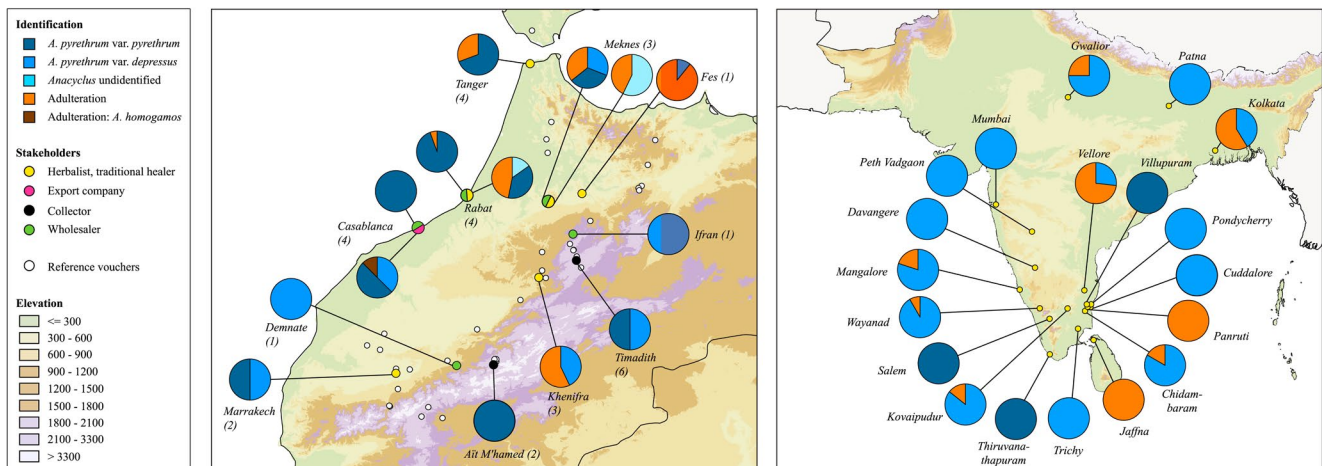

FIGURE 5 Species identification of market samples. Sample locations are shown with coloured circles according to the type of stakeholder. A pie-chart with the proportions of adulteration and identified species is represented for each location in (a) Morocco (native range) and (b) India (exported material)

through target capture (Figure S13). The 10 discarded samples presented very fragmented plastome assemblies with average 89% missing data in the nuclear matrices. These samples had a very low DNA integrity.

The plastome sequences enabled identification of seven roots to the species level, with the remainder identified as *Anacyclus* sp. (Figure 1, Table S7). Neither ITS nor any of the standard barcode regions were able to discriminate any of these samples below the genus level (Figure 1, Figures S4–S5, Figures S8–S11).

The nuclear markers gave much higher resolution within *Anacyclus* (Tables S2, S7). For plant material traded within Morocco, our analyses of six individual root samples from four rural community collectors identified three *A. pyrethrum* var. *pyrethrum* and three var. *depressus*. Our analysis of five samples from three wholesaler "middle-men" in Morocco identified two *A. pyrethrum* var. *pyrethrum* and three var. *depressus*. Our sequences from 19 samples from 10 herbalists revealed 12 *A. pyrethrum* var. *pyrethrum* and seven var. *depressus*. Likewise, our 11 samples from six traditional healer sources identified seven *A. pyrethrum* var. *pyrethrum*, and four var. *depressus*. For material traded in international markets, our 17 sequenced samples from three export companies in Morocco identified nine *A. pyrethrum* var. *pyrethrum*, six var. *depressus*, and two *A. homogamous*. Our analysis of 30 samples from 17 herbalists in India identified three *A. pyrethrum* var. *pyrethrum* and 27 var. *depressus*.

3.4 | Geographical source

Market samples in Morocco originate from various Moroccan areas as well as Algeria, and material from all of these populations of origin can be found in Indian market samples (Table S1, Figures 3 and 6). Of the 99 nonadulterated roots identified to species level using target capture, we were able to associate 67% to a specific geographic region (Figures 1 and 3, Table S1). Using phylogenetic analysis, root samples of *A. pyrethrum* var. *pyrethrum* clustered with reference material from the High Atlas (Figure 6, case 4), and *A. pyrethrum* var. *depressus* roots clustered with reference material from different regions in Morocco, including the Rif Mountains, the High Atlas and the Middle Atlas (Figure 6, case 2 and 3). Evidence for the international trade from Algeria to Morocco is highlighted by a distinctive clade that includes traded roots collected from west Algeria (Figure 6, case 1). The geographic origin was only resolvable with target capture data; standard barcoding regions, ITS and plastome data lacked variation, resolution or both (Figure 1).

4 | DISCUSSION

This study illustrates the potential for target-capture based DNA barcoding to form the next wave of standard plant DNA barcoding tools and provide greatly needed species-level resolution. A key rate-limiting step for the standard plant barcodes is that they are fundamentally recovering data from just one or two independent

loci (plastid DNA and ITS), which often show trans-specific polymorphism and barcode sharing among related species (Hollingsworth et al., 2011). The use of complete plastid genome sequences suffers from the same problem, as the data are all physically linked in a single nonrecombinant uniparentally inherited locus. Several recent hybridization events have occurred in the *Anacyclus* genus and entire plastid genomes or plastid barcode markers, and/or ITS provide limited resolution below genus level (Figure 1). In contrast, our target capture approach using hundreds of nuclear markers yields significantly higher molecular identification success and more accurate resolution at species and even population level (Figure 1).

4.1 | Species identification, species in trade, and geographic origins of *Anacyclus*

These data provide new insights into trade of *A. pyrethrum* and highlight the extent of adulteration and scarcity of *A. pyrethrum* var. *pyrethrum* (Figures 2–3 and 6). Only a small proportion of the tested samples from herbalists and traditional healers were the potent *A. pyrethrum* var. *pyrethrum*, with the Indian market in particular dominated by var. *depressus*. In both Morocco and India some individual sellers had entirely or almost entirely adulterated products. Most of the non-*Anacyclus* roots we sequenced were identified as *Plantago* spp. by shotgun sequencing, despite being sampled from six different localities including Morocco and India. As *Plantago* roots are similar in appearance to those of *Anacyclus*, it is possible that they are deliberately added as a "difficult to identify" adulterant which may go unnoticed by nonspecialists.

Our analysis of samples from collectors, wholesalers and export companies detected much less adulteration at this point in the supply chain (Figure 4, Table S6). Collection of *A. pyrethrum* var. *pyrethrum* is carried out by professional harvesters employed by export companies who travel across the country and are considered poachers by local communities (Ouarghidi et al., 2017). Local harvesters have increasing difficulty to supply local trade chains (Ouarghidi et al., 2012, 2017), which may finally result in increased adulteration rates in the poorly-governed, national value chains (Figure 4), as has also been observed elsewhere (Booker et al., 2012). Our results also identify previously unreported international trade in North Africa prior to export to the Indian subcontinent (Figures 3 and 6). We provide evidence that export companies in Morocco source material not only in this country, but also from neighbouring Algeria (Figure 6). Applying this molecular identification approach enables us to distinguish samples at population level and uncover these hidden international sourcing channels.

4.2 | Conservation of *A. pyrethrum*

High national and international demand for *A. pyrethrum* probably encourages its overharvesting and adulteration. As the plant is a remedy of the Indian pharmacopoeia, its demand is likely to
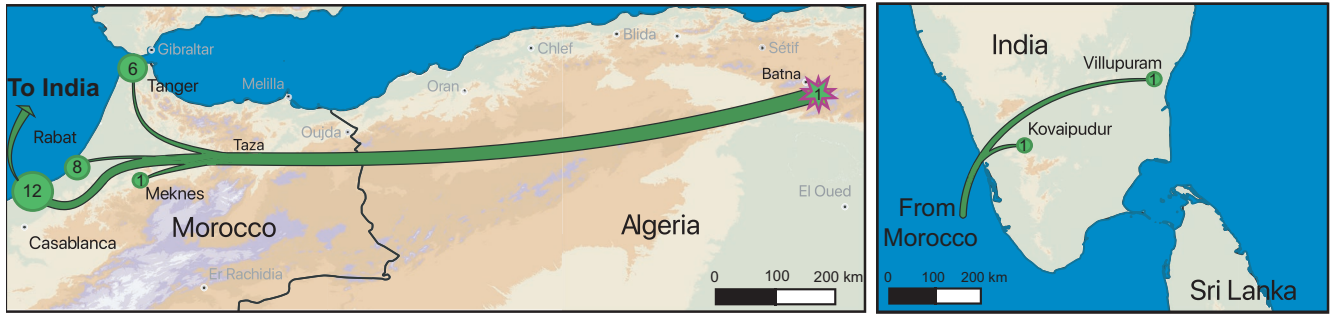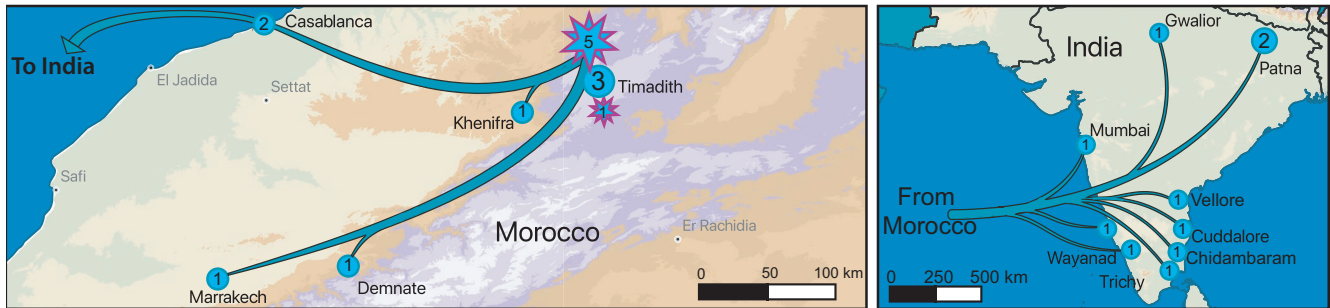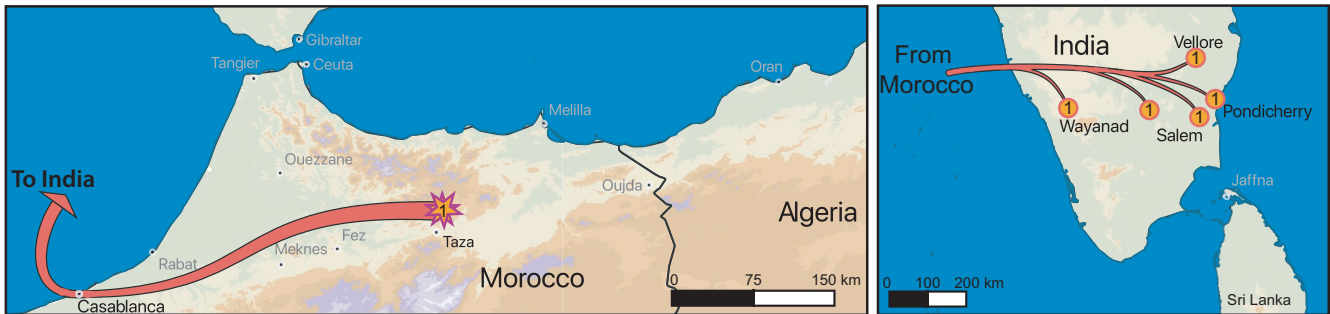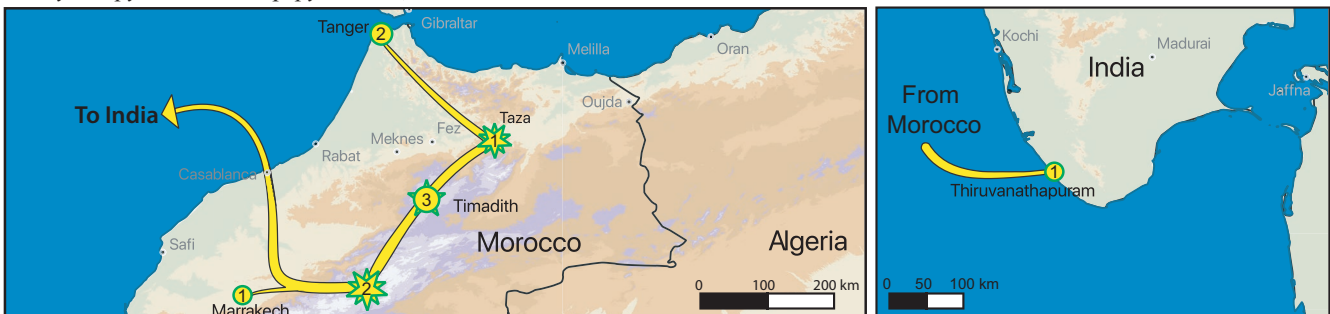
Anacyclus pyrethrum subsp. *pyrethrum*, case 1

Anacyclus pyrethrum subsp. *depressus*, case 2

Anacyclus pyrethrum subsp. *depressus*, case 3

Anacyclus pyrethrum subsp. *pyrethrum*, case 4

FIGURE 6 Map of the origin of the traded samples. The figure relates three supported clades mentioned in the Figure 3. The stars indicate reference samples and the dots the traded samples. The number in the circle or the start corresponds to the number of individuals

increase along with that of other Ayurvedic medicines (Kala et al., 2006). Although *A. pyrethrum* has been assessed to be vulnerable globally (Rankou et al., 2015) and endangered in Morocco on the IUCN Red List, the plant is not listed in the CITES appendices and its international trade is not regulated. Nonetheless, continued overharvesting is driving wild populations to critical levels of depletion and conservation policies are necessary. Common strategies to conserve overharvested medicinal plants include collection and trade restrictions as well as cultivation (Schippmann et al., 2002). Cultivation is often proposed as a solution to both conservation

issues and sourcing high quality, appropriately identified material (Hamilton, 2004; Schippmann et al., 2002, 2006). Cultivation necessitates engagement with local communities that depend on plant harvest, as well as monitoring of professional trading networks. Both promotion of sustainable harvesting for livelihood security as well as restriction of unsustainable professional trade may be needed. Only with fine-grained mapping of sourcing areas and supply chains, as our results highlight for *A. pyrethrum*, can meaningful conservation action be achieved. With the implementation of target enrichment, we identified the origin of the harvested populations of medicinal or traded populations and indicate where conservation efforts could be initially implemented. We also reveal previously undocumented harvesting and trade in Algeria (Figure 6, case 1), and show that *Anacyclus* is harvested at a national scale in Morocco (Figure 6, case 2–4). This study gives scientific evidence to support conservation initiatives like Global Diversity Foundation's High Atlas Cultural Landscapes Programme and potentially encourage future *Anacyclus* conservation projects.

4.3 | Future prospects for plant DNA barcoding

Key criteria for developing new DNA barcoding approaches for plants include resolving power (differentiating between taxa and discovering their geographical origins), recoverability (enabling use on a wide diversity of tissue sources), and cost and efficiency (enabling scaling over very large sample sets).

In terms of resolving power, the target capture approach used here offers substantial improvement compared to plant barcodes based on plastid sequences and ITS. The key enabling step is access to multiple nuclear markers, as this reduces sensitivity of the identification to predominance by one or two loci as is the case for barcodes from rDNA or the plastid genome. In this study, we show high resolution from nuclear markers retrieved by target capture for identification of species in a genus that has undergone recent hybridization events (Manzanilla, 2018). The successful recovery of these markers from low quality input DNA is also important and the combination of these observations makes the case that this method provides a viable solution for future plant barcoding. In contrast, shotgun sequencing requires a substantial sequencing effort to retrieve the same suite of nuclear markers for species with large genomes. Thus for *Anacyclus* it would require 14 HiSeq 3000 lanes to obtain the same coverage of the loci we have used here. Overall, and regardless of approach, our findings provide empirical evidence to support predictions made in previous review papers (Hollingsworth et al., 2016) that barcoding based on nuclear markers should outperform standard barcoding methods.

The successful recovery of sequences via target capture depends on how closely related the sampled species are to the reference set on which the baits were designed, and/or the level of variation in the loci that form the bait set (McCormack et al., 2016; Paijmans et al., 2016). In the current study, the baits were designed from *Anacyclus* and related genera. This optimized their specificity to our study

group and enabled their successful high-resolution application for assessing trade. A clear challenge for wider use of target capture is applicability over much greater phylogenetic distances. The recently published universal angiosperm bait kit (Buddenhagen et al., 2016; Johnson et al., 2018), designed to recover 353 loci from a wide diversity of flowering plants, offers great potential here as they can provide species and population-level resolution (Van Andel et al., 2019). There is a need for a more general evaluation of when, how and at what scale to most effectively combine taxon-specific bait sets with universal bait sets to simultaneously obtain very high resolution and sequencing success over wide phylogenetic distances.

Another important aspect of recoverability is efficacy with degraded DNA. Drying, storage, and transportation affect the quality of plant material in trade and can cause extensive DNA degradation (Booker et al., 2014). In consequence, traded samples have similar challenges to working with ancient DNA, herbarium samples or archaeological remains. Target capture is particularly well suited to this challenge (McCormack et al., 2016; Paijmans et al., 2016). Although shotgun sequencing can also be very effective on degraded material (Alsos et al., 2020; Bakker, 2017; Zeng et al., 2018), our recovery rate in this study was greater for target capture than shotgun sequencing (Figure 1), and we recovered data from 100% of the samples used for establishing the reference library and over 70% of the samples in trade. This success rate for hundreds of nuclear markers providing high resolution from suboptimal tissue of traded samples is noteworthy. The other mainstream approach for highly degraded DNA is sequencing a portion of the chloroplast *trnL* (UAA) intron, specifically the P6 loop (10–143 bp) (Taberlet et al., 2007), which has been highly successful in recovering sequence data from degraded samples (Parducci et al., 2012; Willerslev et al., 2014). However, this short region of the plastid genome has a low variation at the species level and does not typically discriminate among con-generic species (Taberlet et al., 2007).

The target capture methodology presented here, including library construction and sequencing, cost 70 USD per sample in 2016, a price similar to those presented by Hale et al. (2020). With the optimisations suggested in Hale et al. (2020), prices drop to 22 USD per sample. Thus, this target enrichment approach is quickly becoming affordable for large-scale biomonitoring projects. With a growing interest and investment in DNA-based identification solutions in medicine and industry (Afshinnekoo et al., 2017; Menegon et al., 2017), ongoing work is expected to continue to optimize protocols and drive these costs down even further, reaching those of standard barcoding approaches (Hebert et al., 2018).

5 | CONCLUSIONS

In plants, the frequent sharing of plastid and ribosomal sequences among con-generic species, coupled with the difficulty of routinely accessing multiple nuclear markers, has acted as a constraint on the resolution of DNA barcoding approaches. Current advances in sequencing technology and bioinformatics are removing this

constraint and offer the potential for a new wave of high-resolution identification tools for plants. These approaches, such as target capture, have the capability to distinguish species and populations, providing insights into diversity and ecology, as well as the multitude of societal applications which require information on the identification and provenance of biological materials.

AUTHOR CONTRIBUTIONS

The project was coordinated by Anneleen Kool, Gary J. Martin, Hugo J. de Boer and Vincent Manzanilla. Vincent Manzanilla did the design of the study and performed data analysis. Anneleen Kool, Hugo J. de Boer, Irene Teixidor-Toneu, Peter M. Hollingsworth and Vincent Manzanilla wrote the manuscript. All authors provided contributions to data analysis and interpretation of the results. All authors have read and approved the final version of the manuscript.

OPEN RESEARCH BADGES

This article has earned an Open Data badge for making publicly available the digitally-shareabledata necessary to reproduce the reported results. The data is available at https://osf.io/9bh3p/.

DATA AVAILABILITY STATEMENT

For reproducibility purposes, all the scripts used during the data processing have been made available on the OSF work repository https://osf.io/9bh3p/. New sequencing data have been deposited under a single NCBI BioProject accession PRJNA631886.

REFERENCES

Adams, M., Alther, W., Kessler, M., Kluge, M., & Hamburger, M. (2011). Malaria in the Renaissance: remedies from European herbals from the 16th and 17th century. *Journal of Ethnopharmacology*, *133*(2), 278–288. https://doi.org/10.1016/j.jep.2010.10.060

Afshinnekoo, E., Chou, C., Alexander, N., Ahsanuddin, S., Schuetz, A. N., & Mason, C. E. (2017). Precision metagenomics: rapid metagenomic analyses for infectious disease diagnostics and public health surveillance. *Journal of Biomolecular Techniques: JBT*, *28*(1), 40. https://doi.org/10.7171/jbt.17-2801-007

Alsos, I. G., Lavergne, S., Merkel, M. K. F., Boleda, M., Lammers, Y., Alberti, A., Pouchon, C., Denoeud, F., Pitelkova, I., Puşcaş, M., Roquet, C., Hurdu, B.-I., Thuiller, W., Zimmermann, N. E., Hollingsworth, P. M., & Coissac, E. (2020). The treasure vault can be opened: Large-scale genome skimming works well using herbarium and silica gel dried material. *Plants*, *9*(4), 432. https://doi.org/10.3390/plants9040432

Bakker, F. T. (2017). Herbarium genomics: skimming and plastomics from archival specimens. *Webbia*, *72*(1), 35–45. https://doi.org/10.1080/00837792.2017.1313383

Benarba, B. (2016). Medicinal plants used by traditional healers from South-West Algeria: An ethnobotanical study. *Journal of Intercultural Ethnopharmacology*, *5*(4), 320. https://doi.org/10.5455/jice.20160814115725

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Booker, A., Frommenwiler, D., Johnston, D., Umealajekwu, C., Reich, E., & Heinrich, M. (2014). Chemical variability along the value chains of turmeric (*Curcuma longa*): a comparison of nuclear magnetic resonance spectroscopy and high performance thin layer chromatography. *Journal of Ethnopharmacology*, *152*(2), 292–301. https://doi.org/10.1016/j.jep.2013.12.042

Booker, A., Johnston, D., & Heinrich, M. (2012). Value chains of herbal medicines-Research needs and key challenges in the context of ethnopharmacology. *Journal of Ethnopharmacology*, *140*(3), 624–633. https://doi.org/10.1016/j.jep.2012.01.039

Buddenhagen, C., Lemmon, A. R., Lemmon, E. M., Bruhl, J., Cappa, J., Clement, W. L., & Kortyna, M. (2016). Anchored phylogenomics of angiosperms I: Assessing the robustness of phylogenetic estimates. *BioRxiv*, 086298.

de Boer, H. J., Ouarghidi, A., Martin, G., Abbad, A., & Kool, A. (2014). DNA barcoding reveals limited accuracy of identifications based on folk taxonomy. *PLoS One*, *9*(1), e84291. https://doi.org/10.1371/journal.pone.0084291

De Vos, P. (2010). European materia medica in historical texts: longevity of a tradition and implications for future use. *Journal of Ethnopharmacology*, *132*(1), 28–47. https://doi.org/10.1016/j.jep.2010.05.035

Degnan, J. H., & Rosenberg, N. A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, *24*(6), 332–340. https://doi.org/10.1016/j.tree.2009.01.009

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Esquinas-Alcazar, J. (2004). International treaty on plant genetic resources for food and agriculture. *Plant Genetic Resources Newsletter*, *11*, 1.

Garcia, S., Hidalgo, O., Jakovljević, I., Siljak-Yakovlev, S., Vigo, J., Garnatje, T., & Vallès, J. (2013). New data on genome size in 128 Asteraceae species and subspecies, with first assessments for 40 genera, 3 tribes and 2 subfamilies. *Plant Biosystems - an International Journal Dealing with All Aspects of Plant Biology*, *147*(4), 1219–1227. https://doi.org/10.1080/11263504.2013.863811

Garnatje, T., Canela, M. Á., Garcia, S., Hidalgo, O., Pellicer, J., Sánchez-Jiménez, I., Siljak-Yakovlev, S., Vitales, D., & Vallès, J. (2011). GSAD: A genome size in the Asteraceae database. *Cytometry Part A*, *79A*(6), 401–404.

Hahn, C., Bachmann, L., & Chevreux, B. (2013). Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Research*, *41*(13), e129. https://doi.org/10.1093/nar/gkt371

Hale, H., Gardner, E. M., Viruel, J., Pokorny, L., & Johnson, M. G. (2020). Strategies for reducing per-sample costs in target capture

sequencing for phylogenomics and population genomics in plants. *Applications in plant sciences*, 8(4), e11337.

Hamilton, A. (2004). Medicinal plants, conservation and livelihoods. *Biodiversity & Conservation*, 13(8), 1477–1517. https://doi.org/10.1023/B:BIOC.0000021333.23413.42

Hebert, P. D. N., Braukmann, T. W. A., Prosser, S. W. J., Ratnasingham, S., deWaard, J. R., Ivanova, N. V., Janzen, D. H., Hallwachs, W., Naik, S., Sones, J. E., & Zakharov, E. V. (2018). A Sequel to Sanger: amplicon sequencing that scales. *BMC Genomics*, 19(1), 219. https://doi.org/10.1186/s12864-018-4611-3

Hollingsworth, P. M., Graham, S. W., & Little, D. P. (2011). Choosing and Using a Plant DNA Barcode. *PLoS One*, 6(5), e19254. https://doi.org/10.1371/journal.pone.0019254

Hollingsworth, P. M., Li, D.-Z., van der Bank, M., & Twyford, A. D. (2016). Telling plant species apart with DNA: from barcodes to genomes. *Philosophical Transactions of the Royal Society B*, 371(1702), 20150338. https://doi.org/10.1098/rstb.2015.0338

Humphries, C. J. (1979). A revision of the genus *Anacyclus* L. (Compositae: Anthemideae). *Bulletin of the British Museum (Natural History). Historical Series*, 7, 83–142.

ISE Code of Ethics Online (2018). *International Society of Ethnobiology*. Retrieved http://www.ethnobiology.net/what-we-do/core-programs/ise-ethics-program/code-of-ethics/code-in-english/ Accessed March 3, 2018

Johnson, M. G., Pokorny, L., Dodsworth, S., Botigué, L. R., Cowan, R. S., Devault, A., Eiserhardt, W. L., Epitawalage, N., Forest, F., Kim, J. T., Leebens-Mack, J. H., Leitch, I. J., Maurin, O., Soltis, D. E., Soltis, P. S., Wong, G.-S., Baker, W. J., & Wickett, N. J. (2018). A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Systematic Biology*, 68(4), 594–606. https://doi.org/10.1093/sysbio/syy086

Kala, C. P., Dhyani, P. P., & Sajwan, B. S. (2006). Developing the medicinal plants sector in northern India: challenges and opportunities. *Journal of Ethnobiology and Ethnomedicine*, 2(1), 32. https://doi.org/10.1186/1746-4269-2-32

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772–780. https://doi.org/10.1093/molbev/mst010

Leclerc, L. (1877). *Traité des simples (Vol. 1)*. Imprimerie nationale.

Lemmon, E. M., & Lemmon, A. R. (2013). High-throughput genomic data in systematics and phylogenetics. *Annual Review of Ecology, Evolution, and Systematics*, 44, 99–121. https://doi.org/10.1146/annurev-ecolsys-110512-135822

Lenzen, M., Moran, D., Kanemoto, K., Foran, B., Lobefaro, L., & Geschke, A. (2012). International trade drives biodiversity threats in developing nations. *Nature*, 486(7401), 109. https://doi.org/10.1038/nature11145

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). Genome project data processing S. The sequence alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079.

Mamanova, L., Coffey, A. J., Scott, C. E., Kozarewa, I., Turner, E. H., Kumar, A., Howard, E., Shendure, J., & Turner, D. J. (2010). Target-enrichment strategies for next-generation sequencing. *Nature Methods*, 7(2), 111. https://doi.org/10.1038/nmeth.1419

Manouze, H., Bouchatta, O., Gadhi, A. C., Bennis, M., Sokar, Z., & Ba-M'hamed, S. (2017). Anti-inflammatory, antinociceptive, and antioxidant activities of methanol and aqueous extracts of Anacyclus pyrethrum roots. *Frontiers in Pharmacology*, 8, 598. https://doi.org/10.3389/fphar.2017.00598

Manzanilla, V. (2018). *Raising the bar—Molecular identification of plants (Doctoral thesis, University of Oslo)*. University of Oslo, Oslo. Retrieved from https://www.duo.uio.no/handle/10852/61935?show=full

Manzanilla, V., Kool, A., Nguyen Nhat, L., Van Nong, H., Le Thi Thu, H., & de Boer, H. J. (2018). Phylogenomics and barcoding of Panax: toward the identification of ginseng species. *BMC Evolutionary Biology*, 18(1), 44. https://doi.org/10.1186/s12862-018-1160-y

McCormack, J. E., Tsai, W. L., & Faircloth, B. C. (2016). Sequence capture of ultraconserved elements from bird museum specimens. *Molecular Ecology Resources*, 16(5), 1189–1203. https://doi.org/10.1111/1755-0998.12466

Menegon, M., Cantaloni, C., Rodriguez-Prieto, A., Centomo, C., Abdelfattah, A., Rossato, M., Bernardi, M., Xumerle, L., Loader, S., & Delledonne, M. (2017). On site DNA barcoding by nanopore sequencing. *PLoS One*, 12(10), e0184741. https://doi.org/10.1371/journal.pone.0184741

Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, 2010(6), pdb.prot5448. https://doi.org/10.1101/pdb.prot5448

Newmaster, S., 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/10.1186/1741-7015-11-222

Ouarghidi, A., Martin, G., Powell, B., Esser, G., & Abbad, A. (2013). Botanical identification of medicinal roots collected and traded in Morocco and comparison to the existing literature. *Journal of Ethnobiology and Ethnomedicine*, 9(1), 59.

Ouarghidi, A., Powell, B., Martin, G. J., & Abbad, A. (2017). Traditional sustainable harvesting knowledge and distribution of a vulnerable wild medicinal root (*A. pyrethrum* var. *pyrethrum*) in Ait M'hamed Valley, Morocco. *Economic Botany*, 71(1), 83–95.

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

Ouelbani, R., Bensari, S., Mouas, T. N., & Khelifi, D. (2016). Ethnobotanical investigations on plants used in folk medicine in the regions of Constantine and Mila (North-East of Algeria). *Journal of Ethnopharmacology*, 194, 196–218. https://doi.org/10.1016/j.jep.2016.08.016

Paijmans, J. L., Fickel, J., Courtiol, A., Hofreiter, M., & Förster, D. W. (2016). Impact of enrichment conditions on cross-species capture of fresh and degraded DNA. *Molecular Ecology Resources*, 16(1), 42–55. https://doi.org/10.1111/1755-0998.12420

Parducci, L., Jorgensen, T., Tollefsrud, M. M., Elverland, E., Alm, T., Fontana, S. L., Bennett, K. D., Haile, J., Matetovici, I., Suyama, Y., Edwards, M. E., Andersen, K., Rasmussen, M., Boessenkool, S., Coissac, E., Brochmann, C., Taberlet, P., Houmark-Nielsen, M., Larsen, N. K., … Willerslev, E. (2012). Glacial survival of boreal trees in northern scandinavia. *Science*, 335(6072), 1083–1086. https://doi.org/10.1126/science.1216043

Parks, M., Cronn, R., & Liston, A. (2009). Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biology*, 7(1), 84. https://doi.org/10.1186/1741-7007-7-84

Pillon, Y., Johansen, J., Sakishima, T., Chamala, S., Barbazuk, W. B., Roalson, E. H., & Stacy, E. A. (2013). Potential use of low-copy nuclear genes in DNA barcoding: a comparison with plastid genes in two Hawaiian plant radiations. *BMC Evolutionary Biology*, 13(1), 35. https://doi.org/10.1186/1471-2148-13-35

Pittle, K. D. (2005). *Continuity and change in islamic ethnopharmacological practice: New methods for cognitive dialectometry*. Florida State University.

Rankou, H., Ouhammou, A., Taleb, M., Manzanilla, V., & Martin, G. (2015). Anacyclus pyrethrum. The IUCN Red List of Threatened Species 2015: E.T202924A53798702, 2015(2015). Retrieved from http://dx.doi.org/10.2305/IUCN.UK.2015-4.RLTS.T202924A53798702.en

Rosato, M., Álvarez, I., Feliner, G. N., & Rosselló, J. A. (2017). High and uneven levels of 45S rDNA site-number variation across wild populations of a diploid plant genus (Anacyclus, Asteraceae). PLoS One, 12(10), e0187131. https://doi.org/10.1371/journal.pone.0187131

Schippmann, U., Leaman, D. J., & Cunningham, A. (2002). Impact of cultivation and gathering of medicinal plants on biodiversity: global trends and issues. Biodiversity and the Ecosystem Approach in Agriculture, Forestry and Fisheries.

Schippmann, U., Leaman, D., & Cunningham, A. (2006). A comparison of cultivation and wild collection of medicinal and aromatic plants under sustainability aspects. Frontis, 47, 777–780.

Schmickl, R., Liston, A., Zeisek, V., Oberlander, K., Weitemier, K., Straub, S. C. K., & Suda, J. (2016). Phylogenetic marker development for target enrichment from transcriptome and genome skim data: the pipeline and its application in southern African Oxalis (Oxalidaceae). Molecular Ecology Resources, 16(5), 1124–1135. https://doi.org/10.1111/1755-0998.12487

Soltis, P. S., & Soltis, D. E. (2009). The role of hybridization in plant speciation. Annual Review of Plant Biology, 60(1), 561–588. https://doi.org/10.1146/annurev.arplant.043008.092039

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics, 22(21), 2688–2690. https://doi.org/10.1093/bioinformatics/btl446

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann, C., & Willerslev, E. (2007). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. Nucleic Acids Research, 35(3), e14. https://doi.org/10.1093/nar/gkl938

Talavera, G., & Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Systematic Biology, 56(4), 564–577. https://doi.org/10.1080/10635150701472164

Tiwari, N. N., Poudel, R. C., & Uprety, Y. (2004). Study on domestic market of medicinal and aromatic plants (MAPs) in Kathmandu Valley. Winrock International.

Ved, D., & Goraya, G. (2007). Demand and supply of medicinal plants in India. NMPB, New Delhi & FRLHT, Bangalore, India, 18.

Veldman, S., Gravendeel, B., Otieno, J. N., Lammers, Y., Duijm, E., Nieman, A., Bytebier, B., Ngugi, G., Martos, F., van Andel, T. R., & de Boer, H. J. (2017). High-throughput sequencing of African chikanda cake highlights conservation challenges in orchids. Biodiversity and Conservation, 26(9), 2029–2046. https://doi.org/10.1007/s10531-017-1343-7

WHO (2004). WHO guidelines on safety monitoring of herbal medicines in pharmacovigilance systems. WHO.

Willerslev, E., Davison, J., Moora, M., Zobel, M., Coissac, E., Edwards, M. E., Lorenzen, E. D., Vestergård, M., Gussarova, G., Haile, J., Craine, J., Gielly, L., Boessenkool, S., Epp, L. S., Pearman, P. B., Cheddadi, R., Murray, D., Bråthen, K. A., Yoccoz, N., ... Taberlet, P. (2014). Fifty thousand years of Arctic vegetation and megafaunal diet. Nature, 506(7486), 47–51. https://doi.org/10.1038/nature12921

Wood, T. E., Takebayashi, N., Barker, M. S., Mayrose, I., Greenspoon, P. B., & Rieseberg, L. H. (2009). The frequency of polyploid speciation in vascular plants. Proceedings of the National Academy of Sciences USA, 106(33), 13875–13879. https://doi.org/10.1073/pnas.0811575106

Wysoker, A., Tibbetts, K., & Fennell, T. (2015). Picard tools. Broad Institute.

Zeng, C.-X., Hollingsworth, P. M., Yang, J., He, Z.-S., Zhang, Z.-R., Li, D.-Z., & Yang, J.-B. (2018). Genome skimming herbarium specimens for DNA barcoding and phylogenomics. Plant Methods, 14(1), 43. https://doi.org/10.1186/s13007-018-0300-0

Zhang, C., Sayyari, E., & Mirarab, S. (2017). ASTRAL-III: Increased scalability and impacts of contracting low support branches (pp. 53–75). Springer.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.