TRACKING THE GLOBAL SUPPLY CHAIN OF HERBAL MEDICINES WITH NOVEL GENOMIC DNA BARCODES

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, Sars gate 1, 0562 Oslo, Norway

[2] Global Diversity Foundation, 37 St. Margarets Street, Canterbury, Kent CT1 2TU, England

[3] Royal Botanic Garden Edinburgh, Inverleith Row, Edinburgh EH3 5LR, Scotland

Corresponding author: Vincent Manzanilla

Email: vincent.manzanilla@gmail.com

SIGNIFICANCE

Unsustainable exploitation of natural resources is a major driver of biodiversity loss. Up to a third of the world's biodiversity is considered threatened by trade, but a lack of traceability methods for traded products impedes evaluation of international supply chains and the global impacts of trade on biodiversity. In this study, we pioneer the use of target capture-based genomic DNA barcoding. We compare this with standard DNA barcodes and complete plastid genome sequences for the identification of plants species in trade and for tracing their geographic origin. The target-capture barcoding approach described here presents a major advance for tracing the geographic origin of plant-based food and medicines and establishing the identity of illegally traded species. It enables better understanding and targeting of conservation action, and enhances capacity to assess the quality, safety and authenticity of traded products.

ABSTRACT

Uncontrolled and unsustainable trade in natural resources is an increasingly important threat to global biodiversity. In recent years, molecular identification methods have been proposed as tools to monitor global supply chains, to support regulation and legislative protection of species in trade, and enhancing consumer protection by establishing whether a traded product contains the species it is supposed to contain. However, development of an effective assay that routinely provides species-level identification and information on geographical origin of plants remains elusive, with standard plant DNA barcodes often providing only 'species group' or genus-level resolution. Here, we demonstrate the efficacy of target-capture genomic DNA barcoding, based on 443 nuclear markers, for establishing the identity and geographic origin of samples traded as the red-listed medicinal plant *Anacyclus pyrethrum* (L.) Lag. We also use this approach to provide insights into product adulteration and substitution in national and international supply chains. Compared with standard plant DNA barcodes and entire plastid genome sequences, the target capture approach outperforms other methods, and works with DNA from degraded samples. This approach offers the potential to meet the 'holy-grail' of plant DNA barcoding, namely routine species-level DNA-based identification, and also providing insights into geographic origin. This represents a major development for biodiversity conservation and for supporting the regulation and monitoring of trade in natural plant products.

Keywords: DNA barcoding, sustainability, hybrid-capture, genomic barcoding, endangered species, medicinal plants, supply chain, international trade

Human exploitation of natural resources is a major challenge for biodiversity conservation and sustainable development. Global trade and consumer demand for natural products provide increasing threats to species (1, 2), and at the same time lead to markets where regulation and authentication is extremely difficult (3–5). Most traded plant products are from wild populations and over-harvested: increasing scarcity results in higher prices, and incentivises adulteration, substitution and poaching (6–9). In recent decades, international conventions including the Convention on Biological Diversity (CBD) and the Convention on International Trade in Endangered Species (CITES) have sought to regulate the trade of threatened species. In parallel, the World Health Organisation has developed guidelines for safety monitoring of herbal medicines in pharmacovigilance systems (4) and the Food and Agriculture Organisation regulates international trade of cultivated plant products (10). However, implementing these regulations and guidelines is hampered by difficulties identifying plant products in trade. Multiple, complex and interacting supply chains can co-exist for a single plant product (11, 12). Traded plants are often not identifiable to species by their morphology or chemistry, as they may be dried, powdered, processed, or commercialised in mixtures with other products. In addition, the design, implementation and enforcement of successful conservation actions, and/or assessments of product quality and authenticity, often require the identification of the geographic origin of species in trade. This is difficult as the development of efficient methods to trace and identify traded products to their geographical area of origin are still in their infancy.

DNA barcoding is highly effective for species-level identification in animals using a portion of the mitochondrial marker Cytochome Oxidase 1 (COI). In plants, standard DNA barcoding involve using varying combinations of one to four plastid DNA regions (*rbcL, matK, trnH-psbA, trnL*) and/or the internal transcribed spacers of nuclear ribosomal DNA (nrDNA ITS). Although these markers are very informative in many cases, no single marker or combination of these markers routinely provides complete species-level resolution, especially in species-rich groups (13, 14), let alone provide population level assignment.

The development of high throughput sequencing (HTS) using new sequencing chemistries and platforms offer opportunities to extend the concept of DNA barcoding in plants (15–17). In addition to sequencing standard barcoding loci in a more cost-effective fashion, two major approaches have been proposed for increasing the resolution (and coverage) of plant DNA barcoding. Shallow pass shotgun sequencing (genome skimming) is now frequently used to recover organellar genomes and nuclear ribosomal DNA sequences, increasing the amount of data per sample, leading to some increases in resolution (18–22), and workflows and bioinformatic pipelines are becoming increasing refined for this approach (23–25). However, at present, cost constraints mean that most genome skimming barcoding projects only have sufficient sequencing depth to recover comparative data from multiple samples for multi-copy regions such as plastid genomes and ribosomal DNA; these regions represent a limited number of independent loci, ultimately constraining resolving power (26–28). Target capture sequencing offers the potential to overcome this, 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 (29). Like genome skimming, target capture is successful in sequencing degraded DNA samples (30, 31), and can be used to sequence hundreds of samples at the same time (32). It can also be designed to recover standard DNA barcodes in the same assay (33).

88 Although target capture has been advocated as a powerful tool for molecular identification of
89 plants, its utility in a barcoding context remains untested to-date (16, 22, 34).

90 We evaluated the power of target-capture DNA barcoding with an investigation of the
91 traceability of plant products, focusing on an internationally traded and vulnerable medicinal
92 plant species, *Anacyclus pyrethrum* (L.) Lag., widely used in traditional Arabic and Ayurvedic
93 medicine. This exemplar case includes a well-established international trade chain (35), and
94 classic challenge for plant molecular identification such as a recent radiation, frequent
95 hybridization (36, 37) and a large genome size (38, 39).
96 The genus *Anacyclus* (Asteraceae) comprises 12 species of annual and perennial weedy
97 herbs with partly overlapping geographic ranges around the Mediterranean basin (36, 40, 41).
98 Some species are abundant and have wide geographical ranges (for example, *A. clavatus*
99 (Desf.) Pers. and *A. radiatus* Loisel.), whereas others are rare and have restricted ranges (for
100 example, *A. maroccanus* (Ball) Ball and *A. pyrethrum* (L.) Lag.). Only *A. pyrethrum* is used in
101 traditional herbal medicine.
102 *Anacyclus pyrethrum* is endemic to Morocco, Algeria and southern Spain (36, 41), and
103 has a long history of use in Islamo-Arabic, European, and Indian Ayurvedic medicine (42–45).
104 In the 13th century, Ibn al-Baytār wrote that the plant was "*known across the world*" and traded
105 from the Maghreb to *all* other areas (46). Its popularity as a medicinal plant stems from the
106 many pharmacological activities of its roots (47). In Morocco, two varieties of *A. pyrethrum*
107 are distinguished, var. *pyrethrum* and var. *depressus*, the first being more potent and up to ten
108 times more expensive than the second (48). Today, both varieties are harvested from the wild
109 and used extensively for the treatment of pain and inflammatory disorders across Morocco (48–
110 52) and Algeria (53–55), as well as the Middle East (42) and the Indian sub-continent (56, 57).
111 *A. pyrethrum* is still traded today from the Maghreb to India (58, 59) and Nepal (56), and is
112 known to be over-harvested and is increasingly difficult to find in local markets in Morocco
113 (48, 50, 52, 60). Collectors are proficient in identifying the plant and its two varieties, but
114 material is possibly misidentified and adulterated along the chain of commercialisation (48, 50,
115 61, 62).

116 Here we apply target-capture genomic barcoding to distinguish *Anacylus* species and
117 geographical races, to identify traded *Anacylus* root samples in the national and international
118 supply chains, and compare this novel approach with plastid genome and nrDNA ITS barcodes.

119

120 **RESULTS**

121 We constructed a reference database of DNA sequences from fresh and herbarium specimens,
122 consisting of 83 individuals of 10 *Anacyclus* species, and 5 individuals representing outgroup
123 species. Four datasets were retrieved from our shotgun and target capture sequencing methods,
124 standard barcode markers (*matK*, *trnH-psbA*, the *trnL* intron and *rbcL*), nrDNA (ITS), along
125 with complete plastid genomes (from shotgun genome skimming) and hundreds of nuclear
126 markers (from target capture). We then used this reference database to assess the identity and
127 geographic origins of 110 root samples acquired from traded materials. We show that the target
128 capture approach is the most powerful method to identify plant species in trade and their
129 geographic origin.

130 **Data recovery for plastome, standard barcode loci, nrDNA and skimming and target**
131 **capture data**

132 The data recovery based on target capture outperforms shotgun sequencing (Figure 1). On
133 average, 2.3 million reads were obtained from shotgun sequencing per sample after quality
134 control filtering and 1.6 million reads for target capture. The shotgun data produced, ITS,
135 plastome and standard barcodes, with retrieval rates of 93% of samples for ITS, 56% for the
136 plastid genomes and 46% to 55% for the standard barcodes (Figure 1); 73% of the samples
137 were retrieved using target capture with 303X coverage for the targeted nuclear loci (Table
138 S3). The resulting aligned matrices for each of the datasets were 633 bp from ITS (including
139 5.8S), 4408 bp from the standard barcoding markers (1523 bp *matK,* 1438 bp *rbcL,* 500 bp
140 *trnH*-psbA, 947 bp *trnL*), and 289,236 bp from 443 nuclear loci recovered from the target
141 capture approach (for the standard barcodes, we included the full length of the coding regions
142 of *rbcL* and *matK*). The bioinformatics workflow for data analyses is described in Figure S1.

143 **Comparative levels of species discrimination using different approaches**

144 The data from shotgun sequencing (nrDNA, plastome phylogenies, standard barcodes)
145 highlight the complex evolutionary history of *Anacyclus*. The nrDNA phylogeny lacks
146 resolution in general (Figure S2-3). The outgroups *Tanacetum*, *Matricaria*, *Achillea*,
147 *Othanthus* and *Tripleurospermum* have well-supported bootstrap values, but within the genus
148 *Anacyclus,* only *A. atlanticus* Litard. & Maire, *A. maroccanus* and *A. radiatus* are highly
149 supported. The plastome phylogeny shows very good support at genus level for the *Anacyclus*
150 node, and at species level for the outgroups. The lack of variation in the plastid genome within
151 the genus *Anacyclus* results in little phylogenetic support with no species-specific clusters
152 recovered (Figure S4-5). The standard barcode markers, *matK*, *rbcL, trnH-psbA* and *trnL*
153 (Figure S6-9) displayed low levels of resolution at the species level.

154 The 443 loci recovered by target capture, led to a well-resolved phylogeny and high
155 levels of species discrimination: all the genera in the Matricariinae tribe and all interspecific
156 relationships are well-supported, with most nodes showing posterior probabilities (PP) of 1
157 (Figure S10-11). Within *Anacyclus,* all species, sub-species and varieties are well supported.
158 PP are lower for *A. monanthos* (PP = 0.75). The complex of hybrid species composed of *A.*
159 *clavatus*, *A. homogamos*, and *A. valentinus* is polyphyletic and shows signs of hybridization
160 and incomplete lineage sorting. Intraspecific nodes have PP varying between 0.27 and 1,
161 mostly depending on species population structure.

162
163
164 **Assessment of *Anacyclus* trade**
165
166 Interviews with 39 harvesters, middlemen, retailers, and wholesalers in various
167 Moroccan cities, indicate that the national and international trade of *Anacyclus pyrethrum*

168 follow two separate supply chains. Retailer herbalists in Moroccan cities are supplied by
169 middlemen who acquire the plant from local harvesters from rural communities. These retailers
170 typical hold between a few hundred grams to one kilogram of the plant material in their shops.
171 In contrast, wholesalers who export the plant internationally, hire professional harvesters who
172 travel across the species geographical range to collect plant material. Harvested roots are
173 brought directly from the wild to the export companies in Rabat, Casablanca and Tangier, from
174 where they enter the international market, including supply of samples to India. According to
175 informants from export companies, between 3-10 tons of the plant product can be stocked at a
176 time.
177 Our examination of material in trade involved screening a total of 62 bags each
178 containing an average of 25g of dry roots. Initial morphological examination of these samples
179 identified obvious non-*Anacyclus* adulterants in 39/66 batches. The adulterants were present
180 with a proportion from 3% to 100% with an average of 42%. The non-*Anacyclus* adulterants
181 were found at high frequency in collections from traditional healers and herbalists, less so from
182 collectors, wholesalers and export companies (Figures 2-3).
183 We selected at random 110 individual roots for DNA analysis from the 62 root batches.
184 Of these 98 had a morphology consistent with *Anacyclus,* and 12 which were classed as similar
185 to *Anacyclus* but likely to be non-*Anacyclus* based on their morphology. We recovered shotgun
186 sequence data from nine of these 12 non-*Anacylus* roots and using sequences queries against
187 GenBank, the ITS and plastid sequences obtained a genus-level identification (= *Plantago sp.*).
188 Of the 98 *Anacyclus* roots, 18 had no recoverable DNA sequences via any of our
189 methods, with 80/80 of the remaining samples working for target capture vs 61/80 for shotgun
190 sequencing). Extracting plastome sequences from the shotgun sequencing enabled
191 identification of 7 roots to the species level, with the remainder identified as *Anacyclus* sp.
192 (Figure 1). Neither ITS nor any of the standard barcodes were able to discriminate any of these
193 samples below the genus level (Figure S6-9).

194 The target capture data gave much higher resolution within *Anacyclus*. In our
195 investigation of the national market in Morocco, our analyses of six individual root samples
196 from four rural community collectors identified three *Anacyclus* var. *pyrethrum* and three var.
197 *depressus*. Our analysis of five samples from three wholesaler 'middle-men' in Morocco
198 identified two *Anacyclus* var. *pyrethrum* and three var. *depressus*. Our sequences from 18
199 samples from 8 herbalists revealed 11 *Anacyclus* var. *pyrethrum* and seven var. *depressus*.
200 Likewise, our seven samples from four traditional healer sources identified four *Anacyclus* var.
201 *pyrethrum*, and three var. *depressus*. In the international market, our 16 sequenced samples
202 from three export companies in Morocco identified eight *Anacyclus* var. *pyrethrum*, six var.
203 *depressus*, and two *A. homogamous*. Our analysis of 28 samples from 17 herbalists in India
204 identified three *Anacyclus* var. *pyrethrum* and 25 var. *depressus*.

205 **Geographical source**

206 Market samples in Morocco originate both from various Moroccan areas as well as Algeria,
207 and material from all of these populations of origin can be found in Indian market samples

(Table S3, Figure S11). Of the 80 non-adulterated roots which we identified to species level using target capture, we were able to associate 58% to a specific geographic region (Figure 2, Table S2). Using phylogenetic analysis, root samples of *A. pyrethrum* var. *pyrethrum* clustered with reference material from the High Atlas, 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. Evidence for the international trade from Algeria to Morocco is highlighted by a distinctive clade that includes traded roots collected from west Algeria (Figure S12). The geographic origin was only resolvable with target capture data; standard barcoding markers, nrDNA and plastome data lacked variation, resolution or both.

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 the 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 nrDNA), which often show trans-specific polymorphism and barcode sharing among related species (63). Even using complete plastid genome sequences suffers from the same problem, as the data are all physically linked in a single non-recombinant uni-parentally inherited locus (22). In this study, using entire plastid genomes or plastid barcode markers, and/or ITS we recovered very limited resolution below the genus level (Figure 1, Figure S4). In contrast, our target capture approach using hundreds of nuclear loci yields significantly higher molecular identification success and more accurate resolution to species and even population level (Figure 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 the scarcity of *A. pyrethrum* var. *pyrethrum* (Figures 2-3). 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 product. All of the non-*Anacyclus* roots we sequenced were identified as *Plantago* spp. by shot-gun sequencing, despite being sampled from six different localities including Morocco and India. As *Plantago* roots are similar in appearance to *Anacyclus,* it is possible that they are deliberately added as a 'difficult to identify' adulterant which may go unnoticed by non-specialists.

Our analysis of samples from collectors, wholesalers and export companies detected much less adulteration at this point in the supply chain. 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 (64). Local harvesters have increasing difficulty to supply local trade chains (48, 64), which may finally result in increased adulteration rates in the poorly-governed, national value chains (Figure 2), as has

247 also been observed elsewhere (11, 12). Our results also identify previously unreported
248 international trade in North Africa prior to export to the Indian sub-continent (Figure 2,3, S11).
249 We provide evidence that export companies in Morocco source material not only in this
250 country, but also from neighbouring Algeria. Applying this molecular identification approach
251 enables us to distinguish samples at population level and uncover these hidden international
252 sourcing channels.

253 High national and international demand for *A. pyrethrum* likely encourages its
254 overharvesting and adulteration. As the plant is a remedy of the Indian pharmacopoeia, its
255 demand is likely to increase along with that of other Ayurvedic medicines (65). Although *A.*
256 *pyrethrum* has been assessed to be vulnerable internationally (35) and endangered in Morocco
257 on the IUCN Red List, the plant is not listed in the CITES appendices and its international trade
258 is not regulated. Nonetheless, continued overharvesting is driving wild populations to critical
259 levels and conservation policies are necessary. Common strategies to conserve overharvested
260 medicinal plants often include collection and trade restrictions as well as cultivation (6).
261 Cultivation is often proposed as a solution to both conservation issues and sourcing high
262 quality, appropriately identified material (6, 7, 66). However, trade of cultivated plants has not
263 been as successful as expected due to beliefs that plants harvested from the wild are more
264 potent and thus preferable (67). This kind of demand requirements necessitate trade monitoring
265 and restrictions. Only with fine-grained mapping of sourcing areas and supply chains, as our
266 results highlight for *A. pyrethrum*, trade monitoring and conservation can be achieved.

267

268 *Future prospects for plant DNA barcoding*

269 Key criteria for developing new DNA barcoding approaches include resolving power (telling
270 species apart), recoverability (enabling use on a wide diversity of tissue sources), and cost and
271 efficiency (enabling scaling over very large sample sets).

272 In terms of resolving power, the target capture approach used here provides a step-
273 change over plant barcodes based on plastid sequences and ITS. The key enabling step is access
274 to multiple nuclear markers, removing the sensitivity of the identification to introgression of
275 one or two loci as is the case for barcodes from rDNA or the plastid genome.

276 In terms of recoverability, the successful recovery of sequences via target capture
277 depends on how closely related the sampled species are to the reference set on which the baits
278 were designed, and/or the level of variation in the loci that form the bait set (68, 69). In the
279 current study, the baits were designed from *Anacyclus* and related genera. This optimizes their
280 specificity to our study group and enabled their successful high-resolution application for
281 assessing trade. A clear challenge for wider use in DNA barcoding, is applicability over much
282 greater phylogenetic distances. The recently published universal angiosperm baits (70, 71),
283 designed to recover 353 loci from a wide diversity of flowering plants offer great potential
284 here. An immediate priority is to assess the degree to which this universal bait set (71) is
285 variable enough to routinely provide species-level resolution. In addition, there is a need for a
286 more general evaluation of when / how and at what scale to most effectively to combine taxon-

287 specific bait sets (as used here) with universal bait sets, to simultaneously obtain very high
288 resolution and sequencing success over wide phylogenetic distances.

289 Another important aspect to recoverability is efficacy with degraded DNA. Drying,
290 storage, and transportation affect the quality of plant material in trade and can cause extensive
291 DNA degradation (12, 72, 73). In consequence, traded samples have similar challenges to
292 working with ancient DNA, herbarium samples or archaeological remains. Target capture is
293 particularly well suited to this challenge (68, 69, 74). Although both shotgun sequencing can
294 also be very effective on degraded material (30, 75, 76), in the current study, our recovery rate
295 was greater for target capture than shotgun sequencing (Figure 1), and we recovered data from
296 100% of the samples used for establishing the reference library and over 70% of the samples
297 in trade. This 70% success rate for hundreds of nuclear loci providing high resolution from
298 suboptimal tissue of traded samples is noteworthy. The other mainstream approach for highly
299 degraded DNA is a portion of the chloroplast *trnL* (UAA) intron, specifically the P6 loop (10–
300 143 bp) (77). This has been highly successful in recovering sequence data from degraded
301 samples (78–84). However, this short region of the plastid genome has a low variation at the
302 species level and does not typically discriminate among con-generic species (77).

303 In terms of costs, although sequencing costs continue to fall, the current protocols for
304 target capture (including library construction and sequencing) are still expensive, and equate
305 to 70 USD per sample. This is manageable for well-resourced projects and high value
306 applications, but still prohibitively expensive for many large-scale biomonitoring projects or
307 less well-resourced projects. Ongoing work is required to optimize protocols to drive these
308 costs down, as has been done for standard barcoding approaches (85).

309

310 **Conclusion**

311 In plants, the frequent sharing of plastid and ribosomal sequences among con-generic
312 species, coupled with the difficulty of routinely accessing multiple nuclear loci, has acted as a
313 constraint on the resolution of DNA barcoding approaches. Current advances in sequencing
314 technology and bioinformatics are removing this constraint, and offer the potential for a new
315 wave of high-resolution identification tools for plants. These approaches, such as target capture
316 have the capability to distinguish species and populations, providing insights into diversity and
317 ecology, as well as the multitude of societal applications which require information on the
318 identification and provenance of biological materials.

319

320 **Materials and Methods**

321 **Sample collection.** Forty-two wild populations (n=72 individuals) of *Anacyclus* were sampled
322 in Morocco and Spain to build a reference database. The sampling was complemented with
323 eleven herbarium voucher specimens of species occurring elsewhere in the Mediterranean. The
324 outgroups included two species of *Matricaria*, two species of *Achillea*, and one species of
325 *Otanthus*. The specimen origins and vouchers number are listed in Table S4 and Figure S12.
326 One hundred and ten trade vouchers consisting of 50g of roots were bought from collectors,

327 herbal shops, middle-men, traditional healer, wholesalers, and export companies in Morocco
328 and India (Table S3).

329 **Trade information.** During the acquisition of samples from collectors and traders in Morocco
330 and India, semi-structured interviews were conducted to enquire about the trade of *Anacyclus*.
331 In total, 39 informants were interviewed and asked where the plant material was sourced, to
332 whom it was sold and in what quantities, for what price, and if there were several qualities of
333 this medicinal plant. The quantity of this product was weighted in herbalist shops and estimated
334 in export companies, estimates were given by informants. Interviews were carried out
335 anonymously and followed the International Society of Ethnobiology Code of Ethics (86).

336 **Extraction and Library preparation.** DNA from reference and traded vouchers was extracted
337 from approximately 40 mg of dry leaf or root material using the DNeasy Plant Mini Kit
338 (Qiagen). Total DNA (0.2-1.0 µg) was sheared to 500 bp fragments using a Covaris S220
339 sonicator (Woburn, MA, USA). Dual indexed libraries were prepared using the Meyer and
340 Kircher protocol (87) for shotgun sequencing and target capture.

341 **Target capture.** We designed 872 low-copy nuclear markers and associated RNA probes by
342 following the Hyb-Seq pipeline, with minor adjustments (SI). For target capture enrichment,
343 twelve equimolar pools were prepared with ten to 24 samples and an average 300 ng of input
344 DNA per pool. The RNA probes were hybridized for 16 hours before target baiting, and 14
345 PCR cycles were carried out after enrichment following the MyBaits v3 manual. The enriched
346 libraries and shotgun libraries were sequenced on two Illumina HiSeq 3000 lanes (150bp
347 paired-end).

348 **Data Processing.** The sequencing runs were trimmed and quality filtered using Trimmomatic
349 (88). Low-copy nuclear markers and their alleles were retrieved for each sample. First, the
350 reads were mapped against the selected low-copy nuclear loci (SI) using BWA v0.7.5a-r405
351 (89). Duplicate reads were removed using Picard v2.10.4 (90). Alleles were phased for each
352 marker and individual using SAMtools v1.3.1 (91). The last step of the pipeline combined the
353 retrieved alleles into single gene matrices. Plastome and nrDNA sequences were recovered by
354 pooling shotgun and target enrichment sequencing data. Plastid genomes were build using
355 MITOBim v1.8 (23). nrDNA sequences were recovered using BWA by mapping the reads to
356 the reference nrDNA of *Anacyclus pyrethrum* (KY397478) for *Anacyclus* species and traded
357 samples, to the reference nrDNA of *Achillea pyrenaica* Sibth. ex Godr. (AY603247) for
358 *Otanthus* and *Achillea*, and to the reference nrDNA of *Matricaria aurea* (Loefl.) Sch.Bip.
359 (KT954177) for *Matricaria* samples.

360 **Phylogenomics.** The recovered matrices (nuclear genes, nrDNA and plastomes) were aligned
361 with MAFFT (92), refined with MUSCLE (93) and filtered with Gblocks (94). Phylogenies
362 were inferred using RAxML v8.0.26 (95), with 1000 bootstrap replicates under the
363 GTRGAMMA model. For the low-copy nuclear genes, the species tree was inferred from the
364 individual nuclear gene trees under the multi-species coalescence (MSC) framework with
365 ASTRAL-III v5.5.9 (96). The multi-alleles option in ASTRAL-III was used for reconciliation
366 of the independent evolutionary histories of the alleles. The molecular identification of traded
367 roots was assessed from the MSC tree and posterior probabilities (PP) greater than 0.95.

368 Results from morphological identification of the traded roots (described in (SI)) were combined
369 with results from molecular identification, and samples were identified according to their
370 position in the supply chain (Figure 2) and geographical origin (Figure 3).
371

DECLARATIONS

Authors' contributions

The project was coordinated by AK, GM, HdB and VM. VM did the design of the study and performed data analysis. AK, HdB, ITT, PH and VM wrote the manuscript. All authors provided useful contributions to data analysis and interpretation of the results. All authors have read and approved the final version of the manuscript.

Acknowledgments

Maxime Borry for the DNA extraction; Julie Hawkins for hosting V.M. in Reading during the phylogenomics analyses; the Global Diversity Foundation team in Morocco for their logistic support during the field work. The PET group for their useful discussions and recommendations during the project. This work was performed on the Abel Cluster, owned by the University of Oslo and the Norwegian metacenter for High Performance Computing (NOTUR), and operated by the Department for Research Computing at USIT, the University of Oslo IT-department, http://www.hpc.uio.no/.

Reproducibility

For reproducibility purposes, all the script used during the data processing are available on the OSF work repository https://osf.io/9bh3p/ and github.

Funding

This project was supported by the European Union's Seventh Framework Programme for research, technological development and demonstration under the Grant agreement no. 606895 to the FP7-MCA-ITN MedPlant, "Phylogenetic Exploration of Medicinal Plant Diversity".

References

1. Lenzen M, et al. (2012) International trade drives biodiversity threats in developing nations. *Nature* 486(7401):109.

2. Díaz S, et al. (2019) Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. *U N Paris Fr*:1–39.

3. Barnes J (2003) Pharmacovigilance of herbal medicines. *Drug Saf* 26(12):829–851.

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

5. Newmaster S, Grguric M, Shanmughanandhan D, Ramalingam S, Ragupathy S (2013) DNA barcoding detects contamination and substitution in North American herbal products. *BMC Med* 11(1):222.

6. Schippmann U, Leaman DJ, Cunningham A (2002) Impact of cultivation and gathering of medicinal plants on biodiversity: global trends and issues. *Biodivers Ecosyst Approach Agric For Fish*.

7. Hamilton A (2004) Medicinal plants, conservation and livelihoods. *Biodivers Conserv* 13(8):1477–1517.

8. Williams V, Victor J, Crouch N (2013) Red listed medicinal plants of South Africa: status, trends, and assessment challenges. *South Afr J Bot* 86:23–35.

9. Veldman S, et al. (2017) High-throughput sequencing of African chikanda cake highlights conservation challenges in orchids. *Biodivers Conserv* 26(9):2029–2046.

10. Vandecandelaere E, et al. (2018) *Strengthening sustainable food systems through geographical indications An analysis of economic impacts* (FAO, Roma).

11. Booker A, Johnston D, Heinrich M (2012) Value chains of herbal medicines-Research needs and key challenges in the context of ethnopharmacology. *J Ethnopharmacol* 140(3):624–633.

426 12. Booker A, Johnston D, Heinrich M (2016) The welfare effects of trade in
427 phytomedicines: a multi-disciplinary analysis of turmeric production.
428 *World Dev* 77:221–230.

429 13. Ross HA, Murugan S, Sibon Li WL (2008) Testing the reliability of genetic
430 methods of species identification via simulation. *Syst Biol* 57(2):216–230.

431 14. Hollingsworth PM (2011) Refining the DNA barcode for land plants. *Proc*
432 *Natl Acad Sci* 108(49):19451–19452.

433 15. Lemmon EM, Lemmon AR (2013) High-throughput genomic data in
434 systematics and phylogenetics. *Annu Rev Ecol Evol Syst* 44:99–121.

435 16. Coissac E, Hollingsworth PM, Lavergne S, Taberlet P (2016) From
436 barcodes to genomes: extending the concept of DNA barcoding. *Mol Ecol*
437 25(7):1423–1428.

438 17. Hollingsworth PM, Li D-Z, van der Bank M, Twyford AD (2016) Telling
439 plant species apart with DNA: from barcodes to genomes. *Phil Trans R Soc*
440 *B* 371(1702):20150338.

441 18. Parks M, Cronn R, Liston A (2009) Increasing phylogenetic resolution at
442 low taxonomic levels using massively parallel sequencing of chloroplast
443 genomes. *BMC Biol* 7(1):84.

444 19. Nock CJ, et al. (2011) Chloroplast genome sequences from total DNA for
445 plant identification. *Plant Biotechnol J* 9(3):328–333.

446 20. Kane N, et al. (2012) Ultra-barcoding in cacao (*Theobroma* spp.;
447 Malvaceae) using whole chloroplast genomes and nuclear ribosomal DNA.
448 *Am J Bot* 99(2):320–329.

449 21. Manzanilla V, et al. (2018) Phylogenomics and barcoding of Panax: toward
450 the identification of ginseng species. *BMC Evol Biol* 18(1):44.

451 22. Ruhsam M, et al. (2015) Does complete plastid genome sequencing
452 improve species discrimination and phylogenetic resolution in *Araucaria*?
453 *Mol Ecol Resour* 15(5):1067–1078.

454 23. Hahn C, Bachmann L, Chevreux B (2013) Reconstructing mitochondrial
455 genomes directly from genomic next-generation sequencing reads—a

456 baiting and iterative mapping approach. *Nucleic Acids Res* 41(13):e129–
457 e129.

458 24. Dierckxsens N, Mardulyn P, Smits G (2016) NOVOPlasty: de novo
459 assembly of organelle genomes from whole genome data. *Nucleic Acids Res*
460 45(4):e18–e18.

461 25. Ankenbrand M, et al. (2018) chloroExtractor: extraction and assembly of
462 the chloroplast genome from whole genome shotgun data. *J Open Source*
463 *Softw* 3:464.

464 26. Wood TE, et al. (2009) The frequency of polyploid speciation in vascular
465 plants. *Proc Natl Acad Sci U A* 106. doi:10.1073/pnas.0811575106.

466 27. Soltis PS, Soltis DE (2009) The Role of Hybridization in Plant Speciation.
467 *Annu Rev Plant Biol* 60(1):561–588.

468 28. Soltis PS, Marchant DB, Van de Peer Y, Soltis DE (2015) Polyploidy and
469 genome evolution in plants. *Curr Opin Genet Dev* 35(Supplement C):119–
470 125.

471 29. Degnan JH, Rosenberg NA (2009) Gene tree discordance, phylogenetic
472 inference and the multispecies coalescent. *Trends Ecol Evol* 24(6):332–340.

473 30. Staats M, et al. (2013) Genomic treasure troves: complete genome
474 sequencing of herbarium and insect museum specimens. *PLoS One*
475 8(7):e69189.

476 31. Särkinen T, Staats M, Richardson JE, Cowan RS, Bakker FT (2012) How
477 to open the treasure chest? Optimising DNA extraction from herbarium
478 specimens. *PloS One* 7(8):e43808.

479 32. Mamanova L, et al. (2010) Target-enrichment strategies for next-generation
480 sequencing. *Nat Methods* 7(2):111.

481 33. Schmickl R, et al. (2016) Phylogenetic marker development for target
482 enrichment from transcriptome and genome skim data: the pipeline and its
483 application in southern African *Oxalis* (Oxalidaceae). *Mol Ecol Resour*
484 16(5):1124–1135.

485 34. Pillon Y, et al. (2013) Potential use of low-copy nuclear genes in DNA
486 barcoding: a comparison with plastid genes in two Hawaiian plant
487 radiations. *BMC Evol Biol* 13(1):35.

488 35. Rankou H, Ouhammou A, Taleb M, Manzanilla V, Martin G (2015)
489 *Anacyclus pyrethrum*. *IUCN Red List Threat Species 2015*
490 *ET202924A53798702* 2015(2015). Available at:
491 http://dx.doi.org/10.2305/IUCN.UK.2015-
492 4.RLTS.T202924A53798702.en.

493 36. Humphries CJ (1979) A revision of the genus *Anacyclus* L. (Compositae:
494 Anthemideae). *Bull Br Mus Nat Hist Hist Ser* (7):83–142.

495 37. Agudo García A, Tomasello S, Alvarez I, Oberprieler C (2016) *A*
496 *phylogenetic hypothesis for the genus* Anacyclus *L. (Anthemideae,*
497 *Compositae)* doi:10.13140/RG.2.2.26382.05442/1.

498 38. Humphries CJ (1981) Cytogenetic and cladistic studies in *Anacyclus*
499 (Compositae: Anthemideae). *Nord J Bot* 1(1):83–96.

500 39. Garcia S, et al. (2013) New data on genome size in 128 Asteraceae species
501 and subspecies, with first assessments for 40 genera, 3 tribes and 2
502 subfamilies. *Plant Biosyst - Int J Deal Asp Plant Biol* 147(4):1219–1227.

503 40. Oberprieler C, et al. (2009) Anthemideae. *Systematics, Evolution, and*
504 *Biogeography of Compositae.*, eds Funk VA, Susanna A, Stuessy TF, Bayer
505 RJ (International Association for Plant Taxonomy, Vienna., Vienna), pp
506 631–666.

507 41. Rosato M, Álvarez I, Feliner GN, Rosselló JA (2017) High and uneven
508 levels of 45S rDNA site-number variation across wild populations of a
509 diploid plant genus (*Anacyclus*, Asteraceae). *PloS One* 12(10):e0187131.

510 42. Pittle KD (2005) Continuity and Change in Islamic Ethnopharmacological
511 Practice: New Methods for Cognitive Dialectometry. Dissertation (Florida
512 State University).

513 43. De Vos P (2010) European materia medica in historical texts: longevity of
514 a tradition and implications for future use. *J Ethnopharmacol* 132(1):28–
515 47.

516 44. Adams M, Alther W, Kessler M, Kluge M, Hamburger M (2011) Malaria
517 in the Renaissance: remedies from European herbals from the 16th and 17th
518 century. *J Ethnopharmacol* 133(2):278–288.

519 45. Staub P, Casu L, Leonti M (2016) Back to the roots: A quantitative survey
520 of herbal drugs in Dioscorides' De Materia Medica (ex Matthioli, 1568).
521 *Phytomedicine* 23(10):1043–1052.

522 46. Leclerc L (1877) *Traité des simples* (Imprimerie nationale).

523 47. Manouze H, et al. (2017) Anti-inflammatory, antinociceptive, and
524 antioxidant activities of methanol and aqueous extracts of Anacyclus
525 pyrethrum roots. *Front Pharmacol* 8:598.

526 48. Ouarghidi A, Powell B, Martin GJ, De Boer H, Abbad A (2012) Species
527 substitution in medicinal roots and possible implications for toxicity of
528 herbal remedies in Morocco. *Econ Bot* 66(4):370–382.

529 49. Merzouki A, Ed-derfoufi F, Molero Mesa J (2000) Contribution to the
530 knowledge of Rifian traditional medicine. II: Folk medicine in Ksar Lakbir
531 district (NW Morocco). *Fitoterapia* 71(3):278–307.

532 50. Ouarghidi A, Martin G, Powell B, Esser G, Abbad A (2013) Botanical
533 identification of medicinal roots collected and traded in Morocco and
534 comparison to the existing literature. *J Ethnobiol Ethnomedicine* 9(1):59.

535 51. Jamila F, Mostafa E (2014) Ethnobotanical survey of medicinal plants used
536 by people in Oriental Morocco to manage various ailments. *J*
537 *Ethnopharmacol* 154(1):76–87.

538 52. Rhafouri R, et al. (2014) Ethnobotanical study of medicinal plants in Ifran's
539 National Park (Morocco).

540 53. Benarba B (2016) Medicinal plants used by traditional healers from South-
541 West Algeria: An ethnobotanical study. *J Intercult Ethnopharmacol*
542 5(4):320.

543 54. Benarba B, et al. (2015) Ethnobotanical study of medicinal plants used by
544 traditional healers in Mascara (North West of Algeria). *J Ethnopharmacol*
545 175:626–637.

546 55. Ouelbani R, Bensari S, Mouas TN, Khelifi D (2016) Ethnobotanical
547 investigations on plants used in folk medicine in the regions of Constantine
548 and Mila (North-East of Algeria). *J Ethnopharmacol* 194:196–218.

549 56. Tiwari NN, Poudel RC, Uprety Y (2004) Study on domestic market of
550 medicinal and aromatic plants (MAPs) in Kathmandu Valley. *Kathmandu*
551 *Winrock Int*.

552 57. Ministry of Health and Family Welfare (2007) *The ayurvedic*
553 *pharmacopoeia of India* (Govt. of India, Ministry of Health and Family
554 Welfare, Dept. of ISM & H.) Available at:
555 https://books.google.no/books?id=55xFAAAAYAAJ.

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

558 59. Ghosh S, General FDD (2013) Promotion of Medicinal and Aromatic Plants
559 in the Asia-Pacific Region. *Expert Consult Promot Med Aromat Plants*
560 *Asia-Pac Reg Proc*:33.

561 60. Taleb MS (2017) Aromatic and Medicinal Plants in Morocco: Diversity and
562 Socio-Economic Role. *World Acad Sci Eng Technol Int J Biol Biomol Agric*
563 *Food Biotechnol Eng* 11(12):742–746.

564 61. Kool A (2012) *Desert Plants and Deserted Islands [Elektronisk resurs] :*
565 *Systematics and Ethnobotany in Caryophyllaceae* (Acta Universitatis
566 Upsaliensis, Uppsala) Available at: http://uu.diva-
567 portal.org/smash/get/diva2:546726/FULLTEXT01.pdf
568 http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-179853.

569 62. de Boer HJ, Ouarghidi A, Martin G, Abbad A, Kool A (2014) DNA
570 barcoding reveals limited accuracy of identifications based on folk
571 taxonomy. *PLoS ONE* 9(1):e84291.

572 63. Hollingsworth PM, Graham SW, Little DP (2011) Choosing and Using a
573 Plant DNA Barcode. *PLoS ONE* 6(5):e19254.

574 64. Ouarghidi A, Powell B, Martin GJ, Abbad A (2017) Traditional Sustainable
575 Harvesting Knowledge and Distribution of a Vulnerable Wild Medicinal

576 Root (*A. pyrethrum* var. *pyrethrum*) in Ait M'hamed Valley, Morocco. *Econ*
577 *Bot* 71(1):83–95.

578 65. Kala CP, Dhyani PP, Sajwan BS (2006) Developing the medicinal plants
579 sector in northern India: challenges and opportunities. *J Ethnobiol*
580 *Ethnomedicine* 2(1):32.

581 66. Schippmann U, Leaman D, Cunningham A (2006) A comparison of
582 cultivation and wild collection of medicinal and aromatic plants under
583 sustainability aspects. *Frontis*:75–95.

584 67. Alam G, Belt J (2009) Developing a medicinal plant value chain: Lessons
585 from an initiative to cultivate Kutki (*Picrorhiza kurrooa*) in Northern India.
586 *KIT Work Pap Ser* (WPS. C5).

587 68. McCormack JE, Tsai WL, Faircloth BC (2016) Sequence capture of
588 ultraconserved elements from bird museum specimens. *Mol Ecol Resour*
589 16(5):1189–1203.

590 69. Paijmans JL, Fickel J, Courtiol A, Hofreiter M, Förster DW (2016) Impact
591 of enrichment conditions on cross-species capture of fresh and degraded
592 DNA. *Mol Ecol Resour* 16(1):42–55.

593 70. Buddenhagen C, et al. (2016) Anchored phylogenomics of angiosperms I:
594 Assessing the robustness of phylogenetic estimates. *BioRxiv*:086298.

595 71. Johnson MG, et al. (2018) A universal probe set for targeted sequencing of
596 353 nuclear genes from any flowering plant designed using k-medoids
597 clustering. *Syst Biol* 68(4):594–606.

598 72. Bombardelli E (1991) Technologies for the processing of medicinal plants.
599 *Med Plant Ind*:85–98.

600 73. Booker A, et al. (2014) Chemical variability along the value chains of
601 turmeric (*Curcuma longa*): a comparison of nuclear magnetic resonance
602 spectroscopy and high performance thin layer chromatography. *J*
603 *Ethnopharmacol* 152(2):292–301.

604 74. Hart ML, Forrest LL, Nicholls JA, Kidner CA (2016) Retrieval of hundreds
605 of nuclear loci from herbarium specimens. *Taxon* 65(5):1081–1092.

606 75. Zeng C-X, et al. (2018) Genome skimming herbarium specimens for DNA
607 barcoding and phylogenomics. *Plant Methods* 14(1):43.

608 76. Bakker FT (2017) Herbarium genomics: skimming and plastomics from
609 archival specimens. *Webbia* 72(1):35–45.

610 77. Taberlet P, et al. (2007) Power and limitations of the chloroplast trnL
611 (UAA) intron for plant DNA barcoding. *Nucleic Acids Res* 35(3):e14.

612 78. Valentini A, Pompanon F, Taberlet P (2009) DNA barcoding for ecologists.
613 *Trends Ecol Evol* 24(2):110–117.

614 79. Jørgensen T, et al. (2012) A comparative study of ancient sedimentary
615 DNA, pollen and macrofossils from permafrost sediments of northern
616 Siberia reveals long-term vegetational stability. *Mol Ecol* 21(8):1989–2003.

617 80. Parducci L, et al. (2012) Glacial Survival of Boreal Trees in Northern
618 Scandinavia. *Science* 335(6072):1083–1086.

619 81. Hibert F, et al. (2013) Unveiling the diet of elusive rainforest herbivores in
620 next generation sequencing era? The tapir as a case study. *PLoS One*
621 8(4):e60799.

622 82. Soininen EM, et al. (2013) Shedding new light on the diet of Norwegian
623 lemmings: DNA metabarcoding of stomach content. *Polar Biol*
624 36(7):1069–1076.

625 83. Boessenkool S, et al. (2014) Use of Ancient Sedimentary DNA as a Novel
626 Conservation Tool for High-Altitude Tropical Biodiversity. *Conserv Biol*
627 28(2):446–455.

628 84. Willerslev E, et al. (2014) Fifty thousand years of Arctic vegetation and
629 megafaunal diet. *Nature* 506(7486):47–51.

630 85. Hebert PD, et al. (2018) A Sequel to Sanger: amplicon sequencing that
631 scales. *BMC Genomics* 19(1):219.

632 86. ISE Code of Ethics Online (2018) *Int Soc Ethnobiol*. Available at:
633 http://www.ethnobiology.net/what-we-do/core-programs/ise-ethics-
634 program/code-of-ethics/code-in-english/ [Accessed March 3, 2018].

635 87. Meyer M, Kircher M (2010) Illumina Sequencing Library Preparation for
636 Highly Multiplexed Target Capture and Sequencing. *Cold Spring Harb*
637 *Protoc* 2010(6):pdb.prot5448.

638 88. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer
639 for Illumina sequence data. *Bioinformatics* 30(15):2114–2120.

640 89. Li H, Durbin R (2009) Fast and accurate short read alignment with
641 Burrows–Wheeler transform. *Bioinformatics* 25(14):1754–1760.

642 90. Wysoker A, Tibbetts K, Fennell T (2015) *Picard tools* (Broad Institute).

643 91. Li H, et al. (2009) Genome project data processing S. The sequence
644 alignment/Map format and SAMtools. *Bioinformatics* 25.
645 doi:10.1093/bioinformatics/btp352.

646 92. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment
647 software version 7: improvements in performance and usability. *Mol Biol*
648 *Evol* 30(4):772–780.

649 93. Edgar RC (2004) MUSCLE: multiple sequence alignment with high
650 accuracy and high throughput. *Nucleic Acids Res* 32(5):1792–1797.

651 94. Talavera G, Castresana J (2007) Improvement of phylogenies after
652 removing divergent and ambiguously aligned blocks from protein sequence
653 alignments. *Syst Biol* 56(4):564–577.

654 95. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based
655 phylogenetic analyses with thousands of taxa and mixed models.
656 *Bioinformatics* 22(21):2688–2690.

657 96. Zhang C, Sayyari E, Mirarab S (2017) ASTRAL-III: Increased Scalability
658 and Impacts of Contracting Low Support Branches (Springer), pp 53–75.

659

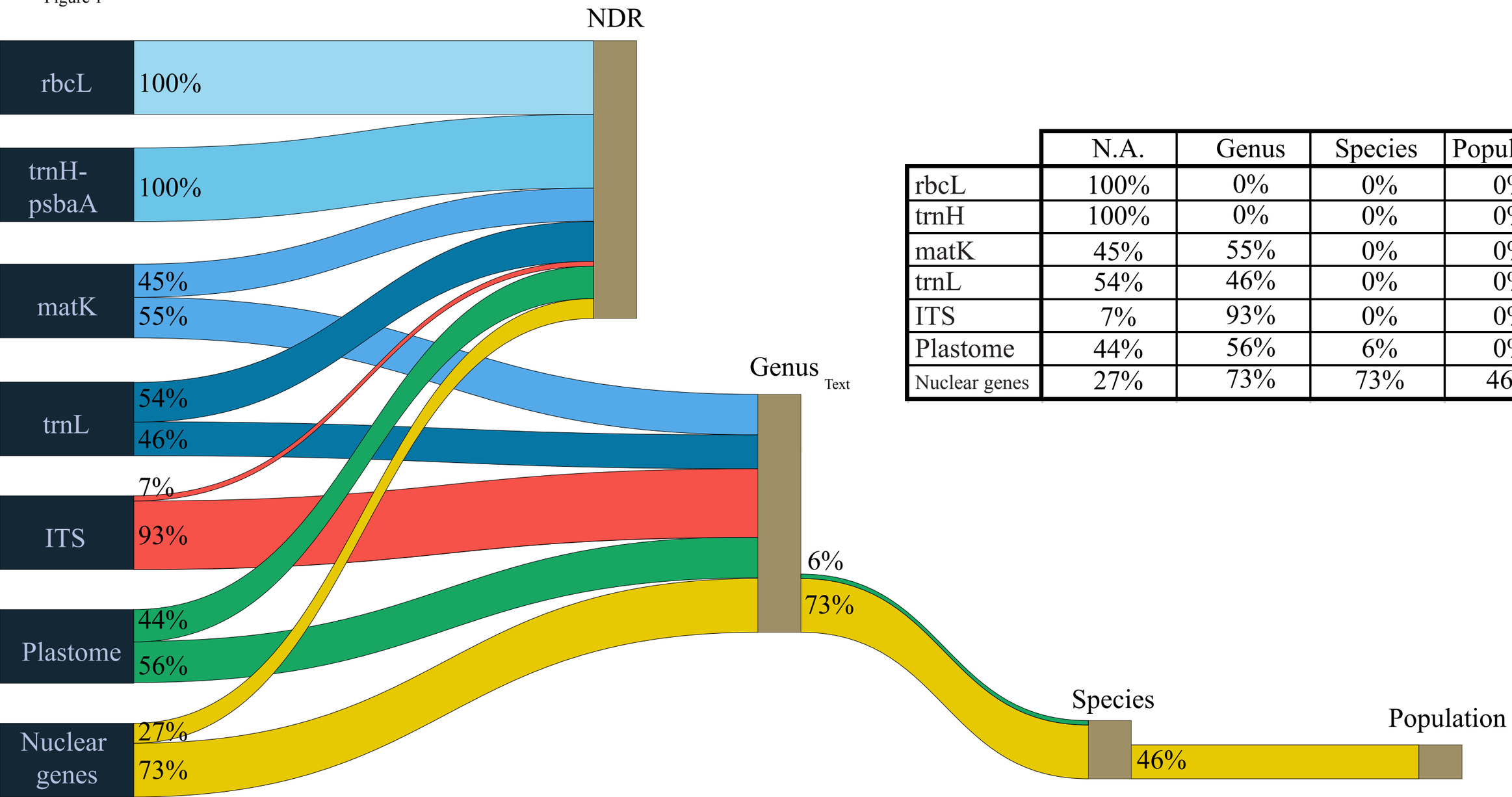
660

Figure 1: Sequencing recovery and identification success for the traded samples. The figure shows the percentage of samples for which useful marker sequences were successfully retrieved for molecular identification by ITS, plastid barcodes, plastomes and nuclear markers. Where no sequence data was recovered or identification above the genus level, samples are classed as NDR (no data recovered). For the samples that produced useable sequence data, the proportion of samples that resulted in identification at the genus, species and population levels is given.

Figure 2: 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/middle-men in India or professional collectors in Morocco (indicated by square boxes).

Figure 3: 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), adulterated roots are shown in orange, adulteration with *A. homogamos* in in brown, and the two varieties of *A. pyrethrum* in blue.

Figure 1



	N.A.	Genus	Species	Population
rbcL	100%	0%	0%	0%
trnH	100%	0%	0%	0%
matK	45%	55%	0%	0%
trnL	54%	46%	0%	0%
ITS	7%	93%	0%	0%
Plastome	44%	56%	6%	0%
Nuclear genes	27%	73%	73%	46%

Figure 2

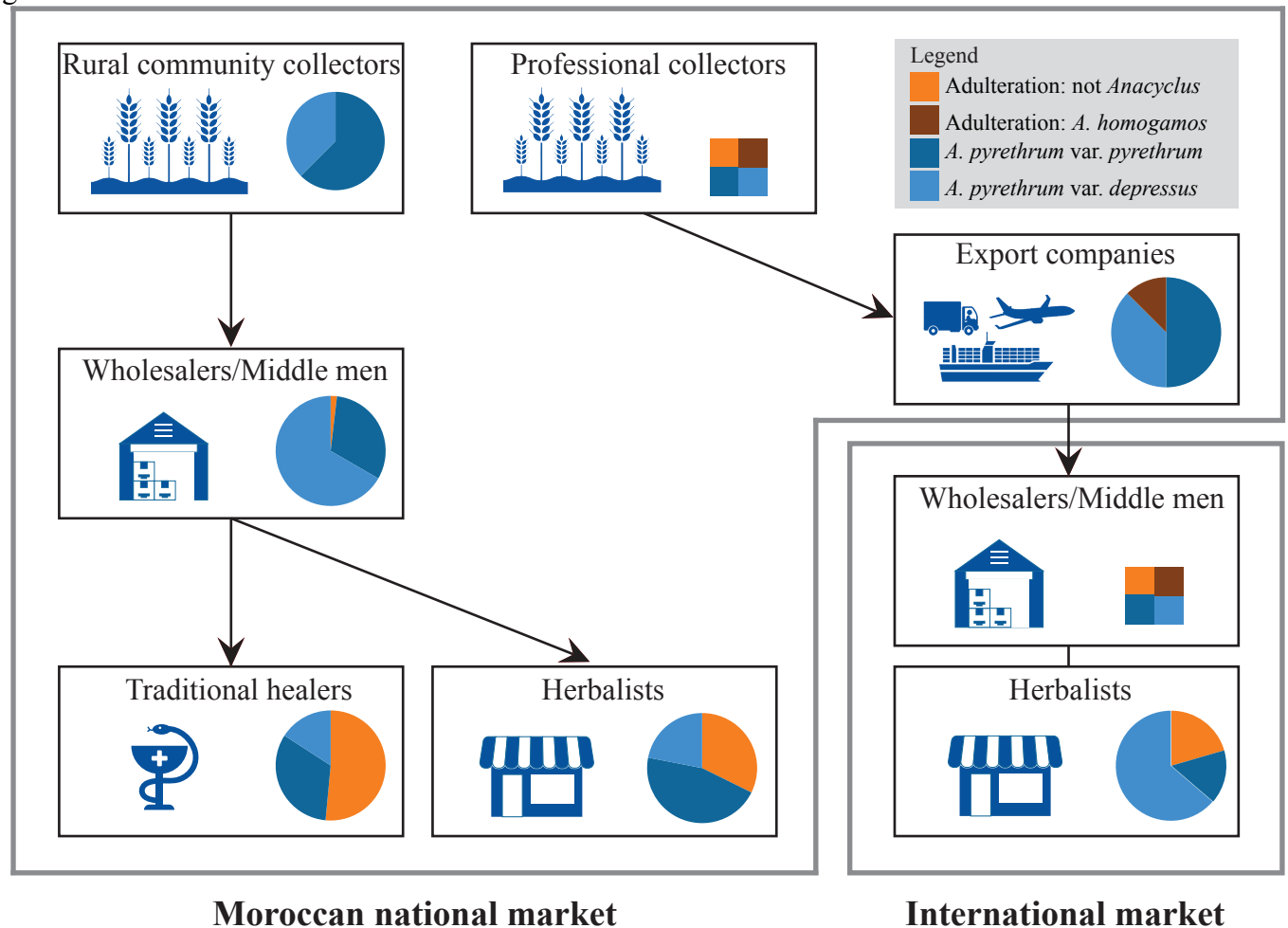
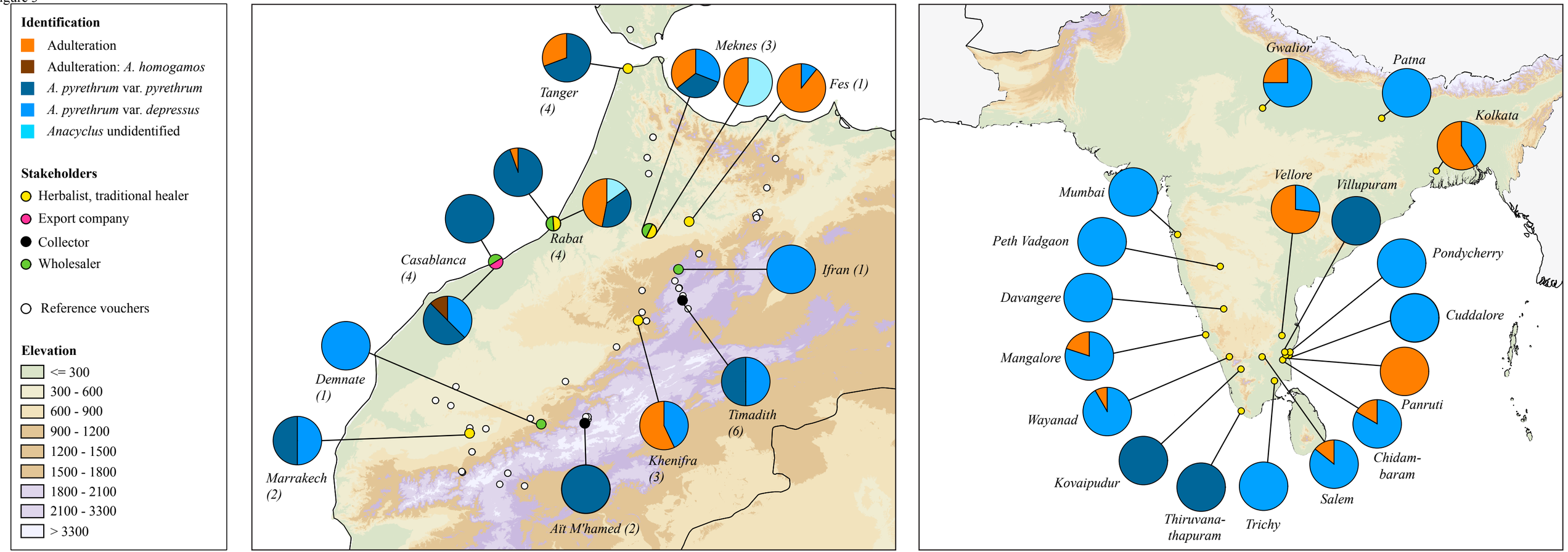

Rural community collectors
Professional collectors
Legend
Adulteration: not *Anacyclus*
Adulteration: *A. homogamos*
A. pyrethrum var. *pyrethrum*
A. pyrethrum var. *depressus*
Export companies
Wholesalers/Middle men
Wholesalers/Middle men
Traditional healers
Herbalists
Herbalists
Moroccan national market
International market

Figure 3


Identification
Adulteration
Adulteration: A. homogamos
A. pyrethrum var. pyrethrum
A. pyrethrum var. depressus
Anacyclus undidentified
Stakeholders
Herbalist, traditional healer
Export company
Collector
Wholesaler
Reference vouchers
Elevation
<= 300
300 - 600
600 - 900
900 - 1200
1200 - 1500
1500 - 1800
1800 - 2100
2100 - 3300
> 3300
Tanger (4)
Meknes (3)
Fes (1)
Rabat (4)
Casablanca (4)
Ifran (1)
Demnate (1)
Timadith (6)
Khenifra (3)
Marrakech (2)
Aït M'hamed (2)
Gwalior
Patna
Kolkata
Mumbai
Vellore
Villupuram
Peth Vadgaon
Pondycherry
Davangere
Cuddalore
Mangalore
Panruti
Wayanad
Chidam-baram
Kovaipudur
Salem
Thiruvana-thapuram
Trichy