Molecular Ecology Resources (2009) doi: 10.1111/j.1755-0998.2009.02547.x

DNA BARCODING

DNA barcoding discriminates the noxious invasive plant species, floating pennywort (*Hydrocotyle ranunculoides* L.f.), from non-invasive relatives

C. C. M. VAN DE WIEL,*J. VAN DER SCHOOT,*J. L. C. H. VAN VALKENBURG,†H. DUISTERMAAT‡ and M. J. M. SMULDERS*

*Wageningen UR Plant Breeding, PO Box 16, NL-6700 AA Wageningen, The Netherlands, †Plant Protection Service of The Netherlands (PPS), PO Box 9102, NL-6700 HC Wageningen, The Netherlands, ‡Nationaal Herbarium Nederland, PO Box 9514, NL-2300 RA Leiden, The Netherlands

Abstract

Floating pennywort (*Hydrocotyle ranunculoides* L.f.), a member of the plant family Araliaceae originating from North America, is an example of an invasive aquatic species posing serious problems to the management of waterways outside of its original distribution area in Australia and Western Europe. As a consequence, its import was banned in the Netherlands. It can be difficult to distinguish *H. ranunculoides* from other species of the genus on a morphological basis. In this regard, DNA barcoding may become a good alternative once this could be performed on a routine basis. In this study, we show that it is possible to distinguish *H. ranunculoides* from a series of closely related congeners by using a single plastid DNA sequence, *trnH-psbA*.

Keywords: DNA barcode, *Hydrocotyle*, invasive plant species, species identification

Received 1 September 2008; revision accepted 4 December 2008

Introduction

DNA barcoding of living organisms increasingly attracts attention for its potential to routinely identify species from all sorts of samples, provided that DNA can be isolated of sufficient quality to allow polymerase chain reaction (PCR) amplification and sequencing of a target locus (for a review, see Hajibabaei *et al*. 2007). This locus must be variable enough to separate closely related species, but preferably should not show variability within species. Chase *et al*. (2005) recently reviewed the scope and purposes of DNA barcoding for land plants.

Barcoding is obviously useful for various types of biodiversity studies, but there are also promising highly practical applications. An important worldwide problem threatening biodiversity is the widespread occurrence of invasive exotic species (Darling & Blum 2007). Such invasive species can disrupt local plant communities, but can also be a nuisance to managers of roadsides and waterways, for example by obstructing water flow. One way of preventing the spread of known invasive species would be to ban their import. To make this feasible, one must be able to unequivocally distinguish them from related, non-invasive species. This can be a problem at, for example ports of entry, when morphological differences between species are only subtle or, as is common, species are traded as small vegetative plants that do not display the morphological differences used for species distinction (which are often related to flowers). In this case, DNA barcoding would be an attractive aid, even more so when it would become routinely applicable (cf. Chase *et al*. 2005).

Floating pennywort (*Hydrocotyle ranunculoides* L.f.) is member of the plant family Araliaceae and originates from North America. It is an example of an invasive aquatic species posing serious problems to the management of waterways outside of its original distribution area in Australia and Western Europe, e.g. the Netherlands. As a result of its prolific development in eutrophic environments, it can obstruct water flow and thus necessitate costly cleaning

Correspondence: C. C. M. van de Wiel, Fax: +31 317483457;
E-mail: clemens.vandewiel@wur.nl

operations. For this reason, its import has been banned in the Netherlands as an ornamental for garden ponds. The species is, however, not easily distinguished from other *Hydrocotyle* species occurring on several continents outside of Europe. Moreover, there is also a related species native to the Netherlands, marsh pennywort (*H. vulgaris* L.). The *Hydrocotyle* example thus appears as a highly relevant one for testing the possibilities of DNA barcoding for species identification.

There has been some difficulty to find reliable universal sequences for the purpose of DNA barcoding in plants, since there is much variation among plant taxa in proposed regions as to their amount of sequence polymorphism and their amplification efficiency. Variation in amplification efficiency can arise due to differences in length of the PCR product and/or sequence variation in primer annealing sites. As a consequence, different combinations of regions have been proposed: Kress & Erickson (2007) selected a combination of two loci, *rbcL* and *trnH-psbA*, whereas Chase *et al*. (2007) selected two combinations of three for further testing, *rpoC1/matK* with either *rpoB* or *trnH-psbA*. Both groups advocated *trnH-psbA* because of its high level of sequence variation, but this is accompanied by enormous length variation, which may interfere with efficient amplification. Another disadvantage of this high variability in relation to database formation is that (automated) sequence alignments across larger groups will often not be possible. For identification purposes within morphologically well-distinguishable species groups, such as the genus *Hydrocotyle*, this is not an insurmountable issue, as the main problems will be to discriminate between the closely related taxa. Lahaye *et al*. (2008) suggested *matK* as the probably most useful single region for barcoding, based on analysis of > 1600 samples from two biodiversity hotspots (Mesoamerica and southern Africa). Where more refinement may prove necessary, *trnH-psbA* was seen as most useful additional locus. Most recently, in a survey of 92 species from 32 genera across the land plants, Fazekas *et al*. (2008) found little difference in species resolution between the three-locus combinations suggested. Again, the most discriminatory single loci were *matK* and *trnH-psbA*. In the present study, sequencing of *matK* and *trnH-psbA* was tested for barcoding of *H. ranunculoides* and a representative series of related species from all over the world.

Materials and methods

Plant material

Plant material was freshly collected from several sites in the Netherlands and obtained from other sources or herbariums. An overview of accessions, their species assignments and origins is presented in Table 1. For all newly collected material, voucher specimens have been stored at the National Herbarium NHN in Leiden (L). The fresh leaf material was immediately dried and stored on silica gel in bags.

DNA analysis

DNA was extracted from dry leaves after tissue disruption in a Retsch MM300 bead mill using the DNeasy Plant Mini kit (QIAGEN) according to the manufacturer's instructions. PCR amplification was performed in a 20-μL solution containing 10 mM Tris-HCl, pH 9.0, 1.5 mM $MgCl_2$, 50 mM KCl, 0.1% Triton X-1, 2 pM of each dNTP, 4 pM of each primer, 20 ng of template DNA and 0.4 U of SuperTaq DNA polymerase (HT Biotechnology Ltd). We used the following PCR protocol: one cycle for 5 min at 94 °C, 30 cycles (30 s at 94 °C, 30 s at 55 °C, 120 s at 72 °C), followed by 10 min at 72 °C. Direct sequencing of (diluted) PCR products was performed using the BigDye 3.1 software on a 3100 sequencer, both from Applied Biosystems. Primers used were according to the Kew barcoding protocols for *matK* (http://www.kew.org/barcoding/) and to Kress & Erickson (2007) for *trnH-psbA* (see Table 2). Sequences were aligned by ClustalW and on the basis of this alignment, dendrograms were constructed by neighbour-joining, using options p-distance and pairwise-deletion of indels, in the program MEGA version 4. All sequences have been submitted to the EMBL database (Accession nos FM207049–FM207086, see Table 1) (Tamura *et al.* 2007).

Results

PCR amplification and *trnH-psbA* sequencing proved possible for all samples, except for two herbarium specimens of *H. sibthorpioides* Lam. (Murata *et al*. 1704 and de Wilde & de Wilde-Duyfjes 16437), probably due to poor DNA quality. The two samples of *Centella asiatica* (L.) Urb. differed too much in sequence to be aligned with the *Hydrocotyle* sequences, despite both genera being members of the order Apiales. The results from aligning the sequences of the *Hydrocotyle* samples are visualized by the dendrogram in Fig. 1. All 13 samples of *H. ranunculoides* had identical sequences and could be clearly distinguished from all other *Hydrocotyle* species tested. *H. vulgaris*, the species native to the Netherlands, and *H. verticillata* Thunb. could also be distinguished, as could all three unidentified samples and the one sample of *H. novae-zelandiae*. The situation for *H. leucocephala* Cham. & Schltdl. and *H. sibthorpioides* was less clear: the samples of *H. leucocephala* were divided over two clusters, one of them also including a sample of *H. sibthorpioides* from Japan. Two other samples of *H. sibthorpioides*, of var. *sibthorpioides* from Dutch cultivation, however, did cluster separately. The one remaining unidentified species, VV 3346 from cultivation, clustered with the *H. sibthorpioides/leucocephala* subgrouping.

Table 1 Overview of accessions used in this study with species identifications, voucher information and collection data

Species	Voucher	EMBL Accession†	Collection		
			Year	Country	Origin
H. ranunculoides	Duistermaat *et al*. HDN 417	FM207049	2007	The Netherlands	Eindhoven
H. ranunculoides	Rayner, Reeves & Peirce sample 1	FM207050	2007	Australia	Perth, Bannister Creek
H. ranunculoides	Rayner, Reeves & Peirce sample 2	FM207051	2007	Australia	Perth, Bannister Creek
H. ranunculoides	Van Valkenburg 3280	FM207052	2007	The Netherlands	Boskoop
H. ranunculoides	Van Valkenburg & van Veen 3276	FM207053	2007	The Netherlands	Boskoop, 'Zuidwijk'
H. ranunculoides	Van Valkenburg & Duistermaat 3261	FM207054	2007	The Netherlands	Groot Wilnis-Vinkeveen polder
H. ranunculoides	Duistermaat 424	FM207055	2007	The Netherlands	Voorschoten, near Duivenvoorde estate, along N 447
H. ranunculoides	Duistermaat 425	FM207056	2007	The Netherlands	Voorschoten, Starrenburg polder, along N447, near grower Groot & Zn.
H. ranunculoides	Cabrera Walsh, G. s.n., 5 July	FM207057	2007	Argentina	Buenos Aires prov., Otamendi
H. ranunculoides	Cabrera Walsh, G. s.n., 13 July	FM207058	2007	Argentina	Buenos Aires prov., José Leon Soarez
H. ranunculoides	Duistermaat HDN 430	FM207059	2007	The Netherlands	Zuid Holland, Nieuw-Lekkerland polder, Tiendwetering
H. ranunculoides	Van Valkenburg, Pot & Rotteveel 3331	FM207060	2007	The Netherlands	Drenthe province, Verlengde Hoogeveensche vaart, Oosterhesselerbrug
H. ranunculoides	Cabrera Walsh, G. s.n., 20 July	FM207061	2007	Argentina	Formosa province, Colonia Cano
H. leucocephala	Van Valkenburg & van Veen 3279	FM207062	2007	The Netherlands-cultivation	Schiphol Amsterdam Airport, origin Singapore grower
H. leucocephala	Van Valkenburg & Den Hartog 3338	FM207063	2007	The Netherlands-cultivation	Pijnacker, origin Kota Tinggi, Malaysia grower
H. leucocephala	TM Pedersen 9276	FM207064	1969	Paraguay	Dep. Central, Villa Elisa
H. leucocephala	TM Pedersen 9253	FM207065	1969	Argentina	Prov. Corrientes, Dep. Santo Tome, Estancia 'Garruchos'
H. leucocephala	Van Valkenburg & Den Hartog 3345	FM207066	2007	The Netherlands-cultivation	Maasdijk, locally produced
H. sibthorpioides	Murata 19692	FM207067	1965	Japan	Loc. Honshu, Pref. Kyoto: Katsura
H. sibthorpioides	de Wilde & de Wilde-Duyfjes 16437		1975	Indonesia	Sumatra, Aceh, Gunung Leuser Nature Reserve, 1000 m alt.
H. sibthorpioides	Murata *et al*. 1704		1978	Indonesia	Java, W., between Cianjul and Punciac-pass, 800 m alt.
H. sibthorpioides var. *sibthorpioides*	Van Valkenburg & Duistermaat 3270	FM207068	2007	The Netherlands-cultivation	Vinkeveen, grower
H. sibthorpioides var. *sibthorpioides*	Van Valkenburg & Zijlstra 3252	FM207069	2007	The Netherlands-cultivation	Amstelveen, grower, origin Costa Rica
H. sp.	Van Valkenburg 3255	FM207070	2007	The Netherlands-cultivation	Amstelveen, grower, origin Costa Rica
H. sp.	Van Valkenburg & Zijlstra 3332	FM207071	2007	The Netherlands-cultivation	Aalsmeer, Las Palmas, with Phoenix rotboellii from Costa Rica
H. sp.	Van Valkenburg & Den Hartog 3346	FM207072	2007	The Netherlands-cultivation	Boskoop, locally produced from initially purchased stock
H. novae-zelandiae var. *montana*	Van Valkenburg & Den Hartog 3347	FM207073	2007	The Netherlands-cultivation	Boskoop, locally produced from initially purchased stock
H. verticillata	Van Valkenburg 3282	FM207074	2007	The Netherlands-cultivation	Vinkeveen, grower
H. verticillata	Van Valkenburg & Zijlstra 3251A	FM207075	2007	The Netherlands-cultivation	Amstelveen, grower
H. verticillata	Van Valkenburg 3319	FM207076	2007	The Netherlands-cultivation	Belfeld, origin Singapore
H. verticillata	Van Valkenburg & Den Hartog 3344	FM207077	2007	The Netherlands-cultivation	Maasdijk, locally produced
H. verticillata	Van Valkenburg 3359	FM207078	2007	The Netherlands	Nijkerk, escape from garden pond
H. vulgaris	Van Valkenburg 3307	FM207079	2007	The Netherlands-cultivation	Bleiswijk, grower
H. vulgaris	Van Valkenburg 3318	FM207080	2007	The Netherlands-cultivation	Belfeld, grower
H. vulgaris	Duistermaat HDN 428	FM207081	2007	France	Bretagne, Finistere, N-coast nr Curnic, Greve du Vougot
H. vulgaris	Duistermaat HDN 429	FM207082	2007	France	Bretagne, Morbihan, c.4 km N of Carnac, nr Le Hohon
H. vulgaris	Duistermaat HDN 431	FM207083	2007	The Netherlands	Twente, Haaksbergerveen
H. vulgaris	Van Valkenburg 3358	FM207084	2007	The Netherlands	'Erica' estate, 'Het Paradijs', Kallenbroek
Centella asiatica	Van Valkenburg 3297	FM207085	2007	The Netherlands-cultivation	Bleiswijk, grower
Centella asiatica	Gwee, A.T. s.n.	FM207086	2007	Singapore	Botanic Gardens

Table 2 PCR primers used for amplification of plastid DNA sequences in this study

Plastid locus	Primer name	Sequence
matK	Kew_*matK*_ 2.1F	ATCCATCTGGAAATCTTAGTTC
	Kew_*matK*_5R	GTTCTAGCACAAGAAAGTCG
trnH-psbA	*trnH-psbA*_1F	ACTGCCTTGATCCACTTGGC
	*trnH-psbA*_1R	CGAAGCTCCATCTACAAATGG

Table 3 shows the distinguishing features of the *trnH-psbA* haplotypes. *H. ranunculoides*, the *H. leucocephala* subgrouping from cultivated material, *H. sibthorpioides* var. *sibthorpioides* and *H. vulgaris* can be positively distinguished (i.e. by unique character states), each by just one sequence position. The only species with more than one unique character states (i.c. 5) was *H. verticillata*. The remaining part, the *H. leucocephala*/*sibthorpioides*/species cluster (cf. Fig. 1) and the three other *Hydrocotyle* spp., did not show unique single nucleotide polymorphisms (SNPs), but they were distinguished by large indels. Thus, indels set apart the *H.* sp. VV3332 and VV3255, respectively, from the three accessions of *H. leucocephala* with which they clustered (Fig. 1). Likewise, an indel set apart one sample from *H. verticillata* (VV3319) from the rest of the *H. verticillata* cluster; this indel also led to the removal of one of the unique characters for the other samples from the species, an A at position 40. One of the unidentified samples, VV3346, was unique in having an additional T at position 206. By this character and an additional A at position 53, it could be distinguished from the other members of the *H. leucocephala*/*sibthorpioides*/species cluster (cf. Fig. 1). At two positions, SNPs could not be scored unequivocally in one or two samples due to the impossibility to locate gaps unequivocally: at position 111 for *H.* species VV3332 and VV3255, and at position 124 for *H. novae-zelandiae* VV3347 (see Table S1, Supporting Information, for the alignment used in the analysis).

In addition to the *trnH-psbA* region, *matK* was also tested for usefulness, but amplification and sequencing were not reliable among samples, which is in line with findings of, for example, Kress & Erickson (2007) and Fazekas *et al*. (2008) that *matK* is one of the most troublesome chloroplast loci in PCR amplification. Therefore, *matK* was omitted from the study.

Discussion

This study shows that sequencing one variable chloroplast locus, *trnH-psbA*, sufficed to distinguish an important invasive plant species, *Hydrocotyle ranunculoides*, from a series of closely related congeners. Thus, a limited investment in DNA barcoding can generate an identification tool for plant material, specifically useful for cases where morphology is inadequate to assess species identity. This is in line with the study by Kress *et al*. (2005), which showed *trnH-psbA* to be the region that combined the highest variability and the highest PCR amplification success in a comparison of 99 angiosperm species across 53 families. In a test on closely related species from the genus *Compsoneura* (Myristicaceae), Newmaster *et al*. 2008 also found most variation with the *trnH-psbA* locus.

Although the greater part of the *Hydrocotyle* species included in the study, namely *H. vulgaris*, *H. verticillata*, and several unidentified samples could be clearly distinguished, this did not apply to all *H. leucocephala* and *H. sibthorpioides* samples. The *H. leucocephala* samples fell separately into two clusters, one of which included an *H. sibthorpioides* sample and an unidentified species (VV 3346), and the other cluster harbouring two unidentified samples that looked morphologically similar to *H. leucocephala* (VV3332 and VV3255). Two other samples, *H. sibthorpioides* var. *sibthorpioides*, formed a separate cluster and thus could be well distinguished from all other species. One possible explanation for problems in positioning *H. leucocephala* and *H. sibthorpioides*, namely interspecific hybridization, is not very likely in this case, since the species have totally separate distribution areas (South America and tropical Africa/Asia, respectively). To clarify this issue, more samples of *H. leucocephala* and particularly, *H. sibthorpioides*, (two samples of the latter failed in this study) would need to be sequenced.

Consistent with other reports on the *trnH-psbA* region (e.g. Kress *et al*. 2005; Chase *et al*. 2007), large indels were found in our *Hydrocotyle* sequences, although they were derived from closely related species. Even within one of the species, *H. verticillata*, a 60-bp insertion was found caused by the imperfect extension of a 14-bp minisatellite repeat in one sample. The indels were helpful in providing additional clues for identification, but also interfered with sequence alignment. As a consequence, not all distinguishing sequence positions were fully reliable: in two cases, the possibility of scoring a specific position depended on the arbitrary placement of an indel (see Table 3). Clearly, this severely limits the usefulness of *trnH-psbA* for phylogenetic reconstruction (cf. Kress *et al*. 2005). However, the primary barcoding goal, species distinction, was not seriously hampered by this phenomenon, except for some samples from *H. leucocephala* and *H. sibthorpioides* (cf. Fig. 1). The haplotypes varied in 15 positions overall, although most of them were really unique for a species at only one of these positions.

The *trnH-psbA* sequences of our outgroup, *Centella asiatica*, could not be aligned with those of the *Hydrocotyle* spp., although they are coming from related families (Apiaceae and Araliaceae, respectively). With a routine identification system, this may be troublesome when one would like to be able to assign any DNA sample from scratch, by

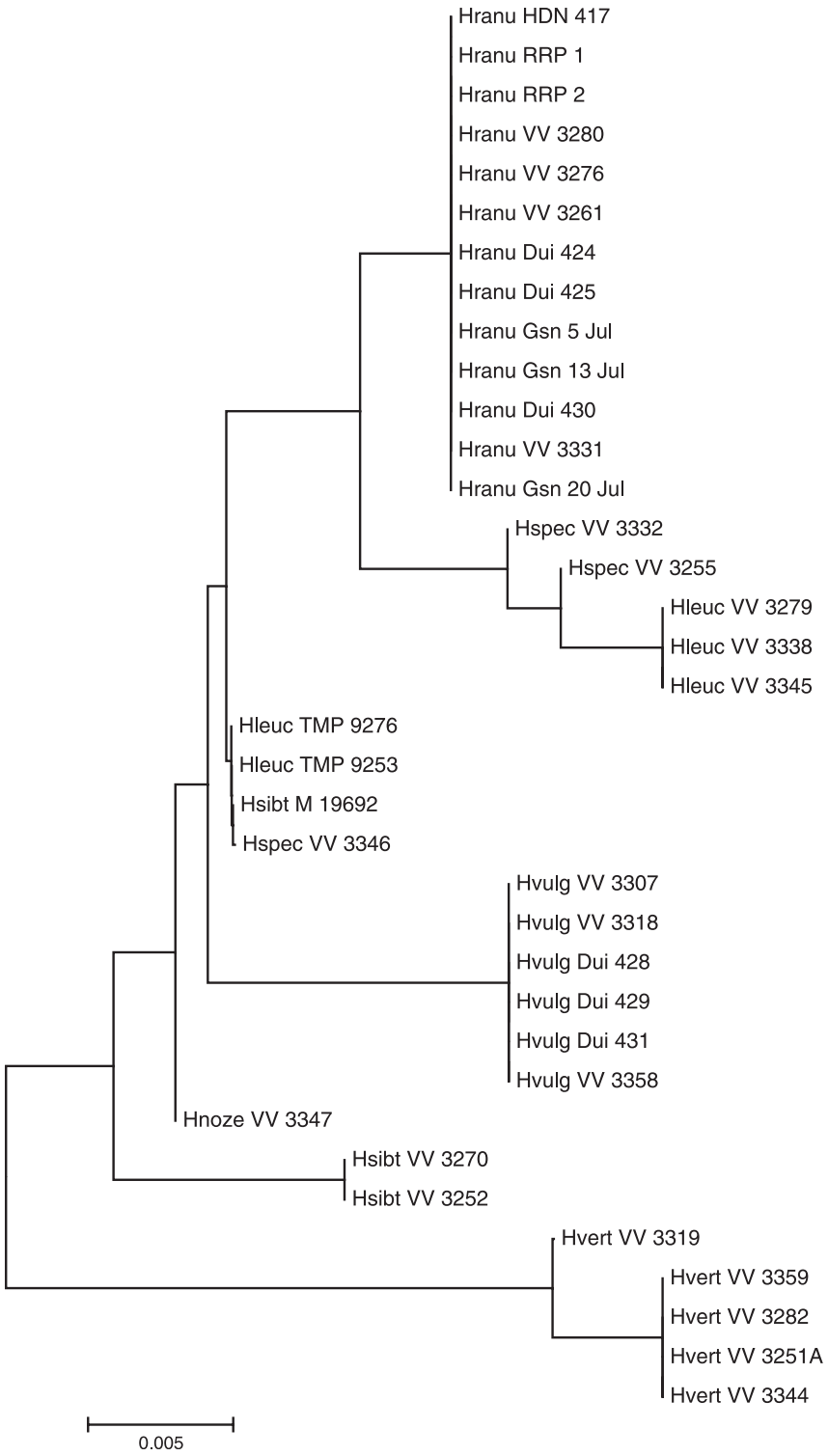

Fig. 1 Neighbour-joining tree (unrooted) based on ClustalW alignment of plastid sequence *trnH-psbA* from accessions of several species of the genus *Hydrocotyle*, using MEGA version 4.

comparison to DNA barcoding databases (cf. Fazekas *et al*. 2008). However, for the purpose of this study, that is, being able to routinely separate living plant materials from the clearly recognizable *Hydrocotyle* genus into either the undesirable *H. ranunculoides* or its unproblematic congeners, this would not pose any problems.

In conclusion, sequencing a single plastid DNA region, *trnH-psbA*, enables to distinguish a notorious invasive plant species, *Hydrocotyle ranunculoides*, from its most closely related congeners. This could be helpful with enforcing a ban on import of such invasives, such as is already in place in the Netherlands. This will become even more so once

Table 3 Haplotypes of groups of genotypes observed in Fig. 1. Unique variants are indicated by grey shading

Sequence position	10	21	40	47	53	60	66	73	99	111	119	124	140	203	206	210	311
Consensus	A	C	C	G	–	T	C	A	A	A	A	T	A	G	–	T	C
H. ranunculoides	A	C	C	G	–	T	G	A	A	A	A	T	A	A	–	T	C
H. leucocephala cultivated	A	A	C	G	–	T	C	A	C	A	A	T	C	A	–	T	C
H. sp. Van Valkenburg 3255 and 3332	A	–	–	–	–	–	–	–	–	(A)*	A	T	C	A	–	T	C
H. sp. Van Valkenburg 3346	A	A	C	G	A	T	C	A	A	A	A	T	A	G	T	T	C
H. leucocephala South America/*H. sibthorpioides* Japan	A	C	C	G	–	T	C	A	A	A	A	T	A	G	–	T	C
H. novae-zelandiae var. *montana*	A	C	C	G	–	T	C	A	A	A	-	(T)*	A	G	–	T	C
H. sibthorpioides var. *sibthorpioides*	T	C	C	G	A	T	C	A	A	A	A	T	A	G	–	G	C
H. vulgaris	A	A	C	G	–	G	C	T	A	A	A	T	A	G	–	T	C
H. verticillata	T	C	A†	T	–	G	C	A	A	T	G	C	A	G	–	T	G

*depending on indel alignment.
†except for *H. verticillata* Van Valkenburg 3319.

DNA barcoding would be turned into machinery routinely operable by nonspecialists in botany and/or molecular genetics (cf. Chase *et al*. 2005).

Acknowledgements

The research was funded by the Netherlands' Ministry of Agriculture, Nature and Food Quality, FES-programme 'Versterking Infrastructuur Plantgezondheid'. We thank Dr F.T. Bakker (Wageningen UR Biosystematics Group) for advice on barcoding sequences.

References

Chase MW, Salamin N, Wilkinson M *et al*. (2005) Land plants and DNA barcodes: short-term and long-term goals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **360**, 1889–1895.

Chase MW, Cowan RS, Hollingsworth PM *et al*. (2007) A proposal for a standardised protocol to barcode all land plants. *Taxon*, **56**, 295–299.

Darling JA, Blum MJ (2007) DNA-based methods for monitoring invasive species: a review and prospectus. *Biological Invasions*, **9**, 751–765.

Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG, Husband BC, Percy DM, Hajibabaei M, Barrett SCH (2008) Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS ONE* **3**, e2802.

Hajibabaei M, Singer GAC, Hebert PDN, Hickey DA (2007) DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics*, **23**, 167–172.

Kress WJ, Erickson DL (2007) A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS ONE*, **2007**, e508.

Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005) Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences, USA*, **102**, 8369–8374.

Lahaye R, Van der Bank M, Bogarin D *et al*. (2008) DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences, USA*, **105**, 2923–2928.

Newmaster SG, Fazekas AJ, Steeves RAD, Janovec J (2008) Testing candidate plant barcode regions in the Myristicaceae. *Molecular Ecology Resources*, **8**, 480–490.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596–1599.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1 Alignment of plastid sequence trnH-psbA of *Hydrocotyle* species

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.