

RESEARCH ARTICLE

Open Access

# Evaluating the feasibility of using candidate DNA barcodes in discriminating species of the large Asteraceae family

Ting Gao<sup>1,2</sup>, Hui Yao<sup>1</sup>, Jingyuan Song<sup>1</sup>, Yingjie Zhu<sup>1</sup>, Chang Liu<sup>1</sup>, Shilin Chen<sup>1\*</sup>

## Abstract

**Background:** Five DNA regions, namely, *rbcl*, *matK*, ITS, ITS2, and *psbA-trnH*, have been recommended as primary DNA barcodes for plants. Studies evaluating these regions for species identification in the large plant taxon, which includes a large number of closely related species, have rarely been reported.

**Results:** The feasibility of using the five proposed DNA regions was tested for discriminating plant species within Asteraceae, the largest family of flowering plants. Among these markers, ITS2 was the most useful in terms of universality, sequence variation, and identification capability in the Asteraceae family. The species discriminating power of ITS2 was also explored in a large pool of 3,490 Asteraceae sequences that represent 2,315 species belonging to 494 different genera. The result shows that ITS2 correctly identified 76.4% and 97.4% of plant samples at the species and genus levels, respectively. In addition, ITS2 displayed a variable ability to discriminate related species within different genera.

**Conclusions:** ITS2 is the best DNA barcode for the Asteraceae family. This approach significantly broadens the application of DNA barcoding to resolve classification problems in the family Asteraceae at the genera and species levels.

## Background

Asteraceae is the largest family of flowering plants in the world. The family includes over 1,600 genera and 23,000 individual species. Many members of the Asteraceae family are important for medicinal, ornamental, and economic purposes.

Approximately 300 Asteraceae species are already used for medicinal purposes in China. For example, *Artemisia annua* and its derivatives are effective in treating malaria [1]. *Saussurea involucrate*, an endangered species, possesses anti-fatigue, anti-inflammation, anti-tumor and free radical scavenging properties [2]. *Echinacea* also has immuno-modulatory properties with its ability to reduce inflammation, speed up wound healing and boost the immune system in response to bacterial or viral infection [3]. Commercially important plants of the Asteraceae family include the food crops *Lactuca*

*sativa* (lettuce), *Cichorium intybus* (chicory), *Cynara scolymus* (globe artichoke), *Smallanthus sonchifolius* (yacon), *Helianthus tuberosus* (jerusalem artichoke), and so on. Aside from consumption, the seeds of *Helianthus annuus* (sunflower), and those of *Carthamus tinctorius* (safflower), another Asteraceae member, can be used for the production of cooking oil. Other commercially important species of the family Asteraceae are members of the *Tanacetum*, *Chrysanthemum* and *Pulicaria* genera, which have insecticidal properties. *Eupatorium adenophorum* is also one of the more noxious invasive plants worldwide, and it does have a significant effect on local ecosystems.

The wide variety of plants in the family Asteraceae often makes identification at the species level difficult [4]. Given the many valuable members of Asteraceae described above, an easy and accurate method of authenticating an Asteraceae species is indispensable for ensuring the drug and food safety of internationally traded herbs.

\* Correspondence: slchen@implad.ac.cn

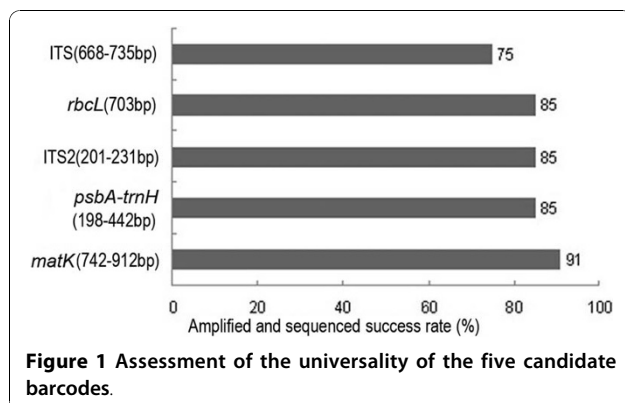
<sup>1</sup>Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, PR China  
Full list of author information is available at the end of the article

DNA barcoding is a process that uses a short piece of DNA sequence from a standard locus as a species identification tool [5]. DNA barcode regions have already been adopted for animal use [6,7] and several regions have previously been recommended for plant use [8-17]. The Plant Working Group of the Consortium for the Barcode of Life (CBOL) proposed *rbcL* and *matK* as the core DNA barcodes for plants [18]. A previous study by Kress et al. [9] tested 7 promising barcodes. However, only 15 sequences of Asteraceae, representing 14 species distributed among only 9 genera, were analyzed in the study. The CBOL Plant Working Group [18] also evaluated the performance of the leading barcoding loci in species identification, but the sequences of Asteraceae used included just 75 samples, consisting of 38 species belonging to 19 genera. Chen et al. [12] likewise compared the practicality of using the suggested barcode sequences against a large number of medicinal plants. However, the study included no more than 450 sequences of Asteraceae derived from 306 species from 50 genera. The researchers did not provide sufficient evidence that the recommended DNA barcode regions are suitable for species identification in the family Asteraceae, which includes a large number of closely related species. Thus, this issue is addressed in our study by comparing the feasibility of using each of these five proposed DNA barcodes (*rbcL*, *matK*, ITS, ITS2, and *psbA-trnH*) in the Asteraceae family.

## Results and Discussion

### Assessment of the universality of the five candidate barcodes

A universal DNA barcode is required to be tractable for use in a wide range of species. Therefore barcode regions must be relatively short in length to facilitate DNA extraction, amplification and sequencing [9]. As shown in Figure 1, for the selected samples, three regions (ITS2, *psbA-trnH* and *rbcL*) were amplified using a single pair of universal primers for each locus



that results in high amplification and a sequencing efficiency of 85%. In comparison, ITS had a relatively lower efficiency at 75%. We used two pairs of *matK* primers exhibiting different universalities for the members of the family Asteraceae. The primers Kim3F/1R and 390F/1326R achieved amplifying and sequencing efficiencies of 91% and 25%, respectively.

### Measurement of inter- versus intra-specific genetic divergence at each locus

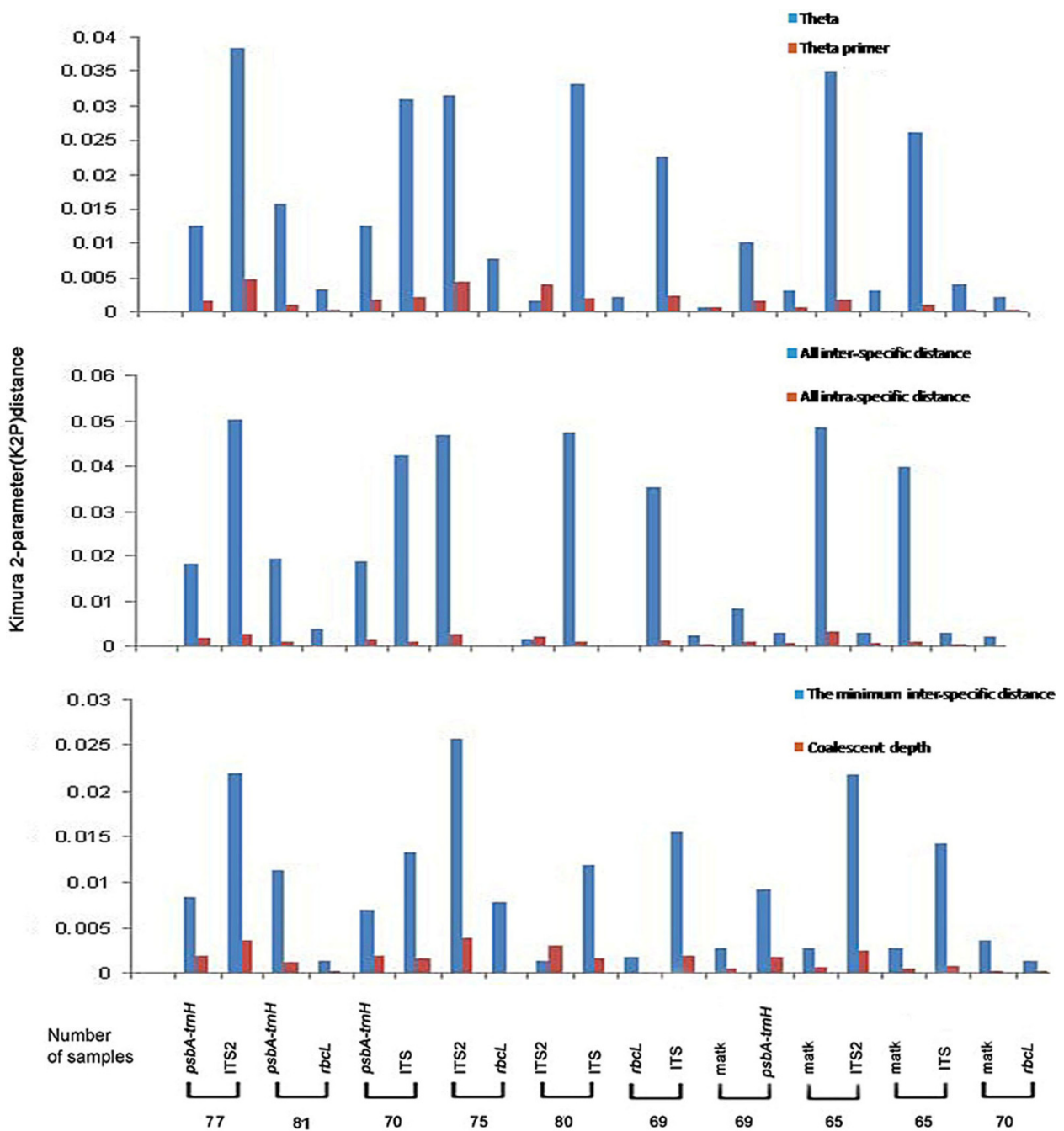
Six metrics were employed to characterize inter- versus intra-specific variation (Figure 2) [11,12,19-21]. A favorable barcode should possess a high inter-specific divergence to distinguish different species. ITS2 and ITS both exhibited significantly higher levels of inter-specific discriminatory ability than *psbA-trnH* and *matK*. The lowest divergence between conspecific individuals, as determined by all inter-specific calculations was exhibited by *rbcL*. Wilcoxon signed-rank tests affirmed that ITS2 had the highest divergence at the inter-specific level, whereas *rbcL* had the lowest (Table 1). The results of the intra-specific differences were similar, with ITS2 contributing the largest and *rbcL* the smallest variations (Figure 2).

### Testing the efficacy of authentication

BLAST1 and Distance methods were used to test the ability of the potential barcoding sequences in assigning unique species identities to the given samples [12,22]. The results from the two methods revealed a clear pattern (Figure 3), demonstrating that the ITS region exhibits the highest identification efficiency. ITS2 and ITS performed well at the genus level using both methods, and at the species level using the Distance method. Using the BLAST1 method, ITS2 was less efficient (2.5%) than ITS at the species level, while *rbcL* was the lowest performer. In addition, except for the combination of *matK* and *psbA-trnH*, which improved the correct identification rates by 1.4%, using one sequence rather than a combination of two markers didn't improve the rates of identification.

The meta-analysis of markers, ITS, *psbA-trnH*, *matK* and *rbcL*, was also performed in parallel with the analysis on ITS2 using GenBank data (see Additional file 1: Identification efficiency of the five regions evaluated in a large pool of Asteraceae samples from GenBank). The correct identification rates were significantly higher for ITS2 than for other markers except ITS. The GenBank data analyses were consistent with our experimental data results. Compared with single markers alone, combinations of markers could improve the rate of correct species identification (<5%).

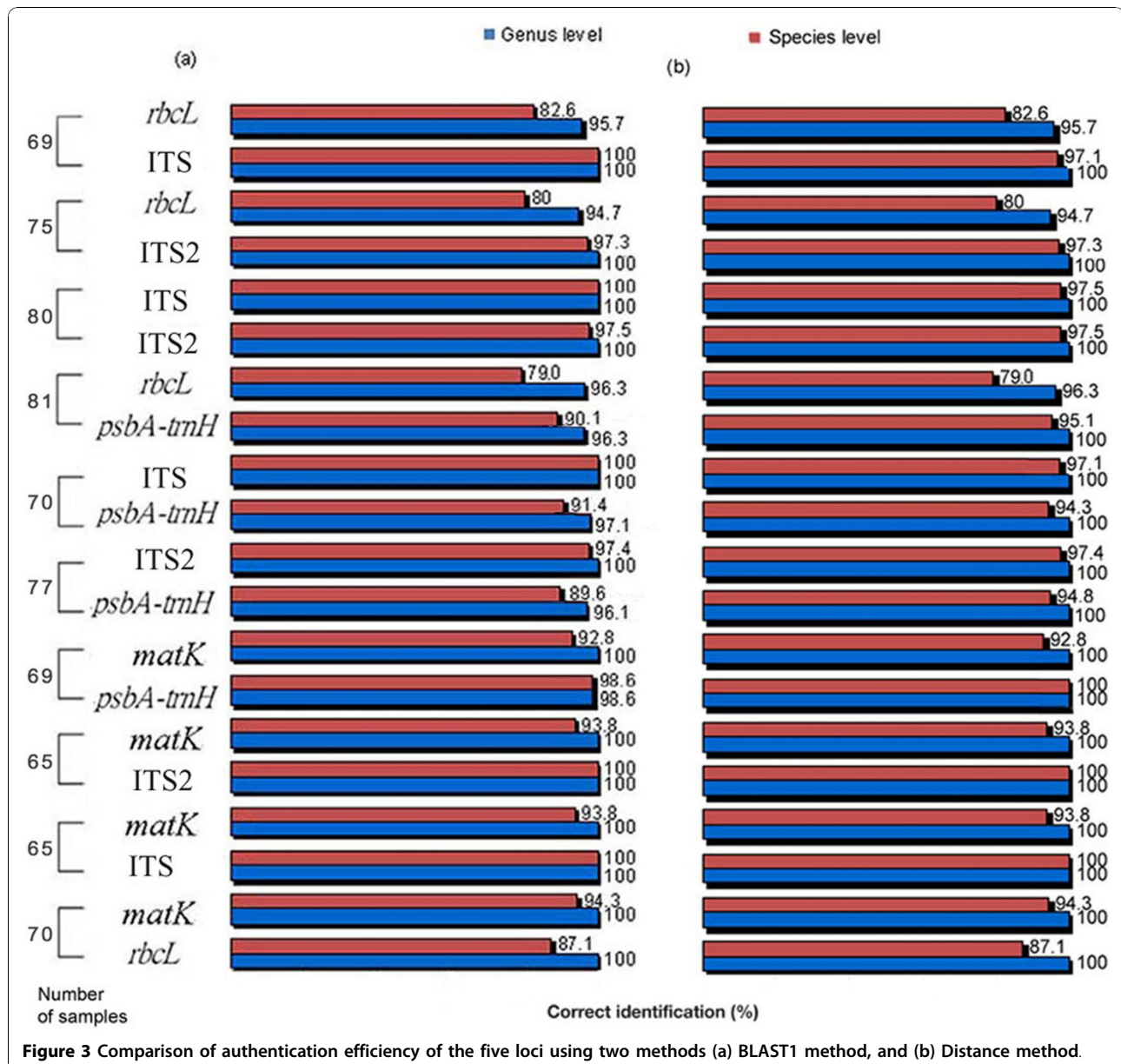
Overall, our study demonstrates that ITS2 is the most successful region in terms of universality, the specific



**Figure 2 Analyses of the inter-specific divergence between congeneric species and intra-specific variation of the five loci.** First, three parameters were used to characterize inter-specific divergence: (i) average inter-specific distance (K2P distance) between all species in each genus with at least two species; (ii) average theta prime ( $\theta'$ ), where theta prime is the mean pairwise distance within each genus with more than one species, thus eliminating biases associated with different numbers of species among genera; and (iii) smallest inter-specific distance, i.e., the minimum inter-specific distance within each genus with at least two species. Second, three additional parameters were used to determine intra-specific variation: (i) average intra-specific difference (K2P distance), that between all samples collected within each species with more than one individual; (ii) theta ( $\theta$ ), where theta is the mean pairwise distance within each species with at least two representatives;  $\theta$  eliminates biases associated with unequal sampling among a species; and (iii) average coalescent depth, which is the maximum intra-specific distance within each species with at least two individuals.

**Table 1 Wilcoxon signed rank test of the inter-specific divergences among the five loci**

W +	W-	inter Relative Ranks, n, P value	Result
ITS2	<i>psbA-trnH</i>	W + = 92, W - = 13, n = 14, P < 0.0132	ITS2 > <i>psbA-trnH</i>
<i>psbA-trnH</i>	<i>rbcl</i>	W + = 89, W - = 2, n = 13, P < 0.0024	<i>psbA-trnH</i> > <i>rbcl</i>
ITS	<i>psbA-trnH</i>	W + = 65, W - = 13, n = 12, P < 0.0414	ITS > <i>psbA-trnH</i>
<i>psbA-trnH</i>	<i>matK</i>	W + = 21, W - = 0, n = 6, P < 0.0277	<i>psbA-trnH</i> > <i>matK</i>
ITS2	<i>rbcl</i>	W + = 36, W - = 0, n = 8, P < 0.0117	ITS2 > <i>rbcl</i>
ITS2	ITS	W + = 355, W - = 23, n = 27, P < 6.6389 × 10 <sup>-5</sup>	ITS2 > ITS
ITS2	<i>matK</i>	W + = 45, W - = 0, n = 9, P < 0.0076	ITS2 > <i>matK</i>
ITS	<i>rbcl</i>	W + = 45, W - = 0, n = 9, P < 0.0076	ITS > <i>rbcl</i>
<i>rbcl</i>	<i>matK</i>	W + = 3, W - = 7, n = 4, P < 0.4615	<i>rbcl</i> = <i>matK</i>
ITS	<i>matK</i>	W + = 45, W - = 0, n = 9, P < 0.0075	ITS > <i>matK</i>



**Figure 3 Comparison of authentication efficiency of the five loci using two methods (a) BLAST1 method, and (b) Distance method.**

genetic divergence, and discrimination between species among the five markers examined. ITS is also proven as a valuable marker for authenticating species in Asteraceae. However, its low amplification efficiency limits its potential for broad taxonomic use. Although *matK*, *rbcL*, and *psbA-trnH* have effective primers for the amplification, the three markers are less powerful than ITS and ITS2 in species discrimination in the family Asteraceae. Moreover, theoretically, the regions based on nuclear DNA are much more informative than barcodes based on organellar DNA [23].

To evaluate further the ability of the ITS2 region to authenticate a wide range of Asteraceae species, it was also tested against a larger database that includes 3,490 samples sequences derived from 2,315 different species (Table 2). The ITS2 region performed well, with a 76.4% (BLAST1 method) and 69.7% (Distance method) successful identification rate at the species level and a 97.4% (BLAST1 method) and 96.2% (Distance method) successful identification rate at the genus level.

#### ITS2 is suitable, but not ideal

Our research displayed a similar trend to that of Chen et al. [12] and demonstrated that ITS2 is a promising barcode for authenticating plant species. In accordance with the criteria outlined by Kress et al. [9], the ITS2 region has several advantages that make it a promising candidate for DNA barcoding. It has been proposed as a candidate marker for taxonomic classification and barcoding of medicinal plants because it has both high

correct identification rate and high amplification efficiency [12,24-27]. As the ITS2 region is one of the most common regions used for phylogenetic analyses [28-30], a vast amount of sequencing data has already been deposited in GenBank and is ready for immediate use.

The presence of multiple copies of ITS2 sequences is challenging [13]. However, Coleman [27] proposed that the repeats displayed a high degree of similarity. Coleman also suggested that the PCR-amplified copies could represent the information of the ITS2 region in individuals and that ITS2 could be considered a single locus in most cases.

Among the six large genera (number of species > 50) in the Asteraceae family (Table 2), the utility of ITS2 for species authentication varied and could only be analyzed individually, not as a group. For the genus *Brachyscome*, with 57 sequences representing 55 species, ITS2 worked well with a 96.5% successful identification rate. Satisfactory results were also obtained for the genus *Erigeron*, where >80.5% of the sequences were correctly identified. In contrast, ITS2 had lower identification efficiency for the genera *Centaurea* and *Artemisia* (48.0% and 59.3%, respectively). And in two other genera (*Senecio* and *Stevia*), ITS2 was relatively powerful for taxonomic classification, precisely authenticating 73.4% and 76.9% of the samples, respectively. The identification efficiencies of ITS2 in dataset 2 are listed in Additional file 2 (Authentication efficiency of ITS2 using different methods for the genera in dataset 2 containing more than one species).

To improve identification accuracy within a particular genus, using combinations of DNA barcodes may be necessary. Therefore, ITS is proposed for use as a complementary barcode for differentiating species within the Asteraceae family.

#### Application and meaning of DNA barcoding

The selected DNA barcode for Asteraceae, ITS2, is not perfect, especially for taxonomists and phylogenetic experts. However, even an imperfect barcode can have a major effect on many areas of research and be sufficient for many applications [13]. For instance, ITS2 might be a suitable DNA barcode for public users, such as customs officials, forensic examiners, food-processing individuals, and research organizations. Considering that ITS2 has a strong ability to group plant samples into their correct genus and has a relatively high accuracy for grouping samples into their correct species, it is of great practical value to individuals without adequate taxonomic training. Compared with ITS2, ITS or the chloroplast genome is better equipped to deal with the biological complexities of species distinctions, a major focus of taxonomists and phylogenetic experts [13].

**Table 2 Identification efficiency of the ITS2 locus for the family and six large genera in dataset 2 using different methods**

Category	Method	No. of samples	No. of species	Success identification (%) at the species level/ at the genus level	
Genus	<i>Centaurea</i>	BLAST1	204	157	48.0/-
		Distance	204	157	45.6/-
	<i>Senecio</i>	BLAST1	203	157	73.4/-
		Distance	203	157	70.9/-
	<i>Artemisia</i>	BLAST1	91	74	59.3/-
		Distance	91	74	51.6/-
	<i>Stevia</i>	BLAST1	91	75	76.9/-
		Distance	91	75	71.4/-
	<i>Erigeron</i>	BLAST1	87	65	80.5/-
		Distance	87	65	75.9/-
	<i>Brachyscome</i>	BLAST1	57	55	96.5/-
		Distance	57	55	96.5/-
Family	Asteraceae	BLAST1	3,490	2,315	76.4/97.4
		Distance	3,490	2,315	69.4/96.2

**Table 3 Number of DNA sequences used in the study**

Dataset	Markers	Total No. of sequences	No. of sequences belonging to genera containing more than one species	No. of sequences belonging to species containing more than one samples
Dataset 1	<i>psbA-trnH</i>	93 (56)	72 (34)	59 (21)
	ITS2	93 (58)	70 (35)	55 (20)
	ITS	83 (55)	60 (32)	48 (20)
	<i>rbcl</i>	93 (57)	68 (32)	56 (20)
	<i>matK</i>	80 (47)	20 (11)	53 (20)
Dataset 2	ITS2	3,490 (2,315)	2,973 (1,877)	1,748 (583)

The numbers of species to which these sequences belong are shown in parentheses.

## Conclusions

Altogether, our results support the claim that ITS2 is a valuable locus for differentiating species within Asteraceae and that DNA barcoding is a useful tool for classification and identification of individual species. We propose applying DNA barcoding technology to resolve classification problems in the family Asteraceae at the genera and species levels.

## Methods

### Sampling of plant materials

Dataset 1, which consists of 110 samples from 63 species representing 48 genera of Asteraceae (see Additional file 3: Samples in dataset 1 for testing the potential barcodes and accession numbers in GenBank) was gathered from a large geographical area in China from July 2007 to January 2008. Great effort was made to ensure that the samples represent the major lineages of Asteraceae. Furthermore, the maximum number of samples belonging to closely related species was collected (Table 3). Plant samples in dataset 1 were spread across two subfamilies (Carduoideae and Cichorioideae) and encompassed a total of 11 tribes of the family Asteraceae.

### DNA extraction, PCR amplification, and sequencing

DNA extraction, PCR amplification, and sequencing were performed as described previously [12].

### Data Acquisition from GenBank

First, all sequences involving the five markers of Asteraceae were downloaded from GenBank. Certain gene regions of the five barcoding markers based on GenBank annotations were then obtained. Sequences <100 bp in length, with ambiguous bases (more than 15'Ns'), or those belonging to unnamed species (i.e. sequences with 'sp.' in the species name) were filtered out. Finally, to avoid contamination with fungal sequences existing in ITS2 sequences, a Hidden Markov Model (HMM) [31] based on well-curated fungal sequences was used to search for downloaded ITS2 sequences to remove the sequences possibly contaminated with fungi. The meta-

analysis was performed using the remaining sequences (see Additional file 4: Accession numbers of the five loci sequences from GenBank for the meta-analysis). The ITS2 sequences were also used to construct dataset 2, which is comprised of 3,490 sequences from 2,315 Asteraceae species downloaded from GenBank (see Additional file 5: Accession numbers of ITS2 sequences used in dataset 2). Many closely related species were also included in dataset 2 (Table 3).

### Sequence alignment and analysis

Consensus sequences and contig generation were accomplished using CodonCode Aligner V 3.5 (CodonCode Co., USA). The sequences of the candidate DNA barcodes were aligned using Clustal W and the genetic distances were calculated using the Kimura 2-Parameter (K2P) model. The average intra-specific distance, theta, and coalescent depth were calculated to evaluate the intra-specific variation [12,19]-[21]. The average inter-specific distance, the minimum inter-specific distance, and Theta primer were used to represent inter-specific divergences [11,12,20,21]. Wilcoxon signed-rank tests were used as previously described [10-12]. Two methods of species identification, namely BLAST1 and the nearest distance method, were performed as described previously [12,22]. The traffic light approach [32] was used to identify the combination of markers, as long as the sequences could be identified by one of the markers in combination, the combination would have identification power. If any of the sequences were identified unsuccessfully for any marker in combination, the combination would be incapable of identifying that sequence.

## Additional material

**Additional file 1: Identification efficiency of the five regions evaluated in a large pool of Asteraceae samples from GenBank.** For each marker and marker combination, number of samples used for identification and the correct identification rates at the species and genus levels are shown.

**Additional file 2: Authentication efficiency of ITS2 using different methods for the genera in dataset 2 containing more than one**

**species.** For each genus, number of samples and species used for identification and the correct identification rates using different methods are shown.

**Additional file 3: Samples in dataset 1 for testing the potential barcodes and accession numbers in GenBank.** For each samples in dataset 1, the latin name and accession numbers in GenBank are shown.

**Additional file 4: Accession numbers of the five loci sequences from GenBank for the meta-analysis.** For each samples used for the meta-analysis, the accession numbers in GenBank are shown.

**Additional file 5: Accession numbers of ITS2 sequences used in dataset 2.** For each samples used in dataset 2, the accession numbers in GenBank are shown.

#### Acknowledgements

We want to thank YL Lin for the morphological confirmation of plant species, and we would like to thank KL Chen, XC Jia, K Luo for providing plant samples. We are also grateful to LC Shi for helping analyze the data. This work is supported by the Special Founding for Healthy Field (No. 200802043) and National Natural Science Foundation (No.30970307).

#### Author details

<sup>1</sup>Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, PR China. <sup>2</sup>College of Life Sciences, Qingdao Agricultural University, Qingdao, PR China.

#### Authors' contributions

TG, JYS, SLC, CL conceived of the project and designed experiments. TG and HY performed the experiments. TG and YJZ analyzed the data. TG, HY wrote the paper. All authors read and approved the final manuscript.

Received: 18 March 2010 Accepted: 26 October 2010

Published: 26 October 2010

#### References

- Mueller MS, Karhagomba IB, Hirt HM, Wemakor E: The potential of *Artemisia annua* L. as a locally produced remedy for malaria in the tropics: agricultural, chemical and clinical aspects. *J Ethnopharmacol* 2000, **73**:487-493.
- Wu W, Qu Y, Gao HY, Yang JY, Xu JG, Wu LJ: Novel ceramides from aerial parts of *Saussurea involucreata* Kar. et. Kir. *Arch Pharm Res* 2009, **32**:1221-1225.
- Haddad PS, Azar GA, Groom S, Boivin M: Natural health products modulation of immune function and prevention of chronic diseases. *Complement Altern Med* 2005, **2**:513-520.
- Bayer RJ, Starr JR: Tribal interrelationships and phylogeny of the Asteraceae. *Ann Mo Bot Gard* 1998, **85**:242-256.
- Hebert PD, Cywinska A, Ball SL, de Waard JR: Biological identifications through DNA barcodes. *Proc Biol Sci* 2003, **270**:313-321.
- Evans KM, Wortley AH, Mann DG: An assessment of potential diatom "barcode" genes (*cox1*, *rbcl*, 18S and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). *Protist* 2007, **158**:349-364.
- Hogg ID, Hebert PD: Biological identification of springtails (Hexapoda: Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes. *Can J Zool* 2004, **82**:749-754.
- Song JY, Yao H, Li Y, Li XW, Lin YL, Liu C, Han JP, Xie CX, Chen SL: Authentication of the family Polygonoaceae in Chinese pharmacopoeia by DNA barcoding technique. *J Ethnopharmacol* 2009, **124**:434-439.
- Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH: Use of DNA barcodes to identify flowering plants. *Proc Natl Acad Sci USA* 2005, **102**:8369-8374.
- Kress WJ, Erickson DL: A two-locus global DNA barcode for land plants: the coding *rbcl* gene complements the non-coding *trnH-psbA* spacer region. *PLoS One* 2007, **2**:e508.
- Lahaye R, van der BM, Bogarin D, Warner J, Pupulin F, Gigot G: DNA barcoding the floras of biodiversity hotspots. *Proc Natl Acad Sci USA* 2008, **105**:2923-2928.
- Chen SL, Yao H, Han JP, Liu C, Song JY, Shi LC, Zhu YJ, Ma XY, Gao T, Pang XH, Luo K, Li Y, Li XW, Jia XC, Lin YL, Leon C: Validation of the ITS2 Region as a novel DNA barcode for identifying medicinal plant species. *PLoS One* 2010, **5**:e8613.
- Chase MW, Cowan RS, Hollingsworth PM, van den Berg C, Madrinan S, Petersen G, Seberg O, Jorgensen T, Cameron KM, Carine M, Pedersen N, Hedderson TAJ, Conrad F, Salazar GA, Richardson JE, Hollingsworth ML, Barraclough TG, Kelly L, Wilkinson M: A proposal for a standardized protocol to barcode all land plants. *Taxon* 2007, **56**:295-299.
- Gonzalez MA, Baraloto C, Engel J, Mori SA, Pétronelli P, Riéra B, Roger A, Thébaud C, Chave J: Identification of amazonian trees with DNA barcodes. *PLoS One* 2009, **4**:e7483.
- Yao H, Song JY, Ma XY, Liu C, Li Y, Xu HX, Han JP, Duan LS, Chen SL: Identification of *Dendrobium* species by a candidate DNA barcode sequence: the chloroplast *psbA-trnH* intergenic region. *Planta Med* 2009, **75**:667-669.
- Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG, Husband BC, Percy DM, Hajibabaei M, Barrett SCH: Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS One* 2008, **3**:e2802.
- Kress WJ, Erickson DL, Jones FA, Swenson NG, Perez R, Sanjurjo O, Bermingham E: Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. *Proc Natl Acad Sci USA* 2009, **106**:18621-18626.
- CBOL Plant Working Group: A DNA barcode for land plants. *Proc Natl Acad Sci USA* 2009, **106**:12794-12797.
- Meier R, Zhang GY, Ali F: The use of mean instead of smallest inter-specific distances exaggerates the size of the "barcoding gap" and leads to misidentification. *Syst Biol* 2008, **57**:809-813.
- Meyer CP, Paulay G: DNA barcoding: error rates based on comprehensive sampling. *PLoS Biol* 2005, **3**:2229-2238.
- Pang XH, Song JY, Zhu YJ, Xu HX, Huang LF, Chen SL: Applying plant DNA barcodes for Rosaceae species identification. *Cladistics* 2010, **26**.
- Ross HA, Murugan S, Li WLS: Testing the reliability of genetic methods of species identification via simulation. *Syst Biol* 2008, **57**:216-230.
- Thomas C: Plant barcode soon to become reality. *Science* 2009, **325**:526.
- Coleman AW: ITS2 is a double-edged tool for eukaryote evolutionary comparisons. *Trends Genet* 2003, **19**:370-375.
- Coleman AW: Pan-eukaryote ITS2 homologies revealed by RNA secondary structure. *Nucleic Acids Res* 2007, **35**:3322-3329.
- Chiou SJ, Yen JH, Fang CL, Chen HL, Lin TY: Authentication of medicinal herbs using PCR-amplified ITS2 with specific primers. *Planta Med* 2007, **73**:1421-1426.
- Coleman AW: Is there a molecular key to the level of "biological species" in eukaryotes? A DNA guide. *Mol Biol Evol* 2009, **50**:197-203.
- Linder CR, Goertzen LR, Heuvel BV, Francisco-Ortegac J, Jansen RK: The complete external transcribed spacer of 18S-26S rDNA: amplification and phylogenetic utility at low taxonomic levels in Asteraceae and closely allied families. *Mol Phylogenet Evol* 2000, **14**:285-303.
- Englund M, Pornpongrueng P, Gustafsson MHG, Anderberg AA: Phylogenetic relationships and generic delimitation in Inuleae (Asteraceae) based on ITS and cpDNA sequence data. *Cladistics* 2009, **25**:319-352.
- Pelser PB, Nordenstam B, Kadereit JW, Watson LE: An ITS phylogeny of tribe Senecioneae (Asteraceae) and a new delimitation of *Senecio* L. *Taxon* 2007, **56**:1062-1077.
- Keller A, Schleicher T, Schultz J, Müller T, Dandekar T, Wolf M: 5.8S-28S rRNA interaction and HMM-based ITS2 annotation. *Gene* 2009, **430**:50-57.
- Chase MW, Salamin N, Wilkinson M, Dunwell JM, Kesanakurthi RP, Haidar N, Savolainen V: Land plants and DNA barcodes: short-term and long-term goals. *Phil Trans R Soc B* 2005, **360**:1889-1895.

doi:10.1186/1471-2148-10-324

Cite this article as: Gao et al.: Evaluating the feasibility of using candidate DNA barcodes in discriminating species of the large Asteraceae family. *BMC Evolutionary Biology* 2010 **10**:324.