Phytomedicine xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Phytomedicine

journal homepage: www.elsevier.com/locate/phymed

Evaluation of DNA barcoding coupled high resolution melting for discrimination of closely related species in phytopharmaceuticals

Q1 Maslin Osathanunkul[a,*], Chatmongkon Suwannapoom[b,c], Kitisak Osathanunkul[d], Panagiotis Madesis[e], Hugo de Boer[f,g]

[a] Department of Biology, Faculty of Science, Chiang Mai University, 239 Huay Kaew Rd., Suthep, Muang, Chiang Mai 50200, Thailand
[b] State Key Laboratory of Genetic Resources and Evolution State, and Yunnan Laboratory of Molecular Biology of Domestic Animals, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China
[c] School of Agriculture and Natural Resources, University of Phayao, Phayao 56000, Thailand
[d] Department of Computer Science, Faculty of Science, Maejo University, Chiang Mai 50290, Thailand
[e] Institute of Applied Biosciences, Centre for Research & Technology Hellas (CERTH), Thessaloniki, Greece
[f] Department of Organismal Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-75236 Uppsala, Sweden
[g] The Natural History Museum, University of Oslo, P.O. Box 1172, NO-0318 Oslo, Norway

ARTICLE INFO

Article history:
Received 6 August 2015
Revised 11 October 2015
Accepted 12 November 2015
Available online xxx

Keywords:
DNA barcoding
Herbal pharmacovigilance
High resolution melting
Medicinal plants
Authentication

ABSTRACT

Background: Phytopharmaceuticals are increasingly popular as alternative medicines, but poorly regulated in many countries. The manufacturers of these products should be subject to strict controls regarding each product's quality and constituents. Routine testing and identification of raw materials should be performed to ensure that the raw materials used in pharmaceutical products are suitable for their intended use.
Hypothesis/purpose: We have applied DNA Barcoding – High Resolution Melting (Bar-HRM), an emerging method for identifying of medicinal plant species based on DNA dissociation kinetics and DNA barcoding, for the authentication of medicinal plant species.
Study design: Commonly commercialized Thai medicinal plants that are widely used for medicinal purposes were used in this study. Publicly available sequences of four plastid markers were used for universal primer design. Species discrimination efficiency of the designed primers was evaluated as single and multi-locus analyses by using the primers sets.
Methods: HRM analysis was performed in triplicate on each of the 26 taxa to establish the T_m for each primer set (*mat*K, *rbc*LA, *rbc*LB, *rbc*LC, *rpo*C1, and *trn*L). The shapes of the melting curves were analyzed to distinguish the different plant species. Bar-HRM species identification success rates were assessed for each single-locus as well as for multi-locus combinations to establish the optimal combination of primer sets.
Results: In single locus analysis the *rpo*C1 primer set gave the highest discrimination (58%), and in multi locus analysis this could be increased from 87% to 99% depending on the total number of regions included. Different combinations proved to be more or less effective at discrimination, depending on the genus or family examined.
Conclusions: Bar-HRM has proven to be a cost-effective and reliable method for the identification of species in this study of Thai medicinal plants, and results show an identification success rate of 99% among species in the test set.

Introduction

Traditional medicine has played an important role in Asian culture for centuries and medicinal plants remain important today (cf. Inta et al. 2008; Sivasankari et al. 2014). A combination of the increasing scientific evidence of the safety and efficacy of traditional herbal remedies (de Boer and Cotingting 2014) along with a shift towards popular skepticism of synthetic pharmaceuticals has

Abbreviations: HRM, high resolution melting; T_m, melting temperature; *mat*K, maturase K; *rbc*L, ribulose-bisphosphate carboxylase; *rpo*C1, RNA polymerase Cl.

* Corresponding author. Tel.: +66 53 943348; fax: +66 53 892259.
E-mail address: omaslin@gmail.com (M. Osathanunkul).

http://dx.doi.org/10.1016/j.phymed.2015.11.018
0944-7113/

enabled steady growth in the global herbal medicine market. A multitude of diverse medicinal plant products are commercialized and are readily available from pharmacies and online retailers. Commercial herbal products are primarily sold in processed or modified forms such as dried herbs, powders, tablets, capsules, and teas, which make it challenging to accurately identify the constituent plant species.

In parallel with increasing popularity and accessibility of medicinal products there is a need for an adequate regulatory mechanism of the market. Furthermore, the manufacturers of these products should be subject to strict controls regarding each product's quality and ingredients. Routine testing or identifying of raw materials should be performed to ensure that the raw materials used in pharmaceutical products are suitable for their intended use. The identification of ingredients in each product can be challenging. This is because many medicinal plants have similar macro-structural morphology among species within the same genus, whereas others are under-differentiated using vernacular names (i.e. where the same name is applied to multiple species within the same genus) (Berlin 1992). Relying solely on morphological characters or vernacular names can lead to confusion in species identification (de Boer et al. 2014), and subsequent substitution, either accidental or intentional, during the manufacturing process. Moreover, commercialized medicinal plants are often sold in processed forms in local markets and on the internet as capsules, tablets, or dried parts, which further complicates or makes impossible species phenotypic identification and adds complexity to the process of regulation and enforcement of product safety (Veldman et al. 2014).

Traditionally, quality control and standardization of raw and processed medicinal plants were authenticated by its physical features, for instance, shape, color, taste, texture and scent – all based on human sensory organs which is measurably unreliable. Owing to this, the modern herbal drug production employs a number of reliable techniques as the following. Morphological examination using voucher specimens as references is practiced in macro and microscopic levels (Serrano et al. 2010; Zhao et al. 2011). It is fast and cost effective but requires taxonomical expertise and may be impractical for structurally damaged samples thus it is impossible in doing so with finely ground plant material samples (Franz et al. 2007; Hebert et al. 2003a). Many chemical analyses are also used, e.g. thin layer chromatography (TLC), high-performance liquid chromatography (HPLC), gas chromatography (GC), infrared spectroscopy (IR), or nuclear magnetic resonance spectroscopy (NMR) (Chan 2003; Mukherjee et al. 2010; Siow et al. 2005). However variations in chemical constituents affected by environmental and genetic factors which may still be problematic in the mentioned techniques and thus make it difficult for plant species identification (Kamboj 2012).

The development and application of reliable methods for species identification and authentication of medicinal plants and their derived products is critical for the enforcement of good manufacturing practice and to avoid safety and efficacy issues, such as adverse herbal drug reactions. Morphology independent methods such as DNA barcoding are suitable for the identification of commercialized medicinal plants and raw materials for industrial manufacture. DNA barcoding has a high discrimination resolution and is applicable for a wide range of substrates, including these processed materials.

Molecular identification through DNA barcoding is a powerful method for the identification of plant species, including medicinal plants and products (de Boer et al. 2015). A considerable amount of literature has been published on this subject, and these studies show the potential for DNA barcoding to effectively distinguish among medicinal plants, as well as to identify component species found in herbal medicines (Chen et al. 2010; Coghlan et al. 2012; Kool et al. 2012; Newmaster et al. 2013). Several reviews have highlighted the increasing and diverse applications of medicinal plant barcoding (Li et al. 2011; Techen et al. 2014). However, DNA barcoding in plants does have limitations, including the inability to amplify marker regions due to degraded DNA in processed samples (Särkinen et al. 2012), limited binding site universality (Sass et al. 2007; Piredda et al. 2011; Kool et al. 2012), low rates of discrimination capabilities with certain markers (Kool et al. 2012; Stoeckle et al. 2011), overlapping intraspecific and intraspecific genetic variation in some groups of plants (Fazekas et al. 2009), and low applicability of chloroplast markers for identification of species of hybrid origin (Fazekas et al. 2009). Another limitation of DNA barcoding is its cost, as it requires a molecular laboratory, costly equipment, chemicals and disposables, and DNA sequencing facilities. It is especially the frequent lack of access to DNA sequencing facilities that hinders the wider implementation of DNA barcoding in developing countries. Developing and validating sequencing-free methods that are reliable, but faster and more economical than DNA barcoding is challenging, but will be beneficial for the advancement of herbal product identification routines in developing countries.

High resolution melting (HRM) is an emerging method for monitoring DNA dissociation ("melting") kinetics, and is a powerful technique for the detection of point mutations, indels, and methylated DNA (Reja et al. 2010). In addition to standard PCR equipment and reagents, HRM requires a generic DNA intercalation fluorescent dye. The dye is added to previously amplified PCR products and as the double-stranded DNA samples dissociate with increasing temperature the dye is progressively released and fluorescence diminishes. The denaturation thermodynamics of individual double-stranded DNA to single strands are based on the binding affinities of individual nucleotide pairs, and melting pattern will vary due to indels, mutations and methylations. These differences are inferred in terms of varying melting temperatures (T_m). Fluorescent measurements are collected at standard temperature increments and plotted as a "melting curve". The curve's shape and peak are characteristic for each sample, allowing for comparison and discrimination among samples. Through this method, even a single base change between samples can be readily detected and identified (Ririe et al. 1997; Wittwer et al. 2003).

In this study, we have applied DNA Barcoding – High Resolution Melting (Bar-HRM) analysis for the authentication of medicinal plant species. Four plastid markers, *mat*K, *rbc*L, *rpo*C1 and *trn*L were used for universal primer design. Here, we have used tailored reduced amplicons from these barcode regions in combination with HRM analysis to identify and authenticate commonly used Thai medicinal plants. To optimize cost-efficiency, data from GenBank (NCBI) was mined to build a DNA reference library in order to create suitable primers for the HRM analysis and to predict HRM melting curves. Previous studies have only reported on the use of Bar-HRM to evaluate herbal medicine substitution among defined groups of plant species, such as three medicinal species of Acanthaceae (Osathanunkul et al. 2015a), *Thunbergia* species (Singtonat and Osathanunkul 2015) and 12 *Croton* species (Osathanunkul et al. 2015b). The main aim of this study is to validate the universality of developed Bar-HRM primers for species identification of species from diverse groups.

Materials and methods

DNA mining of barcode regions for primer design

Ninety-six common Thai medicinal plants were selected as a basis for primer design, and sequences for selected plastid regions

Please cite this article as: M. Osathanunkul et al., Evaluation of DNA barcoding coupled high resolution melting for discrimination of closely related species in phytopharmaceuticals, Phytomedicine (2016), http://dx.doi.org/10.1016/j.phymed.2015.11.018

Table 1
Plants with voucher number included in this study are all commonly found in Thai markets.

Family	Botanical name[a]	Voucher number
Acanthaceae	*Acanthus ebracteatus* Vahl	QBG-MO2100
Acanthaceae	*Andrographis paniculata* (Burm.f.) Nees	QBG-MO2101
Acanthaceae	*Barleria lupulina* Lindl.	QBG-MO2102
Acanthaceae	*Clinacanthus nutans* (Burm.f.) Lindau	QBG-MO2103
Acanthaceae	*Rhinacanthus nasutus* (L.) Kurz	QBG-MO2104
Acanthaceae	*Thunbergia laurifolia* Lindl.	QBG-MO2105
Araliaceae	*Schefflera leucantha* R.Vig.	QBG-MO2106
Compositae	*Pluchea indica* (L.) Less.	QBG-MO2107
Compositae	*Cyanthillium cinereum* (L.) H.Rob.	QBG-MO2108
Cucurbitaceae	*Momordica charantia* L.	QBG-MO2109
Euphorbiaceace	*Phyllanthus amarus* Schumach. & Thonn.	QBG-MO2110
Lamiaceae (Labiatae)	*Orthosiphon aristatus* (Blume) Miq.	QBG-MO2133
Leeaceae	*Leea macrophylla* Roxb. ex Hornem.	QBG-MO2134
Leguminosae	*Senna siamea* (Lam.) H.S.Irwin & Barneby	QBG-MO2111
Leguminosae	*Senna alata* (L.) Roxb.	QBG-MO2112
Menispermaceae	*Tinospora crispa* (L.) Hook.f. & Thomson	QBG-MO2113
Menispermacea	*Tinospora sinensis* (Lour.) Merr.	QBG-MO2114
Moringaceae	*Moringa oleifera* Lam.	QBG-MO2115
Papilioneae	*Derris scandens* (Roxb.) Benth.	QBG-MO2116
Piperaceae	*Piper sarmentosum* Roxb.	QBG-MO2117
Vitaceae	*Cissus quadrangularis* L.	QBG-MO2118
Zingiberaceae	*Amomum verum* Blackw.	QBG-MO2119
Zingiberaceae	*Boesenbergia rotunda* (L.) Mansf.	QBG-MO2120
Zingiberaceae	*Curcuma longa* L.	QBG-MO2121
Zingiberaceae	*Curcuma zedoaria* (Christm.) Roscoe.	QBG-MO2122
Zingiberaceae	*Zingiber montanum* (J.Koenig) Link ex A. Dietr.	QBG-MO2123
Zingiberaceae	*Zingiber officinale* Roscoe	QBG-MO2124
Zingiberaceae	*Zingiber ottensii* Valeton	QBG-MO2125
Zingiberaceae	*Zingiber zerumbet* (L.) Roscoe ex Sm.	QBG-MO2126
Poaceae (Gramineae)	*Cymbopogon citratus* Stapf.	QBG-MO2127
Poaceae (Gramineae)	*Cymbopogon nardus* Rendle	QBG-MO2128
Apiaceae (Umbelliferae)	*Centella asiatica* (L.) Urban.	QBG-MO2129

[a] Plant nomenclature following The Plant List (www.theplantlist.org).

(*mat*K, *rbc*L, *rpo*C1 and *trn*L) were retrieved from the GenBank (NCBI, http://www.ncbi.nlm.nih.gov/) for each of the species (Supplementary Data 1). Low quality sequences and accessions without specified herbarium vouchers were discarded. Primers for high resolution melting analysis were designed using multiple sequence alignments using SeqMan Pro (DNASTAR, Inc., Madison, WI, USA), ClustalX2 (Larkin et al. 2007), and MEGA5 (Tamura et al. 2011). Amplicons were designed to include variable characters, lengths, and GC content. Primers for HRM were developed based on the following criteria: (i) primer pairs should yield an amplicon of <300 bp; and (ii) primer pairs should yield amplicons with a sufficient number of variable sites to enable species discrimination.

Plant materials and DNA isolation

The plant species used in this study are commonly commercialized Thai medicinal plants and all are widely used for medicinal purposes (Table 1). Plant material was collected from the Materia Medica garden at the Faculty of Pharmacy, Chiang Mai. Plant samples were ground with liquid nitrogen, and 100 mg of the material were used for DNA extraction using a Nucleospin Plant® II (Macherey-Nagel, Germany) kit following the manufacturer's instructions. The DNA concentration was estimated by standard spectrophotometric methods at 260 nm and 280 nm UV lengths by an Eppendorf BioPhotometer and the integrity by gel electrophoresis in a 0.8% agarose gel. Samples were then diluted to 25 ng/μl work concentration and stored at −20 °C for further use.

Table 2
Oligonucleotides of designed primers used for HRM analysis and identification.

Primer HRM	5′→3′	T_m (°C)	Expected size (bp)
HRM_rpoC11F	CCSATTGTATGGGAAATACTT	57	170
HRM_rpoC11R	CTTACAAACTAATGGATGTAA		
HRM_matK1F	CTTCTTATTTACGATTAACATCTTCT	57	170
HRM_matK1R	TTTCTTTGATATCGAACATAATG		
HRM_trnL1F	TGGGCAATCCTGAGCCAAATC	57	120
HRM_trnL1R	AACAGCTTCCATTGAGTCTCTGCACCT		
HRM_rbcLAF	GCAGCATTCCGAGTAACTCCTCA	57	100
HRM_rbcLAR	TCCACACAGTTGTCCATGTACC		
HRM_rbcLBF	GGTACATGGACAACTGTGTGGA	57	150
HRM_rbcLBR	ACAGAACCTTCTTCAAAAAGGTCTA		
HRM_rbcLCF	TAGACCTTTTTGAAGAAGGTTCTGT	57	150
HRM_rbcLCR	TGAGGCGGRCCTTGGAAAGTT		

High resolution melting (HRM) method

DNA amplification using real-time PCR was performed using the Eco™ Real-Time PCR system (Illumina®, San Diego, USA) in order to establish characteristic melting temperatures (T_m) to enable the distinction of the different medicinal plants. The reaction mixture for the real-time PCR and HRM analysis consisted of a total volume of 10 µl. This contained 5 µl of 2× THUNDERBIRD® SYBR qPCR Mix, 0.2 µM forward primer, 0.2 µM reverse primer, and 1 µl of 25 ng/µl DNA. The primer pairs used for each region are listed in Table 2. SYBR fluorescence dye was used to monitor the accumulation of the amplified product during PCR and the HRM process to derive the T_m value.

The real-time PCR amplification was conducted in a 48-well Helixis plate using an initial denaturing step at 95 °C for 5 min followed by 35 cycles of 95 °C for 30 s, 57 °C for 30 s, and 72 °C for 20 s. Fluorescence data was acquired at the end of each extension step during the PCR cycles. Subsequently, the PCR amplicons were denatured for HRM at 95 °C for 15 s, and then annealed at 50 °C for 15 s to form random DNA duplexes. The RT-PCR HRM protocol collected fluorescence data at 0.1 °C temperature increments. EcoStudy Software v 5.0 was used to plot a normalized curve of decreasing fluorescence with increasing temperature. The negative derivative of fluorescence (F) over the temperature (T) (dF/dT) curve gives the T_m. To generate normalized melting curves and difference melting curves (Wittwer et al. 2003), pre- and post-melt normalization regions were set to define the main temperature boundaries of the normalized and difference plots with *Phyllanthus amarus* Schumach. & Thonn. (Phyllanthaceae) set as the reference species.

In silico *measurement of taxonomic coverage of the designed primer pairs*

The six primer pairs derived from chloroplast regions (*matK*, *rbcLA*, *rbcLB*, *rbcLC*, *rpoC* and *trnL*) were evaluated for their taxonomic coverage. A reference database of representative chloroplast genomes was created from retrieved GenBank data. The taxonomic coverage and performance of each designed primer pair was analyzed through an *in silico* PCR on the reference genomes (cf. Ficetola et al. 2010).

Statistical analysis

Discriminating power of designed primers

The HRM T_m values derived for each amplicon were analyzed using discriminant model analysis classification as implemented in

Table 3
Characteristics of sequences and derived designed primers for high resolution melting analysis.

Markers	rpoC1	trnL	rbcLA	rbcLB	rbcLC	matK
Available species/total (%)	1/96 (1.04)	28/96 (29.17)	53/96 (55.21)	47/96 (48.95)		
Average total product, excluding primer (SD)	150, 108 (0)	97, 48 (8.93)	100, 55 (0)	149, 102 (0)	145, 99 (0)	158, 109 (4.3)
Characters (bp)	108	77	55	102	99	129
Variable characters (%)	58 (53.70)	53 (68.83)	19(34.54)	43(42.16)	38(38.38)	99(76.74)
Average distance (SE)	0.1145 (0.0175)	0.0358 (0.0194)	0.1137 (0.0316)	0.1250 (0.0222)	0.1179 (0.0229)	0.2968 (0.0376)
Conserved forward primer/total (%)	12/21 (57.14)	20/21 (95.24)	20/23 (83.96)	18/22 (81.82)	22/25 (88)	4/26 (15.38)
Conserved reverse primer/total (%)	16/21 (76.19)	25/27 (92.59)	18/22 (81.82)	22/25 (88)	16/21 (76.19)	10/23 (43.48)
Average %GC content (SD)	46.42 (1.93)	30.99 (6.81)	59.61 (2.85)	44.42 (2.04)	41.85 (1.46)	35.02 (3.02)

SPSS version 14 (Coakes and Steed 2009), to evaluate the species discrimination efficiency of the designed HRM primer pairs. Discriminant analysis is a predictive model of group membership based on observed characteristics of each case (Ye et al. 2004). The procedure generates a discriminant function (or, for more than two groups, a set of discriminant functions) based on linear combinations of the predictor variables that provide the best discrimination between the groups. In this study, Leave-one-out cross-validation (LOOCV) was used, and the discrimination percentage was calculated based on the number of cases that were correctly predicted by the discriminant function. The combinations were taken in combinations ranging from two to six loci (15 sets for two loci combinations, 20 sets for three loci combinations, 15 sets for four loci combinations, six sets for five loci combinations, and one set for six loci combinations) in order to determine which combination(s) discriminate among the species most effectively. The scatter plots of these combinations were created using the first two discriminant functions.

Results and discussion

DNA mining and primer design

GenBank accessions were mined to assemble DNA barcodes of medicinal plants that are difficult to identify either as fresh materials or in their processed forms. Data was present for most markers of the 96 target species, except for *rpo*C1. The total number of sequences retrieved for the respective markers were: *rbc*L 79 of 96 species (82.3%); *mat*K 78.1%, *trn*L 76.0%, and *rpo*C1 1.0%. The absence of *rpo*C1 sequences for the target species was resolved by selecting 60 random *rpo*C1 sequences from GenBank for primer design, which is supported by the high universality of *rpo*C1 (Kool et al. 2012; Kress et al. 2005). The sequence data for the 96 medicinal plant species were extracted and aligned in order to create six universal primer sets for identification through high resolution melting (HRM) analysis.

A single set of primers was designed for each of *mat*K, *trn*L, and *rpo*C1, and three sets of primers were designed for *rbc*L, which yielded amplicons ranging from 100 to 150 bp. Reed and Wittwer (2004) found that amplicons suitable for HRM analysis should be 300 bp or less for optimal results. The *trn*L and *mat*K primer sets yielded amplicons of variable length with high standard deviation (SD) values, whereas the *rbc*L and *rpo*C1 primer sets yielded amplicons of consistent size (Table 3).

Both the sequence length and the nucleotide variation within sequences influence the dissociation energy of the base pairs and result in different T_m values. The *mat*K and *trn*L amplicon sequences were observed to have higher nucleotide variation than the amplicons of the other regions, at 76.7% and 68.8%, respectively. The relative nucleotide variation and thus the discrimination power within amplicons were found to be as follows: *mat*K > *trn*L > *rpo*C1 > *rbc*LB > *rbc*LC > *rbc*LA (Table 3). The forward and reverse *mat*K primers matched the consensus sequence of the target species at the binding sites in only 4 out of 24 sites (15.38%) and 10 out of 23 sites (43.48%), respectively (Table 3). High universality at the initial bases of the primer site is crucial for primer annealing and subsequent elongation initiation by the DNA polymerase.

The *mat*K locus is one of the most variable plastid coding regions and has high interspecific divergence and good discriminatory power. However, it can be difficult to amplify with the standard barcoding primers due to high substitution rates at the primer sites (CBOL Plant Working Group 2009; Hollingsworth 2011). The *trn*L primer pair designed in this study was expected to be a suitable primer for HRM analysis for discrimination between the tested plant species. These *trn*L primers were nearly identical in base similarity to the mined consensus sequence, and the primer binding sites were designed in the conserved Q and R regions in the secondary structure of tRNA leucine (UAA) to optimize amplification success (Taberlet et al. 2007). The *trn*L amplicon included the variable P6 loop in order to increase differentiation among the tested species during HRM. High primer binding success rates, short amplicon length, and high sequence variation make *trn*L an ideal marker for HRM based species discrimination in plants. The combination of these three characteristics has also made *trn*L the marker of choice for ancient DNA and ancient sedimentary DNA metabarcoding studies (Boessenkool et al. 2014; Jørgensen et al. 2012; Parducci et al. 2012; Taberlet et al. 2012; Willerslev et al. 2014). Recently, the application of *trn*L barcoding coupled with HRM has successfully detected adulteration in agricultural products (Madesis et al. 2012), has been used in forensics (Madesis et al. 2013), and has been used for taxonomic identification as well as hybrid tests (Ganopoulos et al. 2013).

The average %GC content of amplicons was calculated in order to predict variation in melting curves for the different markers. *trn*L had the lowest average %GC content, with 30.99%, followed by *mat*K, *rbc*LC, *rbc*LB, *rpo*C1 and *rbc*LA, with 35.02%, 41.85%, 44.42%, 46.42% and 59.61% respectively (Table 3). The standard deviations (SD) values of the %GC content in each amplicon from the different primers reflect the wide sequence variation among the target species.

In silico *amplicon identification rates*

An *in silico* PCR for each primer pair was performed to test sequence variation and putative HRM species discrimination rates (Ficetola et al. 2010). The resolution of each HRM barcode at different taxonomic ranks was analyzed (Fig. 1). The input sequence file contained 231 sequence records representing the same number of species. The *mat*K primer amplicon could unambiguously identify only 43 of the 231 taxa from the taxonomic database. This low discrimination rate, in combination with the poor primer fit, indicate that *matK* is unsuitable for HRM analysis. Conversely, *trn*L gave the best rate of species level discrimination using HRM curves for identification (115 out of 231). This indicates that *trn*L is a good candidate for HRM due to: (i) high SD within amplicons lengths;

Please cite this article as: M. Osathanunkul et al., Evaluation of DNA barcoding coupled high resolution melting for discrimination of closely related species in phytopharmaceuticals, Phytomedicine (2016), http://dx.doi.org/10.1016/j.phymed.2015.11.018

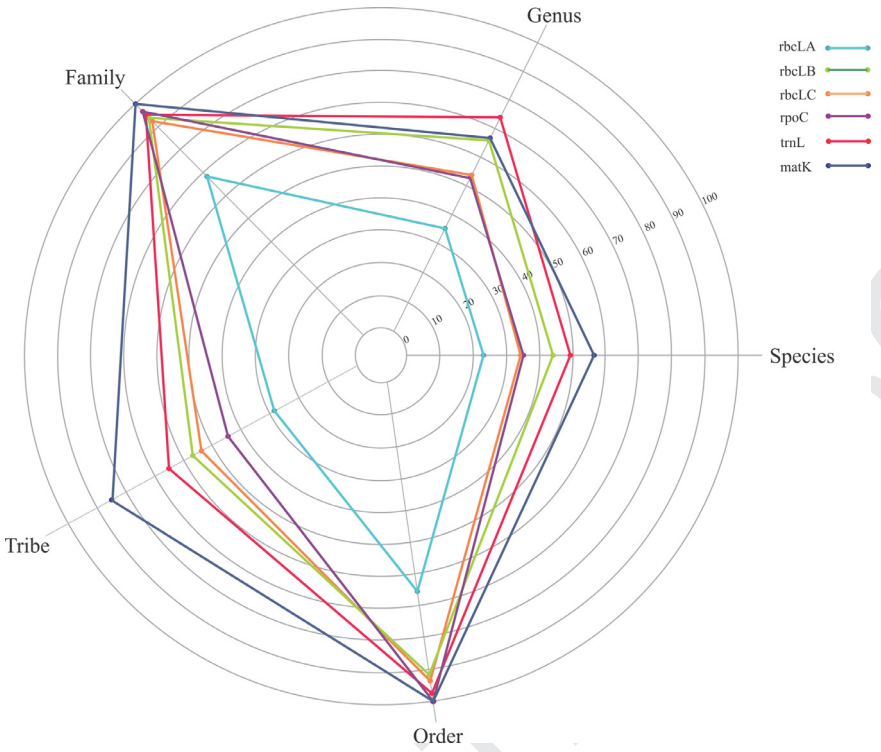

Fig. 1. Results from *in silico* analysis of each primer pair. The resolution in percentage of each HRM barcode on the chloroplast genome at different five taxonomic ranks including order, tribe, family, genus and species are shown. Testing is based on sequence records retrieved from GenBank.

(ii) high rate of character variation; (iii) high conservation in the primer sites in the consensus sequence; and (iv) high SD of the average %GC content.

Evaluation of primers and amplicons for medicinal plant species discrimination by Bar-HRM

The six primers pairs designed to amplify sections of four chloroplast markers were tested with 26 medicinal plant species from 14 families using HRM analysis (Table 2). These six primer sets amplified products from *mat*K, *rbc*LA, *rbc*LB, *rbc*LC, *rpo*C1, and *trn*L and yielded amplicons of 170 bp, 100 bp, 150 bp, 150 bp, 170 bp, and 120 bp, respectively. HRM analysis was performed in triplicate on each of the 26 taxa to establish the T_m for each primer set. The shapes of the melting curves were analyzed using EcoStudy Software v 5.0 to distinguish the different plant species. Bar-HRM species identification success rates were assessed for each single-locus as well as for multi-locus combinations to establish the optimal combination of primer sets.

Single-locus analyses

The HRM primers sets were used for the amplification of DNA-fragments from all 26 Thai medicinal plant species, and the resulting amplicons were analyzed using HRM to define T_m. The melting profiles of all amplicons are illustrated in Figs. 2A–F and 3A–F. The analysis is presented by means of conventional derivative plots, which show that the T_m value of each species is represented by a peak. The melting temperature peaks of all the 26 medicinal plant species are calculated as T_m and presented in Supplementary Data 2.

The mean of the melting temperatures obtained from each primer pair was used to measure species discrimination for each locus. These species discriminations for each locus ranged from 15% (*mat*K) to 58% (*rpo*C1), with *trn*L, *rbc*LA, *rbc*LB, and *rbc*LC, providing 40%, 32%, 32%, and 28% discriminatory power, respectively (Fig. 4). Although *mat*K has been proposed as one of the best plant barcodes in terms of species discrimination (Fazekas et al. 2008; Hollingsworth et al. 2009) we found that the section of *mat*K amplified by our novel primer set for HRM has a low success rate in PCR amplification and also a low species discrimination rate among the taxa that were successfully amplified. Conversely, the high species discrimination for *rpo*C1 (58%), contrasts with previous barcoding studies that indicate that this marker has relatively low discriminatory power (Fazekas et al. 2008; Hollingsworth et al. 2009; Kool et al. 2012). Not all amplicons from the different species yielded distinctive HRM profiles (Fig. 2A–F), but all could be discriminated at the family level, with the exception of *mat*K. Furthermore, each primer set gave different rates of species discrimination in different families. For example, none of the primer sets could clearly distinguish among species in Zingiberaceae. However, the *rbc*LB primer pair enabled division of Zingiberaceae into two groups: (i) *Curcuma longa* L. and *Curcuma zedoaria* (Christm.) Roscoe and (ii) *Zingiber montanum* (J.Koenig) Link ex A.Dietr., *Zingiber officinale* Roscoe and *Boesenbergia rotunda* (L.) Mansf. (Fig. 5). Zingiberaceae is known to be a difficult group to identify using molecular methods, and recent studies have shown that the nuclear ribosomal marker ITS2 is the best single marker for species identification (Chen et al. 2014; Shi et al. 2011). Other families were easier to identify, such as the family Acanthaceae, which in this dataset had a high rate of species discrimination for two of the loci, *rpo*C1 (66.7%) and *trn*L (50%) (Fig. 4).

Results from the present study are similar to those from previous studies suggesting that each single marker has limited resolution and that combinations of two or more markers are needed for plant species identification (Hollingsworth 2011). Here we find that no single marker identifies more than 58% of all species (*rpo*C1). Nevertheless, it could be that other single markers suggested by the results of other studies, such as ITS1, ITS2, *trn*H-*psb*A, could

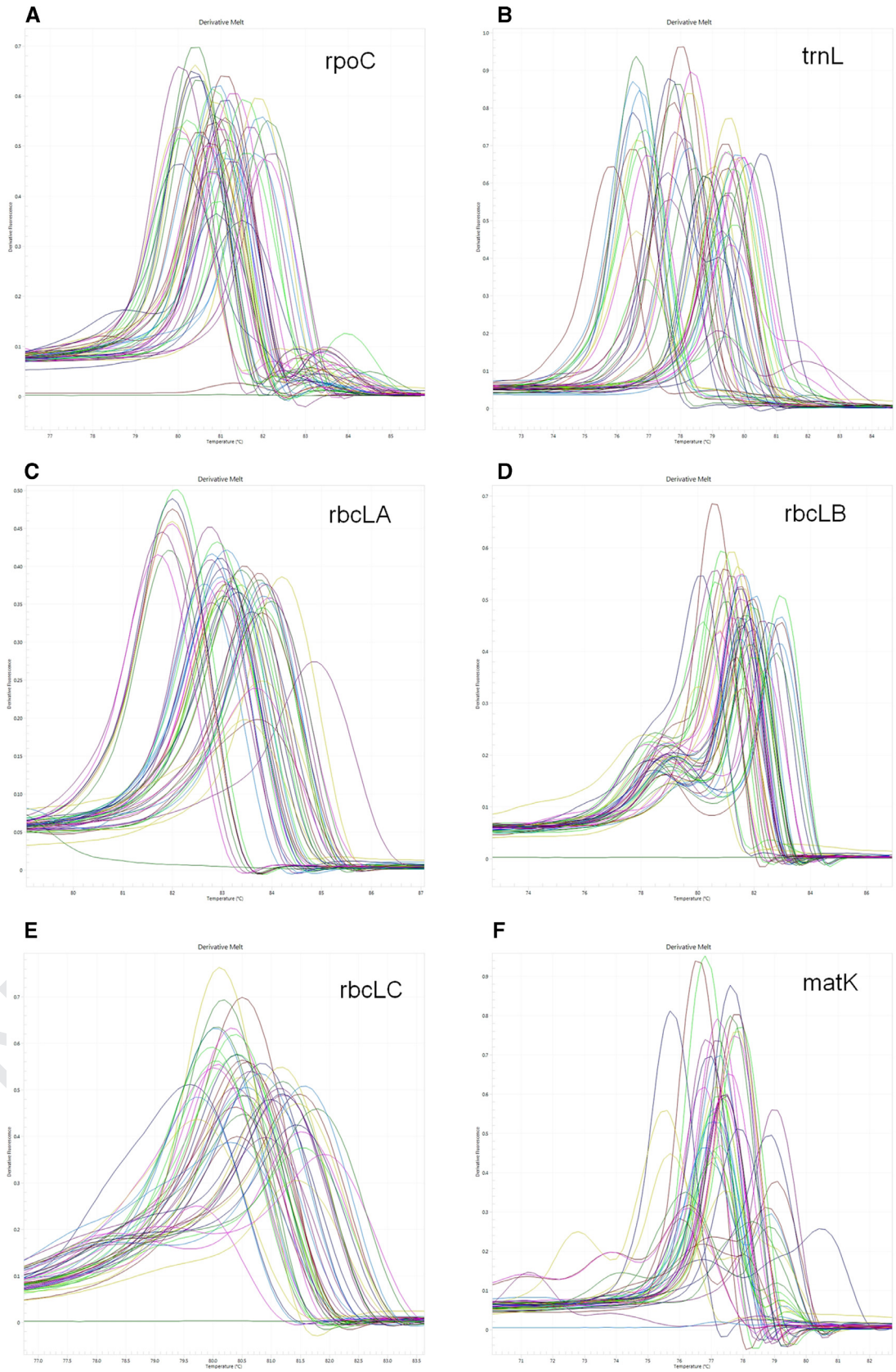

Fig. 2. Melting curve profiles of amplicons obtained from each primer set: *rpoC*1 (A), *trnL* (B), *rbcLA* (C), *rbcLB* (D), *rbcLC* (E) and *matK* (F).

Please cite this article as: M. Osathanunkul et al., Evaluation of DNA barcoding coupled high resolution melting for discrimination of closely related species in phytopharmaceuticals, Phytomedicine (2016), http://dx.doi.org/10.1016/j.phymed.2015.11.018

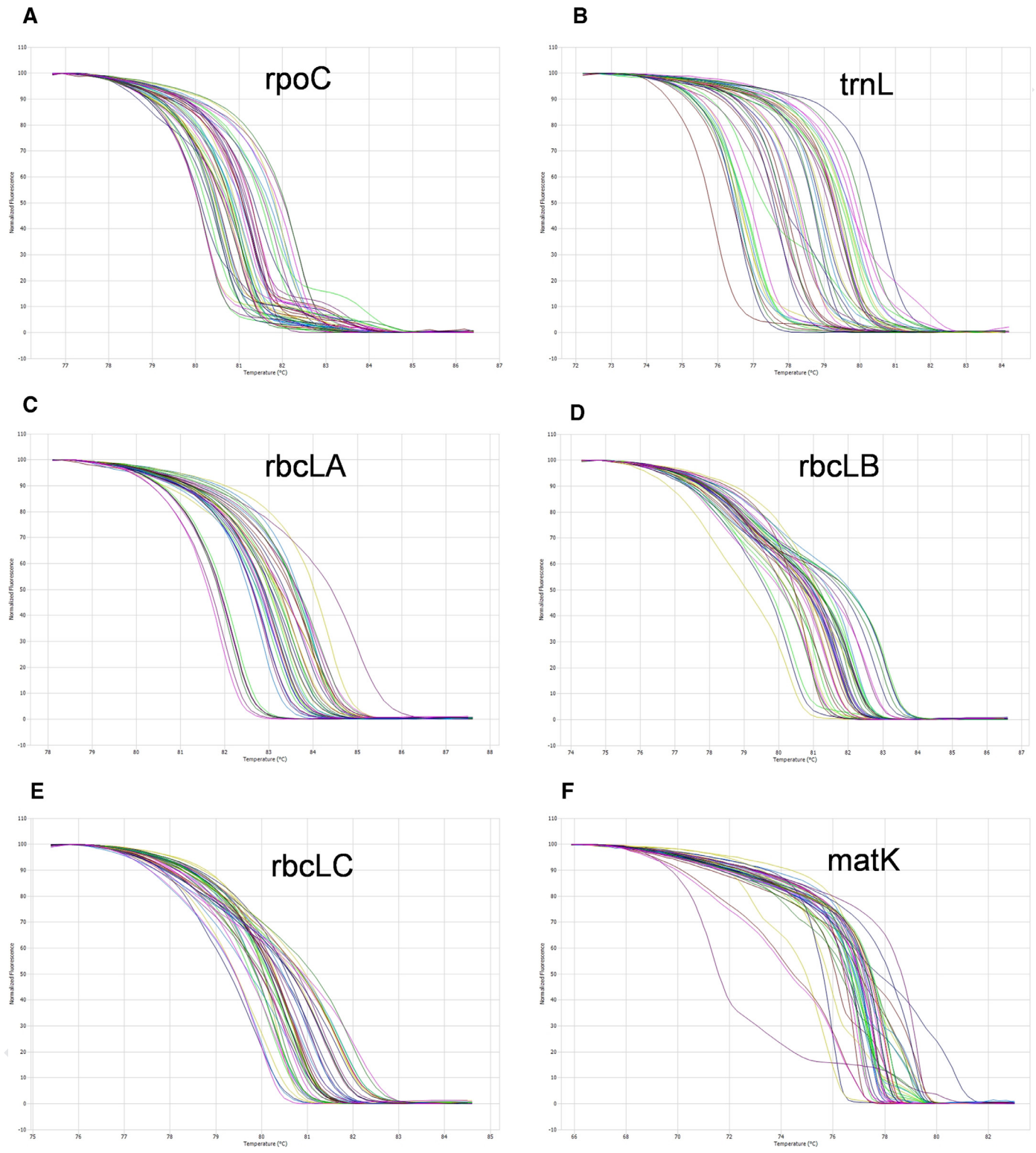

Fig. 3. The normalized plot of each primer pair: *rpoC*1 (A), *trnL* (B), *rbcLA* (C), *rbcLB* (D), *rbcLC* (E) and *matK* (F) show the differentiation of melting temperature (T_m) of each amplicon from each species, generated by high resolution melting (HRM) analysis.

be used to produce HRM primer sets with even higher resolution (Chen et al. 2010; de Boer et al. 2014; Fazekas et al. 2008; Gao et al. 2010; Kool et al. 2012; Kress et al. 2005; Li et al. 2011).

Multi-locus analyses

Simultaneously employing multiple standard loci to identify taxa, a method known as multi-locus DNA barcoding, reduces the overall negative effects of lacking amplicons due to variable primer-sites within taxa, while it increases discriminating power in plant identification. The use of multi-locus DNA barcoding has been recommended from the inception of barcoding as a method for identification in plants (CBOL Plant Working Group 2009; Fazekas et al. 2008; Kress and Erickson 2007; Kress et al. 2005). For example, Kress et al. (2005) recommended combining ITS and

Please cite this article as: M. Osathanunkul et al., Evaluation of DNA barcoding coupled high resolution melting for discrimination of closely related species in phytopharmaceuticals, Phytomedicine (2016), http://dx.doi.org/10.1016/j.phymed.2015.11.018

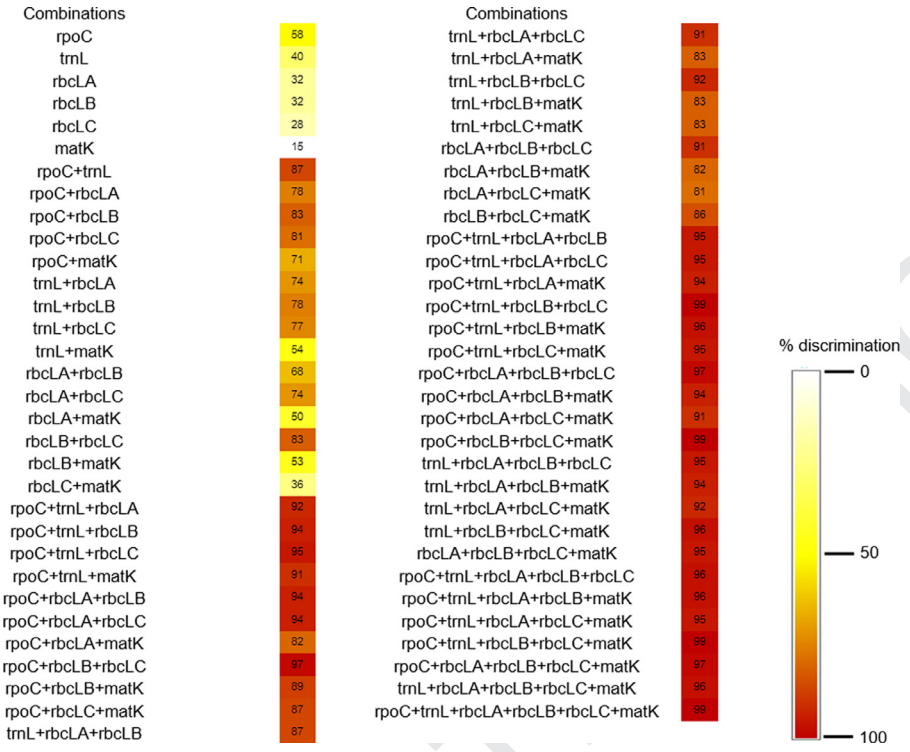

Fig. 4. Discriminatory power of six loci from four barcoding regions (*rbcL*, *trnL*, *rpoC*1 and *matK*) for species identification using high resolution melting (HRM) analysis and their combinations (2–6 combinations) expressed as percentage of accurately discriminated species.

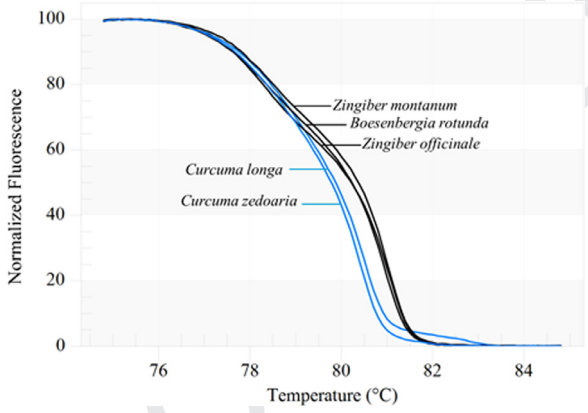

Fig. 5. HRM analyses using the designed *rbcL*B primers. Normalized curves of the five Zingiberaceae species. Two groups of species can be divided using the curves, the first group (blue lines) includes *Curcuma longa* and *Curcuma zedoaria*, while the other group (black lines) includes *Boesenbergia rotunda*, *Zingiber montanum* and *Zingiber officinale*). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Q5

*trn*H-*psb*A, Kress and Erickson (2007) recommended *trn*H-*psb*A and *rbc*L, Fazekas et al. (2008) evaluated eight different plastid markers, and the CBOL Plant Working Group (2009) evaluated seven plastid markers and recommended a combination of *rbc*L and *mat*K, possibly supplemented with ITS (Hollingsworth 2011). To find an optimal combination of HRM primer sets all combinations of the six studied loci were evaluated, in relation to the differences among the obtained T_m values. The discriminatory power of the primer sets was tested using discriminant analyses.

In single locus analysis the T_m profiles from the *rpo*C1 primer set gave the highest discrimination (58%), and in multi locus analysis this could be increased from 87% to 99% depending on the total number of regions included. When using the two primer pairs *rpo*C1 + *trn*L discrimination reached 87%, and with three pairs 95% and 97% discrimination could be made with *rpo*C1 + *trn*L + *rbc*LC and *rpoC1* + *rbc*LB + *rbc*LC, respectively. In four loci analyses species resolution ranged from 91% (*rpo*C1 + *rbc*LA + *rbc*LC + *mat*K) to 99% (*rpo*C1 + *trn*L + *rbc*LB + *rbc*LC and *rpo*C1 + *rbc*LB + *rbc*LC + *mat*K). The success of the species resolution reached an upper limit of 99% with four markers, and this was not further improved with five or six markers (Fig. 4). Due to the low universality of the forward *mat*K primer (HRM_matK1F), we recommend the use of the *rpo*C1 + *trn*L + *rbc*LB + *rbc*LC combination over the *rpo*C1 + *rbc*LB + *rbc*LC + *mat*K combination.

In this study, identification of all Zingiberaceae species was not possible, even when using data from all six loci (Fig. 5). Among the Acanthaceae species, a combination of *rpo*C1 and *trn*L identified 94.4% of species, whereas the combination of *rpo*C1 and *rbc*LA identified 100% of species. These results indicate that taxa in different plant groups may be more or less readily discriminated by different combinations of markers. Thus, when identification is focused on specific families or genera is it essential to determine the optimal combination of markers for species discrimination. These marker combinations can be estimated based on available sequence data in public repositories, and used for the design of unique and reliable HRM primer sets for species identification.

Conclusions

Bar-HRM has proven to be a cost-effective and reliable method for the identification of species in this study of Thai medicinal plants. The Bar-HRM primer sets developed here are not only useful for identification of medicinal plant vouchers, but can also be used for species discrimination, authentication, and detection of adulteration in samples lacking diagnostic morphological characters, such as single ingredient herbal products. Bar-HRM is dependable, fast, and sensitive enough to distinguish between species, and results show an identification success rate of 99% among species in the test set. Single locus analyses alone yielded

Please cite this article as: M. Osathanunkul et al., Evaluation of DNA barcoding coupled high resolution melting for discrimination of closely related species in phytopharmaceuticals, Phytomedicine (2016), http://dx.doi.org/10.1016/j.phymed.2015.11.018

insufficient identification success rates, but combinations of three to four markers were nearly perfect. Species discrimination success rates reached an upper limit of 99% with four markers, and further combinations of five or six marker did not yield a significant improvement in identification. The phylogenetic affinities of the species tested also affect the discriminatory power, as each multi-locus combination has a different potential for the discrimination among taxa. Although the four marker combinations, *rpo*C1 + *trn*L + *rbc*LB + *rbc*LC and *rpo*C1 + *rbc*LB + *rbc*LC + *mat*K were found to be optimal in this study, it is likely that other markers and combinations thereof might perform better in other plant groups.

Conflict of interest

The authors declare that no competing interests exist.

Uncited references

Li et al. (2011), Skalli et al. (2002), Zaim et al. (2008).

Acknowledgments

This research was financially supported by Thailand Research Fund New Researcher Grant TRG5780027 awarded to MO. HdB acknowledges support from the Norway–Romania EEA Research Programme operated by the MECS-ANCSI PO under the EEA Financial Mechanism 2009–2014 Project Contract no. 2SEE/2014. We thank late Mr. James F. Maxwell (Curator of CMUB herbarium, Chiang Mai University) and Mr. Samphan Wongthep (Medicinal Plant Garden, Chiang Mai University) for providing, identifying and collecting the samples.

References

Berlin, B., 1992. Ethnobiological Classification: Principles of Categorization of Plants and Animals in Traditional Societies. Princeton University Press, Princeton, NJ.
De Boer, H.J., Cotingting, C., 2014. Medicinal plants for women's healthcare in southeast Asia: a meta-analysis of their traditional use, chemical constituents, and pharmacology. J. Ethnopharmacol. 151, 747–767.
De Boer, H.J., Ouarghidi, A., Martin, G., Abbad, A., Kool, A., 2014. DNA barcoding reveals limited accuracy of identifications based on folk taxonomy. PLoS One 9, e84291.
De Boer, H.J., Ichim, M.C., Newmaster, S.G., 2015. DNA barcoding and pharmacovigilance of herbal medicines. Drug Saf. 38, 611–620.
Boessenkool, S., Mcglynn, G., Epp, L.S., Taylor, D., Pimentel, M., Gizaw, A., Nemomissa, S., 2014. Use of ancient sedimentary DNA as a novel conservation tool for high-altitude tropical biodiversity. Conserv. Biol. 28, 446–455.
CBOL Plant Working Group, 2009. A DNA barcode for land plants. Proc. Natl. Acad. Sci. USA 106, 12794–12797.
Chan, K., 2003. Some aspects of toxic contaminants in herbal medicines. Chemosphere 52, 1361–1371.
Chen, J., Zhao, J., Erickson, D.L., Xia, N., Kress, W.J., 2014. Testing DNA barcodes in closely related species of *Curcuma* (Zingiberaceae) from Myanmar and China. Mol. Ecol. Resour. 15, 337–348.
Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., Zhu, Y., 2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS One 5, 1–8.
Coakes, S.J., Steed, L., 2009. SPSS: Analysis Without Anguish using SPSS Version 14.0 for Windows. John Wiley & Sons, Inc Available at: http://dl.acm.org/citation.cfm?id=1804538 (accessed 05.01.15.) .
Coghlan, M., Haile, J., Houston, J., Murray, D., White, N., Moolhuijzen, P., Bellgard, M., Bunce, M., 2012. Deep sequencing of plant and animal DNA contained within traditional chinese medicines reveals legality issues and health safety concerns. PLoS Genet. 8, e1002657.
Fazekas, A.J., Burgess, K.S., Kesanakurti, P.R., Graham, S.W., Newmaster, S.G., Husband, B.C., Percy, D.M., 2008. Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. PLoS One 3, e2802.
Fazekas, A.J., Kesanakurti, P.R., Burgess, K.S., Percy, D.M., Graham, S.W., Barrett, S.C.H., Newmaster, S.G., 2009. Are plant species inherently harder to discriminate than animal species using DNA barcoding markers? Mol. Ecol. Resour. 9, 130–139.
Ficetola, G.F., Coissac, E., Zundel, S., Riaz, T., Shehzad, W., Bessière, J., Taberlet, P., Pompanon, F., 2010. An *in silico* approach for the evaluation of DNA barcodes. BMC Genom. 11, 434.
Franz, C., Klier, B., Reich, E., Novak, J., 2007. Identification and authentication of herbal substances. Planta Med. 73, 819–820.
Ganopoulos, I., Aravanopoulos, F., Madesis, P., Pasentsis, K., Bosmali, I., Ouzounis, C., Tsaftaris, A., 2013. Taxonomic identification of Mediterranean pines and their hybrids based on the high resolution melting (HRM) and trnL approaches: from cytoplasmic inheritance to timber tracing. PLoS One 8, e60945.
Gao, T., Yao, H., Song, J., Liu, C., Zhu, Y., Ma, X., Pang, X., 2010. Identification of medicinal plants in the family Fabaceae using a potential DNA barcode ITS2. J. Ethnopharmacol. 130, 116–121.
Hollingsworth, P.M., 2011. Refining the DNA barcode for land plants. Proc. Natl. Acad. Sci. USA 108, 19451–19452.
Hollingsworth, P.M., Forrest, L.L., Spouge, J.L., Hajibabaei, M., Ratnasingham, S., Van Der Bank, M., Chase, M.W., 2009. A DNA barcode for land plants. Proc. Natl. Acad. Sci. USA 106, 12794–12797.
Inta, A., Shengji, P., Balslev, H., Wangpakapattanawong, P., Trisonthi, C., 2008. A comparative study on medicinal plants used in Akha's traditional medicine in China and Thailand, cultural coherence or ecological divergence? J. Ethnopharmacol. 116, 508–517.
Jørgensen, T., Kjaer, K.H., Haile, J., Rasmussen, M., Boessenkool, S., Andersen, K., Coissac, E., 2012. Islands in the ice: detecting past vegetation on Greenlandic nunataks using historical records and sedimentary ancient DNA Meta-barcoding. Mol. Ecol. 21, 1980–1988.
Kool, A., de Boer, H.J., Krüger, Å., Rydberg, A., Abbad, A., Björk, L., Martin, G., 2012. Molecular identification of commercialized medicinal plants in Southern Morocco. PLoS One 7, e39459.
Kress, W.J., Erickson, D.L., 2007. A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. PLoS One 2, e508.
Kress, W.J., Wurdack, K.J., Zimmer, E.A., Weigt, L.A., Janzen, D.H., 2005. Use of DNA barcodes to identify flowering plants. Proc. Natl. Acad. Sci. USA 102, 8369–8374.
Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., et al., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948.
Li, D.Z., Gao, L.M., Li, H.T., Wang, H., Ge, X.J., Liu, J.Q., Chen, Z.D., 2011. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. Proc. Natl. Acad. Sci. USA 108, 19641–19646.
Li, M., Cao, H., BUT, P.P.H., SHAW, P.C., 2011. Identification of herbal medicinal materials using DNA barcodes. J. Syst. Evol. 49, 271–283.
Madesis, P., Ganopoulos, I., Anagnostis, A., Tsaftaris, A., 2012. The application of Bar-HRM (Barcode DNA-High Resolution Melting) analysis for authenticity testing and quantitative detection of bean crops (Leguminosae) without prior DNA purification. Food Control 25, 576–582.
Madesis, P., Ganopoulos, I., Bosmali, I., Tsaftaris, A., 2013. Barcode high resolution melting analysis for forensic uses in nuts: a case study on allergenic hazelnuts (*Corylus avellana*). Food Res. Int. 50, 351–360.
Mukherjee, P.K., Pitchairajan, V., Murugan, V., Sivasankaran, P., Khan, Y., 2010. Strategies for revitalization of traditional medicine. Chin. Herb. Med. 2, 1–15.
Newmaster, S.G., Grguric, M., Shanmughanandhan, D., Ramalingam, S., Ragupathy, S., 2013. DNA barcoding detects contamination and substitution in North American herbal products. BMC Med. 11, 222.
Osathanunkul, M., Madesis, P., de Boer, H.J., 2015a. Bar-HRM for authentication of plant-based medicines: evaluation of three medicinal products derived from Acanthaceae species. PLoS One 10, e0128476.
Osathanunkul, M., Suwannapoom, C., Ounjai, S., A Rora, J., Madesis, P., de Boer, H.J., 2015b. Refining DNA barcoding coupled high resolution melting for discrimination of 12 closely related *Croton* species. PLoS One 10, e0138888.
Parducci, L., Jørgensen, T., Tollefsrud, M.M., Elverland, E., Alm, T., Fontana, S.L., Bennett, K.D., 2012. Glacial survival of boreal trees in Northern Scandinavia. Science 335, 1083–1086.
Piredda, R., Simeone, M.C., Attimonelli, M., Bellarosa, R., Schirone, B., 2011. Prospects of barcoding the Italian wild dendroflora: oaks reveal severe limitations to tracking species identity. Mol. Ecol. Resour. 11, 72–83.
Reed, G.H., Wittwer, C.T., 2004. Sensitivity and specificity of single-nucleotide polymorphism scanning by high-resolution melting analysis. Clin. Chem. 50, 1748–1754.
Reja, V., Kwok, A., Stone, G., Yang, L., Missel, A., Menzel, C., Bassam, B., 2010. ScreenClust: advanced statistical software for supervised and unsupervised high resolution melting (HRM) analysis. Methods 50, S10–S14.
Ririe, K.M., Rasmussen, R.P., Wittwer, C.T., 1997. Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. Anal. Biochem. 245, 154–160.
Särkinen, T., Staats, M., Richardson, J.E., Cowan, R.S., Bakker, F.T., 2012. How to open the treasure chest? Optimising DNA extraction from herbarium specimens. PLoS One 7, e43808.
Sass, C., Little, D.P., Stevenson, D.W., Specht, C.D., 2007. DNA barcoding in the cycadales: testing the potential of proposed barcoding markers for species identification of cycads. PLoS One 2, e1154.
Shi, L.C., Zhang, J., Han, J.P., Song, J.Y., Yao, H., Zhu, Y.J., Li, J.C., 2011. Testing the potential of proposed DNA barcodes for species identification of Zingiberaceae. J. Syst. Evol. 49, 261–266.
Singtonat, S., Osathanunkul, M., 2015. Fast and reliable detection of toxic *Crotalaria spectabilis* Roth. in *Thunbergia laurifolia* Lindl. herbal products using DNA barcoding coupled with HRM analysis. BMC Complement. Altern. Med. 15, 162.
Siow, Y.L., Gong, Y., Au-Yeung, K.K., Woo, C.W., Choy, P.C., O, K., 2005. Emerging issues in traditional Chinese medicine. Can. J. Physiol. Pharmacol. 83, 321–334.

Sivasankari, B., Anandharaj, M., Gunasekaran, P., 2014. An ethnobotanical study of indigenous knowledge on medicinal plants used by the village peoples of Thoppampatti, Dindigul district, Tamil Nadu, India. J. Ethnopharmacol. 153, 408–423.
Skalli, S., Alaoui, I., Pineau, A., Zaid, A., Soulaymani, R., 2002. L'intoxication par le chardon à glu (Atractylis gummifera L.): à propos d'un cas clinique. Bull. Soc. Pathol. Exot. 95, 284–286.
Stoeckle, M.Y., Gamble, C.C., Kirpekar, R., Young, G., Ahmed, S., Little, D.P., 2011. Commercial teas highlight plant DNA barcode identification successes and obstacles. Sci. Rep. 1, 1–7.
Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., Willerslev, E., 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular ecology 21, 2045–2050.
Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., 2007. Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. Nucleic Acids Res. 35, e14.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.
Techen, N., Parveen, I., Pan, Z., Khan, I.A., 2014. DNA barcoding of medicinal plant material for identification. Curr. Opin. Biotechnol. 25, 103–110.
Veldman, S., Otieno, J., Gravendeel, B., van Andel, T., de Boer, H.J., 2014. Conservation of endangered wild harvested medicinal plants: use of DNA barcoding. In: Gurib-Fakim, A. (Ed.), Novel Plant Bioresources: Applications in Food, Medicine and Cosmetics, pp. 81–88.
Willerslev, E., Davison, J., Moora, M., Zobel, M., Coissac, E., Edwards, M.E., Lorenzen, E.D., 2014. Fifty thousand years of Arctic vegetation and megafaunal diet. Nature 506, 47–51.
Wittwer, C.T., Reed, G.H., Gundry, C.N., Vandersteen, J.G., Pryor, R.J., 2003. High-resolution genotyping by amplicon melting analysis using LCGreen. Clin. Chem. 49, 853–860.
Ye, J., Li, T., Xiong, T., Janardan, R., 2004. Using uncorrelated discriminant analysis for tissue classification with gene expression data. IEEE/ACM Trans. Comput. Biol. Bioinform. 1, 181–190.
Zaim, N., Guemouri, L., Lamnaouer, D., Benjouad, A., 2008. Étude de quatre cas d'intoxication par *Atractylis gummifera* L. au Maroc. Thérapie 63, 49–54.

Please cite this article as: M. Osathanunkul et al., Evaluation of DNA barcoding coupled high resolution melting for discrimination of closely related species in phytopharmaceuticals, Phytomedicine (2016), http://dx.doi.org/10.1016/j.phymed.2015.11.018