Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

DNA barcoding of medicinal plant material for identification

Natascha Techen[1], Iffat Parveen[1], Zhiqiang Pan[2] and Ikhlas A Khan[1,3]

Because of the increasing demand for herbal remedies and for authentication of the source material, it is vital to provide a single database containing information about authentic plant materials and their potential adulterants. The database should provide DNA barcodes for data retrieval and similarity search. In order to obtain such barcodes, several molecular methods have been applied to develop markers that aid with the authentication and identification of medicinal plant materials. In this review, we discuss the genomic regions and molecular methods selected to provide barcodes, available databases and the potential future of barcoding using next generation sequencing.

Addresses
[1] National Center for Natural Products Research and Research Institute of Pharmaceutical Sciences, Department of Pharmacognosy, School of Pharmacy, University of Mississippi, P.O. Box 1848, MS 38677, USA
[2] USDA-ARS-NPURU, P.O. Box 8048, University, MS 38677, USA
[3] School of Pharmacy, King Saud University, Saudi Arabia

Corresponding author: Khan, Ikhlas A (ikhan@olemiss.edu)

Current Opinion in Biotechnology 2014, **25**:103–110

This review comes from a themed issue on **Analytical biotechnology**

Edited by **Frank L Jaksch** and **Savaş Tay**

0958-1669/$ – see front matter,

http://dx.doi.org/10.1016/j.copbio.2013.09.010

Introduction

The traditional system of medicine utilizes medicinal plants to cure various ailments but the herbal industry suffers from substitution and adulteration of medicinal herbs with closely related species. The efficacy of the drug decreases if it is adulterated, and in some cases, can be lethal if it is substituted with toxic adulterants. Hence, the correct formulation is important for the medicinal herb to be effective. The traditional methods of medicinal plant identification include organoleptic methods (identification by the senses: taste, sight, smell, touch), macroscopic and microscopic methods (identification by shape, colour, texture) and chemical profiling (e.g. TLC, HPLC-UV, HPLC-MS). However, neither method can identify the related species easily in processed products because the former method requires trained personal for macroscopic and microscopic examinations. In the latter method, chemical profiles or markers may be affected by physiological and storage conditions. Authentication at the DNA level provides more reliability because, in contrast to RNA, DNA is a stable macromolecule that is not affected by external factors and is found in all tissues. Therefore, development of DNA-based markers is important for authentication of medicinal plants.

The novel technique of identifying biological specimens using short DNA sequences from either nuclear or organelle genomes is called DNA barcoding. The term 'DNA barcode' as taxon identifiers was first proposed by Paul Hebert of University of Guelph in 2003 [1]. He recommended that the 5′ end of cytochrome c oxidase 1 (CO1) from the mitochondrial genome was sufficient to generate DNA barcodes for the identification of animals [1–4]. On the basis of this initial success with animals, CO1 was suggested as the locus that could provide recognition tags for all organisms. They further emphasized that DNA barcoding not only helps in the identification of species but can also define species boundaries, flagging of new species and species delimitation [2,3]. However, in plants the mitochondrial genes are slowly evolving, with very low substitution rates and were not suitable for barcoding. Therefore, the search for plant barcode shifted to chloroplast and nuclear genomes with high substitution rates. Following initial *in-silico* and laboratory-based assessment of different loci from chloroplast and nuclear genomes led to the conclusion that no single locus plant barcode exists, and soon it was realized that multi-locus barcodes are requisite for plant barcoding. Subsequently a number of loci were being tested for their suitability as plant barcodes and many multi-locus combinations were suggested. The Consortium for the Barcode of Life Plant Working Group (CBOL) [5] evaluated seven chloroplast genomic regions across the plant kingdom and proposed a combination of *mat*K and *rbc*L as plant barcodes. High universality but less species resolution is provided by *rbc*L whereas *mat*K affords high resolution but less universality. A combination of these two can help to achieve maximum species discrimination. Nevertheless, in closely related species, the discriminating ability of these two markers is low [6,7]. Therefore, the China Plant BOL Group [8] proposed the addition of nuclear ITS (Internal Transcibed Spacer) to the *mat*K + *rbc*L combination as plant barcode in order to achieve maximum identification rates even in closely related species.

The aim of this review is to assess the progress made so far in the field of DNA barcoding in relation to the identification of botanicals. In the current paper, we review the genomic regions selected as possible barcodes for medicinal plants and the emerging new methods for rapid generation of barcodes. We also discuss the challenges of barcoding and what databases are available to retrieve barcodes of medicinal plants, their substitutes and adulterants.

Loci suggested as plant barcodes

At the Fourth International Barcode of Life Conference (http://www.dnabarcodes2011.org/) the option of a three-locus barcode (matK + *rbc*L + *psb*A-*trn*H) versus a two-locus barcode was discussed. The two-locus barcode was preferred to avoid the increased costs of sequencing three loci rather than two in very large sample sets, and to prevent further delays in implementing a standard barcode for land plants. The barcode combination *rbc*L + *mat*K was the preferred choice.

A search of the literature in SciFinder (a chemical abstracts service database) from 2010 to 2013 resulted in 60 publications (Figure 1 and Table 1). In the literature analyzed in this review, a total of 17 barcode regions (*mat*K, *rbc*L, ITS, ITS2, *psb*A-*trn*H, *atp*F-*atp*H, *ycf*5, *psb*K-I, *psb*M *trn*D, *rps*16, *cox*I, *nad*1, *trn*L-F, *rpo*B, *rpo*C1, *atp*F-*atp*H, *rps*16) of medicinal plants were reported to aid in the authentication and identification of medicinal plant materials. The majority of barcoding regions mentioned in the literature were the ITS region (26 references), *psb*A-*trn*H (21 references), *mat*K (19 references), and *rbc*L (14 references). Further genomic regions used for barcoding were ITS2 (9 references), *rpo*C1 (6 references), *rpo*B (4 references), and *trn*L-F (3 references).

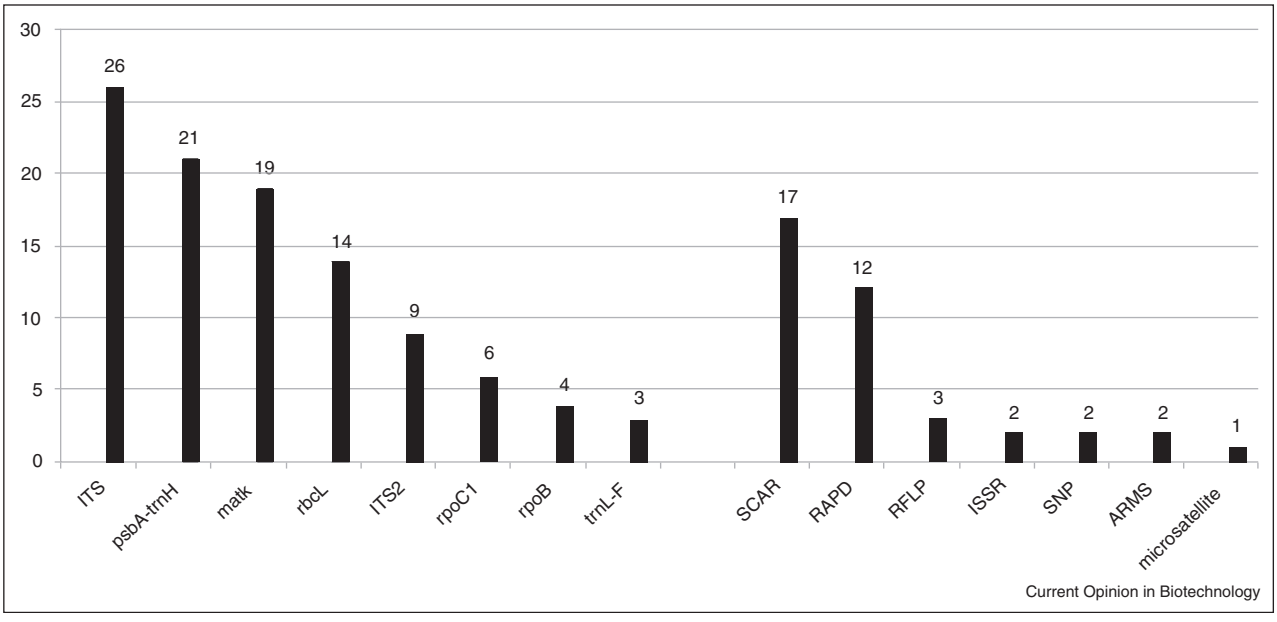
Besides using known genomic regions, other PCR-based methods have been applied to develop markers that help with the authentication and identification of medicinal plant material: RAPD, RFLP, microsatellites, ISSRs, SNPs, and ARMS (12,3,1,2,2,2 publications, respectively). SCAR markers (17 references) have been developed from RAPD, ISSR and a variety of genomic regions.

Zuo *et al.* [14] analyzed 12 genomic regions of 95 ginseng samples, representing all the species in the genus. They demonstrated that the combination of *psb*A-*trn*H and ITS was sufficient for the identification of all species and species clusters in the genus. For various samples analyzed, the combination of up to three genomic regions (*mat*K, *rbc*L, ITS, *psb*A-*trn*H) provided the required information for identification [15,35,38].

Chen *et al.* [11] analyzed >6600 plant samples belonging to 4800 species from 753 distinct genera using the genomic regions *psb*A-*trn*H, *mat*K, *rbc*L, *rpo*C1, *ycf*5, ITS, and ITS2. Their data suggested that the second internal transcribed spacer (ITS2) of nuclear ribosomal DNA represents the most suitable region for DNA barcoding applications. The percentage of successful identification with ITS2 at the species level was 92.7%. He *et al.* and Selvaraj *et al.* [9,18] also analyzed multiple genomic barcode regions and came to a similar conclusion that ITS or ITS2 showed the highest discrimination rate among the samples analyzed. In contrast, the highest discrimination rate was found to be from *psb*A-*trntrn*H [10,12,39] or from *mat*K [41,48].

Theodoridis *et al.* [36] analyzed medicinal plants of the *Labiatae* (*Lamiaceae*) family from Chios Island (Greece) and the adjacent Cesme-Karaburun Peninsula (Turkey) using all three barcode loci *mat*K, *rbc*L, *psb*A-*trn*H and

Figure 1



Literature search in SciFinder (a Chemical Abstracts Service database) using the keywords 'DNA barcod* medicinal plant' (excluding endophytic fungi, including patents) from 2010 to 2013 resulted in 60 publications. Various genomic regions and PCR-based methods were used to develop barcodes.

Table 1

Overview of DNA based methods and technologies investigated for medicinal plant identification.

Ref.	Analyzed material	Genomic regions analyzed/molecular methods applied					
[9]	*Angelica* spp.	*mat*K	*rbc*L	ITS	ITS2	*psb*A-*trn*H	
[10]	*Rhododendron* spp.	*mat*K	*rbc*L	ITS	ITS2	*psb*A-*trn*H	
[11]	753 genera	*mat*K	*rbc*L	ITS	ITS2	*psb*A-*trn*H	
[12]	*Lonerica* spp.	*mat*K	*rbc*L	ITS	*psb*A-*trn*H	*trn*L-F	
[13]	*Solanum* spp and adulterants	*mat*K	*rbc*L	ITS	*psb*A-*trn*H	*trn*L-F	
[14]	*Ginseng* genus	*mat*K	*rbc*L	ITS	*psb*A-*trn*H	*rpo*B	*rpo*C1
[15]	*Astragalus* spp. and adulterants	*mat*K	*rbc*L	ITS			
[16]	Various medicinal roots	*mat*K	ITS	*psb*A-*trn*H	*rpo*C1		
[17]	Various medicinal plants	ITS	ITS2				
[18]	*Boerhavia* spp. *Astragalus* spp. and adulterants	ITS	ITS2	*psb*A-*trn*H			
[19]	*Sedum* spp. *Astragalus* spp. and adulterants	ITS	ITS2				
[20•]	Various (old) medicinal plants	ITS	ITS2				
[21]	*Rubus* spp.	ITS	*psb*A-*trn*H	*trn*L-F	SCAR	SNP	
[22•]	*Hypericum* spp.	ITS	SCAR				
[23]	*Ochradenus* spp.	ITS	*rpo*B	*rpo*C1			
[24]	*Rehmannia* spp.	ITS	RAPD	SCAR			
[25]	Medicinal vines (22 genera)	ITS					
[26]	*Dipsacus* spp.	ITS					
[27]	*Dendrobium* spp.	ITS	ARMS				
[28]	*Paris* spp.	ITS	RFLP				
[29]	*Citrus* spp.	ITS	SNP				
[30]	*Dendrobium* spp.	ITS					
[31]	*Ruta* spp.	ITS	*rpo*B	*rpo*C1	SCAR		
[32]	*Astragalus* spp.	ITS					
[33]	Various medicinal plants	ITS					
[34]	*Meconopsis* spp.	ITS					
[35]	*Scutellaria* spp. *Astragalus* spp. and adulterants	*mat*K	*rbc*L	*psb*A-*trn*H			
[36]	*Lamiaceae*	*mat*K	*rbc*L	*psb*A-*trn*H			
[37]	Various medicinal plants	*mat*K	*rbc*L	*psb*A-*trn*H			
[38]	*Sabia* spp.	*mat*K	*rbc*L	*psb*A-*trn*H			
[39]	*Pteridophytes*	*mat*K	*rbc*L	*psb*A-*trn*H	*rpo*B	*rpo*C1	
[40]	*Vitex* spp.	*mat*K	*psb*A-*trn*H	RFLP			
[41]	*Sideritis* spp.	*mat*K	*psb*A-*trn*H				
[42]	*Paris* spp. and adulterants	*psb*A-*trn*H					
[43]	*Senna* spp.	*psb*A-*trn*H					
[44]	*Smilax* spp.	*psb*A-*trn*H					
[45]	*Phyllanthus* spp.	*psb*A-*trn*H					
[46]	*Cistance* spp.	*psb*A-*trn*H					
[47]	*Galpemia* spp.	*mat*K	*rbc*L	*rpo*C1			
[48]	*Dendrobium* spp.	*mat*K	*rbc*L				
[49]	Rhubarb	*mat*K	SCAR				
[50]	*Pueraria candollei*, *Butea superb*, *Mucuna collettii*	*mat*K	RFLP				
[51]	*Orobanche* spp. and adulterants	ITS2					
[52]	*Taraxacum* and adulterants	ITS2	SCAR				
[53]	*Cuscuta* spp. and adulterants	RAPD	SCAR				
[54]	*Artemisia* spp.	RAPD	SCAR				
[55]	*Liriope* spp, *Ophiopogon* spp.	RAPD	SCAR				
[56]	Rheum spp.	RAPD	SCAR				
[57]	*Glycerrhiza* spp.	SCAR					
[58]	*Zizyphus* spp.	RAPD	SCAR				
[59]	*Crocus* spp. and adulterants	RAPD	SCAR				
[60]	*Prunella* spp.	RAPD	SCAR				
[61]	*Viola* spp. and adulterants	RAPD	SCAR				
[62]	Srocphularia spp.	SCAR	ISSR				
[63]	*Panax ginseng*	RAPD	SCAR				
[64]	*Uncaria* spp.	RAPD					
[65]	*Marsdeniae* and adulterants	RAPD					
[66]	*Phyllanthus* spp.	ISSR					
[67]	*Asparagus* and adulterants	ARMS					
[68]	*Dendrobium* spp.	MS					

MS = microsatellites.

tested them either as single-region or as multi-region barcodes. They showed that *mat*K and *psb*A-*trn*H were as useful in discriminating species of the *Labiatae* as any multi-region combination. For the *Labiatae* species analyzed *mat*K and *psb*A-*trn*H could serve as single-region barcodes. Schori and Showalter [37] analyzed the *rbc*L, *mat*K, *psb*A-*trn*H loci of fourteen species of medicinal plants from Pakistan and found that, depending on plant to be identified, one region was preferred over the other to aid with species identification.

Single barcode-regions for identification have been reported for (i) *mat*K [41,49,50], (ii) ITS [22•,25–30,32–34], (iii) ITS2 [17,20•,51,52], and (iv) *psb*A-*trn*H [42–46].

Not only genomic regions, but also the usefulness of RAPD-based DNA markers has been reported for medicinal plant identification. RAPD and SCAR markers have been developed for the identification of several medicinal plant materials [53–56,58–61,62].

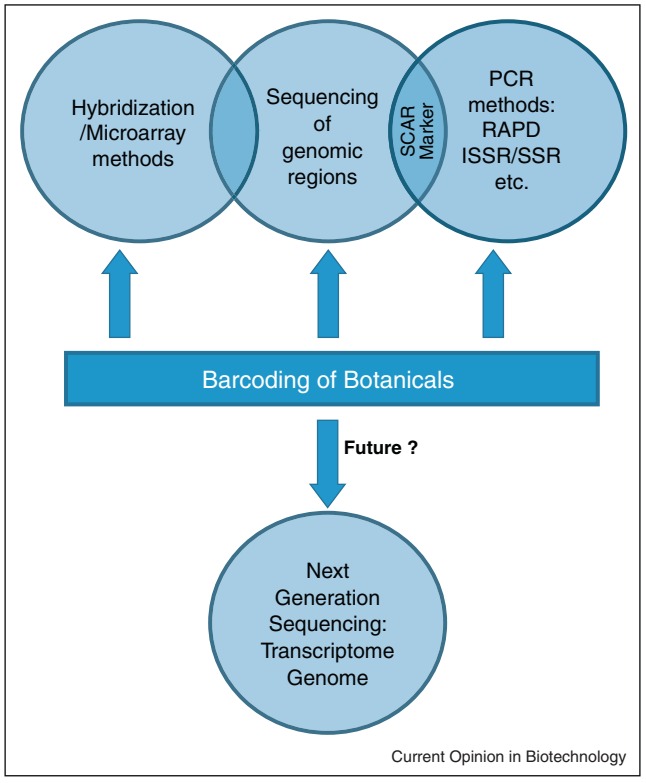
The review by Heubl [69••] focused on the most commonly used DNA-based technologies (RAPD, RFLP, ARMS, CAPS, AFLP, DAF, ISSR, SSR, sequencing, hybridization and microarrays) and new achievements in the field of DNA barcoding and DNA chip technology used to identify traditional Chinese medicinal materials. This particular review provides a critical view of the methods presented concerning sensitivity, reproducibility, and dependability.

Future developments

Although the traditional DNA barcoding techniques remain an effective DNA method for identification of medicinal plants, the more advanced and newly developed high throughput sequencing, specifically next-generation sequencing (NGS) technologies [70], could be adopted and potentially revolutionize the process. Even though DNA barcoding usually targets short regions of DNA molecule within the genome and does not require full genome-scale data, the potential of using NGS to simultaneously identify multiple species or bulk samples of organisms by sequencing DNA barcodes is enormous.

The continuing improvement in NGS technologies and the massive expansion of reference sequence databases have made the NGS approach promising. However, only one publication using NGS for medicinal plant DNA barcoding has appeared to date in which it was utilized to identify potential nuclear genomic regions for barcoding [71]. But most recently, there have been a number of cases published in which NGS was employed, for example, to explore the intragenomic divergence of Dikarya [72], and the taxonomic characterization of freshwater diatom communities [73]. Presently NGS and bioinformatics are used to probe and deduce the transcriptome of medicinal plants to provide a robust resource for gene discovery and downstream metabolic pathway discovery [74–76]. Since medicinal plants tend to have a large genome, the combination of targeted genomic enrichment [77] and NGS could ultimately make the technologies applicable in DNA barcoding of medicinal plant material (Figure 2).

Figure 2



Current and proposed barcoding applications of botanicals.

Challenges and limitations of barcoding

The isolation of pure, high molecular weight DNA is critical for the successful application of molecular methods. This can be quite a challenge since in processed medicinal plant material the DNA is often highly degraded or the plant material contains high amounts of polysaccharides, polyphenols and other secondary metabolites, such as, alkaloids and flavonoids. Various commercial kits and modified traditional methods are available to yield in good quality DNA from raw and powdered medicinal plant material, herbarium specimens, capsules, tablets, or tinctures for downstream applications [22•,78,79••,80,81]. Särkinen *et al.* [79••] found a strong negative correlation between amplicon size and PCR success, indicating that shorter fragments are easier to amplify from herbarium DNA. The relatively large genomic region *rbc*L (670 bp) was amplified only in 10% of the analyzed samples, the medium sized LEAFY intron 3 (260 bp) amplified in 24%, while the small

genomic region *trn*L P6 loop (10–143 bp) was successfully amplified in 78% of the samples. For heavily damaged DNA, it is therefore recommended to develop conditions to produce very short amplicons, which are easier to amplify, or perform DNA 'repair reactions' [79••,82], and that are also available as commercial kits. Särkinen *et al.* [79••] believes that fragmented DNA of medicinal plants is going to be far less of a problem in the near future using the NGS approach since this technique requires fragmented (and ligated) DNA as starting material.

Availability of data

It is desirable to have access to a single barcode library for all medicinal material used (including fungal and animal species). Currently, however several barcode libraries are freely accessible (see also review by Taylor and Harris [83]):

(i) BOLD (The barcode of life data system) [84] (http://www.barcodinglife.com) was created and is maintained by the University of Guelph in Ontario. It offers researchers a way to collect, manage, and analyze DNA barcode data. The goal is, over the next 20 years, to provide a barcode library for all eukaryotic life.

(ii) CBOL (Consortium for the barcode of life) (http://www.barcodeoflife.org/) is a public reference library of species identifiers which could be used to assign unknown specimens to known species. CBOL was founded in 2004 and promotes barcoding through working groups, networks, workshops, conferences, outreach, and training. CBOL has 200 member organizations from 50 countries and operates from a Secretariat Office located in the Smithsonian Institution's National Museum of Natural History in Washington, DC [5].

(iii) iBOL (International Barcode of Life project) (http://www.ibol.org/) consists of a group of hundreds of scientists from 25 nations working together to construct a DNA barcode reference library that will be the foundation for a DNA-based identification system for all multi-cellular life. Their five-year (2010–2015) goal is to barcode five million specimens representing 500,000 species.

(iv) The GenBank online genetic sequence database (http://www.ncbi.nlm.nih.gov/genbank/) [85] is possibly one of the most important repositories of genetic information. GenBank contains over 108 million entries for over 260,000 named organisms and is one of the most frequently used databases for genomic authentication [86]. With the BLAST program [87] an unknown DNA sequence can be rapidly and accurately compared to known and well characterized sequences.

(v) MMDBD (Medicinal Materials DNA Barcode Database) (http://137.189.42.34/mherbsdb/index.php) is a website that includes DNA sequences and information and key references of the medicinal materials recorded in the Pharmacopoeia of the People's Republic of China, American Herbal Pharmacopoeia and other related references. This database, updated in May 2012 with 1658 species and 31,468 sequences available, provides information material for distinguishing medicinal materials (plant, animal, and fungi) from their common substitutes and adulterants [88••].

(vi) The GDR (Genome Database for Rosaceae), founded in 2009, provides genetic markers and ESTs of *Rosaceae*. A large number of species in Rosaceae or rose family have a medicinal value (http://www.rosaceae.org/).

Conclusion

Molecular barcoding methods are reliable tools for the identification of medicinal plants, their substitutes and adulterants at the genus and species level. The methods discussed provide consistent and reliable results regardless of the age, plant part, or environmental factors of the sample.

Based on the literature analyzed in this review, it appears that, although the Barcode of Life Plant Working Group [5] recommends the genomic regions *rbc*L + *mat*K for barcoding, often other genomic regions could be more useful for medicinal material identification. Furthermore, depending on the material analyzed, one or the combination of up to three genomic regions was necessary to provide the required information for identification.

Because of the increasing demand for herbal remedies, authentication of the medicinal plant material is important; therefore it is vital to provide a sole, extensive database with DNA data for easy identification.

Future sequencing advances to analyze large scale nucleic acid sequences have a great potential for genotyping and taxon identification, even for damaged or fragmented template DNA. In the near future, full genomic sequence data from medicinal plants will be available. The challenge will be to provide sufficient storage space for all the data produced.

Acknowledgements

This research was funded in part by the Food and Drug Administration and USDA. We thank Jon Parcher for his revision of the manuscript and suggestions.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Hebert PDN, Cywinska A, Ball SL, de Waard JR: **Biological identifications through DNA barcodes**. *Proc R Soc Biol Sci Ser B* 2003, **270**:313-321.

2. Hebert PDN, Ratnasingham S, de Waard JR: **Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species**. *Proc R Soc Biol Sci Ser B* 2003, **270**:S96-S99.

3. Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W: **Ten species in one: DNA barcoding reveals cryptic species in neotropical skipper butterfly *Astraptes fulgerator***. *Proc Natl Acad Sci USA* 2004, **101**:14812-14817.

4. Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM: **Identification of birds through DNA barcodes**. *PLoS Biol* 2004, **2**:e312.

5. CBOL Plant Working Group: **A DNA barcode for land plants**. *Proc Natl Acad Sci USA* 2009, **106**:12794-12797.

6. Singh HK, Parveen I, Raghuvanshi S, Babbar SB: **Loci recommended as universal barcode for plants on the basis of floristic studies may not work with congeneric species as exemplified by DNA barcoding of *Dendrobium* species**. *BMC Res Notes* 2012, **5**:42.

7. Cameron KM, Chase MW, Whitten WM, Kores PJ, Jarrell DC, Albert VA, Yukawa T, Hills HG, Goldman DH: **A phylogenetic analysis of the *Orchidaceae*: evidence from *rbcL* nucleotide sequences**. *Am J Bot* 1999, **86**:208-224.

8. China Plant BOL Group: Li DZ, Gao LM, Li HT, Wang H, Ge XJ, Liu JQ, Chen ZD, Zhou SL, Chen SL, Yang JB, Fu CX, Zeng CX, Yan HF, Zhu YJ, Sun YS, Chen SY, Zhao L, Wang K, Yang T, Duan GW: **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* 2011, **108**:19641-19646.

9. He Y, Hou P, Fan G, Song Z, Arain S, Shu H, Tang C, Yue Q, Zhang Y: **Authentication of *Angelica anomala* Ave-Lall cultivars through DNA barcodes**. *Mitochondrial DNA* 2012, **23**:100-105.

10. Liu Y, Zhang L, Liu Z, Luo K, Chen S, Chen K: **Species identification of Rhododendron (*Ericaceae*) using the chloroplast deoxyribonucleic acid *psb*A-*trn*H genetic marker**. *Pharmacogn Mag* 2012, **8**:29-36.

11. Chen S, Yao H, Han J, Liu C, Song J, Shi L, Zhu Y, Ma X, Gao T, Pang X, Luo K, Li Y, Li X, Jia X, Lin Y, Leon C: **Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species**. *PLoS ONE* 2010, **5**:e8613.

12. Sun Z, Gao T, Yao H, Shi L, Zhu Y, Chen S: **Identification of *Lonicera japonica* and its related species using the DNA barcoding method**. *Planta Med* 2011, **77**:301-306.

13. Li M, Au KY, Lam H, Cheng L, Jiang RW, But PP, Shaw PC: **Identification of Baiying (Herba Solani Lyrati) commodity and its toxic substitute Xungufeng (Herba Aristolochiae Mollissimae) using DNA barcoding and chemical profiling techniques**. *Food Chem* 2012, **135**:1653-1658.

14. Zuo Y, Chen Z, Kondo K, Funamoto T, Wen J, Zhou S: **DNA barcoding of *Panax* species**. *Planta Med* 2011, **77**:182-187.

15. Guo H, Wang W, Yang N, Guo B, Zhang S, Yang R, Yuan Y, Yu J, Hu S, Sun Q, Yu J: **DNA barcoding provides distinction between Radix Astragali and its adulterants**. *Sci China Life Sci* 2010, **53**:992-999.

16. Kool A, de Boer HJ, Krüger A, Rydberg A, Abbad A, Björk L, Martin G: **Molecular identification of commercialized medicinal plants in Southern Morocco**. *PLoS ONE* 2012, **7**:e39459.

17. Gao T, Yao H, Song J, Liu C, Zhu Y, Ma X, Pang X, Xu H, Chen S: **Identification of medicinal plants in the family *Fabaceae* using a potential DNA barcode ITS2**. *J Ethnopharmacol* 2010, **130**:116-121.

18. Selvaraj D, Shanmughanandhan D, Sarma RK, Joseph JC, Srinivasan RV, Ramalingam S: **DNA barcode ITS effectively distinguishes the medicinal plant *Boerhavia diffusa* from its adulterants**. *Genom Proteom Bioinform* 2012, **10**:364-367.

19. Liu MZ, Luo K, Yao H, Liu P: **Authentication of *Sedum sarmentosum* and its adulterants by application of ITS2 sequences**. *Zhongguo Xiandai Zhongyao* 2011, **13** pp. 29–31, 38.

20. • Han J, Zhu Y, Chen X, Liao B, Yao H, Song J, Chen S, Meng F: **The short ITS2 sequence serves as an efficient taxonomic sequence tag in comparison with the full-length ITS**. *Biomed Res Int* 2013:741476.

Herbarium voucher specimen, some 90 years old, were used to compare amplification success of the ITS (average length 634 bp) and ITS2 (average length 233 bp) region. Efficiency was 23% and 91% respectively. 12,861 plant sequences were used to compare identification efficiency of ITS and ITS2. Results show ITS2 can serve as minibarcode.

21. Yang JY, Jang SY, Kim HK, Park SJ: **Development of a molecular marker to discriminate Korean *Rubus* species medicinal plants based on the nuclear ribosomal DNA internal transcribed spacer and chloroplast *trn*L-F intergenic region sequences**. *J Korean Soc Appl Biol Chem* 2012, **55**:281-289.

22. • Kazi T, Hussain N, Bremner P, Slater A, Howard: **The application of a DNA-based identification technique to over-the-counter herbal medicines**. *Fitoterapia* 2013, **87**:27-30.

This article reports the identification technique for *Hypericum perforatum*, St. John's Wort. A PCR based method has been applied to DNA extracted from capsules, tablets, and tinctures. Extracting good quality DNA from the samples is a challenge. Only four out of 13 products yielded in a PCR product due to low quality of DNA.

23. Khan S, Al-Qurainy F, Nadeem M, Tarroum M: **Development of genetic markers for *Ochradenus arabicus* (Resedaceae), an endemic medicinal plant of Saudi Arabia**. *Genet Mol Res* 2012, **11**:1300-1308.

24. Kim YS, Ryuk JA, Kom BS: **Discrimination of Korean *Rehmannia glutinosa* from Chinese *Rehmannia glutinosa* using sequence-characterized amplified region marker**. *J Korean Soc Appl Biol Chem* 2012, **55**:1-6.

25. Liu Z, Zeng X, Yang D, Ren G, Chu G, Yuan Z, Luo K, Xiao P, Chen S: **Identification of medicinal vines by ITS2 using complementary discrimination methods**. *J Ethnopharmacol* 2012, **141**:242-249.

26. Li D, Wang Z, Liu X, Yuan Y: **Identification of the medicinal plant *Dipsacus asperoides* from three other species in genus *Dipsacus* (*Dipsaceae*) by internal transcribed spacer of ribosomal deoxyribonucleic acid (rDNA ITS)**. *J Med Plant Res* 2012, **6**:289-295.

27. Chiang CH, Yu TA, Lo SF, Kuo CL, Peng WH, Tsay HS: **Molecular authentication of *Dendrobium* species by multiplex polymerase chain reaction and amplification refractory mutation system analysis**. *JASHS* 2012, **137**:438-444.

28. Liu T, Ji Y: **Molecular authentication of the medicinal plant *Paris polyphylla* Smith var. *yunnanensis* (*Melanthiaceae*) and its related species by polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP)**. *J Med Plant Res* 2012, **6**:1181-1186.

29. Wang H, Kim MK, Kim YJ, Lee HN, Jin H, Chen J, Yang DC: **Molecular authentication of the Oriental medicines Pericarpium Citri Reticulatae and Citri Unshius Pericarpium using SNP markers**. *Gene* 2012, **494**:92-95.

30. Takamiya T, Wongsawad P, Tajima N, Shioda N, Lu JF, Wen CL, Wu JB, Handa T, Iijima H, Kitanaka S, Yukawa T: **Identification of dendrobium species used for herbal medicines based on ribosomal DNA internal transcribed spacer sequence**. *Biol Pharm Bull* 2011, **34**:779-782.

31. Al-Qurainy F, Khan S, Tarroum M, Al-Hemaid FM, Ali MA: **Molecular authentication of the medicinal herb *Ruta graveolens* (*Rutaceae*) and an adulterant using nuclear and chloroplast DNA markers**. *Genet Mol Res* 2011, **10**:2806-2816.

32. Xiao WL, Motley TJ, Unachukwu UJ, Lau CB, Jiang B, Hong F, Leung PC, Wang QF, Livingston PO, Cassileth BR, Kennelly EJ: **Chemical and genetic assessment of variability in commercial Radix Astragali (*Astragalus* spp.) by ion trap LC–MS and nuclear ribosomal DNA barcoding sequence analyses**. *J Agric Food Chem* 2011, **59**:1548-1556.

33. Mati E, de Boer H: **Ethnobotany and trade of medicinal plants in the Qaysari Market, Kurdish Autonomous Region, Iraq**. *J Ethnopharmacol* 2011, **133**:490-510.

34. Li R, Dao Z: **Identification of Meconopsis species by a DNA barcode sequence: the nuclear internal transcribed spacer**

(ITS) region of ribosomal deoxyribonucleic acid (DNA). *Afr J Biotechnol* 2011, **10**:15805-15807.

35. Guo X, Wang X, Su W, Zhang G, Zhou R: **DNA barcodes for discriminating the medicinal plant *Scutellaria baicalensis* (Lamiaceae) and its adulterants**. *Biol Pharm Bull* 2011, **34**:1198-1203.

36. Theodoridis S, Stefanaki A, Tezcan M, Aki C, Kokkini S, Vlachonasios KE: **DNA barcoding in native plants of the *Labiatae* (*Lamiaceae*) family from Chios Island (Greece) and the adjacent Cesme-Karaburun Peninsula (Turkey)**. *Mol Ecol Resour* 2012, **12**:620-633.

37. Schori M, Showalter AM: **DNA barcoding as a means for identifying medicinal plants of Pakistan**. *Pak J Bot* 2011, **43(SI)**:1-4.

38. Sui XY, Huang Y, Tan Y, Guo Y, Long CL: **Molecular authentication of the ethnomedicinal plant *Sabia parviflora* and its adulterants by DNA barcoding technique**. *Planta Med* 2011, **77**:492-496.

39. Ma XY, Xie CX, Liu C, Song JY, Yao H, Luo K, Zhu YJ, Gao T, Pang XH, Qian J, Chen SL: **Species identification of medicinal pteridophytes by a DNA barcode marker, the chloroplast psbA-trnH intergenic region**. *Biol Pharm Bull* 2010, **33**:1919-1924.

40. Phoolcharoen W, Sukrong S: **Molecular analysis of *Vitex* species using candidate DNA barcoding and PCR-RFLP of the *mat*K gene for authentication of *Vitex glabrata***. *Nat Prod Commun* 2013, **8**:125-158.

41. Tezcan M, Vlachonasios KE, Aki C: **DNA barcoding study on *Sideritis trojana* Bornm. an endemic medicinal plant of Ida Mountain, Turkey**. *Fresen Environ Bull* 2010, **19**:1352-1355.

42. Yang Y, Zhai Y, Liu T, Zhang F, Ji Y: **Detection of *Valeriana jatamansi* as an adulterant of medicinal *Paris* by length variation of chloroplast *psb*A-*trn*H region**. *Planta Med* 2011, **77**:87-91.

43. Monkheang P, Sudmoon R, Tanee T, Noikotr K, Bletter N, Chaveerach A: **Species diversity, usages, molecular markers and barcode of medicinal Senna species (Fabaceae, Caesalpinioideae) in Thailand**. *J Med Plant Res* 2011, **5**:6173-6181.

44. Kritpetcharat O, Khemtonglang N, Kritpetcharat P, Daduang J, Daduang S, Suwanrungruang K, Bletter N, Sudmoon R, Chaveerach A: **Using DNA markers and barcoding to solve the common problem of identifying dried medicinal plants with the examples of Smilax and Cissus in Thailand**. *J Med Plant Res* 2011, **5**:3480-3487.

45. Srirama R, Senthilkumar U, Sreejayan N, Ravikanth G, Gurumurthy BR, Shivanna MB, Sanjappa M, Ganeshaiah KN, Shaanker RU: **Assessing species admixtures in raw drug trade of *Phyllanthus*, a hepato-protective plant using molecular tools**. *J Ethnopharmacol* 2010, **130**:208-215.

46. Han JP, Song JY, Liu C, Chen J, Qian J, Zhu YJ, Shi LC, Yao H, Chen SL: **Identification of *Cistanche* species (*Orobanchaceae*) based on sequences of the plastid *psb*A-*trn*H intergenic region**. *Yao Xue Xue Bao* 2010, **45**:126-130.

47. Sharma A, Folch JL, Cardoso-Taketa A, Lorence A, Villarreal ML: **DNA barcoding of the Mexican sedative and anxiolytic plant *Galphimia glauca***. *J Ethnopharmacol* 2012, **144**:371-378.

48. Asahina H, Shinozaki J, Masuda K, Morimitsu Y, Satake M: **Identification of medicinal *Dendrobium* species by phylogenetic analyses using *mat*K and *rbc*L sequences**. *J Nat Med* 2010, **64**:133-138.

49. Xu G, Wang X, Liu C, Li W, Wei S, Liu Y, Cheng X, Liu J: **Authentication of official Da-huang by sequencing and multiplex allele-specific PCR of a short maturase K gene**. *Genome* 2013, **56**:109-113.

50. Wiriyakarun S, Yodpetch W, Komatsu K, Zhu S, Ruangrungsi N, Sukrong S: **Discrimination of the Thai rejuvenating herbs *Pueraria candollei* (White Kwao Khruea), *Butea superba* (Red Kwao Khruea), and *Mucuna collettii* (Black Kwao Khruea) using PCR-RFLP**. *J Nat Med* 2013, **67**:562-570.

51. Sun Z, Song J, Yao H, Han J: **Molecular identification of Cistanches Herba and its adulterants based on nrITS2 sequence**. *J Med Plant Res* 2012, **6**:1041-1045.

52. Chiang YC, Chang WT, Chen MD, Lai GH, Chen HJ, Chao J, Lin MK, Chang YS, Chou YM, Lee MS *et al.*: **Rapid identification of the medicinal plant *Taraxacum formosanum* and distinguishing of this plant from its adulterants by ribosomal DNA internal transcribed spacer (ITS) based DNA barcode**. *Afr J Biotechnol* 2011, **24**:4838-4843.

53. Abdin MZ, Mirza KJ, Khan S, Ki U, Ram M, Ahmad P: **Development and detection efficiency of SCAR markers of *Cuscuta reflexa* and its adulterant *Cuscuta chinensis***. *J Food Drug Anal* 2012, **20**:471-477.

54. Doh EJ, Oh SE: **Multiplex PCR method to discriminate *Artemisia iwayomogi* from other *Artemisia* plants**. *Methods Mol Biol* 2012, **862**:149-160.

55. Li G, Park YJ: **SCAR markers for discriminating species of two genera of medicinal plants, Liriope and Ophiopogon**. *Genet Mol Res* 2012, **11**:2987-2996.

56. Abdin MZ, Khan KJMS, Alam P: **SCAR primers and a kit for the authentication of Unani Drug (Rewand Chini) Rheum emodi and its adulterant *Rheum palmatum***. *Indian Pat Appl* 2012. IN 2010DE01948 A 20120302.

57. Abdin MZ, Khan KJMS, Kiran U: **SCAR primers and a kit for the authentication of Unani Drug Mulethi (*Glycerrhiza glabra*) and its adulterant *Abrus precatorious***. *Indian Pat Appl* 2012. IN 2010DE01943 A 20120302.

58. Abdin MZ, Khan KJMS, Ali A: **SCAR primers and a kit for the authentication of Unani drug Unnab (*Zizyphus jujube*) and its adulterant Jhad Beri (*Zizyphus nummularia*)**. *Indian Pat Appl* 2012. IN 2010DE01946 A 20120302.

59. Abdin MZ, Khan KJMS, Saxena P: **SCAR primers and a kit for the authentication of Unani drug Zaafran (*Crocus sativus*) and its adulterant Qurtum (*Carthamus tinctorious*)**. *Indian Pat Appl* 2012. IN 2010DE01947 A 20120302.

60. Sun X, Wei YL, Zhou YF, Guo JL, Hang YY: **Development of species and region-specific random amplification of polymorphic DNA-sequence characterized amplified region (RAPD-SCAR) markers for identification of the genuineness of *Spica prunellae* (Lamiaceae)**. *J Med Plant Res* 2011, **5**:1677-1684.

61. Yu H, Yu H, Liu H: **Identification of medicinal *Viola philippica* from *V. mandshurica* using SCAR markers**. *Int J Botany* 2011, **7**:189-194.

62. Chen C, Duan LN, Zhou XL, Chen BL, Fu CX: **Molecular authentication of geo-authentic *Scrophularia* ningpoensis**. *J Zhejiang Univ Sci B* 2011, **12**:393-398.

63. Ahn CH, Kim YS, Lim S, Yi JS, Choi YE: **Random amplified polymorphic DNA (RAPD) analysis and RAPD-derived sequence characterized amplified regions (SCAR) marker development to identify Chinese and Korean ginseng**. *J Med Plant Res* 2011, **5**:4487-4492.

64. Xu P, Wu YS, Huang RS, Zhou J, Luo Y, Li KZ, Tang YL, Liu Y: **Polymorphism of three kinds of medicinal plants Uncaria from Guangxi by RAPD**. *Shizhen Guoyi Guoyao* 2012, **23**:703-705.

65. Zhang H, Pei ZD, Ni C, Kang TG: **Identification of Marsdeniae tenacissimae caulis and its adulterants by RAPD**. *Zhong Yao Cai* 2011, **34**:1355-1357.

66. Senapati SK, Aparajita S, Rout GR: **Identification of species-diagnostic inter simple sequence repeat markers for ten *Phyllanthus* species**. *Z Naturforsch C* 2011, **66**:167-172.

67. Kumeta Y, Maruyama T, Wakana D, Kamakura H, Goda Y: **Method for identifying the botanical origin of shatavari products and its application for survey analysis of products in the Japanese market**. *Nippon Shokuhin Kagaku Gakkaishi* 2011, **18**:163-167.

68. Yuan YH, Hou BW, Xu HJ, Luo J, Ding XY: **Identification of the geographic origin of *Dendrobium thyrsiflorum* on Chinese herbal medicine market using trinucleotide microsatellite markers**. *Biol Pharm Bull* 2011, **34**:1794-1800.

69. •• Heubl G: **New aspects of DNA-based authentication of Chinese medicinal plants by molecular biological techniques**. *Planta Med* 2010, **76**:1963-1974.
This article provides a critical evaluation of currently applied molecular biology methods in the field of barcoding. It gives a critical view concerning sensitivity, reliability, reproducibility and running costs. Recent achievements in the next generation technologies are briefly outlined.

70. Hamilton JP, Buell CR: **Advances in plant genome sequencing**. *Plant J* 2012, **70**:177-190.

71. Pillon Y, Johansen J, Sakishima T, Chamala S, Barbazuk WB, Roalson EH, Price DK, Stacy E: **Potential use of low-copy nuclear genes in DNA**. *BMC Evol Biol* 2013, **13**:35.

72. Lindner DL, Carlsen T, Nilsson RH, Davey M, Schumacher T, Kauserud H: **Employing 454 amplicon pyrosequencing to reveal intragenomic divergence in the internal transcribed spacer rDNA region in fungi**. *Ecol Evol* 2013, **3**:1751-1764.

73. Kermarrec L, Franc A, Rimet F, Chaumeil P, Humbert JF, Bouchez A: **Next-generation sequencing to inventory taxonomic diversity in eukaryotic communities: a test for freshwater diatoms**. *Mol Ecol Resour* 2013 http://dx.doi.org/10.1111/1755-0998.12105. [Epub ahead of print].

74. Van Moerkercke A, Fabris M, Pollier J, Baart GJ, Rombauts S, Hasnain G, Rischer H, Memelink J, Oksman-Caldentey KM, Goossens A: **CathaCyc, a metabolic pathway database built from *Catharanthus roseus* RNA-seq data**. *Plant Cell Phys* 2013, **54**:673-685.

75. Marques JV, Kim KW, Lee C, Costa MA, May GD, Crow JA, Davin LB, Lewis NG: **Next generation sequencing in predicting gene function in podophyllotoxin biosynthesis**. *J Biol Chem* 2013, **288**:466-479.

76. Pyle BW, Tran HT, Pickel B, Haslam TM, Gao Z, MacNevin G, Vederas JC, Kim SU, Ro DK: **Enzymatic synthesis of valerena-4,7(11)-diene by a unique sesquiterpene synthase from the valerian plant (*Valeriana officinalis*)**. *FEBS J* 2012, **279**:3136-3146.

77. Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, Howard E, Shendure J, Turner DJ: **Target-enrichment strategies for next generation sequencing**. *Nat Methods* 2010, **7**:111-118.

78. Sahu SK, Thangaraj M, Kathiresan K: **DNA extraction protocol for plants with high levels of secondary metabolites and polysaccharides without using liquid nitrogen and phenol**. *ISRN Mol Biol* 2012, **205049**:6.

79. •• Särkinen T, Staats M, Richardson JE, Cowan RS, Bakker FT: **How to open the treasure chest? Optimising DNA extraction from herbarium specimens**. *PloS ONE* 2012, **7**:e43808.
The comparison of various DNA extraction methods from Herbarium samples often stored under DNA degrading conditions. PCR amplification success is measured by using three differentially sized markers. The shorter the PCR product, the higher the amplification success rate.

80. Shahzadi I, Ahmed R, Hassan A, Shah MM: **Optimization of DNA extraction from seeds and fresh leaf tissues of wild marigold (*Tagetes minuta*) for polymerase chain reaction analysis**. *Genet Mol Res* 2010, **9**:386-393.

81. Dehestani A, Kazemi Tabar SK: **A rapid efficient method for DNA isolation from plants with high levels of secondary metabolites**. *Asian J Plant Sci* 2007, **6**:977-981.

82. Pusch CM, Giddings I, Scholz M: **Repair of degraded duplex DNA from prehistoric samples using *Escherichia coli* DNA polymerase I and T4 DNA ligase**. *Nucleic Acids Res* 1998, **26**:857-859.

83. Taylor HR, Harris WE: **An emergent science on the brink of irrelevance: a review of the past 8 years of DNA barcoding**. *Mol Ecol Resour* 2012, **12**:377-388.

84. Ratnasingham S, Hebert PD: **Bold: the barcode of life data system**. *Mol Ecol Notes* 2007, **7**:355-364 http://www.barcodinglife.org.

85. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW: **GenBank**. *Nucleic Acids Res* 2013, **41**:D36-D42.

86. Hennell JR, D'Agostino PM, Lee S, Khoo CS, Sucher NJ: **Using GenBank® for genomic authentication: a tutorial**. *Methods Mol Biol* 2012, **862**:181-200.

87. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool**. *J Mol Biol* 1990, **215**:403-410.

88. •• Lou SK, Wong KL, Li M, But PP, Tsui SK, Shaw PC: **An integrated web medicinal materials DNA database: MMDBD (Medicinal Materials DNA Barcode Database)**. *BMC Genomics* 2010, **11**:402.
This article introduces a new database: Medicinal Materials DNA Barcode Database MMDBD for data retrieval and similarity search. The database contains information about the Chinese medicinal material including known adulterants, photographs, and primers used for obtaining the DNA barcodes.