

REVIEW

Open Access



# DNA barcoding: an efficient technology to authenticate plant species of traditional Chinese medicine and recent advances

Shuang Zhu<sup>1</sup>, Qiaozhen Liu<sup>1</sup>, Simin Qiu<sup>1</sup>, Jiangpeng Dai<sup>1</sup> and Xiaoxia Gao<sup>2\*</sup>

## Abstract

Traditional Chinese medicine (TCM) plays an important role in the global traditional health systems. However, adulterated and counterfeit TCM is on the rise. DNA barcoding is an effective, rapid, and accurate technique for identifying plant species. In this study, we collected manuscripts on DNA barcoding published in the last decade and summarized the use of this technique in identifying 50 common Chinese herbs listed in the Chinese pharmacopoeia. Based on the dataset of the major seven DNA barcodes of plants in the NCBI database, the strengths and limitations of the barcodes and their derivative barcoding technology, including single-locus barcode, multi-locus barcoding, super-barcoding, meta-barcoding, and mini-barcoding, were illustrated. In addition, the advances in DNA barcoding, particularly identifying plant species for TCM using machine learning technology, are also reviewed. Finally, the selection process of an ideal DNA barcoding technique for accurate identification of a given TCM plant species was also outlined.

**Keywords:** Traditional Chinese medicine, DNA barcoding, Authentication, Plant species

## Introduction

Traditional Chinese medicine (TCM), including Chinese herbal medicine, continue to receive international recognition. TCMs have been widely used in the traditional Chinese medical systems and diet therapy. At the same time, TCMs also play an important role in the global traditional health system, not only as food additives, but also as some of the bioactive medical ingredients, such as artemisinin and paclitaxel, etc., which have made a splash in the traditional herbal drug market [1, 2]. In the past decade, the global market for herbal products has expanded, and there has been an increase in the export and import of traditional medicinal products worldwide [3]. Especially, following the outbreak of coronavirus

disease 2019 (COVID-19) in 2019, the National Health Commission of the People's Republic of China recommended a combination of traditional Chinese, such as the Huoxiang Zhengqi capsule [4], Lianhua Qingwen capsule [5], among others, and Western medicine for treating the disease. According to the National Bureau of Statistics of China, the turnover of the Chinese herbal medicine market in 2019 reached 165.3 billion yuan for the domestic market and \$6.175 billion for the international side. The increased demand for natural products has created the need to ascertain the authenticity of TCMs' species.

The authentication of Chinese herbal medicine species began 5000 years ago. For conventional authentication, ancient people generally relied on the flowering and fruiting period of Chinese herbal medicine, as this period is easier and more convenient to authenticate. However, this method faces numerous problems: first, the conventional authentication is limited in species identification without relating to the quality of TCM. Secondly, the

\*Correspondence: gaouxia91@163.com

<sup>2</sup> School of Pharmacy, Guangdong Pharmaceutical University, Guangzhou 510006, China

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

required features are only visible during specific periods and need to be authenticated by experts with extensive personal experience. Currently, understanding plant and animal genetics has facilitated the invention of species authentication technologies. DNA barcoding has become an extremely widely used technology in molecular marker-based species authentication technologies, given its standardization, minimization, and scalability. DNA barcoding is now widely used for the rapid identification of TCM species.

In this study, we aim to: ① discuss the dynamics and application prospects of DNA barcoding and its derivative technologies; ② address the issue of accessing the optimal barcodes to authenticate common Chinese herbal species; ③ outline processes for selecting the appropriate technologies for identifying given traditional Chinese herbal species.

### Prevalent adulteration in current Chinese herbs industry

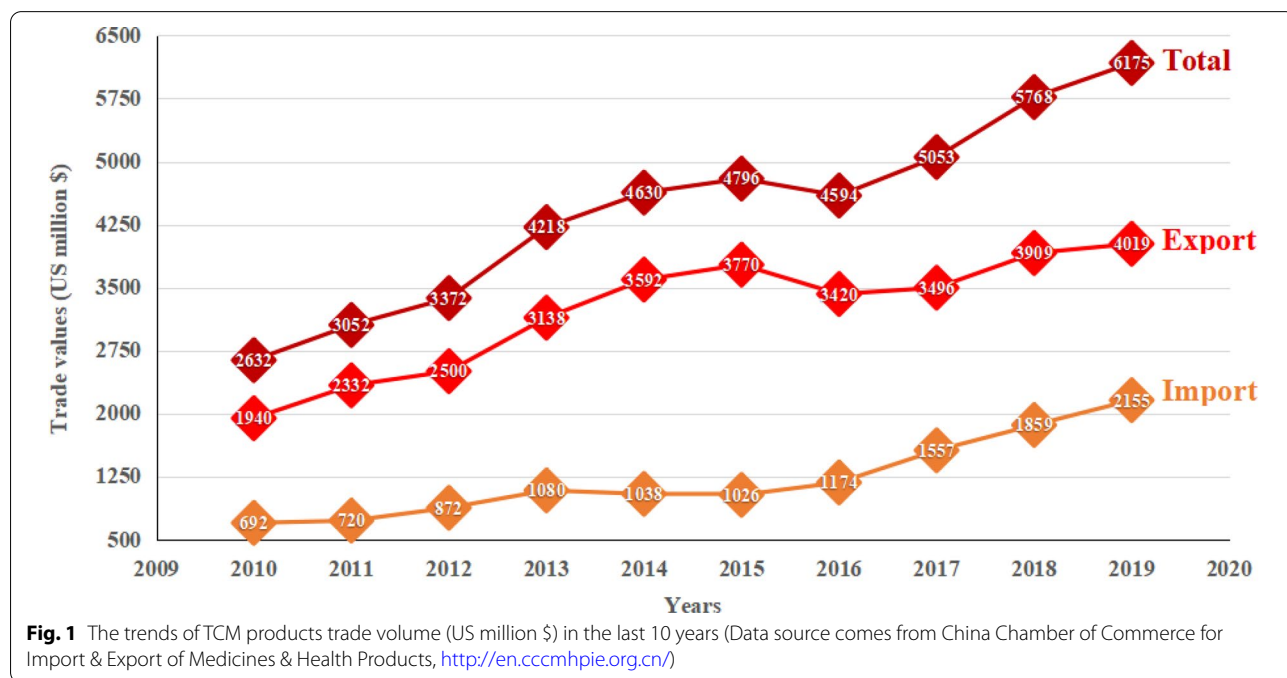
The TCM industry has rapidly expanded over the past years. Accordingly, the competition resulting from the growing demand for TCM is a key factor of concern. The trend of the TCM industry is shown in Fig. 1. Along with the growing TCM market, there has been an increase in poor quality/fake herbal products, as shown in Table 1. Han et al. investigated 1260 valid samples of 295 medicinal species from 7 TCM markets in China and found that about 4.2% were found to be adulterated [6]. The prevalent problem deteriorated in 2018. Another research

investigating 400 seeds for TCM products found that 7.5% of the samples were incorrectly labeled [7].

The emergence of fake and poor-quality TCM has been attributed to the profit-seeking businessmen who gained improper benefits from cheaper and more profitable adulterants with similar shapes or vernacular names that may lead to confusion in species identification, or during the manufacturing process [8–10]. In view of the above, either accidental or intentional, the emergence and increase in the number of fake TCMs on the market are alarming. This problem has an unpredictable impact on the subsequent clinical use and efficacy of Chinese herbal medicine and hinders the progress in the development of precision medicine. Therefore, there is an urgent need for rapid and simple inspection procedures for validating the authenticity of Chinese herbal materials.

### One of the solutions: origins and development of DNA barcoding

The identification of TCM has four major development stages, including sensory evaluation, microscopic identification, physical, and chemical identification (e.g. high-performance liquid chromatography (HPLC) [11]) and DNA-based molecular identification. The former three stages have some limitations in distinguish authentic from fake medicinal materials accurately. The efficacy of the authenticity of TCMs is affected by numerous factors such as the harvesting time, the complexity of the materials, and uncertain bioactive substances. To address this issue, researchers have gradually turned their attention to



**Table 1** Experiment evidence-based existing substitutes or adulterants for Traditional Chinese medical herbs

S.no	Latin name of Traditional Chinese medical herbs	Chinese name	Substitutes or adulterants	References
1	<i>Lonicerae japonicae Flos</i>	Jinyinhua	<i>Eucommiae Folium</i> or <i>Lonicerae Flos</i>	[100]
2	<i>Angelica sinensis</i>	Danggui	<i>Angelica amurensis</i>	[126]
3	<i>Pulsatilla chinensis</i>	Baitouweng	<i>Potentilla chinensis</i>	[127]
4	<i>Ficus hirta</i>	Wuzhimaotao	<i>Gelsemium elegans</i>	[128]
5	<i>Veronica officinalis</i> L	Popona	<i>Veronica chamaedrys</i> L	[11]
6	<i>Cynanchi Atrati Radix et Rhizoma</i>	Cangshu	<i>Ampelopsis japonica</i>	[38]
7	<i>Cynanchi Atrati Radix et Rhizoma</i>	Baiwei	<i>Cynanchum mongolicum</i> or <i>Cynanchum inamoenum</i>	
8	<i>Artemisia annua</i> L	Huanghuahao	other species of <i>Artemisia</i> (e.g. <i>Artemisia argyi</i> )	[78]
9	<i>H. rhamnoides</i> ssp. <i>Sinensis</i>	Shaji	<i>Nitraria tangutorum</i> / <i>Sorbus pohuashanensis</i> / <i>Berberis vulgaris</i>	[107]
10	<i>Zanthoxylum armatum</i>	Zhuyehuajiao	<i>Zanthoxylum schinifolium</i>	[129]
11	<i>Panax notoginseng</i>	Sanqi	<i>Panax vietnamensis</i> var. <i>fuscidicus</i>	[20]
12	<i>Hyoscyami Semen</i> (seeds of <i>Hyoscyamus niger</i> L.)	Tianxianzi	seeds from <i>Hygrophila salicifolia</i> (Vahl) Nees	[130]
13	<i>Cuscuta australis</i> R. Br. and <i>C. chinensis</i> Lam	Tusizi	<i>Cuscuta japonica</i> Choisy	[7]
14	<i>Atractylodis Rhizoma</i>	Cangzhu	<i>Atractylodes koreana</i> (Nakai) Kitamura	[62]
15	<i>Notopterygii Rhizoma et Radix</i>	Qianghuo	<i>Sanguisorbae Radix</i> (Diyu, <i>Sanguisorba officinalis</i> L.)	
16	<i>Tripterygum wilfordii</i>	Leigongteng	<i>Celastrus angulatus</i>	[131]
17	<i>Arisaematis rhizoma</i>	Tiannanxing	<i>Pinellia pedatisecta</i>	[34]
18	<i>Tinospora crispa</i>	Lübaoteng	<i>Tinospora baenzigeri</i>	[61]
19	<i>Rhizoma Paridis</i>	Chonglou	<i>Polygonum paleaceum</i> Wall	[132]
20	<i>Dipsacus asper</i>	Chuan-xuduan	<i>Dipsacus japonicus</i>	[37]
21	<i>Akebiae Caulis</i>	Mutong	<i>Aristolochiae manshuriensis Caulis</i> / <i>Clematis armandi</i>	[62]
22	<i>Clematidis Armandii Caulis</i>	Chuan-mutong	<i>Akebiae Caulis</i>	
23	<i>Alisma orientale</i>	Zexie	<i>Alisma plantago-aquatica</i>	
24	<i>Bupleuri Radix</i>	Chaihu	<i>Bupleurum marginatum</i>	
25	<i>Orthosiphon stamineus</i>	Maoxucao	<i>Clinacanthus nutans</i>	[60]
26	<i>Aquilaria</i> ( <i>Thymelaeaceae</i> )	Chenxiang	<i>Memecylon</i> sp. ( <i>Melastomataceae</i> ) and <i>Strychnos</i> sp. ( <i>Loganiaceae</i> )	[133]
27	<i>Berberis aristata</i>	Xiaobo	<i>Berberis asiatica</i>	[134]
28	<i>Ocimum sanctum</i>	Shengluole	<i>Vitex negundo</i>	[135]
29	<i>Bacopa monnieri</i>	Jiamachixian	<i>Centella asiatica</i>	
30	<i>Inulae Flos</i>	Xuanfuhua	<i>Inula linariifolia</i>	[136]
31	<i>Sophorae Flos</i>	Huaihua	<i>Robinia pseudoacacia</i>	

DNA-based molecular identification of medicinal herbs. DNA barcoding can be used for quality control of Chinese herbal medicine by validating the identity of the corresponding species.

The concept of applying DNA barcodes to identify species was first proposed by Hebert et al. in 2003 [12]. The technique was successfully used in animals and fungi by using the 5' end of the cytochrome oxidase I (COI) from the mitochondrial gene. COI barcode is a haploid, uniparentally-inherited, single-locus gene with high discriminatory power. The gene does not frequently display drastic length variation, strong secondary structure, micro-inversions, or frequent mononucleotide repeats in animals [13]. Combined with well-developed primer sets, the COI barcode method is easy to perform and

accurately identifies animal species. However, the COI barcode is unsuitable for plant identification because mitochondrial genes in plants are slowly evolving with very low substitution rates [14].

Researchers have turned their attention to chloroplast and nuclear genomes to find more powerful barcodes in plant species. In the last two decades, major standard single-locus candidate barcodes have been proposed: ITS, ITS2, *matK*, *rbcL*, *psbA-trnH*, and *trnL-trnE*, which discriminate plants species with high accuracy. However, it was found that a single barcode was not enough to identify all plants, which necessitated the use of multi-locus DNA barcodes. The Consortium for the Barcoding of Life (CBOL) Plant Working Group proposed a combination of *matK* and *rbcL* locus to enhance the accuracy of

species discrimination [14, 15]. Chen et al. then proposed the ITS2 + *psbA-trnH* for the DNA barcoding system for identifying botanical medicinal herbs [16].

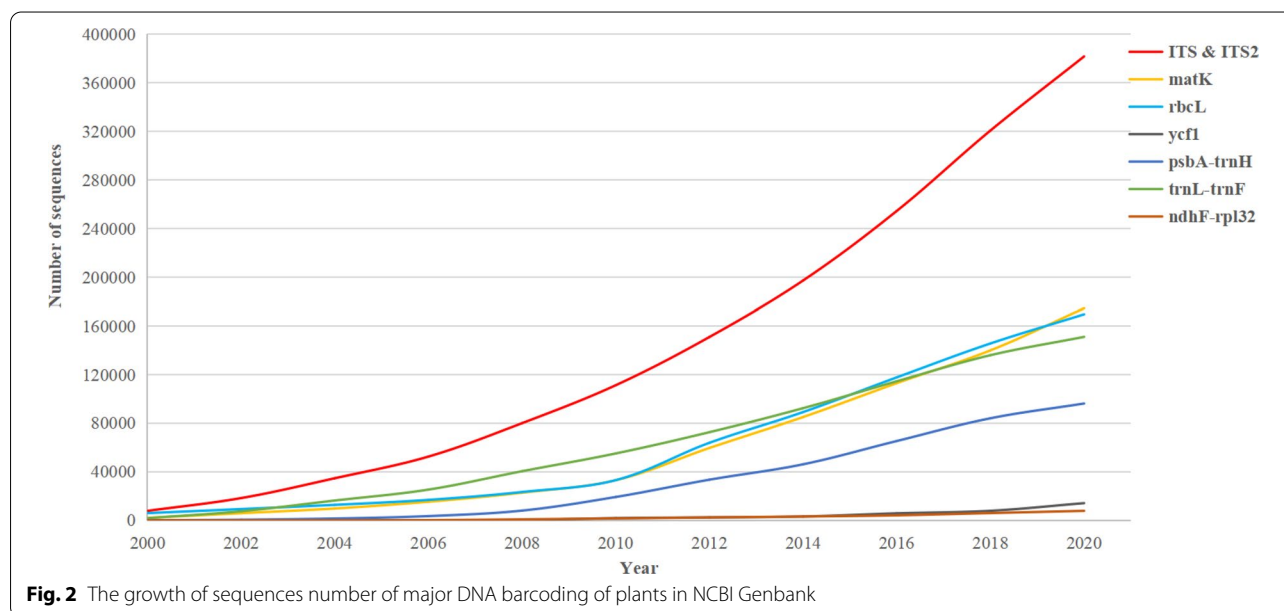
The invention of next-generation sequencing (NGS) technology and the emergence of the third-generation sequencing technology have further enhanced the development of the DNA barcode-derived technologies in identifying Chinese herbal medicine species. The current DNA barcode derivative techniques include super-barcoding, meta-barcoding, and mini-barcoding. For example, (i) mini-barcoding can identify species from highly degraded DNA [17, 18]; (ii) meta-barcoding is useful for species richness analysis in a sample containing a mixture of species [19, 20]; (iii) super-barcoding based on plant chloroplast genome is used for species relatedness [21, 22]. The advance in sequencing technology facilitated the improvement of DNA barcoding from detecting a single herb in Chinese medicine to simultaneously detecting several herbs in a Chinese herbal medicine cocktail [23], influencing the selection and utilization of DNA barcoding. These three DNA barcoding-based technologies have broadened the applications and enhanced the practicality of DNA barcoding. Data mining and analysis tools have strengthened the application of DNA barcoding-based technologies, which could effectively identify biological systems in Chinese herbal medicines [24].

Although DNA barcoding technology's accuracy is increasing daily, this technique faces numerous challenges, such as inadequate standard reference libraries, low success rate of PCR amplification and PCR bias. Despite these problems, the application of DNA barcoding is rising due to its easy operation, high identification

success rate and repeatability. For the quality control technology of TCM materials, especially in plant species of TCM, single technology identification of TCM materials and Chinese patent medicines (CPMs) has certain one-sidedness and, thus, combining several technologies is required [23]. Therefore, we recommend combining several identification methods to achieve comprehensive and accurate identification of TCM with DNA barcoding.

### Standard single-locus DNA barcoding

DNA barcoding technology has been used in TCM identification. The number of publications and sequences of different barcodes are rapidly increasing. According to the CBOL Plant Working Group and the number of publications on DNA barcoding between 2010 and 2020, ITS and ITS2, *rbcL*, *matK*, *trnL-trnF*, *psbA-trnH*, *ycf1*, and *rpoC1* are the seven major plant barcodes that have attracted the most attention [25]. Based on the number of DNA barcode sequences in the NCBI database collected (Fig. 2), we found that: (i) ITS and ITS2 are the predominant barcodes. Since 2010, the number of ITS and ITS2 barcodes have been booming. ITS2 region can not only discriminate plant taxa from different plant families but can also distinguish closely related taxa at the genus and species levels [16, 26]. Accordingly, ITS and ITS2 sequences should be utilized more in the future; (ii) *rbcL* (179,816 items), *matK* (174,431 items), and *trnL-trnF* (159,360 items) are the second most dominant barcodes, possibly because they can be used as multi-locus DNA barcodes. The number of publications on *trnL-trnF* has increased to 10,000 in 2021; (iii) reasons assumed for the slow growth of *rpoC1* (15,387 items) and *ycf1* (16,344



items) might be attributed to the long gene sequences (5709 bp for the *ycf1* gene of *Nicotiana tabacum*) and lower discriminatory power [27–29]. Several genes, including *atpF-atpH*, *ndhF-rpl32*, and *psbK-psbI*, are potential barcoding candidates. These targets are not so popular in recent publications (less than 1%), probably because of their relatively low discrimination ability, poor universality in different taxa, or unsatisfactory amplification rates [28, 29]. Generally, the sequence number of standard single-locus DNA barcodes is still increasing.

### ITS and ITS2

The internal transcribed spacer (ITS) region of the nuclear ribosomal cistron is the most usually sequenced locus for systematic molecular investigations of TCM at the lower-taxa levels, including the genera, species, and subspecies [30]. ITS offers the advantages of generality, simplification, high copy number, interspecific variability, and intraspecific uniformity [31]. ITS has been used as a universal barcode for distinguishing more than 21,722 plant species and is recommended for validating the authenticity of Chinese herbal medicine [32]. However, certain limitations hinder its application for Chinese herbal medicine barcoding: incomplete concerted evolution as well as difficulties of amplification and sequencing [29]. ITS2, a non-coding nuclear DNA between 5.8S rRNA and 25S rRNA genes, can distinguish closely related taxa at the family, genus, and species levels [26, 33].

ITS2 has strengths in variability, sequence quality and high inter-specific and intra-specific divergence power [16, 26, 34]. ITS2 can identify 92.7% of species correctly in more than 6600 samples obtained from 4800 species in 753 genera [16, 26], such as *Cynanchum auriculatum* [35], *Acanthopanax* [36], *Dipsacales* [37], *Xueteng* [38]. Besides, the secondary structures of ITS2 provided additional information that enhances the species' discrimination [39–42]. ITS2 could be used as an alternative mini-barcoding when a full-length ITS is not available and can correctly identify *R. rosea*. [43], and *U. lanosa* [42], among many other species. Currently, effective experimental methods have been developed to avoid fungal contaminants. The Hidden Markov Model (HMM) fungus model proposed by the Florida State University can remove fungal contaminating sequences, enhancing the reliability of the data. Meanwhile, the risk of fungal contamination can be effectively reduced by cleaning the surface of herb roots and scraping off the cortex during sampling. ITS2 could be used to identify herbs in a broader range of plant taxa [26, 44–49], including herbarium specimens with degraded DNA [50]. Accordingly, it is suitable for authentication of traditional Chinese herbal medicine powder.

Although ITS2 has many strengths, it is not ideal for identifying ferns [51, 52]. A major concern is the existence of multiple copies in ITS2 with high levels of within-species and even within-individual sequence differentiation [53]. Furthermore, heterogeneity is an issue for ITS2 due to concerted evolution, which may lead to inaccurate or misleading results [54, 55].

### matK

The high sequence variation and sequencing efficiency rates, evolution, PCR amplification, suitable sequence length, accurate discrimination of angiosperms [53, 56, 57], and the intra and inter-specific divergence distinction in the barcoding gap [58] indicate *matK* is a useful DNA barcode for plants. This barcode has been used for nearly 5 years to accurately identify *Paeonia suffruticosa* [59], *Veronica officinalis* [11], etc. Despite this, there is a need to develop universal primers for the identification of plant species.

### rbcl

As one of the best potential barcode candidates, *rbcl* can discriminate plants at the family and genus level [60]. The remarkable advantages of *rbcl* are high primer versatility, easy amplification and alignment, and high discrimination power [25]. Recent studies used this barcode to identify plants in *Tinospora* [61], *Aceraceae* [62], and *Artemisia* [30] genera, among others.

*rbcl* has a relatively low interspecific identification power and is generally used for genetic variation tests. As a separate candidate sequence, it is unsuitable due to this region evolves slowly, implying that its discriminatory power is restricted [33]. Recently, researchers have indicated that poor discrimination of closely related species limits its utility in detecting ingredient substitutions [62], indicating that it should be used alongside other potential barcodes.

### psbA-trnH

The *psbA-trnH* barcode, one of the fastest evolving regions in the chloroplast genome, is the interval between both *trnH* (H-GUG) sequence ends and both sides of the *psbA* gene. Usually, *psbA-trnH* has better primer universality, a relatively high amplification success rate, and is of good length. Therefore, it can be used to amplify biodegraded samples. These features are especially suitable for the level of species and the higher taxonomic level [63, 64]. *psbA-trnH* regions can accurately discriminate members of *Dendrobium* [65] and medicinal pteridophytes (90.2% of species could be accurately identified) [66], and *Mentha haplocalyx* [67].

Meanwhile, due to the repeated loci, pseudogenes, and high insertion/deletion rate, the length of *psbA-trnH* vary

significantly in different groups [28, 68]. As such, manual correction is required for *psbA-trnH* sequence analysis, making it difficult to compare different genera and species.

### trnL-trnF

The *trnL-trnF* region is located in the large single-copy region of the chloroplast genome, which consists of the *trnL* gene and the *trnL-trnF* intergenic spacer [69]. The *trnL-trnF* region has been considered for accurate discrimination of plants at the lower taxonomic levels. The region has a high nucleotide conversion rate, which causes a relatively high genetic variation and provides sites with more systematic taxonomic information. The *trnL-trnF* region has been used in systematic taxonomic studies of the *Elytrigia lolioides* [70], the Apocynaceae [71], and *Radix et Rhizoma Rhei* [72], among others. Although mononucleotide repeats can impact sequencing reads in some taxa, this barcode is generally simple to sequence [29].

### Other standard single-locus DNA barcoding

Besides standard single-locus DNA barcoding mentioned above, many other DNA sequences, including *ycf1* [73], *rpoC1* [28], *ycf5* [26], *accD* [28], *ndh* [28], and *ndhF-rpl32* [74] have been used for identifying Chinese herbal medicine. This DNA barcoding mentioned above is absent in some major groups of land plants. For instance, *ycf1* is absent in *Poaceae* [27], whereas *ndh* is absent in *Pinus* [75], or it just has lower discriminatory power [25]. Therefore, they are not widely considered accurate plant standard barcodes for identifying Chinese herbal medicine [76].

### Multi-locus DNA barcoding

Several studies have demonstrated the difficulties of discriminating between all plants using a universal DNA barcode [77, 78]. Conflicting results have sometimes been found for related species when using certain barcodes, whereas a single locus barcode does not sufficiently provide the evolutionary distinctions required to distinguish related species. Considering the requirements for accurate discrimination and satisfactory genetic information, multi-locus DNA barcoding is more preferable. Multi-locus DNA barcoding is gradually being accepted for accurate identification of TCM.

Multi-locus DNA barcoding represents a practical solution to reach a trade-off between universality, sequence quality, discrimination, and cost. At first, Kress et al. suggested that ITS + *psbA-trnH* have the potential to discriminate against numerous plant species [33]. The CBOL Plant Working Group evaluated seven chloroplast genomic regions and proposed the 2-locus *matK + rbcL*

plant barcode in an international conference since *matK* provides high resolution but less universality, whereas *rbcL* provides high universality but less species resolution [25]. Researchers believed combining these two barcodes could achieve maximum species discrimination [29]. To achieve higher discrimination in closely related species, the China Plant BOL Group proposed to add the nuclear ITS (internal transcribed spacer) to the *matK + rbcL* combination [79]. Chen et al. first proposed the ITS2 sequence as a universal barcode for medicinal plant identification and the ITS2 + *psbA-trnH* combination as a DNA barcoding system for identifying botanical medicinal herbs [26]. The advantages of multi-locus barcoding are that the results can be mutually verified and complemented and can discriminate among numerous species. This combination demonstrated the excellent reliability for species authentication, and researchers have identified more than 23,262 different species for Chinese, Japanese, Korean, and European herbal medicine [36, 79]. Among the top ten Chinese herbal medicine and decoction of processed materials exported in 2019, five were identified using multi-locus barcoding: *Pinellia hunanensis* using *matK + rbcL* [29], *Panax ginseng* C.A. Meyer and *Radix Astragali* using *psbA-trnH* + ITS [26, 81], *Zizyphus jujube* using ITS2 + *psbA-trnH* [82], *Angelica sinensis* using ITS + *rbcL + matK + psbA-trnH* (slightly better discriminatory power than ITS) [27].

Although it still failed to meet the original goal of the universality of DNA barcoding and the differentiation of closely related complex groups is still uncertain, the multi-locus approach of combining different barcodes has been successful in certain cases, including species discrimination [28, 29]. In general, the discrimination of Chinese herbal medicine species using DNA barcoding is still under research and development.

### Super-barcoding

In 2008 at the Botany without Borders conference, it was pointed out that the chloroplast genome contains about as much information as the short mitochondrial barcode sequence used in animals [83]. With the need for accurate identification of certain closely related species, scholars proposed the concept of super-barcoding (ultra-barcoding), which means sequencing the whole plastid genomes as a barcode [83]. Here, the whole organelle's genome or large (greater than 5 kb) contiguous portions of the nuclear genome are sequenced and assembled [21]. Compared with the nuclear genome, the chloroplast genome is smaller and has a higher interspecific and lower intraspecific divergence [53]. Therefore, sequencing the chloroplast genome is more common.

Super-barcoding is a promising approach for identifying Chinese herbal medicine and has many advantages,

including ① circumventing gene deletion problems, locus choice, and low PCR recovery rate often encountered in the conventional barcoding [84], ② higher resolution, and better versatility [21], and ③ can be supplemented the traditional DNA barcoding. Compared with traditional barcoding, super-barcoding enhances the identification of closely related groups, including accurate discrimination of subspecies. For instance, the super-barcoding was shown to successfully distinguish closely related species such as *Araucaria* spp. (Araucariaceae) [85] and *Echinacea* (Asteraceae) [86], especially for taxonomically complex groups, e.g., *Camellia* spp. (Theaceae) [87], Chinese herbal medicine *Epimedium* spp. (Berberidaceae) [88], *Fritillaria* spp. (Liliaceae) [89] and *Taxus* (Taxaceae) [84]. Super-barcoding often uses high-throughput next-generation sequencing (generally in massively parallel sequencing) to scan the genome and generate a reliable sequence of high copy number regions. It gets more information sites and expands the traditional barcode regions (standard single-locus barcoding) to their full, many-kilobase length [21]. This method increases the density and phylogenetic coverage of the complete plastid genome sequence and is expected to accurately identify traditional Chinese herbal medicines.

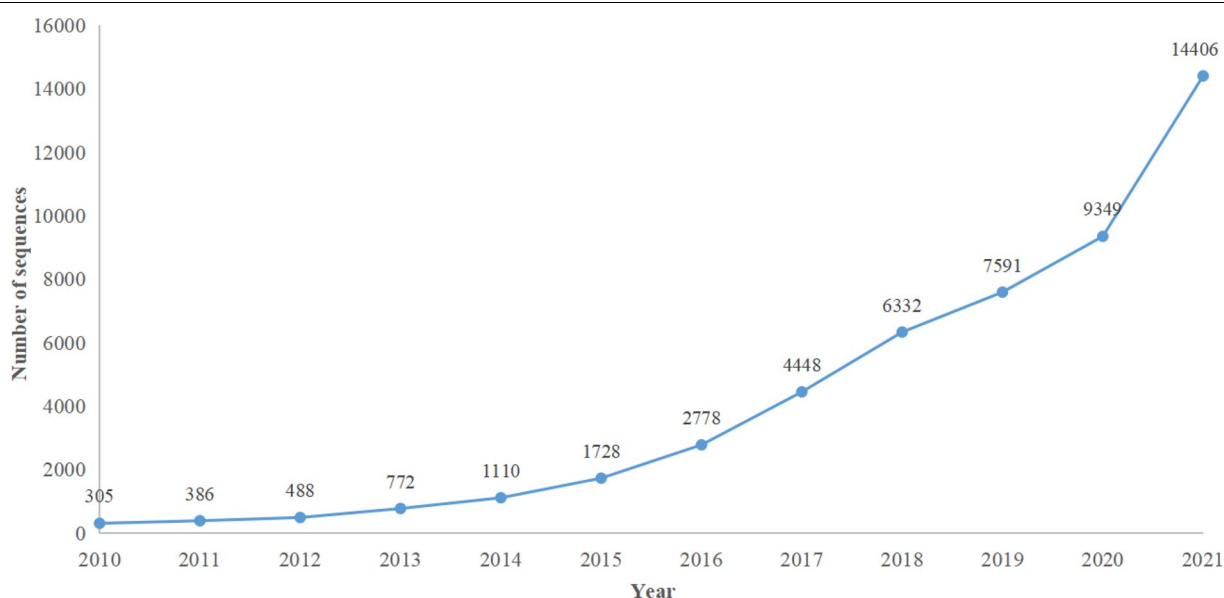
The main stumbling blocks for super-barcoding are the cost and the requirement for high quality and quantity of DNA, large next-generation sequence data generated as well as large amounts of next-generation sequence data needed to deal with [21]. Besides, the variation present over short regions may be too low to distinguish recently

diverged taxa because evolution is generally slow in the plastid genome [90].

With the increasing number of the whole chloroplast genomes in GenBank (Fig. 3), it is foreseeable that the super-barcoding application in TCM herbs will be wider than standard plant DNA barcoding in the coming years. Super-barcoding does not override the need for continued use of traditional barcode methods but rather provides necessary data to examine variation below the species level [21]. Continued advances in sequencing technology may make super-barcoding the choice for plant identification at the intra-species or population levels in the future [32].

### Meta-barcoding

Currently, a new DNA barcoding-based method for rapidly and simultaneously identifying numerous taxa (i.e., different Chinese medical herbs) in a single environmental sample (i.e., multi-ingredient traditional CPMs) has been developed. The emergence of DNA meta-barcoding has been facilitated by the availability of the next-generation sequencing platforms and the need for high-throughput taxon identification. In 2012, meta-barcoding was defined as “designate high-throughput multispecies (or higher-level taxon) identification using total but degraded DNA extracted from an environmental sample (i.e., soil, water, feces, etc.)” [91]. DNA meta-barcoding to identify samples include ① collecting mixed-species environmental DNA samples (obtain raw materials), ② sample processing (DNA extraction and PCR amplified sequences), ③ next-generation



**Fig. 3** The total number of complete plant chloroplast genome sequences to GenBank from 2010 to 2021

sequencing, ④ data analysis (obtain clean data and OTUs from raw data), and ⑤ species identification [92].

The greatest advantage of DNA meta-barcoding is its ability to identify every species in a complex sample or processed mixtures simultaneously. Even so, the application of DNA barcoding and conventional analytical methods are considerably limited [23]. The CPMs' components are complex, and the sample DNA is degraded seriously. Thanks to high accuracy, DNA meta-barcoding can measure the components of CPMs simultaneously with high coverage and, thus, override the aforementioned problems. Thus, meta-barcoding is increasingly used for detecting CPMs' components. For instance, an Australian team identified barcodes for CPMs, including animal and plant medicines, in the form of tablets, capsules, powders, and herbal teas [93]. The potential power of DNA meta-barcoding is the ability to reveal plant species diversity within processed products. For example, it has successfully identified *Veronica* species, and detected substitution or admixture of other *Veronica* species in *V. officinalis* herbal products [11]. The main medicinal plants in the CPMs, including *Lonicera japonica* Thunb., *Forsythia suspensa*, and *Angelica pubescens* have been identified using DNA meta-barcoding [94].

However, the potential applications of DNA meta-barcoding are limited by the PCR success rate and the considerable investment in building comprehensive taxonomic reference libraries [95]. Also, sequencing errors in high-throughput sequencing are still inevitable.

DNA meta-barcoding can simultaneously detect multiple species from complex samples and facilitates species diversity assessment in processed products, which is extremely important for validating the authenticity of products in Chinese medicinal plants [23]. Therefore, this method can rapidly and accurately identify TCM, including Chinese herbal medicine. However, meta-barcoding should be used in combination with other appropriate chemical methods.

### Mini-barcoding

Due to the common DNA degradation in TCM, it is difficult to obtain the full-length sequence data using the traditional standard barcodes. Mini-DNA barcoding technology can override this limitation. Mini-barcoding can utilize incomplete, relatively short sequences from standard DNA barcodes to identify different species, which is useful for degraded DNA preservation. Overall, it improves the identification accuracy of species [96, 97]. One of the most common mini-barcode regions is *trnL* (UAA) intron. The P6 loop of the chloroplast *trnL* (UAA) intron can be robustly amplified with highly conserved primers from degraded DNA samples [95, 98]. Therefore, it can be used to identify the components in processed

medicinal materials up to the species or genus level [18]. Other common mini-barcoding regions include the shorter *ycf1a* and *ycf1b* [99], short region in ITS2 [100–102], and short region in *rbcL* [103].

Mini-barcoding has been successfully used to identify traditional Chinese herbal ingredients such as *Angelica sinensis radix*, *Ligusticum sinense*, and *Notopterygium incisum*, among others [102]. Currently, it has been applied to identify the traditional medicinal plant *Rhodiola* (Crassulaceae) [43], distinguish members of the *Apiaceae* family [104], and discovery of numerous species in *Metazoa* [105] and more natural herbal products [17]. Nonetheless, the few nucleotides often limit taxonomic discrimination using mini-barcoding, resulting in the main limitation of mini-barcoding being the resolution [97, 106]. An acceptable resolution not only depends on the accurate species identification but also on whether reference sequence data is sufficient. To fully maximize the power of mini-barcoding, more reference sequences need to be added to the databases.

Due to the shorter molecular markers of mini-barcode, different physicochemical technologies can be combined to identify Chinese herbal samples rapidly. For example, sea buckthorn (*Hippophae*) were accurately identified in Chinese herbal products using a combination of mini-barcoding and high-resolution dissolution (HRM) [107]. In the future, Mini-barcoding may become a complementary barcoding technique to identify traditional Chinese herbal medicine [18].

### Applications of the current DNA barcoding techniques for authenticating Chinese herbal medicine

We selected 50 common Chinese herbal medicines in the Chinese pharmacopoeia based on the published papers on TCM and DNA barcode identification in recent years. The barcode choices are shown in Table 2. We found that DNA barcoding has been used for large-scale identification of Chinese herbal medicines. We also summarized the preferred barcodes for different families or genera based on published papers (Table 3). Each species has a specific most ideal barcode, called “specific barcode”. A specific barcode may include one of the single-locus barcodes (e.g., *matK* or *psbA-trnH*) or could be based on new markers never used before [52]. Tables 2 and 3 summarize the recent developments in DNA barcoding for identifying Chinese herbal medicine species and the preferred DNA barcode for specific plants.

In recent years, with the continuous development of high-throughput sequencing technology and DNA barcode research, genomics is increasingly being applied to identify Chinese herbal medicine. Genome capture of nuclear markers has attracted researchers' attention,



**Table 2** Applications of DNA barcoding and DNA sequence-based markers in the identification of Chinese herbs

Name	Plant parts	Chinese herbs name (Pinyin)	Barcoding markers used	References
<i>Ephedra</i>	–	mahuang	ITS2	[80]
<i>Salvia miltiorrhiza</i>	Root and rhizome	danshen	ITS2	[51]
<i>Angelica sinensis</i>	Root	danggui	ITS2 and <i>psbA-trnH</i>	[102]
<i>Bupleurum</i>	Root	chaihu	ITS2	[137]
<i>Eucommia</i>	Bark	duzhong	ITS2	[138]
<i>Notoginseng</i>	Root	sanqi	ITS2	[139]
<i>Szechwan lovage</i>	Rhizome	chuanxiong	ITS2 and <i>psbA-trnH</i>	[140]
<i>Schisandra</i>	Fruit	wuweizi (beiwuweizi)	ITS2	[141]
<i>Atractylodes</i>	Rhizome	baizhu	ITS2	[142]
<i>Astragalus mongholicus</i>	Root	huangqi	ITS2	[143]
<i>Baical skullcap</i>	Root	huangqin	<i>psbA-trnH</i>	[144]
<i>Angelica dahurica</i>	Root	baizhi	ITS	[145]
<i>Isatis</i>	Root	banlangen	ITS2	[146]
Peony root, white	–	baishao	ITS2	[59]
Peony root, red	–	chishao	ITS2	[59]
<i>Carthamus tinctorius</i> L.	Flower	honghua	ITS2	[147]
<i>Coptis chinensis</i> Franch	Rhizome	huanglian	ITS2	[26]
Cape jasmine	Fruit	zhizi	<i>matK</i>	[120]
<i>Codonopsis</i>	Root	dangshen	ITS + <i>matK</i>	[32]
<i>Magnolia officinalis</i>	Bark	houpo	ITS2	[138]
<i>Uncaria</i>	Stem	gouteng	ITS and ITS2	[42]
<i>Platycodon</i>	Root	Jiegeng	ITS2	[148]
<i>Angelica pubescens</i>	Root	duhuo	ITS	[32]
<i>Thomson kudzuvine</i>	Root	fenge	ITS2	[26]
<i>Kudzuvine</i>	Root	gegen(yege)	ITS2	[26]
<i>Gastrodia</i>	Rhizome	tianma	<i>matK</i>	[120]
<i>Evodia</i>	Fruit	wuzhuyu	ITS2	[47]
<i>Fleeceflower</i>	Root	heshouwu	<i>matK</i> + <i>rbcl</i> + <i>psbA-trnH</i> + ITS2 or <i>psbA-trnH</i>	[32]
<i>Coix</i>	Seed	yiyiren	ITS2	[7]
<i>Andrographis</i>	–	chuanxinlian	<i>matK</i>	[120]
<i>Lightyellow sophora</i>	Root	kushen	ITS2	[147]
<i>Achyranthes bidentata</i>	Root	niuxi	ITS2	[26]
<i>Anemarrhena asphodeloides</i>	Rhizome	zhimu	<i>trnL-trnF</i>	[32]
<i>Akebia</i>	Stem	mutong	ITS2	[42]
<i>Aucklandia</i>	Root	muxiang	ITS2	[26]
<i>Atractylodes lancea</i>	Rhizome	cangzhu	ITS2	[42]
<i>Lycium barbarum</i> L.	Fruit	gouqizi	ITS	[42]
<i>Corydalis</i>	Rhizome	yanhusuo	ITS2	[26]
<i>Typhae</i>	Pollen	puhuang	ITS2	[147]
<i>Polygonum cuspidatum</i>	Rhizome and root	huzhang	<i>psbA-trnH</i>	[32]
<i>Moutan</i>	Bark	mudanpi	ITS2	[1]
<i>Drynaria</i>	Rhizome	gusuibu	<i>psbA-trnH</i>	[32]
<i>Prunella vulgaris</i> L.	Rhizome and root	xiakucao	ITS2	[51]
<i>Amomum</i>	Fruit	sharen	ITS2	[147]
<i>Belamcanda chinensis</i>	Rhizome	shegan	<i>rbcl</i>	[149]
<i>Piper longum</i> L.	–	bibo	<i>matK</i> + ITS	[149]
<i>Sophora</i>	Flower	huaihua	ITS2	[33]
<i>Piper Nigrum</i> L.	–	hujiao	<i>matK</i> + ITS	[150]
<i>Zanthoxylum bungeanum</i>	Pericarp	huajiao	ITS2	[150]

**Table 2** (continued)

Name	Plant parts	Chinese herbs name (Pinyin)	Barcoding markers used	References
<i>Clematis armandii</i>	Stem	chuanmutong	ITS2	[52]

**Table 3** Preferred loci for family or genera level in plants identification

	Family	References	Family	References	Genus	References	Genus	References
ITS and ITS2	<i>Verbena officinalis</i>	[151]	Caprifoliaceae	[152]	<i>Begoniaceae</i>	[153]	<i>Crassulaceae</i>	[141]
	Rubiaceae	[154]	Apiaceae	[140]	<i>Paris</i>	[155]	<i>Ilex</i>	[156]
	Rutaceae	[157]	Euphorbiaceae	[49]	<i>Amomi Fructus</i>	[158]	<i>Pulsatilla</i>	[127]
	Rosaceae	[49]	Asteraceae	[46]	<i>Viburnum</i>	[159]	<i>Panax</i>	[81]
	Malvaceae	[160]	Aristolochia	[161]	<i>Astragalus</i>	[162]		
	Zingiberaceae	[163]	Gentianaceae	[164]				
<i>matK</i>	Fabaceae	[119]	Juglandaceae	[165]	<i>Gardenia</i>	[125]	<i>Ardisia</i>	[58]
	Araceae	[166]						
<i>psbA-trnH</i>	Polygonaceae	[167]	Lauraceae	[168]	<i>Rhododendron</i>	[169]	<i>Aconitum</i>	[170]
					<i>Cistanche</i>	[171]	<i>Dendrobium</i>	[65]
<i>ycf1</i>	Cymbidium	[27]						
super-barcoding	Asteraceae	[53]	Schisandra chinensis	[172]	<i>Aconitum</i>	[173]	<i>Dendrobium officinale</i>	[174]
	Cymbidium	[87]						
<i>psbK-psbI</i>					<i>Dendrobium</i>	[175]		

and the genome skimming approach can bridge the gap between the standard barcode and genome sequencing [108]. Research on TCM genomics with TCM original species has achieved tremendous success [109–111]. However, the huge workload posed by data processing and sequencing cost is significantly higher than the cost of common barcode sequencing. It is not necessary to use genomics to identify plant species of TCM.

Regarding data mining, some studies suggest that machine learning methods can identify species using DNA multi-locus barcoding or just standard single-locus barcoding [112, 113]. Machine learning is based on building algorithms that receive input data for calibration and statistical analysis of the output value within an acceptable range. The common DNA barcode analysis methods in machine learning include BLOG (Barcoding with LOGic) and WEKA. Currently, eight *Dalbergia* timber species use SMO, a classifier, as part of the WEKA approach [114–116]. This approach resulted in the best (98–100%) discrimination, and the two-locus combination of ITS2 + *psbA-trnH* showed the highest success rate [112]. The character-based DNA barcode method in BLOG 2.0 was applied to classify members of the *Epidendrum* genus. It was found that *psbA-trnH* + ITS and *psbA-trnH* + ITS + *rbcl* exhibited the highest identification ability [117]. Machine learning and DNA barcoding technology are intertwined in two different fields. With

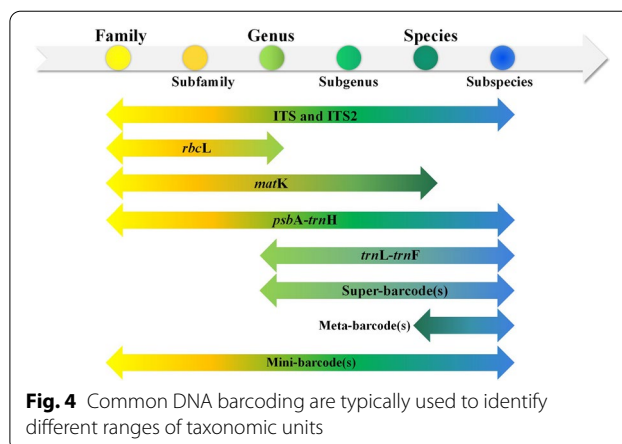
the help of machine learning, the application of DNA barcoding technology in the identification of TCM will be strongly promoted in the future.

The increasing use of DNA barcoding is due to the emergence of more available sequence data and information for machine learning and the regular update of public DNA barcode databases. Currently, DNA barcoding is widely used in authenticating medicinal materials in TCM, inseparable from the continued development of public barcode databases. As one of the most common databases, Chen et al. constructed a large-scale DNA barcode platform (<http://www.tcmbarcoding.cn>), widely used to identify herbal materials for varied needs [79]. This database is a collection of barcode sequences for herbs, including Chinese, Japanese, Korean, Indian, and European pharmacopeia species [7, 62, 80, 118, 119]. This reliable system for DNA barcoding of herbal materials has been established based on a two-locus combination of ITS2 + *psbA-trnH* loci barcode and contains 78,847 sequences for 23,262 species. To be specific, this platform has been used in TCM enterprises for raw herbal material identification [32]. This greatly speeds up the industrial procurement of raw materials and provides a standardized method for industrial identification of Chinese herbal medicine. That aside, a library of genuine Lingnan medical herbs DNA barcodes based on ITS2 has been constructed, containing 1276 sequences from 309

species from southern China [51]. It is used to identify genuine Lingnan medical herbs and the authenticity of the constituent ingredients, improving the standard of the Chinese medicine market. The Chinese University of Hong Kong built a Medicinal Materials DNA Barcode Database (MMDBD, <http://www.cuhk.edu.hk/icm/mmdbd.htm>), encompassing other barcodes such as *rbcL* in seed plant species [120], ITS2 + *psbA-trnH* [32, 43], and *rbcL* + *psbA-trnH* [60, 120]. All these public DNA barcode databases provide a platform for identifying TCM plant species. It is vital to update and maintain a public, standard DNA barcode database. Besides, good practice protocols are needed to ensure such databases provide clear information in this respect [32].

The development of new apparatuses in recent years has also made this technology more practical. Based on the need to automate the identification of TCM, digitize and integrate the identification of herb-based species, the new Chinese herbal DNA barcoding high-throughput gene sequencing machine (HMBI-G30) was developed successfully. This new apparatus can test up to 30,000 samples in a single run with high accuracy and reliability, facilitating one-stop sequence processing. Meanwhile, high-curvature nanostructuring-based electrochemical herb sensor (nanoE-herb sensor) is a direct, sequencing-free method for identifying herbal species accurately [121]. The use of such portable and cheap sensors facilitates rapid identification of other plant species in herbal medicines. NanoE-herb sensor has been for the ITS2 sequence to accurately identify herbal *C. sativus* in a mixture of counterfeit products. The continuous innovation of new instruments based on the DNA barcode principle has facilitated the identification and standardization of Chinese herbal medicine.

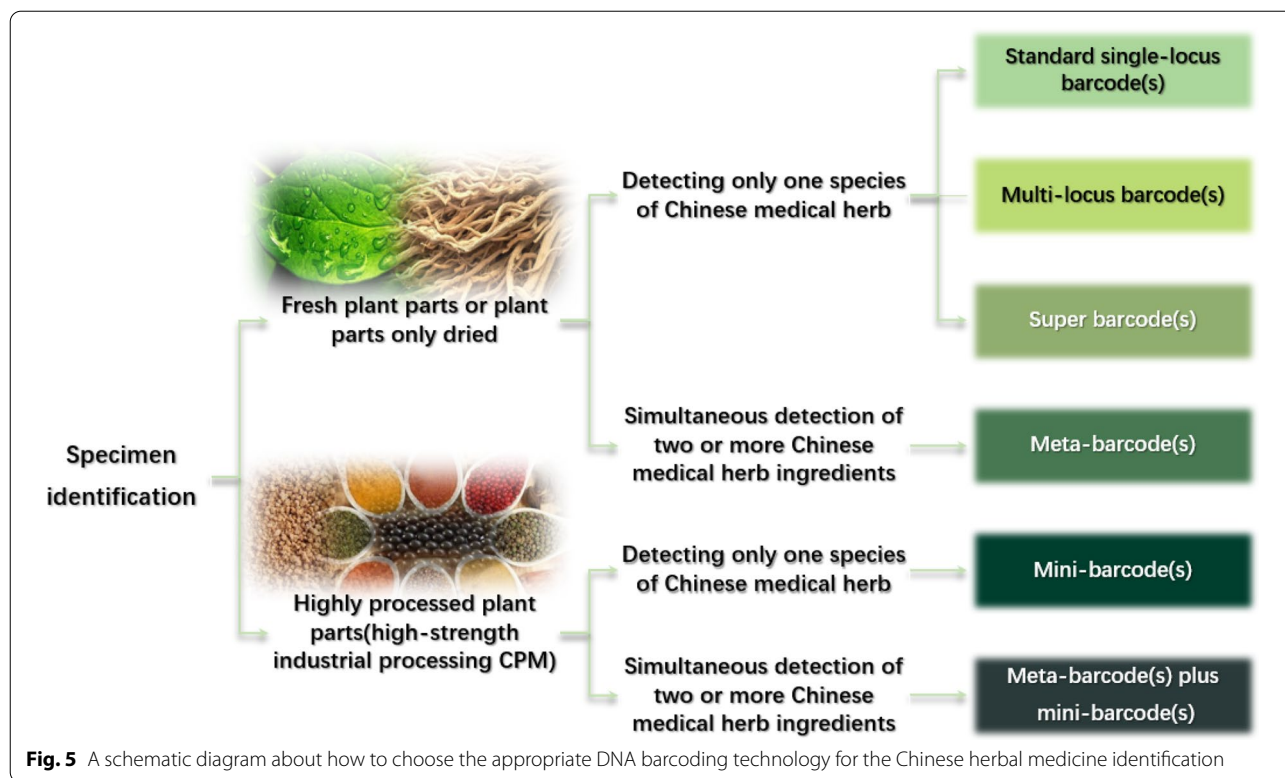
Therefore, we hold the following views regarding the application prospect of DNA barcoding and its derived technologies. From specific species to families and genera, our conclusion is captured in Figs. 4 and 5. Overall, a common DNA barcode can be used for organisms at different taxonomic units (Fig. 4). Since the standard single-locus barcodes *ycf1* and *rpoC1* are ambiguous as described in the papers, they are generally used in combination with other barcodes in the multi-locus barcoding approach. Given that they are not used alone, they are not listed in Fig. 4. It only shows the ranges of common applications but does not exclude the possibility that some barcodes have higher or lower accuracy in identifying certain species or members of a given genus. For the relatively new barcoding technologies, the super-barcoding and meta-barcoding have high accuracy and resolution in the identification of species at lower taxonomic levels. Super-barcoding and



**Fig. 4** Common DNA barcoding are typically used to identify different ranges of taxonomic units

meta-barcoding are rarely used for primary screening but for verification or validation of doubtful results generated by the conventional standard single-locus barcoding or multi-locus barcoding techniques. The mini-barcoding has gained wider recognition because overly degraded DNA is difficult to identify using conventional single-locus barcoding or multi-locus barcoding techniques. Therefore, mini-barcoding is directly used for identifying plant taxonomic groups and, literally, the classification range is wider.

Based on the ranges of application of DNA barcoding shown in Fig. 4 and the characteristics of each derivative technology summarized above, we provide a schematic procedure for selecting the ideal DNA barcode for identifying Chinese herbal medicine (Fig. 5). In this diagram, the high processing includes but is not limited to injections, pills, tablets, granules, powders, plasters, capsules, and other dosage forms. Traditional DNA barcoding is preferred to identify TCM herbs. It is recommended to use the traditional standard simple-locus barcoding in a single sample. If this method fails and accuracy is needed, then super-barcoding should be applied. Meta-barcoding is the technique of choice for the simultaneous identification of multicomponent samples. Overall, meta-barcoding and super-barcoding have become more and more common for identifying species in Chinese herbal medicine [22]. Research on mini-barcoding has broadened the application of DNA barcodes and has broadened the prospects for identifying Chinese herbal medicine materials from highly degraded DNA [99]. However, it is hard to identify all components in Chinese medicinal materials simultaneously using only the mini-barcoding [17]. Despite this, a combination of meta-barcoding and mini-barcoding has become a new trend for identifying proprietary Chinese herbal medicine, which has greatly promoted analyzing the composition of CPMs.



### Future perspectives in DNA barcoding for validating the authenticity of TCM

DNA barcoding and its derivative technologies in combination with other technologies (e.g., machine learning, electrochemical sensors, etc.) have achieved tremendous results in identifying Chinese herbal species. In the present paper, we summarized the development of DNA barcoding, both single and multi-locus barcoding widely used in validating plant species in TCM. Our research mainly focused on the potential development and application of DNA barcoding derivative technologies, including super-barcoding, meta-barcoding, and mini-barcoding. By carefully analyzing the application of the DNA barcoding derivative technologies, we developed a schematic procedure for selecting the ideal DNA barcoding technique for identifying given species in TCM. The DNA barcoding prospects and its derivative technologies were also suggested.

As sequencing technologies evolve, sequencing costs and error rates decrease, whereas the coverage and sensitivity in sequencing increase. Also, the speed of sequencing increases while the quality of data increases. However, it must be acknowledged that, given the complexity of the preparation of Chinese herbal medicines, DNA barcoding is not a panacea for validating the authenticity of TCM. Looking ahead, the following issues need to be refined to advance the development of DNA barcoding technologies: ① Sampling and classification: the sampling protocols

for DNA barcoding should be standardized. For example, the concept of Daodi medicinal materials has been compared to the “terroir” concept, which means that the specific herbs came from designated geographic regions where conditions including climate, soil, and technologies of cultivation in the case of plants [122, 123]. How can medicinal Daodi materials and non-medicinal Daodi counterparts be differentiated despite being sourced from the same species? ② With the development of NGS and its wide use, the DNA barcoding developments of Chinese medicinal materials are gravitating towards genomics, which will contribute to the development of herb genomics [124]. Can these DNA barcode-based technologies potentially upgrade from authenticity validation or detection of adulteration to authentication of herbal medicines’ quality based on epigenomics or epigenetics information? If molecular information like DNA methylation or histone modifications could help authenticate quality, it will widen the application of these DNA barcode-based technologies, which are essential for developing TCM precision medicine. In general, we advocate for the following: ① maintaining and updating the global plant DNA barcode library; ② updating the standardizing protocols for sampling and classifying process and ③ assessing the feasibility of combining genomics and biological technologies such as transcriptomics (specific expression subset analysis) and proteomics (specific proteome) [125]).

Recent reports and scientific studies have highlighted the widespread adulteration and substitution of ingredients in TCM, which threatens the safety of consumers. In this review, we summarized the strengths and limitations of each DNA barcoding technique and its derivative identification technologies as well as recent developments in sequencing technology, data mining, databases, and new tools related to DNA barcoding. The systematic process for selecting the appropriate barcode or derivative technologies analyzing TCM was also outlined. As a fast and effective method of identifying Chinese herbal medicines, DNA barcoding and its derivative technologies can be combined with several other methods. In the near future, these technologies will be used for quality control of TCM at the species level, which promotes the development of precision of TCM and speeds up the standardization and identification of herbal medicine.

#### Acknowledgements

Not applicable.

#### Author contributions

Conceptualization, SZ; Writing-original draft preparation, SZ, QZL, SMQ, and JPD; Writing-review and editing, SZ and XXG. All authors read and approved the final manuscript.

#### Funding

This work was supported by the Guangxi Natural Science Foundation (Grant Number 2018GXNSFAA281056), and the Guangdong Province Science and Technology Plan Project (Grant Number 2017A020213014), the Special Project of International Science and Technology Cooperation Guidance of Guangdong Academy of Sciences (Grant Number 2019GDASYL-0503002), and the Open Fund Project of the State Key Laboratory of Applied Microbiology Southern China (Grant Number SKLAM002-2018).

#### Availability of data and materials

Not applicable.

#### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Guangdong Province Key Laboratory for Biotechnology Drug Candidates, School of Biosciences and Biopharmaceutics, Guangdong Pharmaceutical University, Guangzhou 510006, China. <sup>2</sup>School of Pharmacy, Guangdong Pharmaceutical University, Guangzhou 510006, China.

Received: 29 April 2022 Accepted: 15 August 2022

Published online: 28 September 2022

#### References

- Uzuner H, Bauer R, Fan TP, Guo DA, Dias A, El-Nezami H, et al. Traditional Chinese medicine research in the post-genomic era: good practice priorities challenges and opportunities. *J Ethnopharmacol.* 2012;140:458–68.
- Zhang L, Yan J, Liu X, Ye Z, Yang X, Meyboom R, et al. Pharmacovigilance practice and risk control of Traditional Chinese Medicine drugs in China: current status and future perspective. *J Ethnopharmacol.* 2012;140:519–25.
- Marichamy K, Kumar NY, Ganesan A. Sustainable development in exports of herbals and Ayurveda Siddha Unani and Homeopathy (Ayush) in India. *Sci Park Res J.* 2014;1:23218045.
- Deng YJ, Liu BW, He ZX, Liu T, Zheng RL, Yang AD, et al. Study on active compounds from Huoxiang Zhengqi Oral Liquid for prevention of coronavirus disease 2019 (COVID-19) based on network pharmacology and molecular docking. *Chin Tradit Herb Drugs.* 2020;51:1113–22.
- Yao KT, Liu MY, Li X, Huang JH, Cai HB. Retrospective clinical analysis on treatment of Coronavirus Disease 2019 with Traditional Chinese Medicine Lianhua Qingwen. *Chin J Exp Tradit Med Formulae.* 2020;26:8–12.
- Han J, Pang X, Liao B, Yao H, Song J, Chen S. An authenticity survey of herbal medicines from markets in China using DNA barcoding. *Sci Rep.* 2016;6:18723.
- Xiong C, Sun W, Li J, Yao H, Shi Y, Wang P, et al. Identifying the species of seeds in Traditional Chinese Medicine using DNA barcoding. *Front Pharmacol.* 2018;9:1–8.
- Newmaster SG, Grguric M, Shanmughanandhan D, Ramalingam S, Ragupathy S. DNA barcoding detects contamination and substitution in North American herbal products. *BMC Med.* 2013;11:222.
- De Boer HJ, Ouarghidi A, Martin G, Abbad A, Kool A. DNA barcoding reveals limited accuracy of identifications based on folk taxonomy. *PLoS ONE.* 2014;9:e84291.
- Suesatpanit T, Osathanunkul K, Madesis P, Osathanunkul M. Should DNA sequence be incorporated with other taxonomical data for routine identifying of plant species? *BMC Complement Altern Med.* 2017;17:437.
- Raclariu Ancuta C, Mocan A, Popa MO, Vlase L, Ichim MC, Crisan G, et al. Veronica officinalis product authentication using DNA metabarcoding and HPLC-MS reveals widespread adulteration with Veronica chamaedrys. *Front Pharmacol.* 2017;8:1–13.
- Hebert PD, Cywinska A, Ball SL, DeWaard JR. Biological identifications through DNA barcodes. *Proc R Soc Lond Ser B Biol Sci.* 2003;270:313–21.
- Hollingsworth PM. Refining the DNA barcode for land plants. *Proc Natl Acad Sci USA.* 2011;108:19451–2.
- Techen N, Parveen I, Pan Z, Khan IA. DNA barcoding of medicinal plant material for identification. *Curr Opin Biotechnol.* 2014;25:103–10.
- Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, van der Bank M, et al. A DNA barcode for land plants. *Proc Natl Acad Sci USA.* 2009;106:12794–7.
- Yao H, Song J, Liu C, Luo K, Han J, Li Y, et al. Use of ITS2 region as the universal DNA barcode for plants and animals. *PLoS ONE.* 2010;5:e13102.
- Gao Z, Liu Y, Wang X, Wei X, Han J. DNA mini-barcoding: a derived barcoding method for herbal molecular identification. *Front Plant Sci.* 2019;10:987.
- Song M, Dong GQ, Zhang YQ, Liu X, Sun W. Identification of processed Chinese medicinal materials using DNA mini-barcoding. *Chin J Nat Med.* 2017;15:481–6.
- de Boer HJ, Ichim MC, Newmaster SG. DNA barcoding and pharmacovigilance of herbal medicines. *Drug Saf.* 2015;38:611–20.
- Yang F, Ding F, Chen H, He M, Zhu S, Ma X, et al. DNA barcoding for the identification and authentication of animal species in traditional medicine. *Evid Based Complement Altern Med.* 2018;2018:5160254.
- Kane N, Sveinsson S, Dempewolf H, Yang JY, Zhang D, Engels JMM, et al. Ultra-barcoding in cacao (theobroma spp. malvaceae) using whole chloroplast genomes and nuclear ribosomal DNA. *Am J Bot.* 2012;99:320–9.
- Krawczyk K, Nobis M, Myszczyński K, Klichowska E, Sawicki J. Plastid super-barcodes as a tool for species discrimination in feather grasses (Poaceae: Stipa). *Sci Rep.* 2018;8:1–10.
- Raclariu AC, Heinrich M, Ichim MC, de Boer H. Benefits and limitations of DNA barcoding and metabarcoding in herbal product authentication. *Phytochem Anal.* 2018;29:123–8.

24. Nevill PG, Zhong X, Filippini JT, Byrne M, Hislop M, Thiele K, et al. Large scale genome skimming from herbarium material for accurate plant identification and phylogenomics. *Plant Methods*. 2020;16:1–8.
25. CBOL Plant Working Group. A DNA barcode for land plants. *Proc Natl Acad Sci USA*. 2009;106:12794–7.
26. Chen S, Yao H, Han J, Liu C, Song J, Shi L, et al. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS ONE*. 2010;5:e8613.
27. Dong W, Xu C, Li C, Sun J, Zuo Y, Shi S, et al. ycf1 the most promising plastid DNA barcode of land plants. *Sci Rep*. 2015;5:8348.
28. Hollingsworth ML, Andra Clark A, Forrester LL, Richardson J, Pennington RT, Long DG, et al. Selecting barcoding loci for plants: evaluation of seven candidate loci with species-level sampling in three divergent groups of land plants. *Mol Ecol Resour*. 2009;9:439–57.
29. Hollingsworth PM, Graham SW, Little DP. Choosing and using a plant DNA barcode. *PLoS ONE*. 2011;6:e19254.
30. Li L, Josef BA, Liu B, Zheng S, Huang L, Chen S. Three-dimensional evaluation on ecotypic diversity of traditional Chinese Medicine: a case study of *Artemisia annua* L. *Front Plant Sci*. 2017;8:1225.
31. Selvaraj D, Shanmuganandhan D, Sarma RK, Joseph JC, Srinivasan RV, Ramalingam S. DNA barcode ITS effectively distinguishes the medicinal plant *Boerhavia diffusa* from its adulterants. *Genomics Proteomics Bioinform*. 2012;10:364–7.
32. Chen S, Pang X, Song J, Shi L, Yao H, Han J, et al. A renaissance in herbal medicine identification: from morphology to DNA. *Biotechnol Adv*. 2014;32:1237–44.
33. Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. Use of DNA barcodes to identify flowering plants. *Proc Natl Acad Sci USA*. 2005;102:8369–74.
34. Li Q, Sun Y, Guo H, Sang F, Ma H, Peng H, et al. Quality control of the traditional Chinese medicine Ruyi jinhuang powder based on high-throughput sequencing and real-time PCR. *Sci Rep*. 2018;8:1–10.
35. Guo M, Ren L, Pang X. Inspecting the true identity of herbal materials from *Cynanchum* using ITS2 barcode. *Front Plant Sci*. 2017;8:1945.
36. Zhao S, Chen X, Song J, Pang X, Chen S. Internal transcribed spacer 2 barcode: a good tool for identifying *Acanthopanax cortex*. *Front Plant Sci*. 2015;6:840.
37. Park I, Yang S, Kim WJ, Noh P, Lee HO, Moon BC. Authentication of herbal medicines *Dipsacus asper* and *Phlomis umbrosa* using DNA Barcodes chloroplast genome and Sequence Characterized Amplified Region (SCAR) marker. *Molecules*. 2018;23:1748.
38. Zhou H, Ma S, Song J, Lin Y, Wu Z, Han Z, et al. QR code labeling system for Xueteng-related herbs based on DNA barcode. *Chin Herb Med*. 2019;11:52–9.
39. Müller T, Philippi N, Dandekar T, Schultz J, Wolf M. Distinguishing species. *RNA*. 2007;13:1469–72.
40. Wolf M, Chen S, Song J, Ankenbrand M, Müller T. Compensatory base changes in ITS2 secondary structures correlate with the biological species concept despite intragenomic variability in ITS2 sequences—a proof of concept. *PLoS ONE*. 2013;8:e66726.
41. Zhang W, Yuan Y, Yang S, Huang J, Huang L. ITS2 secondary structure improves discrimination of medicinal “Mu tong” species when using DNA barcoding. *PLoS ONE*. 2015;10:e0131185.
42. Zhu S, Li Q, Chen S, Wang Y, Zhou L, Zeng C, et al. Phylogenetic analysis of *Uncaria* species based on internal transcribed spacer (ITS) region and ITS2 secondary structure. *Pharm Biol*. 2018;56:548–58.
43. Zhu RW, Li YC, Zhong DL, Zhang JQ. Establishment of the most comprehensive ITS2 barcode database to date of the traditional medicinal plant *Rhodiola* (Crassulaceae). *Sci Rep*. 2017;7:1–9.
44. Ashfaq M, Asif M, Anjum ZI, Zafar Y. Evaluating the capacity of plant DNA barcodes to discriminate species of cotton (*Gossypium*: Malvaceae). *Mol Ecol Resour*. 2013;13:573–82.
45. Gao T, Yao H, Song J, Liu C, Zhu Y, Ma X, et al. Identification of medicinal plants in the family Fabaceae using a potential DNA barcode ITS2. *J Ethnopharmacol*. 2010;13:116–21.
46. Gao T, Yao H, Song J, Zhu Y, Liu C, Chen S. Evaluating the feasibility of using candidate DNA barcodes in discriminating species of the large Asteraceae family. *BMC Evol Biol*. 2010;10:324.
47. Luo K, Chen SL, Chen KL, Song JY, Yao H, Ma X, et al. Assessment of candidate plant DNA barcodes using the Rutaceae family. *Sci China Life Sci*. 2010;53:701–8.
48. Pang X, Song J, Zhu Y, Xu H, Huang L, Chen S. Applying plant DNA barcodes for Rosaceae species identification. *Cladistics*. 2011;27:165–70.
49. Pang X, Song J, Zhu Y, Xie C, Chen S. Using DNA barcoding to identify species within euphorbiaceae. *Planta Med*. 2010;76:1784–6.
50. Chiou SJ, Yen JH, Fang CL, Chen HL, Lin TY. Authentication of medicinal herbs using PCR-amplified ITS2 with specific primers. *Planta Med*. 2007;73:1421–6.
51. Gong L, Qiu XH, Huang J, Xu W, Bai JQ, Zhang J, et al. Constructing a DNA barcode reference library for southern herbs in China: a resource for authentication of southern Chinese medicine. *PLoS ONE*. 2018;13:e0201240.
52. Wang FH, Lu JM, Wen J, Ebihara A, Li DZ. Applying DNA barcodes to identify closely related species of ferns: a case study of the Chinese *Adiantum* (Pteridaceae). *PLoS ONE*. 2016;11:e0160611.
53. Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S. Plant DNA barcoding: from gene to genome. *Biol Rev Camb Philos Soc*. 2015;90:157–66.
54. Álvarez I, Wendel JF. Ribosomal ITS sequences and plant phylogenetic inference. *Mol Phylogenet Evol*. 2003;29:417–34.
55. Song J, Shi L, Li D, Sun Y, Niu Y, Chen Z, et al. Extensive pyrosequencing reveals frequent intra-genomic variations of internal transcribed spacer regions of nuclear ribosomal DNA. *PLoS ONE*. 2012;7:e43971.
56. Min XJ, Hickey DA. Assessing the effect of varying sequence length on DNA barcoding of fungi. *Mol Ecol Notes*. 2007;7:365–73.
57. Selvaraj D, Sarma RK, Sathishkumar R. Phylogenetic analysis of chloroplast mat K gene from Zingiberaceae for plant DNA barcoding. *Bioinformation*. 2008;3:24–7.
58. Liu Y, Wang K, Liu Z, Luo K, Chen S, Chen K. Identification of medical plants of 24 *ardisia* species from China using the matK genetic marker. *Pharmacogn Mag*. 2013;9:331–7.
59. Kim WJ, Ji Y, Choi G, Kang YM, Yang S, Moon BC. Molecular identification and phylogenetic analysis of important medicinal plant species in genus *Paeonia* based on rDNA-ITS matK and rbcL DNA barcode sequences. *Genet Mol Res*. 2016. <https://doi.org/10.4238/gmr.15038472>.
60. Tnah LH, Lee SL, Tan AL, Lee CT, Ng KKS, Ng CH, et al. DNA barcode database of common herbal plants in the tropics: a resource for herbal product authentication. *Food Control*. 2019;95:318–26.
61. Osathanunkul M, Osathanunkul R, Madesis P. Species identification approach for both raw materials and end products of herbal supplements from *Tinospora* species. *BMC Complement Altern Med*. 2018;18:111.
62. Xin T, Su C, Lin Y, Wang S, Xu Z, Song J. Precise species detection of traditional Chinese patent medicine by shotgun metagenomic sequencing. *Phytomedicine*. 2018;47:40–7.
63. Newmaster SG, Fazekas AJ, Steeves RAD, Janovec J. Testing candidate plant barcode regions in the Myricaceae. *Mol Ecol Resour*. 2008;8:480–90.
64. Štorchová H, Olson MS. The architecture of the chloroplast psbA-trnH non-coding region in angiosperms. *Plant Syst Evol*. 2007;268:235–56.
65. Yao H, Song J-Y, Ma X-Y, Liu C, Li Y, Xu H-X, et al. Identification of *Dendrobium* species by a candidate DNA barcode sequence: the chloroplast psbA-trnH intergenic region. *Planta Med*. 2009;75:667–9.
66. Ma X-Y, Xie C-X, Liu C, Song J-Y, Yao H, Luo K, et al. Species identification of medicinal pteridophytes by a DNA barcode marker the chloroplast psbA-trnH intergenic region. *Biol Pharm Bull*. 2010;33:1919–24.
67. Cao L, Qin SS, Yuan Y, Zhu XQ. Molecular identification of *Mentha haplocalyx* and *Mentha spicata* with specific primers multi-PCR system. *Zhong yao cai = Zhongyao cai = J Chin Med Mater*. 2014;37:41–5.
68. Stech M, Quandt D. 20000 species and five key markers: the status of molecular bryophyte phylogenetics. *Phytotaxa*. 2014;9:196.
69. Hao DC, Huang BL, Chen SL, Mu J. Evolution of the Chloroplast trnL-trnF region in the gymnosperm lineages taxaceae and cephalotaxaceae. *Biochem Genet*. 2009;47:351–69.
70. Wang L, Jiang Y, Shi Q, Wang Y, Sha L, Fan X, et al. Genome constitution and evolution of *Elytrigia lolioides* inferred from Acc1 EF-G ITS TrnL-F sequences and GISH. *BMC Plant Biol*. 2019;19:1–14.
71. Nazar N, Clarkson JJ, Goyder D, Kaky E, Mahmood T, Chase MW. Phylogenetic relationships in Apocynaceae based on nuclear PHYA

- and plastid trnL-F sequences with a focus on tribal relationships. *Caryologia*. 2019;72:55–81.
72. Yang M, Zhang D, Liu J, Zheng J. A molecular marker that is specific to medicinal rhubarb based on chloroplast trnL/trnF sequences. *Planta Med*. 2001;67:784–6.
  73. Handy SM, Parks MB, Deeds JR, Liston A, De Jager LS, Lucciolli S, et al. Use of the chloroplast gene *ycf1* for the genetic differentiation of pine nuts obtained from consumers experiencing dysgeusia. *J Agric Food Chem*. 2011;59:10995–1002.
  74. Jiao L, Yu M, Wiedenhoef AC, He T, Li J, Liu B, et al. DNA barcode authentication and library development for the wood of six commercial *Pterocarpus* species: the critical role of *Xylarium* specimens. *Sci Rep*. 2018;8:1–10.
  75. Rogers MB, Gilson PR, Su V, McFadden GI, Keeling PJ. The complete chloroplast genome of the chlorarachniophyte *Bigeloviella natans*: evidence for independent origins of chlorarachniophyte and euglenid secondary endosymbionts. *Mol Biol Evol*. 2006;24:54–62.
  76. Kress WJ, Erickson DL. A two-locus global DNA barcode for land plants: the coding *rbcl* gene complements the non-coding *trnH-psbA* spacer region. *PLoS ONE*. 2007;2:508.
  77. Feng T, Li Q, Wang Y, Qiu S, He M, Zhang W, et al. Phylogenetic analysis of *Aquilaria* Lam. (Thymelaeaceae) based on DNA barcoding. *Holzforchung*. 2019;73:517–23.
  78. Li Q, Wu J, Wang Y, Lian X, Wu F, Zhou L, et al. The phylogenetic analysis of *Dalbergia* (Fabaceae: Papilionaceae) based on different DNA barcodes. *Holzforchung*. 2017;71:939–49.
  79. Li D-Z, Gao L-M, Li H-T, Wang H, Ge X-J, Liu J-Q, et al. 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*. 2011;108:19641–6.
  80. Chen X, Xiang L, Shi L, Li G, Yao H, Han J, et al. Identification of crude drugs in the Japanese pharmacopoeia using a DNA barcoding system. *Sci Rep*. 2017;7:1–7.
  81. Zuo Y, Chen Z, Kondo K, Funamoto T, Wen J, Zhou S. DNA barcoding of *Panax* species. *Planta Med*. 2011;77:182–7.
  82. Burgess KS, Fazekas AJ, Kesanakurti PR, Graham SW, Husband BC, Newmaster SG, et al. Discriminating plant species in a local temperate flora using the *rbcl*+*matK* DNA barcode. *Methods Ecol Evol*. 2011;2:333–40.
  83. Kane NC, Cronk Q. Botany without borders: barcoding in focus. *Mol Ecol*. 2008;17:5175–6.
  84. Fu CN, Wu CS, Ye LJ, Mo ZQ, Liu J, Chang YW, et al. Prevalence of isomeric plastomes and effectiveness of plastome super-barcodes in yews (*Taxus*) worldwide. *Sci Rep*. 2019;9:1–11.
  85. Ruhsam M, Rai HS, Mathews S, Ross TG, Graham SW, Raubeson LA, et al. Does complete plastid genome sequencing improve species discrimination and phylogenetic resolution in *Araucaria*? *Mol Ecol Resour*. 2015;15:1067–78.
  86. Zhang N, Erickson DL, Ramachandran P, Ottesen AR, Timme RE, Funk VA, et al. An analysis of Echinacea chloroplast genomes: implications for future botanical identification. *Sci Rep*. 2017;7:1–9.
  87. Yang JB, Yang SX, Li HT, Yang J, Li DZ. Comparative chloroplast genomes of *Camellia* species. *PLoS ONE*. 2013;8:e73053.
  88. Zhang Y, Du L, Liu A, Chen J, Wu L, Hu W, et al. The complete chloroplast genome sequences of five *Epimedium* species: lights into phylogenetic and taxonomic analyses. *Front Plant Sci*. 2016;7:1–12.
  89. Bi Y, Zhang MF, Xue J, Dong R, Du YP, Zhang XH. Chloroplast genomic resources for phylogeny and DNA barcoding: a case study on *Fritillaria*. *Sci Rep*. 2018;8:1–12.
  90. Piredda R, Simeone MC, Attimonelli M, Bellarosa R, Schirone B. Prospects of barcoding the Italian wild dendroflora: oaks reveal severe limitations to tracking species identity. *Mol Ecol Resour*. 2011;11:72–83.
  91. Toegl R, Hofferek G, Greimel K, Leung A, Phan RCW, Bloem R. Towards next-generation biodiversity assessment using DNA metabarcoding. *Mol Ecol*. 2012;33:2289–94.
  92. Kress WJ, Garcia-Robledo C, Uriarte M, Erickson DL. DNA barcodes for ecology evolution and conservation. *Trends Ecol Evol*. 2015;30:25–35.
  93. Coghlan ML, Haile J, Houston J, Murray DC, White NE, Moolhuijzen P, et al. Deep sequencing of plant and animal DNA contained within traditional Chinese medicines reveals legality issues and health safety concerns. *PLoS Genet*. 2012;8:e1002657.
  94. Arulandhu AJ, Staats M, Hagelaar R, Peelen T, Kok EJ. The application of multi-locus DNA metabarcoding in traditional medicines. *J Food Compos Anal*. 2019;79:87–94.
  95. Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, et al. Power and limitations of the chloroplast *trnL* (UAA) intron for plant DNA barcoding. *Nucleic Acids Res*. 2007;35:e14.
  96. Meusnier I, Singer GAC, Landry JF, Hickey DA, Hebert PDN, Hajibabaei M. A universal DNA mini-barcode for biodiversity analysis. *BMC Genomics*. 2008;9:4–7.
  97. Little DP. A DNA mini-barcode for land plants. *Mol Ecol Resour*. 2014;14:437–46.
  98. Valentini A, Miquel C, Nawaz MA, Bellemain E, Coissac E, Pompanon F, et al. New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the *trnL* approach. *Mol Ecol Resour*. 2009;9:51–60.
  99. Dong W, Liu H, Xu C, Zuo Y, Chen Z, Zhou S. A chloroplast genomic strategy for designing taxon specific DNA mini-barcodes: a case study on *singensis*. *BMC Genet*. 2014;15:138–45.
  100. Gao Z, Liu Y, Wang X, Song J, Chen S, Ragupathy S, et al. Derivative technology of DNA barcoding (Nucleotide signature and SNP double peak methods) detects adulterants and substitution in Chinese Patent Medicines. *Sci Rep*. 2017;7:1–11.
  101. Liu Y, Wang X, Wang L, Chen X, Pang X, Han J. A nucleotide signature for the identification of American ginseng and its products. *Front Plant Sci*. 2016;7:1–9.
  102. Wang X, Liu Y, Wang L, Han J, Chen S. A nucleotide signature for the identification of *angelicae sinensis radix* (Danggui) and its products. *Sci Rep*. 2016;6:34940.
  103. Liu Y, Wang XY, Wei XM, Gao ZT, Han JP. Rapid authentication of *Ginkgo biloba* herbal products using the recombinase polymerase amplification assay. *Sci Rep*. 2018;8:1–8.
  104. Parveen I, Techen N, Khan I. Identification of species in the Aromatic Spice family *Apiaceae* using DNA mini-barcodes. *Planta Med*. 2019;85:139–44.
  105. Yeo D, Srivathsan A, Meier R. Mini-barcodes are equally useful for species identification and more suitable for large-scale species discovery in Metazoa than full-length barcodes. *BioRxiv*. 2019;594952.
  106. Little DP. Authentication of *Ginkgo biloba* herbal dietary supplements using DNA barcoding. *Genome*. 2014;57:513–6.
  107. Liu Y, Xiang L, Zhang Y, Lai X, Xiong C, Li J, et al. DNA barcoding based identification of *Hippophae* species and authentication of commercial products by high resolution melting analysis. *Food Chem*. 2018;242:62–7.
  108. Coissac E, Hollingsworth PM, Lavergne S, Taberlet P. From barcodes to genomes: extending the concept of DNA barcoding. *Mol Ecol*. 2016;25:1423–8.
  109. Chao S, Ying L, Wu Q, Luo H, Sun Y, Song J, et al. De novo sequencing and analysis of the American ginseng root transcriptome using a GS FLX Titanium platform to discover putative genes involved in ginsenoside biosynthesis. *BMC Genomics*. 2010;11:262.
  110. Chen S, Jiang X, Chang L, Zhu Y, Nelson DR, Zhou S, et al. Genome sequence of the model medicinal mushroom *Ganoderma lucidum*. *Nat Commun*. 2012;39:13.
  111. Li Y, Luo HM, Sun C, Song JY, Sun YZ, Wu Q, et al. EST analysis reveals putative genes involved in glycyrrhizin biosynthesis. *BMC Genomics*. 2010;11:268.
  112. He T, Jiao L, Yu M, Guo J, Jiang X, Yin Y. DNA barcoding authentication for the wood of eight endangered *Dalbergia* timber species using machine learning approaches. *Holzforchung*. 2019;73:277–85.
  113. Hartvig I, Czako M, Kjær ED, Nielsen LR, Theilade I. The use of DNA barcoding in identification and conservation of rosewood (*Dalbergia* spp.). *PLoS ONE*. 2015;10:e0138231.
  114. Goldberg DE, Holland JH. Genetic algorithms and machine learning. *Mach Learn*. 1988;3:95–9.
  115. Jordan MI, Mitchell TM. Machine learning: trends perspectives and prospects. *Science*. 2015;349:255–60.
  116. Robert C. Machine learning a probabilistic perspective. *Chance*. 2014;27:62–3.

117. Guo M, Xu Y, Ren L, He S, Pang X. A systematic study on DNA barcoding of medicinally important genus *Epimedium* L. (Berberidaceae). *Genes (Basel)*. 2018;9:623.
118. Fan T-P, Zhu Y, Leon C, Franz G, Bender A, Zheng X. Traditional Chinese Medicine Herbal Drugs: From Heritage to Future Developments. In: Sasisekharan R, Lee SL, Rosenberg A, Walker LA, editors. *The Science and Regulations of Naturally Derived Complex Drugs*. Springer, Cham. 2019;59–77.
119. Gao T, Ma X, Zhu X. Use of the psbA-trnH region to authenticate medicinal species of Fabaceae. *Biol Pharm Bull*. 2013;36:1975–9.
120. Bell KL, Loeffler VM, Brosi BJ. An rbcL reference library to aid in the identification of plant species mixtures by DNA metabarcoding. *Appl Plant Sci*. 2017;5:1–7.
121. Shi R, Hu Z, Lu H, Liu L, Xu L, Liu Y, et al. Hierarchical nanostructuring array enhances mid-hybridization for accurate herbal identification via ITS2 DNA barcode. *Anal Chem*. 2020;92:2136–44.
122. Leon C, Lin Y. *Chinese medicinal plants herbal drugs and substitutes: an identification guide*. Richmond: Kew Publishing; 2017. p. 74–6. ISBN 978-1-84-246387-1.
123. Zhao Z, Guo P, Brand E. The formation of daodi medicinal materials. *J Ethnopharmacol*. 2012;140:476–81.
124. Hu H, Shen X, Liao B, Luo L, Xu J, Chen S. Herbgonomics: a stepping stone for research into herbal medicine. *Sci China Life Sci*. 2019;62:913–20.
125. Mishra P, Kumar A, Nagireddy A, Mani DN, Shukla AK, Tiwari R, et al. DNA barcoding: an efficient tool to overcome authentication challenges in the herbal market. *Plant Biotechnol J*. 2016;14:8–21.
126. Jia J, Xu Z, Xin T, Shi L, Song J. Quality control of the Traditional Patent medicine Yimu Wan based on SMRT sequencing and DNA barcoding. *Front Plant Sci*. 2017;8:926.
127. Shi Y, Zhao M, Yao H, Yang P, Xin T, Li B, et al. Rapidly discriminate commercial medicinal *Pulsatilla chinensis* (Bge.) Regel from its adulterants using ITS2 barcoding and specific PCR-RFLP assay. *Sci Rep*. 2017;7:1–12. <https://doi.org/10.1038/srep40000>.
128. Tian E, Liu Q, Ye H, Li F, Chao Z. A DNA barcode-based RPA Assay (BAR-RPA) for rapid identification of the dry root of *Ficus hirta* (Wuzhimaotao). *Molecules*. 2017;22:2261.
129. Feng S, Liu Z, Chen L, Hou N, Yang T, Wei A. Phylogenetic relationships among cultivated *Zanthoxylum* species in China based on cpDNA markers. *Tree Genet Genomes*. 2016;12:1–9.
130. Xiong C, Hu ZG, Tu Y, Liu HG, Wang P, Zhao MM, et al. ITS2 barcoding DNA region combined with high resolution melting (HRM) analysis of *Hyoscyami Semen*, the mature seed of *Hyoscyamus niger*. *Chin J Nat Med*. 2016;14:898–903.
131. Zhang J, Hu X, Wang P, Huang B, Sun W, Xiong C, et al. Investigation on Species authenticity for herbal products of *Celastrus Orbiculatus* and *Tripterygium Wilfordii* from markets using ITS2 barcoding. *Molecules*. 2018;23:967.
132. Duan BZ, Wang YP, Fang HL, Xiong C, Li XW, Wang P, et al. Authenticity analyses of Rhizoma *Paridis* using barcoding coupled with high resolution melting (Bar-HRM) analysis to control its quality for medicinal plant product. *Chin Med*. 2018;13:8.
133. Yin Y, Jiao L, Dong M, Jiang X, Zhang S. Wood resources, identification, and utilization of agarwood in China. In: Mohamed R, editor. *agarwood*. Singapore: Springer; 2016. p. 21–38.
134. Kreuzer M, Howard C, Adhikari B, Pendry CA, Hawkins JA. Phylogenomic approaches to DNA barcoding of herbal medicines: developing clade-specific diagnostic characters for berberis. *Front Plant Sci*. 2019;10:586.
135. Thakur VV, Tiwari S, Tripathi N, Tiwari G. Molecular identification of medicinal plants with amplicon length polymorphism using universal DNA barcodes of the atpF-atpH, trnL and trnH-psbA regions. *3 Biotech*. 2019;9:188.
136. Wei XM, Wang XY, Gao ZT, Cao P, Han JP. Identification of flower herbs in Chinese pharmacopoeia based on DNA barcoding. *Chin Herb Med*. 2019;11:275–80.
137. Chao Z, Zeng W, Liao J, Liu L, Liang Z, Li X. DNA barcoding Chinese medicinal *Bupleurum*. *Phytomedicine*. 2014;21:1767–73.
138. Sun Z, Chen S. Identification of cortex herbs using the DNA barcode nrITS2. *J Nat Med*. 2013;67:296–302.
139. Yang J, Dong L, Wei G, Hu H, Zhu G, Zhang J, et al. Identification and quality analysis of *Panax notoginseng* and *Panax vietnamensis* var *fuscidicus* through integrated DNA barcoding and HPLC. *Chin Herb Med*. 2018;10:177–83.
140. Liu J, Shi L, Han J, Li G, Lu H, Hou J, et al. Identification of species in the angiosperm family Apiaceae using DNA barcodes. *Mol Ecol Resour*. 2014;14:1231–8. <https://doi.org/10.1111/1755-0998.12262>.
141. Zhang JQ, Meng SY, Wen J, Rao GY. DNA barcoding of *Rhodiola* (crassulaceae): a case study on a group of recently diversified medicinal plants from the Qinghai-Tibetan Plateau. *PLoS ONE*. 2015;10:e0119921.
142. Ya-dong Y, Lin-chun S, Xiao-chong M, Wei S, Meng Y, Li X. Identification of *Atractylodis Macrocephalae* Rhizoma and *Atractylodis Rhizoma* from their adulterants using DNA barcoding. *China J Chin Mater Medica*. 2014;39:2194–8.
143. Zheng S, Liu D, Ren W, Fu J, Huang L, Chen S. Integrated analysis for identifying radix *astragali* and its adulterants based on DNA barcoding. *Evid-based Complement Altern Med*. 2014. <https://doi.org/10.1155/2014/843923>.
144. 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;34:1198–203. <https://doi.org/10.1248/bpb.34.1198>.
145. Yuan Q, Zhang BIN, Jiang DAN, Zhang W, Lin T, Chiou S. Identification of species and materia medica within *Angelica* L. (Umbelliferae) based on phylogeny inferred from DNA barcodes. *Mol Ecol Resour*. 2015;15:358–71. <https://doi.org/10.1111/1755-0998.12296>.
146. Chen S, Zhu Z, Ma H, Yang J, Guo Q. DNA barcodes for discriminating the medicinal plant *Isatis indigotica* Fort. (Cruciferae) and its adulterants. *Biochem Syst Ecol*. 2014;57:287–92. <https://doi.org/10.1016/j.bse.2014.08.007>.
147. Pang X, Shi L, Song J, Chen X, Chen S. Use of the potential DNA barcode ITS2 to identify herbal materials. *J Nat Med*. 2013;67:571–5. <https://doi.org/10.1007/s11418-012-0715-2>.
148. Wang DY, Wang Q, Wang YL, Xiang XG, Huang LQ, Jin XH. Evaluation of DNA barcodes in *Codonopsis* (Campanulaceae) and in some large angiosperm plant genera. *PLoS ONE*. 2017;12:1–14. <https://doi.org/10.1371/journal.pone.0170286>.
149. Li M, Cao H, But PPH, Shaw PC. Identification of herbal medicinal materials using DNA barcodes. *J Syst Evol*. 2011;49:271–83. <https://doi.org/10.1111/j.1759-6831.2011.00132.x>.
150. Zhao LL, Feng SJ, Tian JY, Wei AZ, Yang TX. Internal transcribed spacer 2 (ITS2) barcodes: a useful tool for identifying Chinese *Zanthoxylum*. *Appl Plant Sci*. 2018;6:1–8. <https://doi.org/10.1002/aps3.1157>.
151. Ruzicka J, Lukas B, Merza L, Göhler I, Abel G, Popp M, et al. Identification of *Verbena officinalis* based on ITS sequence analysis and RAPD-derived molecular markers. *Planta Med*. 2009;75:1271–6.
152. Liu Z, Chen K, Luo K, Pan H, Chen S. DNA barcoding in medicinal plants *Caprifoliaceae*. *Zhongguo Zhong Yao Za Zhi*. 2010;35:2527–32.
153. Jiao L, Shui Y. Evaluating candidate DNA barcodes among Chinese *Begonia* (Begoniaceae) species. *Plant Divers Resour*. 2013;35:715–24.
154. Li L, Xiao J, Su Z, Huang Y, Tang L. Identification of Li Medicine plants in *Rubiaceae* using ITS2 barcode sequence. *Chin Tradit Herb Drugs*. 2013;44:1814–8. <https://doi.org/10.7501/j.issn.0253-2670.2013.13.021>.
155. Zhu Y, Chen SL, Yao H, Tan R, Song JY, Luo K, et al. DNA barcoding the medicinal plants of the genus *Paris*. *Yao xue xue bao = Acta Pharm Sin*. 2010;45:376–82.
156. Yi F, Han FM, Peng Y. Molecular identification of 15 species of *Ilex* genus based on ITS sequence analysis. *Zhong Yao Cai*. 2014;37:974–6.
157. 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;10:2806–16. <https://doi.org/10.4238/2011.November.10.3>.
158. Doh EJ, Kim JH, Lee G. Identification and monitoring of *Amomi Fructus* and its adulterants based on DNA barcoding analysis and designed DNA markers. *Molecules*. 2019;24:4193.
159. Clement WL, Donoghue MJ. Barcoding success as a function of phylogenetic relatedness in *Viburnum*, a clade of woody angiosperms. *BMC Evol Biol*. 2012;12:73.
160. Wang K, Chen K, Liu Z, Chen S. Screening of universal DNA Barcodes for *Malvaceae* plants. *Chin Bull Bot*. 2011;46:276.



161. Ma XC, Yao H, Wu L, Xiang L, Chen XC, Song JY. Molecular identification of aucklandiae radix, vladimiriae radix, inulae radix, aristolochiae radix and kudsurae radix using ITS2 barcode. *Zhongguo Zhong Yao Za Zhi*. 2014;39:2169–75.
162. Gao T, Pang XH, Chen SL. Authentication of plants in Astragalus by DNA barcoding technique. *Planta Med*. 2009;75:21.
163. Shi LC, Zhang J, Han JP, Song JY, Yao H, Zhu YJ, et al. Testing the potential of proposed DNA barcodes for species identification of Zingiberaceae. *J Syst Evol*. 2011;49:261–6.
164. Liu J, Yan HF, Ge XJ. The use of DNA barcoding on recently diverged species in the genus *Gentiana* (Gentianaceae) in China. *PLoS ONE*. 2016;11:e0153008. <https://doi.org/10.1371/journal.pone.0153008>.
165. Xiang XG, Zhang JB, Lu AM, Li RQ. Molecular identification of species in Juglandaceae: a tiered method. *J Syst Evol*. 2011;49:252–60.
166. Luo K, Chen S, Chen K, Song J, Yao H. Application of DNA barcoding to the medicinal plants of the Araceae family. *Planta Med*. 2009;75:399–457. <https://doi.org/10.1055/s-2009-1216448>.
167. Song J, Yao H, Li Y, Li X, Lin Y, Liu C, et al. Authentication of the family Polygonaceae in Chinese pharmacopoeia by DNA barcoding technique. *J Ethnopharmacol*. 2009;124:434–9. <https://doi.org/10.1016/j.jep.2009.05.042>.
168. Liu Z, Chen S-L, Song J-Y, Zhang S-J, Chen K-L. Application of deoxyribonucleic acid barcoding in Lauraceae plants. *Pharmacogn Mag*. 2012;8:4.
169. Chen K, Liu Y, Zhang L, Liu Z, Luo K, Chen S. Species identification of *Rhododendron* (Ericaceae) using the chloroplast deoxyribonucleic acid *psbA-trnH* genetic marker. *Pharmacogn Mag*. 2012;8:29. <https://doi.org/10.4103/0973-1296.93311>.
170. Jun H, Ka-Lok W, Pang-Chui S, Hong W, De-Zhu L. Identification of the medicinal plants in *Aconitum* L. by DNA barcoding technique. *Planta Med*. 2010;76:1622–8.
171. Han JP, Song JY, Liu C, Chen J, Qian J, Zhu Y, et al. Identification of *Cistanche* species (Orobanchaceae) based on sequences of the plastid *psbA-trnH* intergenic region. *Yao xue xue bao = Acta Pharm Sin*. 2010;45:126–30.
172. Guo H, Liu J, Luo L, Wei X, Zhang J, Qi Y, et al. Complete chloroplast genome sequences of *Schisandra chinensis*: genome structure, comparative analysis, and phylogenetic relationship of basal angiosperms. *Sci China Life Sci*. 2017;60:1286–90.
173. Park I, Kim W, Yang S, Yeo S-M, Li H, Moon BC. The complete chloroplast genome sequence of *Aconitum coreanum* and *Aconitum carmichaelii* and comparative analysis with other *Aconitum* species. *PLoS ONE*. 2017;12:e0184257.
174. Yang P, Zhou H, Qian J, Xu H, Shao Q, Li Y, et al. The complete chloroplast genome sequence of *Dendrobium officinale*. *Mitochondrial DNA Part A*. 2016;27:1262–4.
175. Wang S, Hou F, Zhao J, Cao J, Peng C, Wan D, et al. Authentication of Chinese herbal medicines *Dendrobium* species and phylogenetic study based on nrDNA ITS sequence. *Int J Agric Biol*. 2018;20:369–74. <https://doi.org/10.17957/IJAB/15.0500>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

