DNA Barcoding System for Identifying Herbal Medicines

Caixia Wang, Ph.D

Institue of Chinese Materia Meidica, China Academy of Chinese Medical Sciences

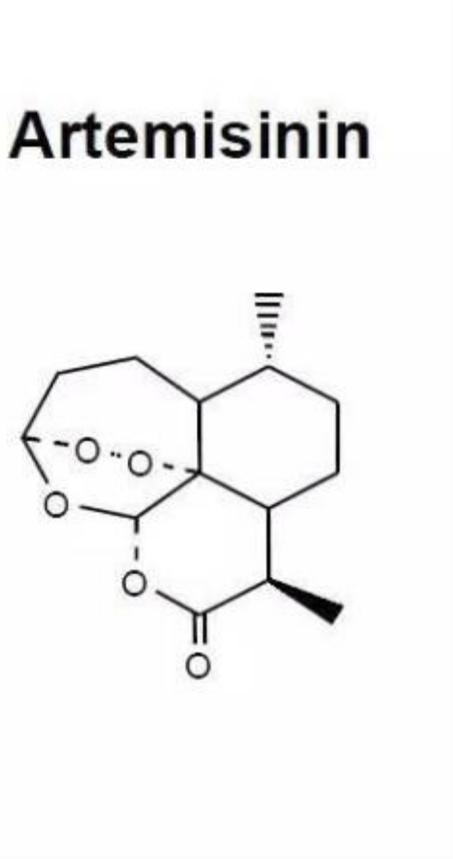
Zhang Bolin, Professor

School of Biological Science & Technology, Beijing Forestry University

2021/01/20

TCM: A Gift To The World

To Youyou,
Institue of Chinese Materia Meidica, China Academy of Chinese Medical Sciences
won the Nobel Prize in Physiology or Medicine in 2015



Outline

- **Introduction of TCM barcode system**
- **Innovations and related technical content**
- **Applications**

Could you identify these medicinal plants and traditional Chinese medicine?













Serious confusion of Chinese herbal medicines

- ✓ **Overall value of Traditional Chinese Medicine industry: 2 trillion/year (2020)**
- ✓ **More than 10,000 Species were used in TCM, some with identification difficult**

Ture（川西獐芽菜） False（抱茎獐芽菜）

The prescription is ture,the medicine is false.

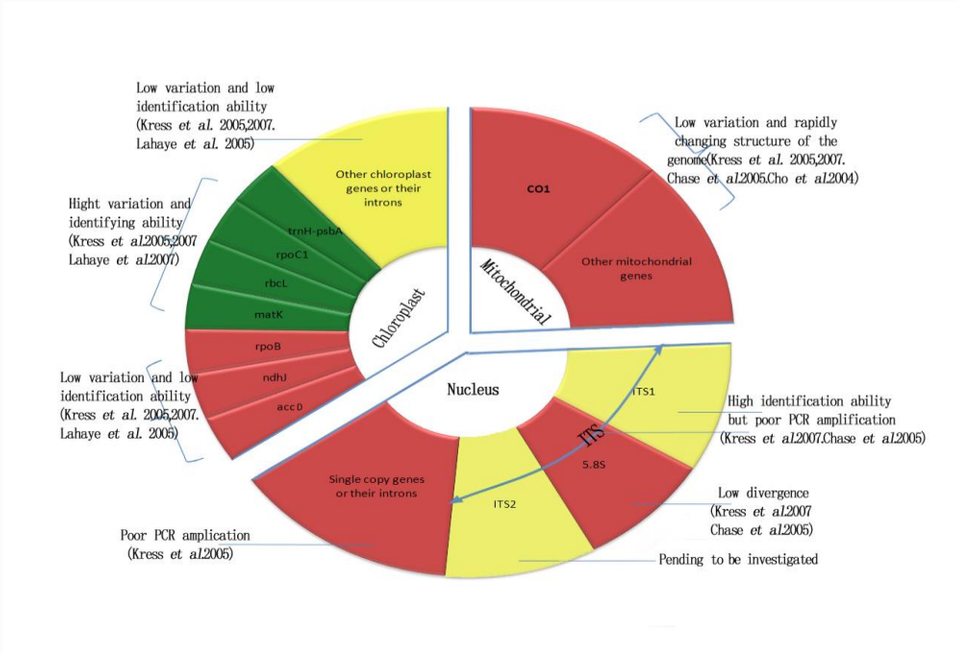
Food Safety News

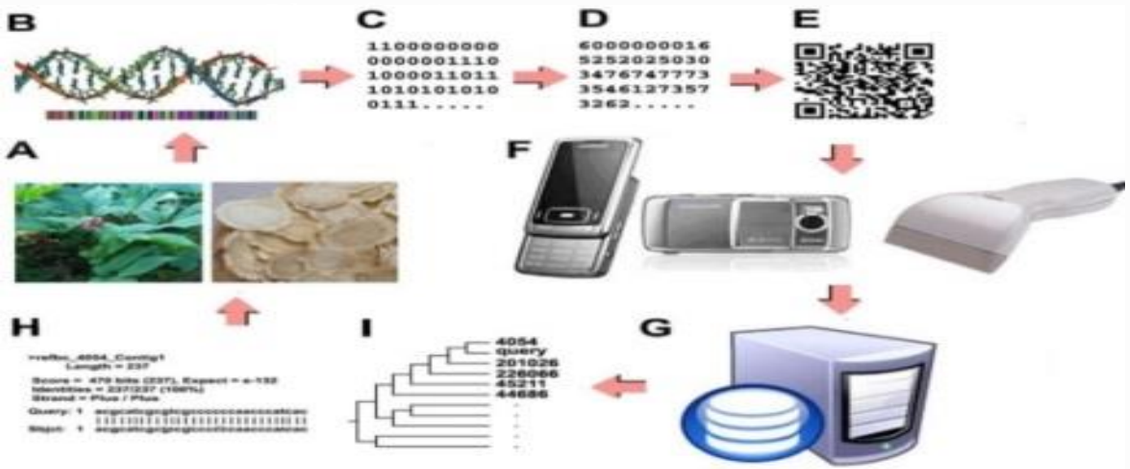
Big Retailers Ordered to Stop Selling 'Adulterated' and 'Mislabeled' Herbal Supplements

西洋参

生晒参

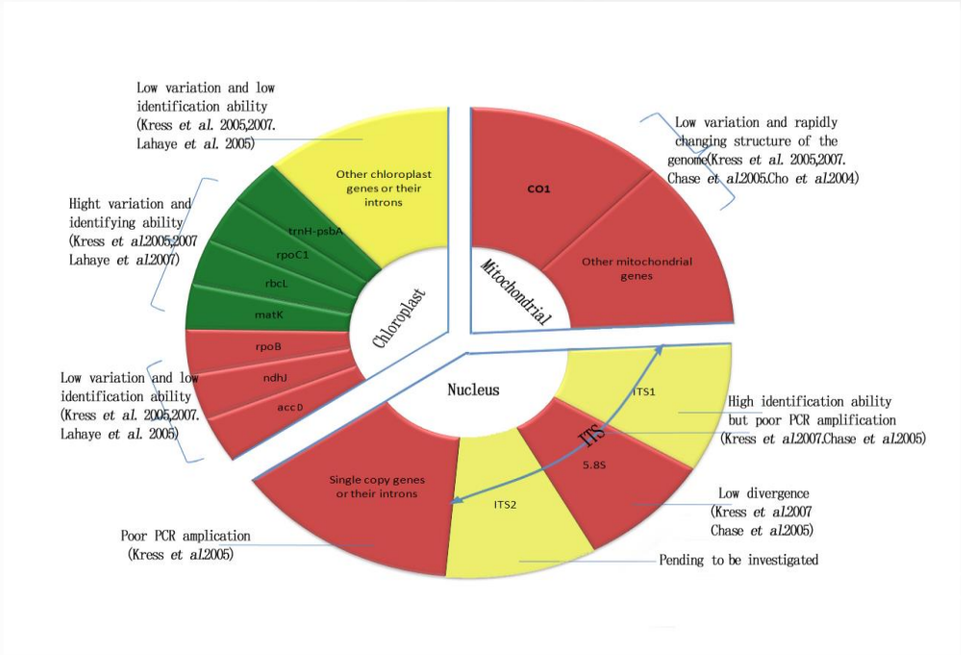
What iS DNA Barcodes

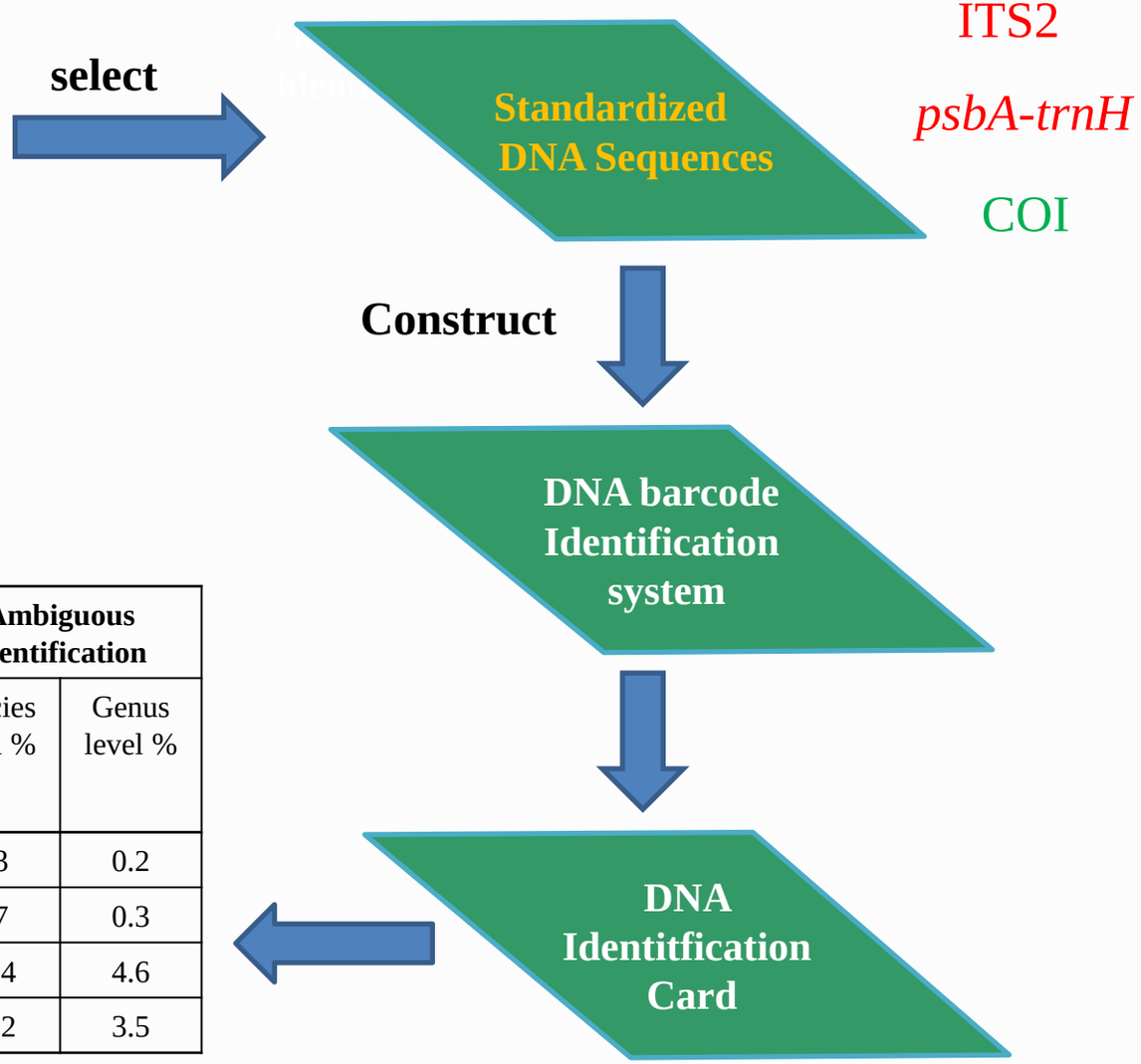




- ✓ DNA barcoding is a new technique that uses standardized DNA sequences of an organism to determine its identity. It uses 400-800bp-long short sequences, which conserved among most species but with enough variability to divide related species. In animal kingdom, Cox1 was the most used sequence, and among plants , MatK and rbcL were the most used barcodes traditionally

DNA Barcodes





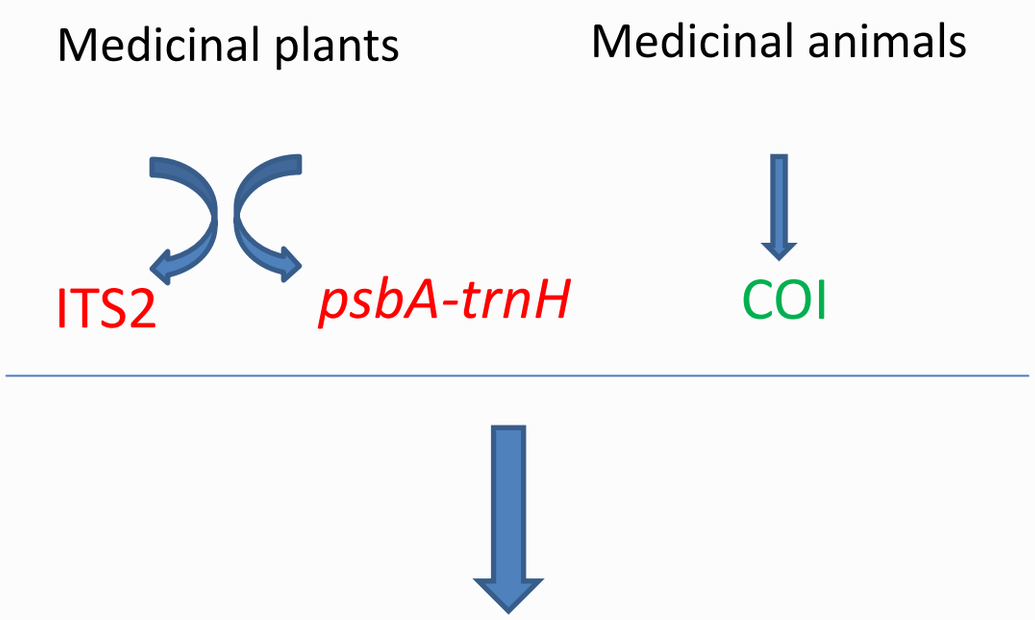
Marker	Method of species identification (Ross et al. 2008 *Syst Biol* 57: 216-230)	Correct identification		Incorrect identification		Ambiguous identification	
		Specis level %	Genus level %	Species level %	Genus level %	Species level %	Genus level %
ITS2	BLAST	92.7	99.8	0	0	7.3	0.2
	Distance	90.3	99.7	0	0	9.7	0.3
psbA-trnH	BLAST	67.6	95.4	0	0	32.4	4.6
	Distance	72.8	96.5	0	0	27.2	3.5

TCMbarcode System

- A publicly available on-line DNA barcoding system for identifying herbal medicine
- Store and collect DNA barcode sequences of herbal materials
- Publish DNA barcode sequences
- Analyze DNA barcode data

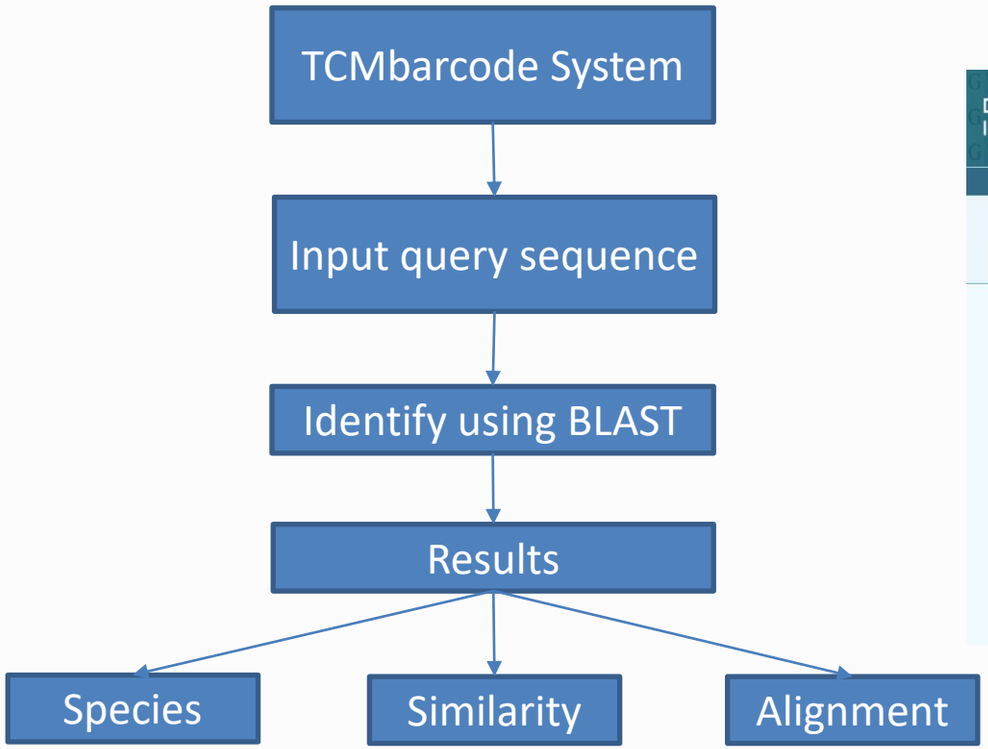
TCMbarcode System

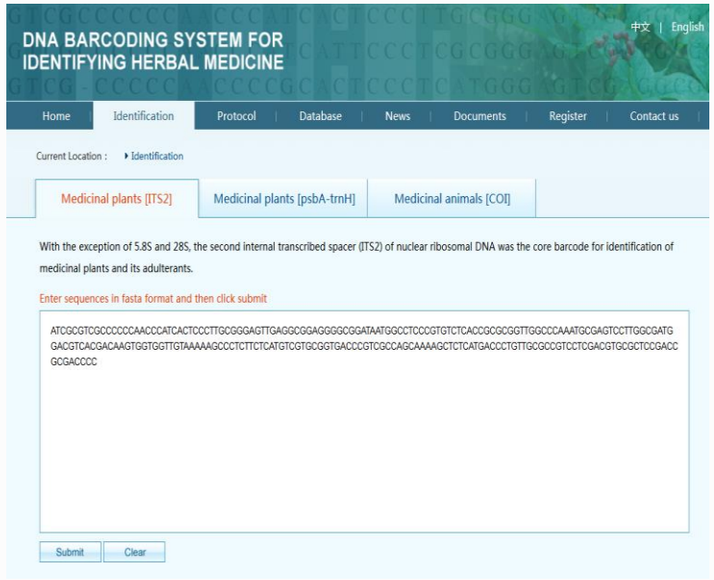
- ✓ 11thousands of species, 40 thousands of standard sequences of TCM are included;
- ✓ ITS2 is selected as the core DNA barcode for medicinal plants and psbA-trnH is selected as the supplementary DNA barcode;
- ✓ COI is for medicinal animals;
- ✓ NOT ONLY medicinal plants and animals, their close relatives, inferior substitutes, adulterants, and counterfeits are also included;



TCMbarcode System

✓ **How to identify species using TCMbarcode system**





Identification page for

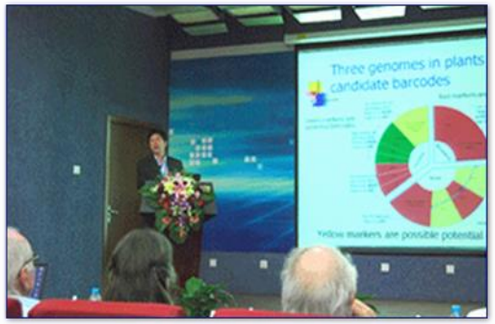
- Medicinal plants (ITS2)
- Medicinal plants (psbA-trnH)
- Medicinal animals (COI)

- Query sequence should be longer than 50bp
- FASTA or only DNA format is accepted
- Case insensitive

First postulated that ITS2 as DNA barcode for identifying medicinal plant species

- First reported in 2009

  CBOL and Third International Conference for DNA barcode.



- First published in 2010

  ITS2 as a novel DNA barcode for identifying medicinal plant species.

OPEN ACCESS Freely available online PLoS one

Validation of the ITS2 Region as a Novel DNA Barcode for Identifying Medicinal Plant Species

Shilin Chen[1*], Hui Yao[1], Jianping Han[1], Chang Liu[2], Jingyuan Song[1*], Linchun Shi[1], Yingjie Zhu[1], Xinye Ma[1], Ting Gao[1], Xiaohui Pang[1], Kun Luo[3], Ying Li[1], Xiwen Li[1], Xiaocheng Jia[1], Yulin Lin[1], Christine Leon[4]

- PNAS 2011

  ITS/ITS2 should be incorporated into the core barcode for seed plants.

PNAS

We therefore propose that ITS/ITS2 should be incorporated into the core barcode for seed plants.

PNAS 2011

ITS2 has high efficiency and stability in the identification of traditional Chinese medicines

Efficient identification of related species

- Rosaceae (1410 samples), Compositae (3490), Leguminosae (1507), Euphorbiaceae (1183), Rutaceae (300).

W+	W-	Relative Ranks, *n*, *P* value	Result
ITS2	*matK*	W+ = 703, W- = 0, $n = 37$, $P \leq 1.139 \times 10^{-7}$	ITS2 > *matK*
ITS2	*rbcL*	W+ = 703, W- = 0, $n = 37$, $P \leq 1.137 \times 10^{-7}$	ITS2 > *rbcL*
ITS2	*rpoC1*	W+ = 703, W- = 0, $n = 37$, $P \leq 1.133 \times 10^{-7}$	ITS2 > *rpoC1*
matK	*rbcL*	W+ = 620, W- = 46, $n = 36$, $P \leq 6.479 \times 10^{-6}$	*matK* > *rbcL*
matK	*rpoC1*	W+ = 655, W- = 11, $n = 36$, $P \leq 4.174 \times 10^{-7}$	*matK* > *rpoC1*
rbcL	*rpoC1*	W+ = 359, W- = 106, $n = 30$, $P \leq 0.009$	*rbcL* > *rpoC1*

Effitively identify adulterants from traditional Chinese medicines

- 6172 samples from roots, rhizomes or skins.
- 4385 samples from whole plants, flowers, fruits or seeds.

The identification of ITS2 is of universal applicability: 48,000 samples

Species diversity

- 11,000 species.
- 217 families, 1,251 genera.

Accurate result

- Confirmed by experts from Kew Garden.
- The sample number is more than 10.

Official calibration

- Control from CAIQ.
- Rechecked by others inspection offices.

北京市药品检验所
Beijing Institute For Drug Control (Beijing Center

深圳市药品检验所深圳市医疗器械检测中心
Shenzhen Institute For Drug Control Shenzhen Testing Center of Medical Devices

上海市食品药品检验所

Multiple Check

Analysis:

- BLAST
- Barcoding Gap
- Phylogenetic Tree

Ten official inspection departments agreed that DNA barcoding is suitable for the identification of TCM.

Current Opinion in Biotechnology (2014) :

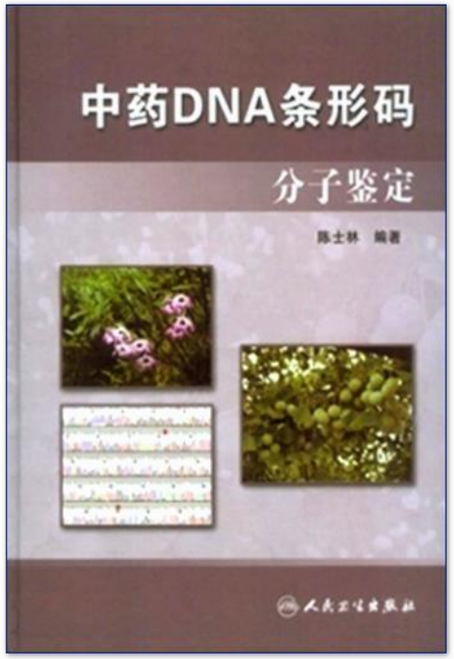
ITS or ITS 2 showed the highest discrimination rate.

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

Construction of DNA barcoding identification system for herbs

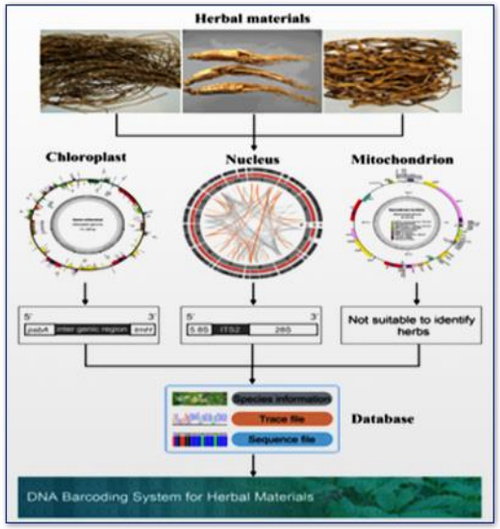
Framework of DNA barcoding identification system

- ITS2 as main sequence
- psbA-trnH as assisted sequence



SOP

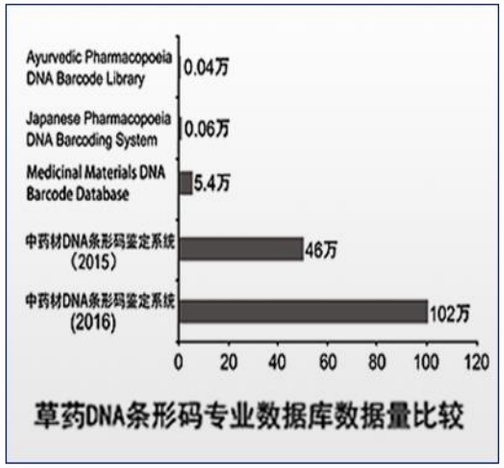
- Pre-processing
- Extraction for DNA
- PCR amplification
- Obtained DNA barcoding sequence.
- Judgement



Identification platform for herbs

- The largest DNA barcoding professional database for herbs in the world(1.02 million sequences, 2016)

www.tcmbarcode.cn



Software system for identification

- Software for assembly
- Software for DNA barcoding of herbs

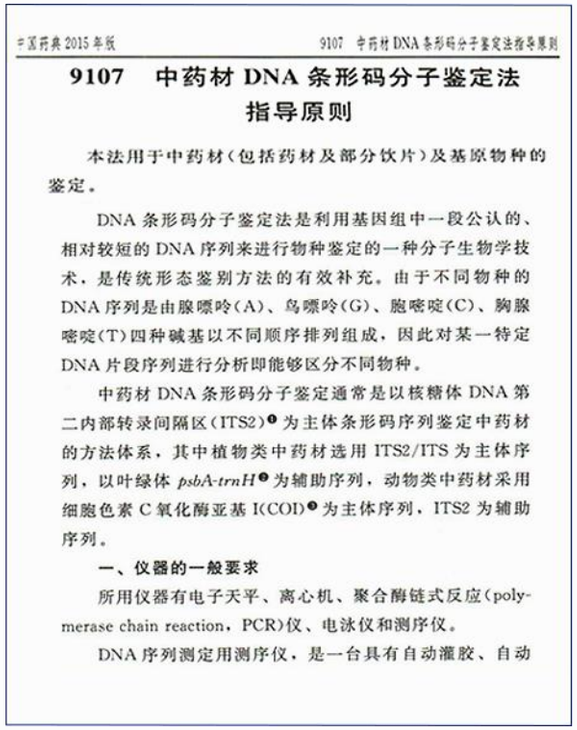
Standard DNA barcodes of Chinese medicinal materials was incorporated into *Chinese Pharmacopoeia*

✓ **As a routine method in Chinese pharmacopoeia**

✓ Published a book ***DNA Barcoding for Chinese Medicinal Materials***

- 《中国药典》2010，2015年版 采用中药材DNA条形码分子鉴定法指导原则

- 《中国药典中药材DNA条形码标准序列》



中国药典 2015 年版　　9107　中药材 DNA 条形码分子鉴定法指导原则

9107　中药材 DNA 条形码分子鉴定法指导原则

本法用于中药材（包括药材及部分饮片）及基原物种的鉴定。

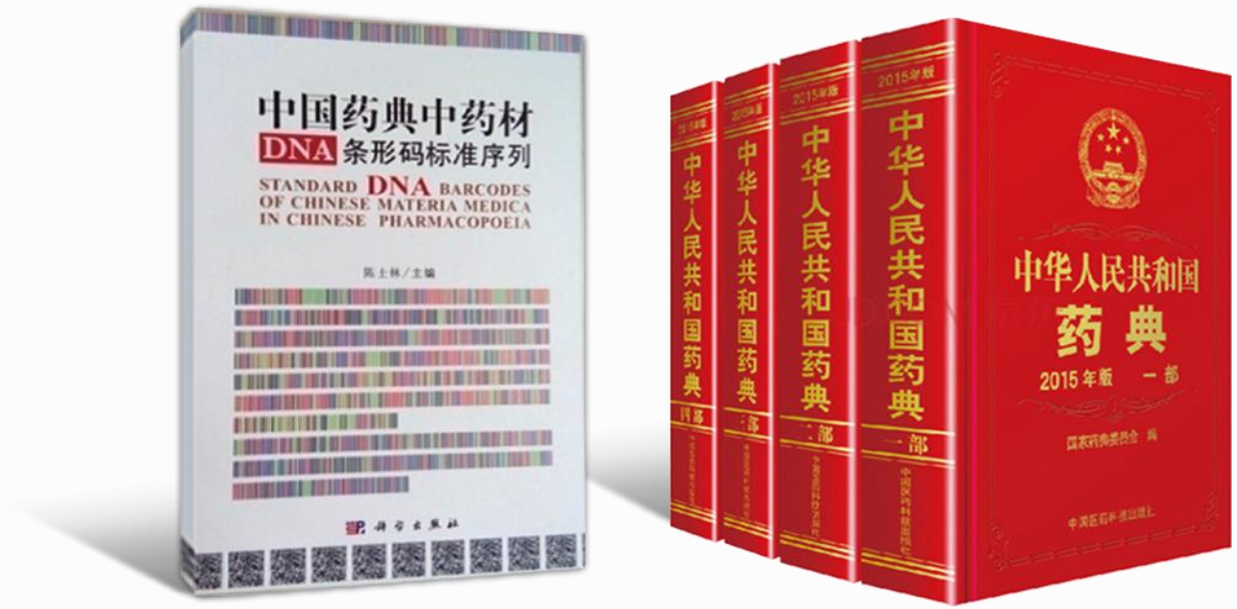
DNA 条形码分子鉴定法是利用基因组中一段公认的、相对较短的 DNA 序列来进行物种鉴定的一种分子生物学技术，是传统形态鉴别方法的有效补充。由于不同物种的 DNA 序列是由腺嘌呤（A）、鸟嘌呤（G）、胞嘧啶（C）、胸腺嘧啶（T）四种碱基以不同顺序排列组成，因此对某一特定 DNA 片段序列进行分析即能够区分不同物种。

中药材 DNA 条形码分子鉴定通常是以核糖体 DNA 第二内部转录间隔区（ITS2）[1] 为主体条形码序列鉴定中药材的方法体系，其中植物类中药材选用 ITS2/ITS 为主体序列，以叶绿体 *psbA-trnH*[2] 为辅助序列，动物类中药材采用细胞色素 C 氧化酶亚基 I（COI）[3] 为主体序列，ITS2 为辅助序列。

一、仪器的一般要求

所用仪器有电子天平、离心机、聚合酶链式反应（polymerase chain reaction，PCR）仪、电泳仪和测序仪。

DNA 序列测定用测序仪，是一台具有自动灌胶、自动



Chinese Pharmacopoeia 2015

Introduction of standard DNA barcodes of Chinese medicinal materials in some developed countries and regions

- Trained experts from British Pharmacopoeia Commission.

British Pharmacopoeia

BP 2017 – New monographs, DNA barcoding chapter and unlicensed medicines information

- Assisted the American Pharmacopoeia Commission in compiling standards for the identification of herbal medicines.

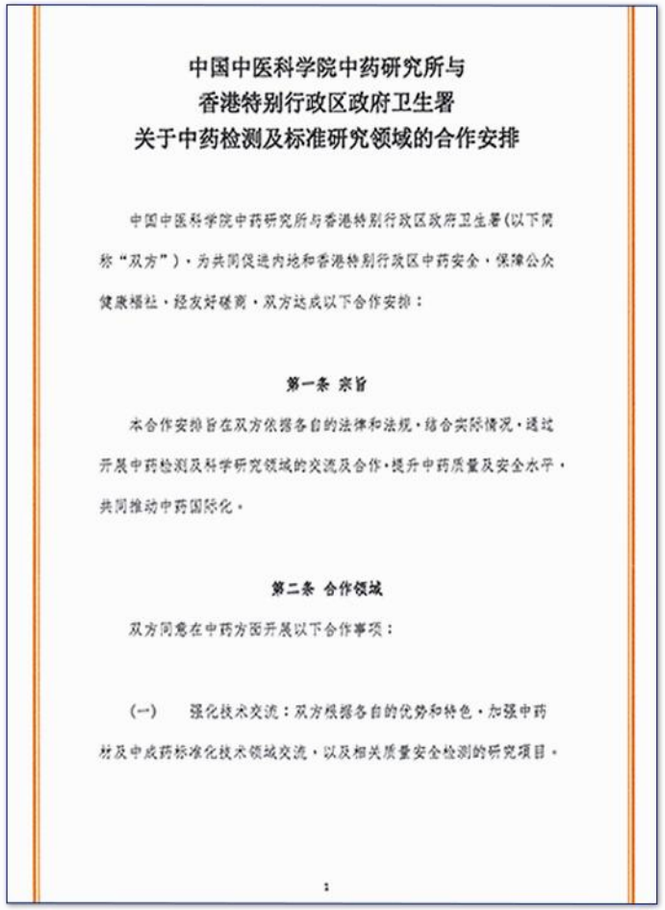
USP Global Expertise Trusted Standards Improved Health

U.S. Pharmacopeial Convention

Roundtable: DNA Methods for the Identification of Botanical Articles – Potential Role of USP Standards

Thursday, May 26, 2016
09:00 a.m. – 04:00 p.m. EST
USP, Rockville, MD

- Trained the staffs of Hongkong Health Department.



中国中医科学院中药研究所与
香港特别行政区政府卫生署
关于中药检测及标准研究领域的合作安排

中国中医科学院中药研究所与香港特别行政区政府卫生署(以下简称“双方”),为共同促进内地和香港特别行政区中药安全,保障公众健康福祉,经友好磋商,双方达成以下合作安排:

第一条 宗旨

本合作安排旨在双方依据各自的法律和法规,结合实际情况,通过开展中药检测及科学研究领域的交流及合作,提升中药质量及安全水平,共同推动中药国际化。

第二条 合作领域

双方同意在中药方面开展以下合作事项:

(一) 强化技术交流:双方根据各自的优势和特色,加强中药材及中成药标准化技术领域交流,以及相关质量安全检测的研究项目。

A renaissance in herbal medicine identification: from morphology to DNA

- A renaissance in herbal medicine identification: from morphology to DNA

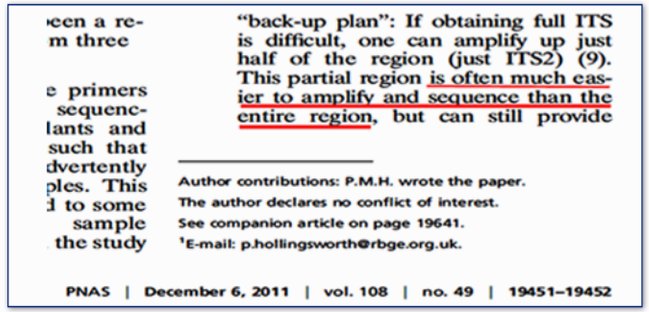
SL Chen et al Biotechnology Advances 2014

Pharmacopeia (Pharm.)	No. of crude herbal drugs		
	In pharm.	In the DNA barcoding database	Percentage(%)
Chinese Pharm.	510	505	99.0
Japanese Pharm.	154	148	96.1
Korean Pharm.	161	159	98.8
Indian Pharm.	52	50	96.2
U.S. Pharm.	42	40	95.2
European Pharm.	187	184	98.4

- PNAS (2011):

ITS2 is often much easier to amplify and sequence than ITS.

P. M. Hollingsworth （CBOL植物组主席）



…een a re-
m three

e primers
sequenc-
lants and
such that
dvertently
ples. This
l to some
sample
the study

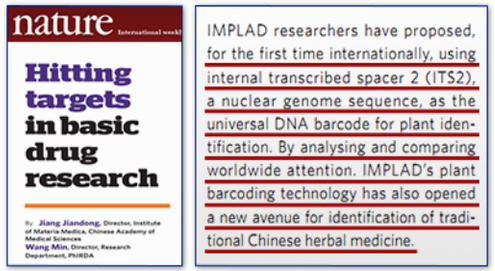
"back-up plan": If obtaining full ITS is difficult, one can amplify up just half of the region (just ITS2) (9). This partial region is often much easier to amplify and sequence than the entire region, but can still provide

Author contributions: P.M.H. wrote the paper.
The author declares no conflict of interest.
See companion article on page 19641.
[1]E-mail: p.hollingsworth@rbge.org.uk.

PNAS | December 6, 2011 | vol. 108 | no. 49 | 19451–19452

- Nature (2017):

IMPLAD's plant barcoding technology has also opened a new avenue for identification of traditional Chinese herbal medicine.



nature

Hitting targets in basic drug research

IMPLAD researchers have proposed, for the first time internationally, using internal transcribed spacer 2 (ITS2), a nuclear genome sequence, as the universal DNA barcode for plant identification. By analysing and comparing worldwide attention. IMPLAD's plant barcoding technology has also opened a new avenue for identification of traditional Chinese herbal medicine.

Nature 2017

- Genetic identity card of herbs safeguard the usage of drug in the clinic application from its origin.

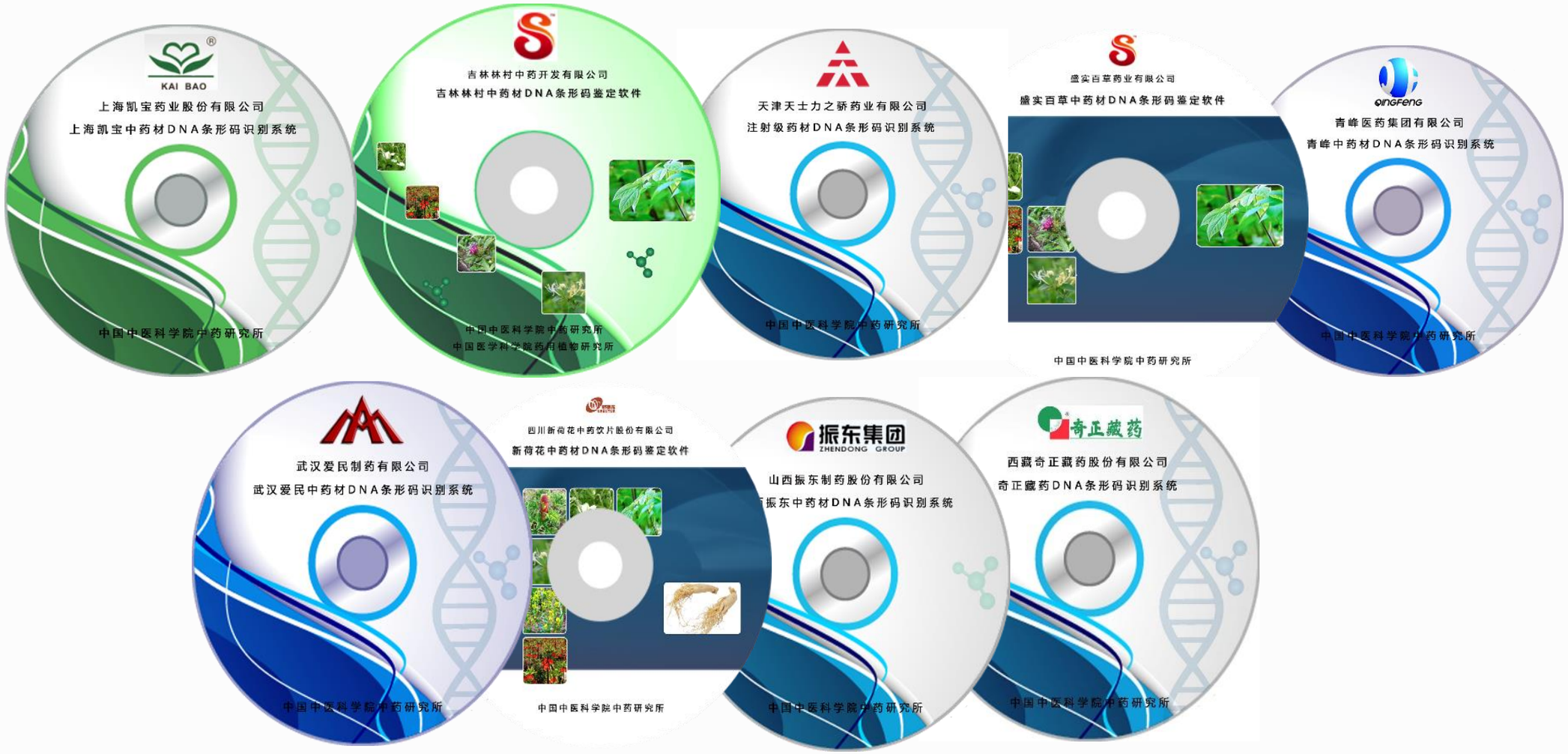
Applications

Applications of TCM barcode system to promote the safe use of herbal medicines

- ✓ Seed and seedling identification for farming: planting correct herbs
- ✓ Herbal supervision :control the sales of adulterants and substitutes in the medicinal material market
- ✓ Material control in the pharmaceutical industry
- ✓ The identification of herbal medicines in pharmacy
- ✓ The supervision of herbal medicines in customs
- ✓ Endangered herb protection

Applications in China

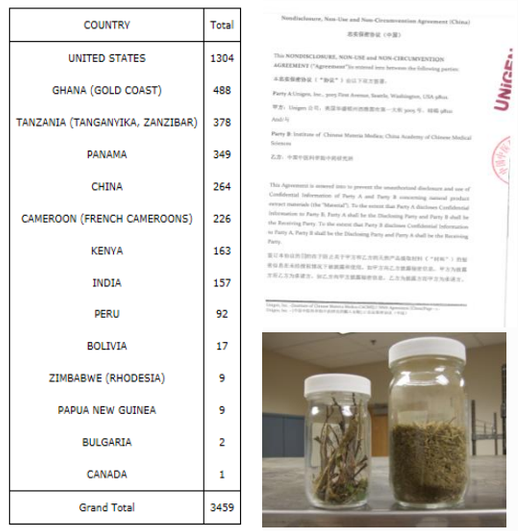
- ✓ Widely used in more than 100 TCM companies in China, including 天士力Group、振东Group 、青峰Group



Applications in world-wide

- Tsumura&Co, Janpan : Identification herbal medicine for TCM
- Unigen, USA : Identification of 7000 herbal samples
- Amway, USA: DNA barcoding system for food supplements in Amway
- British Pharmacopoica 、American Pharmacopoica
- Mississippi University、New York Botanical Garden, University of Macau

COUNTRY	Total
UNITED STATES	1304
GHANA (GOLD COAST)	488
TANZANIA (TANGANYIKA, ZANZIBAR)	378
PANAMA	349
CHINA	264
CAMEROON (FRENCH CAMEROONS)	226
KENYA	163
INDIA	157
PERU	92
BOLIVIA	17
ZIMBABWE (RHODESIA)	9
PAPUA NEW GUINEA	9
BULGARIA	2
CANADA	1
Grand Total	3459

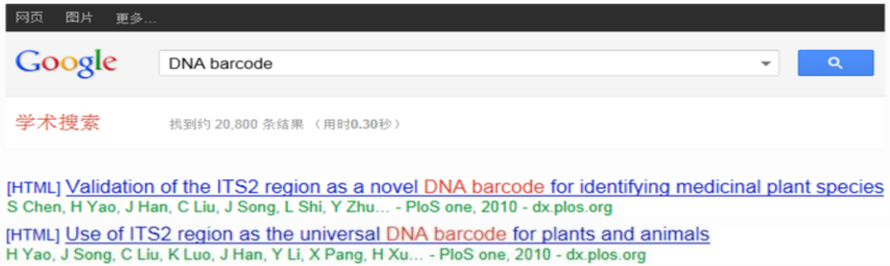


Samples of Unigen

Achievments

- National Science and Technology Progress — Second Prize
- The Natural Science of China association of Chinese Medicine — First Prize
- Ministry of Education Science and Technology Progress Award — First Prize

- Cited by **PNAS**(IF10.7*)、**Mol Biol Evol** (IF 10.5*) 、**Curr Opin Biotechnol**(IF 8.5*)



Acknowledgements