

## 北苍术特异性DNA条形码筛选、种质资源鉴定及遗传多样性分析

尹光耀<sup>1</sup>, 袁林<sup>2</sup>, 王馨<sup>1</sup>, 张志飞<sup>1</sup>, 陈颖<sup>1</sup>, 刘珊珊<sup>1</sup>, 满金辉<sup>1</sup>, 石玥<sup>1</sup>,  
黄钰莹<sup>1</sup>, 张晓芹<sup>1</sup>, 王晓晖<sup>3\*</sup>, 魏胜利<sup>1,4\*</sup>

(1. 北京中医药大学中药学院, 北京 102488; 2. 国家药品监督管理局, 北京 100037; 3. 北京中医药大学北京中医药研究院中药现代研究中心, 北京 102488; 4. 中药材规范化生产教育部工程研究中心, 北京 100102)

**摘要:** 北苍术 *Atractylodes chinensis* 具有重要药用价值和经济价值。本研究利用 Illumina 平台测序获得 4 份不同产地的北苍术的叶绿体全基因组序列, 筛选特异 DNA 条形码, 并利用特异 DNA 条形码对不同产区北苍术样品种质资源鉴定及居群的遗传多样性分析。该 4 份北苍术叶绿体全基因组均为典型的环状四分体结构, 均注释 112 个基因。比较基因组学研究表明 *ccsA* 和 *trnC-GCA\_petN* 是潜在的北苍术种内鉴别的特异 DNA 条形码。选择 *ccsA* 和 *trnC-GCA\_petN* 对来自 9 省 14 产区的 256 份样品进行 PCR 扩增, 扩增效率为 100%。序列分析结果表明 *ccsA* 和 *trnC-GCA\_petN* 分别有 11 和 22 个变异位点, 分别能鉴定 16 和 22 个单倍型, 两段序列联合分析鉴定 39 个单倍型, 命名为 Hap1~Hap39, 其中占比最多和分布最广的基因型为 Hap9。单倍型多样性 ( $H_d$ ) = 0.896, 核苷酸多样性 ( $\Pi$ ) = 0.002 22, 说明北苍术在物种水平上有较高的遗传多样性。各单倍型的遗传距离为 0.000 00~0.004 88, 说明各个单倍型之间有较小的遗传差异。系统进化树分析结果表明, 39 个单倍型有很近的亲缘关系, 并且与除白术外其余同属类群形成明显的两个分支。本研究为后续鉴定北苍术产地来源和后续种质资源保护和育种方面起到重要作用。

**关键词:** 苍术; 叶绿体基因组; DNA 条形码; 单倍型; 遗传多样性

中图分类号: R931 文献标识码: A 文章编号: 0513-4870(2023)06-1693-12

## Screening of specific DNA barcode, identification of germplasm resources, and analysis of genetic diversity of *Atractylodes chinensis*

YIN Guang-yao<sup>1</sup>, YUAN Lin<sup>2</sup>, WANG Xin<sup>1</sup>, ZHANG Zhi-fei<sup>1</sup>, CHEN Ying<sup>1</sup>, LIU Shan-hu<sup>1</sup>, MAN Jin-hui<sup>1</sup>, SHI Yue<sup>1</sup>, HUANG Yu-ying<sup>1</sup>, ZHANG Xiao-qin<sup>1</sup>, WANG Xiao-hui<sup>3\*</sup>, WEI Sheng-li<sup>1,4\*</sup>

(1. School of Chinese Materia Medica, Beijing University of Chinese Medicine, Beijing 102488, China; 2. National Medical Products Administration, Beijing 100037, China; 3. Modern Research Center for Traditional Chinese Medicine, Beijing Institute of Traditional Chinese Medicine, Beijing University of Chinese Medicine, Beijing 102488, China; 4. Engineering Research Center of Good Agricultural Practice for Chinese Crude Drugs of Ministry of Education, Beijing 100102, China)

**Abstract:** *Atractylodes chinensis* has important medicinal and economic values. In this study, the chloroplast genome sequences of four *A. chinensis* samples from different producing areas were sequenced using the Illumina platform. The specific DNA barcodes were screened and the germplasm resources of *A. chinensis* samples from different producing areas and the genetic diversity of the population were analyzed basing on the specific barcodes. The whole chloroplast genomes of the four *A. chinensis* samples had a typical cyclic tetrad structure, with 112

收稿日期: 2022-10-18; 修回日期: 2022-11-29.

基金项目: 北苍术、赤芍精准药材批次分子防伪技术研究项目 (2020110037009381); 北京市科学技术委员会基金项目 (Z201100005420005).

\*通讯作者 Tel: 13683336930, E-mail: wsl7491@126.com;

Tel: 15201166873, E-mail: wangxhui2014@163.com

DOI: 10.16438/j.0513-4870.2022-1118

genes annotated. The comparative genomics results indicated that *ccsA* and *trnC-GCA\_petN* were potential specific DNA barcodes for intraspecific identification of *A. chinensis*. Polymerase chain reaction (PCR) analysis of *ccsA* and *trnC-GCA\_petN* was performed on 256 samples from 14 areas in 9 provinces, and the amplification efficiency was 100%. Sequence analysis showed that *ccsA* and *trnC-GCA\_petN* had 11 and 22 variant positions, which could identify 16 and 22 haplotypes, respectively. The combined sequence analysis identified 39 haplotypes, named Hap1-Hap39, of which the most abundant and widely distributed genotype was Hap9. Haplotype diversity ( $H_d$ ) = 0.896 and nucleotide diversity ( $P_i$ ) = 0.002 22 indicated high genetic diversity at the species level in *A. chinensis*. The genetic distances of the haplotypes were 0.000 00-0.004 88, indicating that there were small genetic differences among the haplotypes. The results of phylogenetic tree analysis showed that 39 haplotypes had very close genetic relationship, and formed two obvious branches with other groups of the same genus except *Atractylodes macrocephala*. This study plays an important role in the identification of the origin of *A. chinensis* and the protection and breeding of germplasm resources.

**Key words:** *Atractylodes chinensis*; chloroplast genome; DNA barcode; haplotype; genetic diversity

苍术是菊科植物北苍术 [*Atractylodes chinensis* (DC) Koidz.] 或茅苍术 [*Atractylodes lancen* (Thunb.) DC.] 的干燥根茎, 具有燥湿健脾、祛风散寒、明目的功效。由于茅苍术多为野生分布, 人工栽培较少, 产量较低, 因此北苍术是目前市场的主流品种。药理学研究表明北苍术具有抗癌、抑菌、抗病毒等药理作用<sup>[1]</sup>, 临床上常用于病毒性疾病的治疗<sup>[2-6]</sup>, 因此在药物研发中频繁使用。近年来苍术药材市场需求量大, 供不应求, 急需扩大北苍术的生产规模。因此对北苍术药材进行种质资源鉴定和遗传多样性分析, 有利于后续北苍术优良品系的选育和保护。

叶绿体是一种多功能细胞器, 具有独立的遗传物质。细胞中叶绿体基因组最常见的结构是双链环状构型, 包括1个小的单拷贝区 (SSC) 和1个大的单拷贝区 (LSC); 这两个区域被一对反向重复区域 (IRA、IRB) 分开, 形成典型的四部分结构。基因组大小从120到160 kb<sup>[7]</sup>。叶绿体基因组结构、大小和基因种类都比较保守, 编码区和非编码区进化速率和模式具有物种鉴别意义, 在研究物种遗传背景、系统进化和物种特异DNA条形码筛选方面具有独特优势。随着叶绿体基因组测序和分析技术的进步, 多种药用植物的叶绿体基因组都已解析, 如南柴胡、北柴胡、长梗秦艽、大黄等物种的叶绿体基因组陆续解析出来。目前苍术属植物北苍术、茅苍术、关苍术、朝鲜苍术、白术等的叶绿体基因组已有研究报道, 但同种不同个体叶绿体比较基因组学的相关研究却鲜有报道。

DNA条形码技术利用生物基因组中一段通用的、具有差异特征的片段进行物种鉴定或种质资源鉴定, 可不受经验和外界环境的影响, 具有准确、客观等特点。如Long等<sup>[8]</sup>的研究结果表明, *trnK-matK*能区分黄精、滇黄精、点花黄精和湖北黄精, 遗传分化指数比较

高, 能区分黄精属物种。Zhang等<sup>[9]</sup>利用大黄的 *matK* 序列能鉴定大黄的41个单倍型, 可将正品大黄的3个基原与混伪品进行区分。比较叶绿体基因组学的发展为物种间和种内鉴定的特异DNA条形码的筛选和应用奠定基础。如通过比较叶绿体基因组学分析和筛选秦艽的特异性DNA片段 *trnT-trnL* 和 *ycf*, 能鉴别长梗秦艽和全萼秦艽<sup>[10]</sup>。

遗传多样性是生物在经历长期的选择后形成的自然属性, 是物种进化的本质。研究种群的遗传多样性, 了解植物遗传多样性的影响因素和变化规律, 对于选育优良植物品种并对其进行保护和开展规范化种植具有重要意义<sup>[11]</sup>。目前苍术属内的叶绿体基因比较分析及近缘关系已有研究报道, 但针对北苍术物种内的叶绿体高变异区域的遗传多样性研究鲜有报道。本研究选取不同产地北苍术样品, 进行比较叶绿体基因组学研究, 筛选特异性DNA条形码, 并基于筛选的特异性DNA条形码对9省14产区256份样品进行种质资源鉴定、种群遗传多样性分析, 为北苍术种质资源评价、保护和优良品种选育提供科学依据。

## 材料与方法

**叶绿体基因组测序材料** 北苍术叶绿体基因组测序样品分别来自内蒙古自治区兴安盟扎赉特旗、河北省保定市安国市、辽宁省朝阳市朝阳县台子镇、陕西省商洛市商州区。

**实验样品材料** 供试的北苍术来源见表1, 所有样品经北京中医药大学中药资源教研室魏胜利教授鉴定, 来自9省14产区的共计256份样品存放于北京中医药大学种质圃中。样品信息见表1。选取生长状况良好的植株中部的健康嫩叶, 用70%乙醇擦去叶片表面灰尘用于总DNA的提取。

**Table 1** The information of samples

Region	Label	Latitude	Longitude	Altitude/m	Number
Zhalaitqi, Neimenggu	NZL	47.31	122.43	270	38
Hulunbeier, Neimenggu	NHL	48.13	123.45	700	15
Chifeng, Neimenggu	NC	39.98	116.42	580	4
Chengde, Hebei	HBCD	38.43	115.33	30	42
Xinxiang, Henan	HNXX	33.19	112.85	69	16
Luoyang, Henan	HNLY	34.70	112.51	144	14
Daxinganling, Heilongjiang	HLJDX	51.93	124.59	625	18
Xi'an, Shaanxi	SXXA	33.86	109.94	1 233	7
Shangluo, Shaanxi	SXSL	34.23	108.93	636	8
Baishan, Jilin	JLBS	41.95	126.42	475	2
Zhangjiakou, Hebei	HBZJK	40.97	115.28	726	23
Zibo, Shandong	SDZB	36.49	117.86	39	17
Miyun, Beijing	BJMY	40.38	116.84	83	10
Linfen, Shanxi	SXLF	33.86	109.94	800	42

**样品基因组 DNA 的提取** 将分别来自内蒙古自治区兴安盟扎赉特旗 (*Atractylodes chinensis* 1)、河北省保定市安国市 (*Atractylodes chinensis* 2)、辽宁省朝阳市朝阳县台子镇 (*Atractylodes chinensis* 4)、陕西省商洛市商州区 (*Atractylodes chinensis* 5) 的北苍术利用十六烷基三甲基溴化铵 (CTAB) 法<sup>[12]</sup>, 从 100 mg 新鲜叶片中提取 DNA 进行高通量测序。对来自 9 省 14 产区的共计 256 份样品 (表 1), 选取干燥叶片, 使用 75% 乙醇擦拭表面, 干燥后放入无菌研钵, 加入适量液氮研磨成细粉, 称取 20 mg 粉末, 用植物基因组提取试剂盒 (Vazyme 公司) 提取北苍术 DNA, 用于后续 PCR 扩增获得特定序列和用于验证拼接的正确性等工作。

**样品叶绿体基因组的获取** 基因组 DNA 的质量和浓度用琼脂糖凝胶电泳和 NanoDrop one 分光光度计 (Thermo Fisher 公司) 进行评价。合格的 DNA 用于文库构建。在 Illumina 平台上进行测序。由于原始测序数据可能包含低质量序列、接头序列等, 为了保证信息分析结果的可靠性, 需经过一系列数据处理过滤这些杂质 raw reads, 从而得到 clean reads, 仍以 FASTQ 格式存储。对原始测序数据进行如下处理: ① 去除质量值连续  $\leq 20$  的碱基数达到 40% 的 reads; ② 去除含 N 的碱基数目总和达到 10% 的 reads; ③ 去除 adapter 污染; ④ 去除 duplication 污染。使用 NOVOPlasty 组装软件对 clean data 进行组装, 以 NCBI 已发表的北苍术叶绿体基因组 (MG874805) 作为参考, 获得各样品组装结果。采用 PGA (Plastid Genome Annotator) 软件进行组装结果的基因预测, 获得基因信息, 分析采用默认参数。获得测序样品基因集后, 需对基因进行数据库比对注释, 确定基因的功能及相关描述信息。能整体反映出样品基因功能分类及方便后续研究找到目的功能基因。对此进行了 GO (gene ontology)、COG (cluster of orthologous groups of proteins)、KEGG (Kyoto

encyclopedia of genes and genomes)、ARDB (antibiotic resistance genes database)、PHI (pathogen host interactions)、CAZy (carbohydrate-active enzymes database)、VFDB (virulence factor database)、T3SS (type III secretion system effector protein) 等 8 个数据库注释。基因注释主要基于氨基酸序列比对。将基因的氨基酸序列与各数据库进行比对, 得到对应的功能注释信息。由于每一条序列比对结果超过一条, 为保证其生物意义, 保留一条最优比对结果作为该基因的注释。所有注释均使用 BLAST 软件结合各个数据库的特点完成, 提供的 BLAST 结果为 M8 格式, 同时还提供部分数据库的注释结果汇总。

**重复序列检测** 利用 REPuter 软件 (<https://bibiserv.cebitec.uni-bielefeld.de/reputer>) 检测叶绿体全基因组序列中的分散重复序列 (dispersed repeats), 参数设置为最小重复序列长度 = 30 bp, 重复序列间的相似度 > 90%<sup>[13]</sup>。串联重复序列 (tandem repeats) 利用 Tandem repeats finder 软件 (<https://tandem.bu.edu/trf/trf.html>) 进行检测, 参数设置选择默认值<sup>[14]</sup>。利用 MISA 软件对本研究新测序获得的北苍术 4 个体叶绿体全基因组序列中的简单重复序列 (simple sequence repeats, SSRs) 分别进行检测, 参数设置均为单核苷酸重复单元不少于 10 个, 二核苷酸重复单元不少于 5 个, 三核苷酸和四核苷酸重复单元不少于 4 个, 五核苷酸和六核苷酸重复单元不少于 3 个, 且两个 SSRs 之间的距离不小于 100 bp<sup>[15]</sup>。并对 SSRs 的类型、数量和分布模式分别进行比较分析。

**叶绿体全基因组比较分析与序列变异比较** 利用 mVISTA 软件在 LAGAN 模型<sup>[16]</sup>下对本研究新测序获得的北苍术个体的叶绿体全基因组序列进行比较分析。同时, 对同种不同个体间叶绿体基因组的 IR/SC 边界位置进行比较, 分析 IRs 区收缩与扩张情况。为

了进一步分析苍术属植物叶绿体全基因组序列间的变异情况,利用 DnaSP 软件<sup>[17]</sup>进行滑动窗口分析 (sliding window analysis), 计算核苷酸多样性 (Pi), 并根据分析结果筛选种间高变异序列。滑动窗口 (windows length) 大小设置为 600 bp, 步长 (step size) 为 200 bp。

**PCR 扩增体系及条件** PCR 扩增体系为 50  $\mu$ L, 包括总 DNA 1.0  $\mu$ L, 10 $\times$ Taq buffer 5.0  $\mu$ L, dNTP Mix (2.5 mmol $\cdot$ L<sup>-1</sup>) 4.0  $\mu$ L, Taq (2.5 U $\cdot$  $\mu$ L<sup>-1</sup>) 0.5  $\mu$ L, 0.1% BSA 1.0  $\mu$ L, 正反向引物各 1.0  $\mu$ L (10  $\mu$ mol $\cdot$ L<sup>-1</sup>), 加 ddH<sub>2</sub>O 至 50.0  $\mu$ L。充分混匀后备用。PCR 扩增结束后, 产物经过纯化采用双向测序。引物名称、序列及扩增条件见表 2。

**数据处理** 利用 DNAMAN 和 Chromas 软件对测序结果进行核对和单倍型汇总。将汇总后的单倍型序列通过 mafft 软件进行比对, 比对后的序列通过 MEGA 5 软件构建邻接法系统进化树, 设置 bootstrap 重复值为 1 000。通过 DnaSP 6 对总体和每个种群的遗传多样性参数、种群间的基因流 ( $N_m$ ) 和分化系数 (Gst) 进行估计。

## 结果与分析

### 1 叶绿体全基因组测序、序列组装拼接与注释

组装拼接后 4 份不同产地的北苍术样品均得到完整的环状四分体结构叶绿体全基因组, 序列总长度分别为 153 111、153 068、153 266 和 153 194 bp。由大单拷贝区 (长度分别为 84 160、84 302、84 249 和 84 279 bp)、小单拷贝区 (18 657、18 658、18 679 和 18 626 bp) 和一对反向重复区 (25 147、25 054、25 149 和 25 147 bp) 组成 (图 1)。4 份北苍术叶绿体全基因组序列及 SSC 和 IRs 区的 GC 含量均相同, 依次为 38.5%、31.5% 和

43.2%, LSC 区的 GC 含量分别为 37.8%、37.8%、37.7% 和 37.7% (表 3)。

在北苍术叶绿体全基因组序列中注释得到 112 个基因, 包括 76 个蛋白编码基因、28 个 tRNA 基因和 8 个 rRNA 基因 (表 4)。其中, 16 个基因 (*ndhA*、*ndhB*、*petB*、*petD*、*atpF*、*rpl16*、*rpl2*、*ps12*、*rps16*、*rpoC1*、*trnA-UGC*、*trnG-UCC*、*trnI-GAU*、*trnK-UUU*、*trnL-UAA* 和 *trnV-UAC*) 含有 1 个内含子、2 个基因 (*clpP*、*ycf3*) 含有 2 个内含子和 18 个基因 (*ndhB*、*rpl2*、*rpl23*、*rps12*、*rps7*、*rrn16*、*rrn23*、*rrn4.5*、*rrn5*、*trnA-UGC*、*trnI-CAU*、*trnI-GAU*、*trnL-CAA*、*trnN-GUU*、*trnR-ACG*、*trnV-GAC*、*ycf15*、*ycf2*) 为多拷贝基因。*rps12* 基因的 5'端和两个重复的 3'端分别位于 LSC 区和 IR 区, 被分成 2 个独立的转录单元, 为反式剪切 (*trans-spliced*) 基因。

### 2 重复序列分析

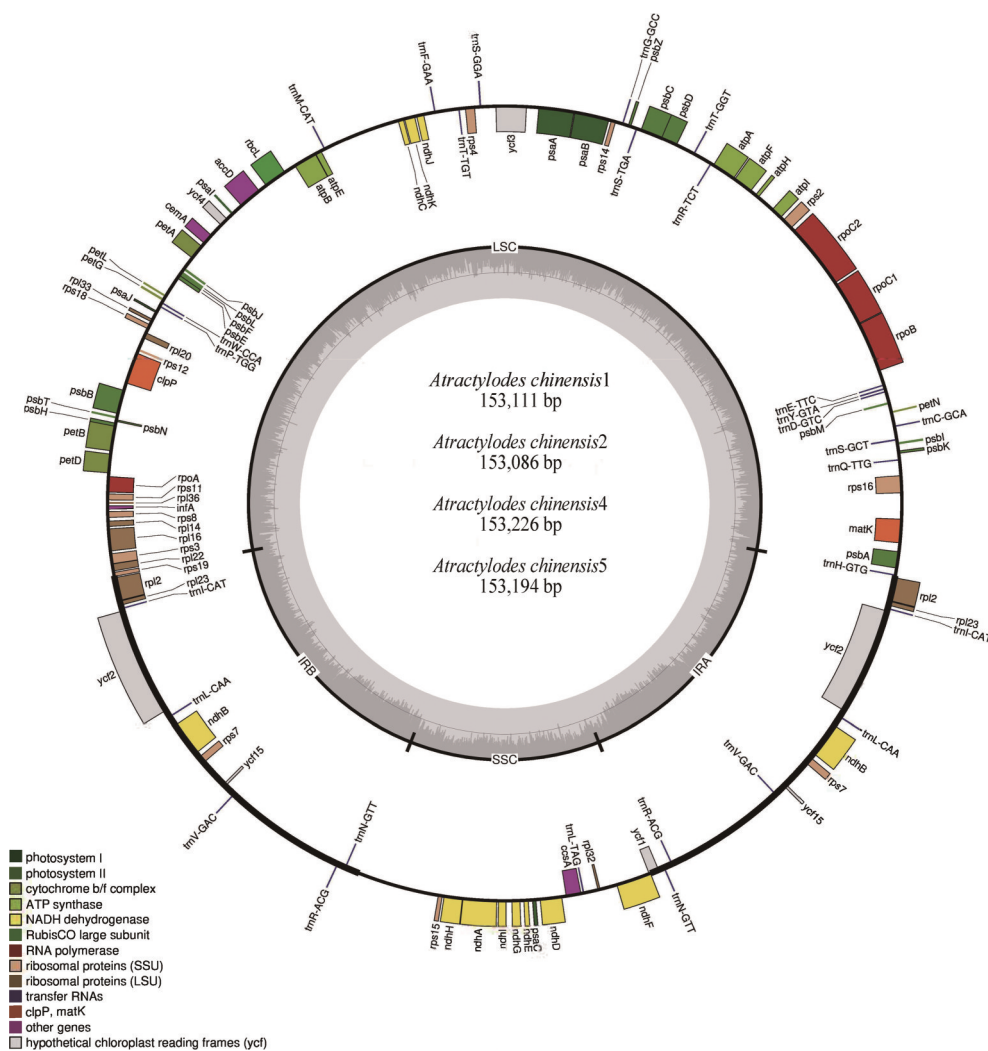
4 份不同产地北苍术植物的叶绿体基因组中分别检测到 49、50、49、44 个 SSR (*Atractylodes chinensis* 1、2、4、5)。SSR 在叶绿体基因组中随机分布。4 份样品的 LSC 区分别有 36、37、34、32 个 SSR; SSC 区分别有 6、6、7、6 个 SSR; IRa 区分别有 2、2、5、3 个 SSR; IRb 区分别有 5、5、4、3 个 SSR。间隔区分别有 32、33、33、30 个 SSR; 内含子中分别有 8、8、8、5 个 SSR; 外显子中分别有 9、9、9、9 个 SSR。在个体基因组中, 大多数 SSR 存在于 LSC 和间隔区。最常见的 SSR 是单核苷酸, 在不同个体中分别占 71%、70%、72%、70%, 其次是四核苷酸, 在不同个体中分别占 10%、12%、12%、13%, 二核苷酸 SSR 在不同个体中分别占 10%、10%、6%、7% (图 2A)。几乎所有的单核苷酸 SSR (99%) 在所有 6 个物种中都由 A 和 T 组成。TA 的二核苷酸重复序列和 TTTC 的四核苷酸重复序列是第二常见的 SSR (图 2B)。

**Table 2** Primer names, sequences, and amplification conditions

DNA fragment	Primer	Sequence (5'-3')	Amplification procedure
<i>ccsA</i>	<i>ccsA-F</i>	ATGATATTTTCAACTTTAGAGCA	Pre-denaturation at 94 °C for 4 min; denaturation at 94 °C for 45 s, annealing at 48.0 °C for 30 s, 72 °C extension of 1 min (35 cycles); finally, at 72 °C for 5 min
	<i>ccsA-R</i>	AAACGAACCATAACTATGTAGCC	
<i>trnC-GCA_petN</i>	<i>trnC-GCA-F</i>	CATCAACAAACGAATCGAAAT	Pre-denaturation at 94 °C for 4 min; denaturation at 94 °C for 45 s, annealing at 49.2 °C for 30 s, 72 °C extension of 1 min (35 cycles); finally, at 72 °C for 5 min
	<i>trnC-GCA-R</i>	AATCAATGGTCCAGAGTCAAA	

**Table 3** Comparison of the plastid genome features of four *Atractylodes chinensis*. LSC: Large single-copy; SSC: Small single-copy; IR: Inverted repeats; CDS: Coding DNA sequence

Species	Genome		LSC		SSC		IR		CDS	
	Length/bp	G+C/%	Length/bp	G+C/%	Length/bp	G+C/%	Length/bp	G+C/%	Length/bp	G+C/%
<i>A. chinensis</i> 1	153 111	38.52	84 160	37.79	18 657	31.53	25 147	43.17	78 789	37.94
<i>A. chinensis</i> 2	153 068	38.52	84 302	37.79	18 658	31.51	25 054	43.16	78 789	37.94
<i>A. chinensis</i> 4	153 226	38.52	84 249	37.71	18 679	31.52	25 149	43.16	78 810	37.95
<i>A. chinensis</i> 5	153 194	38.53	84 279	37.7	18 626	31.56	25 147	43.16	78 750	37.95



**Figure 1** Gene maps of the chloroplast genomes of *Atractylodes chinensis*. Genes on the inside of the large circle are transcribed clockwise and those on the outside are transcribed counterclockwise. The genes are color-coded based on their functions. The dashed area represents the GC composition of the chloroplast genomes

### 3 比较基因组学分析

研究发现 IR/SC 边界区域在 *Atractylodes chinensis* 1、2、4、5 样本中仍相对保守 (图 3)。这 4 个不同产地北苍术叶绿体的比较表明, IRS 区没有明显的扩张和收缩。在 LSC/IR 和 SSC/IR 交界处发现了 *rps19*、*ndhF*、*rpl22*、*ycf1*、*rpl2* 和 *trnH* (Gug) 基因。这些基因在 4 个叶绿体的 SC/IR 交界处表现出相似分布状态。

北苍术植物叶绿体全基因组序列所编码的基因类别、数量和排列顺序在种内均高度一致; 序列间的变异主要发生在非编码基因间区。核苷酸多样性分析结果显示, 种内序列的 Pi 值整体变化范围为 0~0.003 89。同时, 筛选到 7 条 Pi 值较高的种内高变异序列, 分别位于 LSC 区 (*psbA*、*atpI*、*atpH*、*rpl16*、*trnC-GCA*、*petN*、*rps14*) 和 SSC 区 (*rps15*、*ccsA*) (图 4)。此外, 基于 mVISTA 的比较分析在苍术属的 4 个叶绿体基因组中

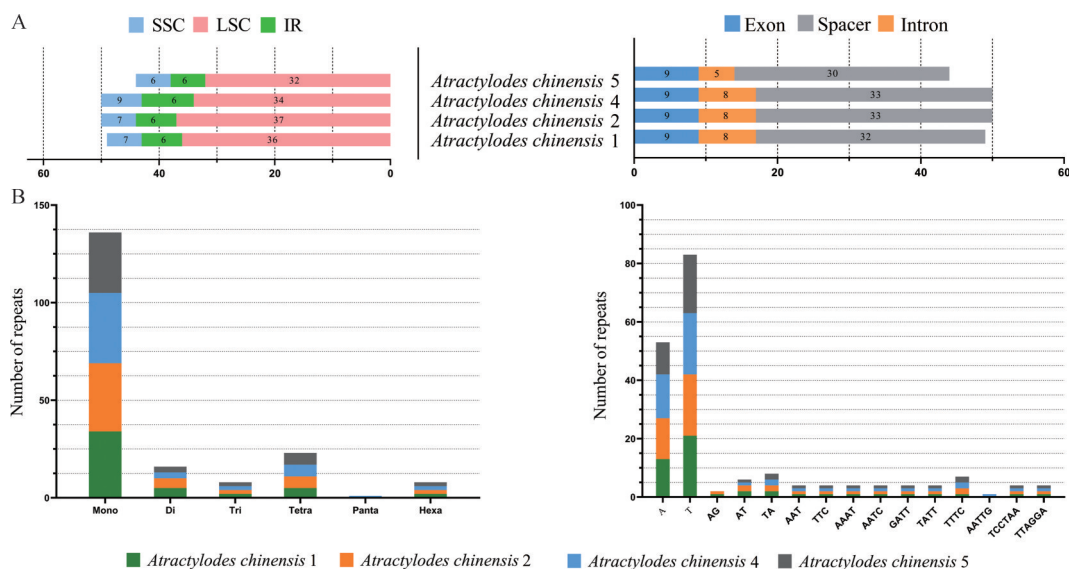
进行, 以确定变异的水平 (图 5), 结果表明整个叶绿体基因组的序列高度相似, 说明叶绿体基因组是高度保守的。IR 区和编码区比单复制区和非编码区更保守。*trnR-TCT*、*ccsA*、*rps16*、*ndhD*、*trnT-GGT* 基因及 *trnC-GCA*、*petN*、*trnN-GTT*、*rps15* 基因间隔区的编码区比其他基因的编码区更具变异性。综上所述, 北苍术种内基因的保守程度很高, 结合 DnaSP 结果中的 Pi 值和 mVISTA 结果, *ccsA*、*trnC-GCA*、*petN* 为潜在的北苍术种内鉴定的特异性 DNA 条形码。

### 4 基于叶绿体基因组高变区域的单倍型分析

基于上述筛选的特异性 DNA 条形码的结果, *ccsA*、*trnC-GCA*、*petN* 大小为 969、821 bp, 设计基因 *ccsA* 及 *trnC-GCA*、*petN* 特异性引物, 提取来自 9 省 14 产区 256 份北苍术样品总 DNA, 以 DNA 为模板进行 PCR 扩增, 并对测序结果进行分析。设计的引物均可对 9 省 14 个

**Table 4** Gene composition in the plastid genomes of *Atractylodes chinensis*. \*: The gene with one intron; \*\*: The gene with two introns; (2): The copy number of the multi-copy gene; NADH: Nicotinamide adenine dinucleotide; ATP: Adenosine triphosphate; CoA: Coenzyme A; ORF: Open reading frame

Category	Gene group	Gene name	
Photosynthesis	Subunits of photosystem I	<i>psaA, psaB, psaC, psaI, psaJ</i>	
	Subunits of photosystem II	<i>psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ</i>	
	Subunits of NADH dehydrogenase	<i>ndhA*, ndhB*(2), ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK</i>	
	Subunits of cytochrome b/f complex	<i>petA, petB*, petD*, petG, petL, petN</i>	
	Subunits of ATP synthase	<i>atpA, atpB, atpE, atpF*, atpH, atpI</i>	
	Large subunit of rubisco	<i>rbcL</i>	
	Subunits photochlorophyllide reductase	–	
	Self-replication	Proteins of large ribosomal subunit	<i>rpl14, rpl16*, rpl2*(2), rpl20, rpl22, rpl23(2), rpl32, rpl33, rpl36</i>
		Proteins of small ribosomal subunit	<i>rps11, rps12, rps12*(2), rps14, rps15, rps16*, rps18, rps19, rps2, rps3, rps4, rps7(2), rps8</i>
		Subunits of RNA polymerase	<i>rpoA, rpoB, rpoC1*, rpoC2</i>
Ribosomal RNAs		<i>rrn16(2), rrn23(2), rrn4.5(2), rrn5(2)</i>	
Transfer RNAs		<i>trnA-UGC*(2), trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA, trnG-GCC, trnG-UCC*, trnH-GUG, trnI-CAU(2), trnI-GAU*(2), trnK-UUU*, trnL-CAA(2), trnL-UAA*, trnL-UAG, trnM-CAU, trnN-GUU(2), trnP-UGG, trnQ-UUG, trnR-ACG(2), trnR-UCU, trnS-GCU, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU, trnV-GAC(2), trnV-UAC*, trnW-CCA, trnY-GUA, trnY-M-CAU</i>	
Other genes		Maturase	<i>matK</i>
		Protease	<i>clpP**</i>
	Envelope membrane protein	<i>cemA</i>	
	Acetyl-CoA carboxylase	<i>accD</i>	
	c-Type cytochrome synthesis gene	<i>ccsA</i>	
	Translation initiation factor	<i>infA</i>	
	Other	–	
Genes of unknown function	Conserved hypothetical chloroplast	<i>yef1, yef15(2), yef2(2), yef3**, yef4</i>	
	ORF		



**Figure 2** The distribution and type of simple sequence repeats (SSRs) in the four *Atractylodes chinensis* chloroplast genome. A: Frequency of SSRs in LSC, SSC and IR regions in the chloroplast genomes of *Atractylodes chinensis* of four different origins; B: Number of SSR repeat types in the chloroplast genome of *Atractylodes chinensis* of four different origins

地区的北苍术 256 份样品样品进行高效扩增。所有实验样本的 PCR 扩增效率为 100%，序列获得率 (有效序列比例) 为 100%，目标条带单一明亮 (图 6A、B)。分析结果显示，基因 *ccsA* 扩增区为 969 bp，GC 含量为 31%，

该序列有 11 个突变位点 (表 5)，分别在 280、310、395、588、596、607、636、693、783、798 和 817 bp 有碱基突变发生，没有碱基的缺失和插入，有 16 个单倍型。*trnC-GCA\_petN* 序列长度为 821 bp，GC 含量为 32%，*trnC-*

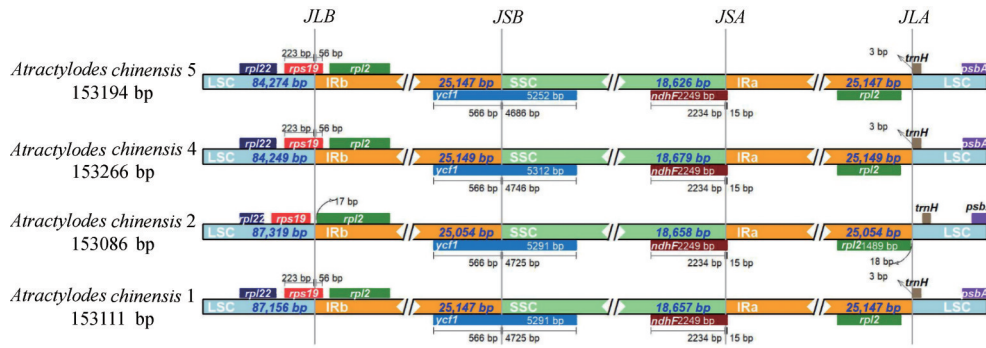


Figure 3 Comparison of the boundaries of LSC, SSC and IRa/b regions among the four *A. chinensis* plastid genomes

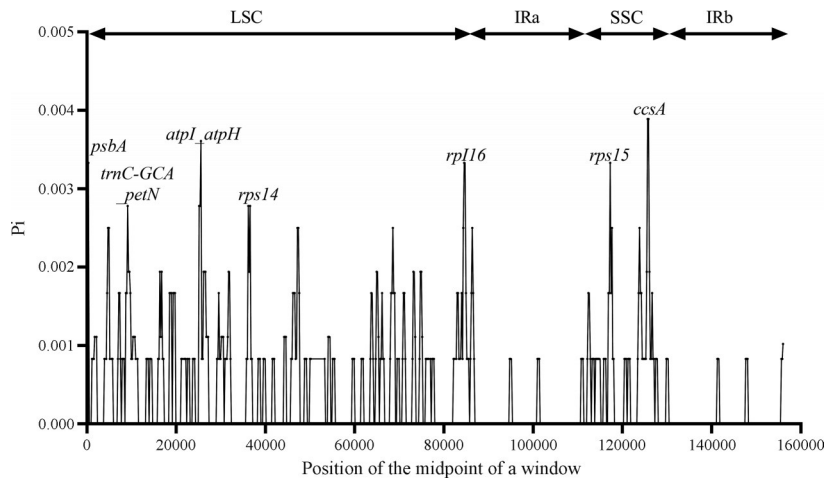


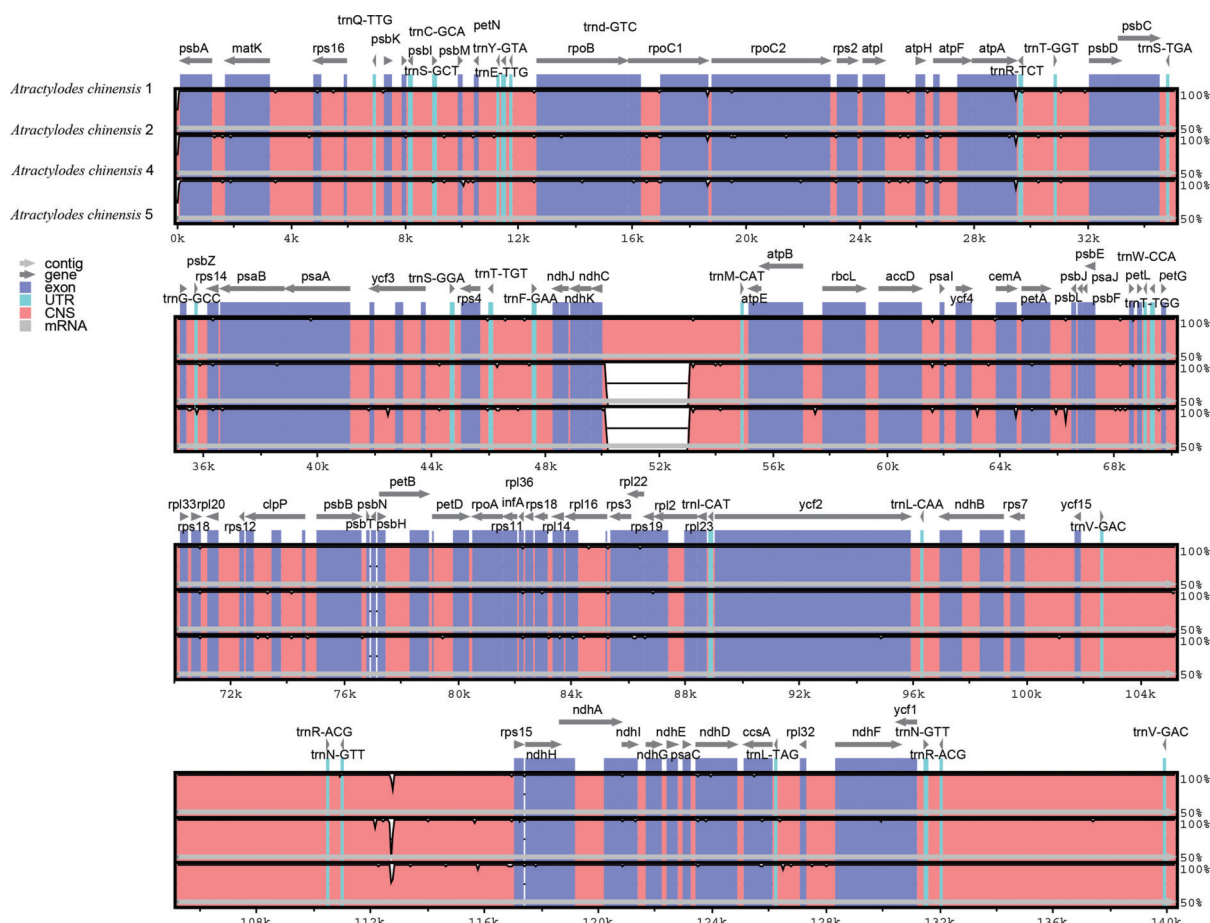
Figure 4 The nucleotide diversity of the four *Atractylodes chinensis* plastid genomes. Window length: 600 bp; Step size: 200 bp. Pi: Nucleotide diversity

*GCA* 序列有 22 个不同碱基, 分别在 138、139、140、154、194、285、419、420、421、423、428、519、569、582、592、610、624、628、678 和 715 bp 有碱基突变发生, 在 395~400 bp 出现 AAGTA 碱基的缺失, 在 452~456 bp 出现 ATATA 插入, 发现 22 个单倍型 (表 6)。通过这两个片段联合分析, 发现 24 个变异位点、39 个单倍型 (图 7)。其中 Hap9 占比最多, 为 26%; Hap30 其次, 为 13%。各个产地单倍型分布如表 7, 从省份来看, 内蒙古、河北所囊括的单倍型种类最多, 分别有 15、17 个单倍型; 其次是河南, 有 13 个单倍型。从产区看, 河北承德单倍型种类最多, 有 13 个单倍型; 其次为内蒙古扎赉特旗地区, 有 9 个单倍型。河北承德 (HBCD) 有 4 个特有单倍型 (Hap24、Hap27、Hap28、Hap29); 河北张家口 (HBZJK) 有 1 个特有单倍型 (Hap3); 内蒙古扎赉特旗 (NZZ) 和呼伦贝尔 (NHL) 分别有 1 个特有单倍型 (Hap38、Hap8); 陕西西安 (SXXA) 和陕西商洛 (SXL) 分别有 1 个特有单倍型 (Hap4、Hap14); 河南新乡 (HNXX) 有 3 个特有单倍型 (Hap3、Hap25、Hap32); 黑龙江大兴安岭 (HLJDX) 有 5 个特有单倍型 (Hap12、Hap16、Hap17、Hap35、Hap36)。

## 5 北苍术遗传多样性及遗传结构分析

用 DnaSP 软件计算北苍术居群变异位点数 (number of segregating sites,  $S$ )、单倍型数目 (number of haplotypes,  $h$ )、单倍型多样性 (haplotype diversity,  $H_d$ )、核苷酸多样性 (nucleotide diversity,  $P_i$ )。将 *ccsA* 和 *trnC-GCA\_petN* 联合分析, 发现北苍术  $H_d$  为 0.896,  $P_i$  为 2.22;  $H_d$  变化范围 0.00~0.972,  $P_i$  变化范围 0.0~2.49; 其中陕西商洛 (SXSL)  $H_d$  最高 (0.972), 内蒙古呼伦贝尔 (NHL)  $P_i$  值最高 (2.49)。河北承德 (HBCD)、黑龙江大兴安岭 (HLJDX)、河南新乡 (HNXX) 等地都有很高的  $H_d$  ( $> 0.9$ ), 说明北苍术总的遗传多样性比较丰富, 大部分产区群居也有较高的遗传多样性。各产区群居遗传多样性如表 8 所示。

基于联合片段基因种群间总的遗传分化系数  $G_{st}$  为 0.036 95 ( $< 0.05$ ), 说明各个群居间存在较小的遗传分化, 各群居间有较大的基因流 ( $N_m = 14.03$ )。其中内蒙古赤峰 (NC) 与陕西西安 (SXXA) 有一定的种间分化 ( $G_{st}$  为 0.085 37); 河北张家口 (HBZJK) 与河南洛阳 (HNLY) 有一定的种间分化 ( $G_{st}$  为 0.051 34); 山西临汾 (SXL) 和陕西商洛 (SXSL) 都与河南洛阳 (HNLY)



**Figure 5** Sequence identity among the plastid genomes of the other 3 parts of *Atractylodes chinensis* from different regions, with *Atractylodes chinensis* 1 as reference. Coding and noncoding regions are colored in blue and red, respectively

有一定的种间分化 ( $G_{st}$  分别为 0.069 93 和 0.078 38)。河北省内承德与张家口的种群间分化较大 ( $G_{st}$  为 0.063 83)。

### 6 遗传距离分析

对 39 个单倍型进行了遗传距离计算, 各单倍型遗传距离范围为 0.000 00~0.004 88。其中 Hap17 与 Hap9、Hap27 与 Hap21、Hap26 与 Hap38、Hap23 与 Hap28、Hap17 与 Hap18 的遗传距离最小为 0.000 00。Hap3、Hap23、Hap28 与 Hap20 遗传距离最大, 为 0.004 88。总体的遗传距离较小, 说明各个单倍型之间有较小的遗传差异, 侧面表明北苍术各群居之间可能存在较大的基因交流, 减少了群居间的遗传分化。

### 7 系统进化分析

对 39 个单倍型及苍术属其他同属植物序列构建 NJ 系统发育进化树 (图 8), 通过联合分析可发现, 外类群朝鲜苍术、关苍术、茅苍术聚为一支, 北苍术 39 个单倍型与白术聚为一支。一方面说明北苍术和外类群相比, 与白术有更近的亲缘关系; 另一方面说明各个单倍型之间亲缘关系很近, 遗传距离较小。

### 讨论

随着叶绿体基因组测序的发展, 多种叶绿体基因组序列被揭示, 而基于叶绿体基因组测序的比较叶绿体基因组学的研究为筛选物种间和种内的特异 DNA 条形码、种质资源鉴定、遗传多样性的研究奠定了基础。目前苍术属的大部分植物的叶绿体基因组已被揭示, 但关于北苍术种内的比较叶绿体基因组学的相关研究鲜有报道。本研究对来自不同产地的 4 份北苍术样品进行叶绿体基因组测序, 测序结果与以前报道的苍术属植物的叶绿体基因组长度相似, 但编码基因的数量略有差别, 可能是由于不同的取样地区造成的差异。

特异性的 DNA 条形码在物种鉴定的过程中起到重要作用。以前的研究经常使用常用的叶绿体高变区如 *matK* 序列进行物种鉴别, 但不同物种叶绿体基因组的高变区可能不同, 随着叶绿体基因组测序技术和比较叶绿体基因组学的进步和发展, 为筛选物种间和种内的特异性 DNA 片段提供了条件。有学者通过比较基因组学分析柴胡属植物叶绿体基因组高变区 *trnK*-

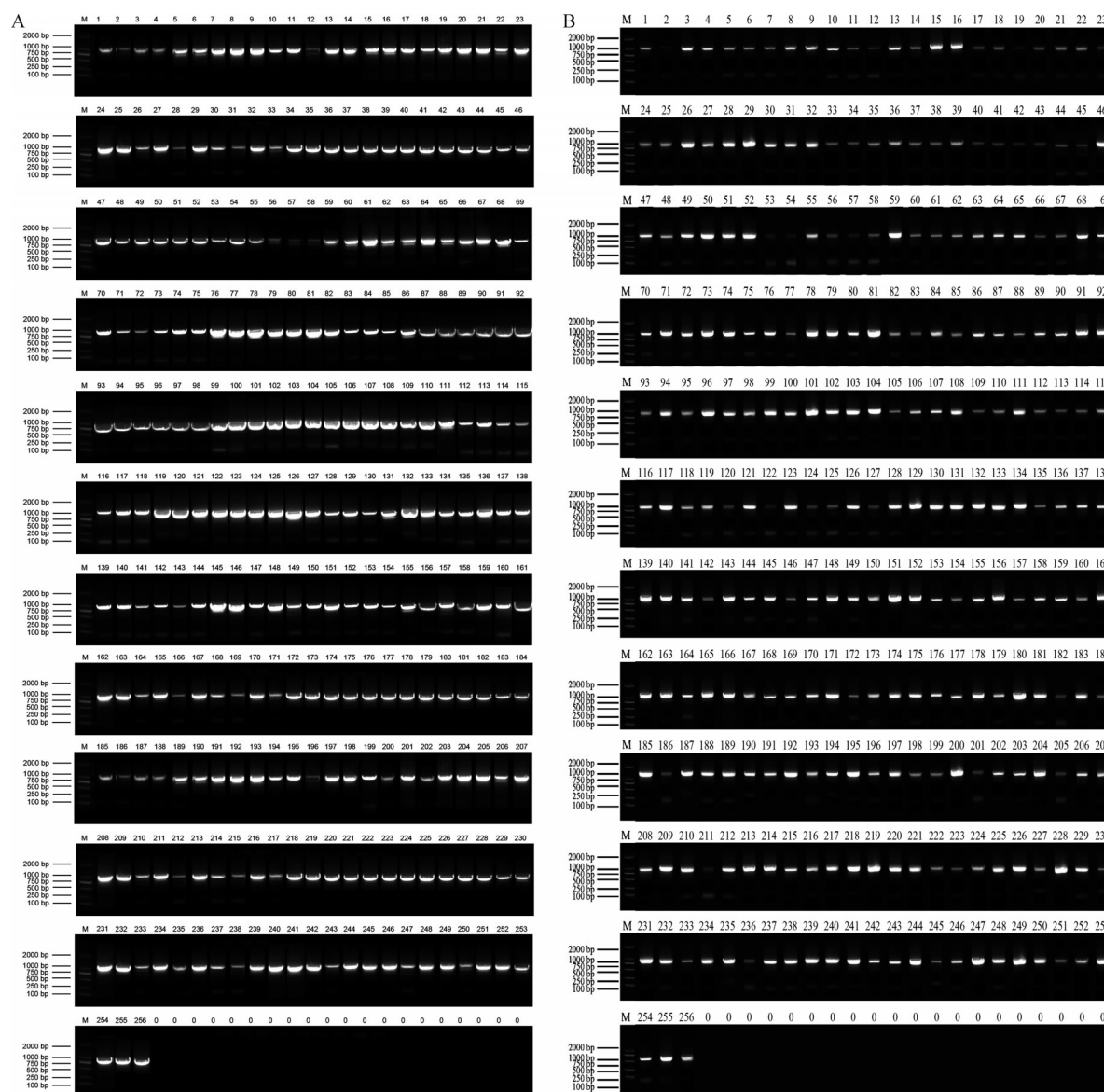


Figure 6 PCR analysis of *ccsA* gene (A) and *trnC-GCA\_petN* gene (B)

Table 5 Nucleotide position of *ccsA*. \*: The same base as the first column

Number	Nucleotide position/bp										
	280	310	39	588	596	607	636	693	783	798	817
HapA1	G	T	A	A	A	A	G	A	G	A	T
HapA2	*	*	*	*	A	*	T	*	*	*	*
HapA3	*	C	G	*	T	*	*	*	*	*	*
HapA4	*	C	*	*	*	*	*	*	T	*	*
HapA5	*	C	*	*	*	*	*	*	*	*	*
HapA6	*	*	*	*	*	*	*	*	*	G	*
HapA7	*	C	*	*	*	T	*	*	*	*	*
HapA8	*	C	*	*	T	*	*	*	*	*	*
HapA9	*	C	*	*	*	*	*	*	*	*	C
HapA10	*	C	*	C	T	*	*	*	*	*	*
HapA11	*	C	*	*	G	*	*	*	*	*	*
HapA12	*	C	*	*	T	*	T	*	*	*	*
HapA13	*	C	*	*	T	C	*	*	*	*	*
HapA14	*	C	*	*	*	*	*	*	*	G	*
HapA15	A	C	*	*	T	*	*	*	*	*	*
HapA16	*	C	*	*	T	*	*	C	*	*	*

*UGG\_rps16*、*petN\_psbM*、*petA\_psbJ*、*ndhF*、*ndhF\_rpl32*、*ycf*为柴胡属种间高变区,但没有相应的实验验证<sup>[18]</sup>。

目前也有少量研究通过比较基因组学筛选的特异DNA片段用于种质资源进行鉴定。如通过对大黄3个基原植物叶绿体基因组进行比较分析,得到高变区 *rps16-trnQ*、*psaA-ycf3*、*rps16-trnQ*、*psbE-petL*、*ndhF-rpl32*、*trnT-trnL*,通过验证发现这些片段也可作为鉴定大黄药材3个基原的特异性DNA条形码<sup>[19]</sup>。但也有相关报道证明利用叶绿体基因组测序获得特异DNA条形码后扩增效果不理想,如比较叶绿体基因组筛选的豆蔻属的特异性DNA片段 *atpH\_atpI* 扩增效果不理想,不适合豆蔻属的鉴定等<sup>[20]</sup>。因此,在利用比较叶绿体基因组筛选后需进行实际验证。近期已有学者研究表明 *rpl22-rps19-rpl2*、*psbM-trnD*、*trnR-trnT\_UGG*、*trnT\_*

**Table 6** Nucleotide position of *trnC-GCA\_petN*. \*: The same base as the first column

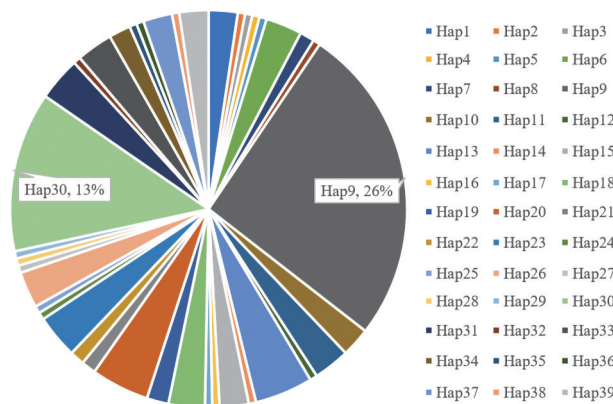
Number	Nucleotide position/bp																						
	138	139	140	154	194	285	395-400	419	420	421	423	428	452-456	519	569	582	592	610	624	628	678	715	
HapC1	G	G	A	A	T	A	AAGTA	A	T	A	T	T	—	A	G	C	C	A	T	A	T	C	
HapC2	*	*	*	*	*	*	*	*	A	*	*	*	—	*	*	*	*	*	*	*	*	*	*
HapC3	*	*	*	*	*	C	*	*	*	*	*	*	—	*	*	*	*	*	*	*	*	*	*
HapC4	*	T	*	*	*	C	*	*	*	*	*	*	—	*	*	*	*	*	*	*	*	*	*
HapC5	*	*	*	*	*	C	*	*	G	*	*	*	—	*	*	*	*	*	*	*	*	*	*
HapC6	*	*	*	*	*	C	*	-	*	*	A	*	—	*	*	*	*	*	*	G	*	*	*
HapC7	*	*	*	*	*	C	*	*	*	*	*	*	—	C	*	*	*	*	*	*	*	*	*
HapC8	*	*	*	*	*	C	*	*	*	*	*	*	—	*	*	*	A	*	*	*	*	*	*
HapC9	*	*	*	*	*	C	*	*	*	*	*	*	—	*	*	*	*	G	*	*	*	*	*
HapC10	*	*	*	*	*	C	*	*	*	C	*	*	—	*	*	*	*	*	*	*	*	*	*
HapC11	*	*	*	*	*	C	*	*	*	*	*	*	—	*	*	*	*	*	*	G	*	*	*
HapC12	*	*	*	*	*	C	*	*	*	*	*	*	—	*	*	*	*	*	*	*	G	*	*
HapC13	*	*	*	*	G	C	*	*	*	*	*	*	—	*	*	*	*	*	*	*	*	*	*
HapC14	T	C	C	*	*	C	*	*	*	*	*	*	—	*	*	*	*	*	*	*	*	*	*
HapC15	*	*	*	C	*	C	*	*	*	*	*	*	—	*	*	*	A	*	*	*	*	*	*
HapC16	*	*	*	*	*	C	—	*	*	*	*	*	—	*	*	*	*	*	*	*	*	*	*
HapC17	*	*	*	*	*	C	*	*	*	*	*	-	—	*	*	*	*	*	G	*	*	*	*
HapC18	*	*	*	*	*	C	*	*	*	*	*	-	—	*	*	*	*	*	*	*	*	*	*
HapC19	*	*	*	*	*	C	*	*	*	*	*	*	ATATA	*	*	*	*	*	*	*	*	*	A
HapC20	*	*	*	*	*	C	*	*	*	*	*	*	*	*	T	*	*	*	*	*	*	*	*
HapC21	*	*	*	*	*	C	*	T	A	*	*	*	*	*	*	*	*	*	G	*	*	*	*
HapC22	*	*	*	*	*	C	*	T	*	*	*	-	*	*	*	*	*	*	G	*	*	*	*

**Table 7** Haplotype distribution table for different regions. \*:

Unique haplotype

Population	Haplotype distribution
NZL	Hap9, Hap13, Hap15, Hap19, Hap30, Hap33, Hap34, Hap37, Hap38*
NHL	Hap6, Hap7, Hap8*, Hap11, Hap20
NC	Hap6, Hap9, Hap21
HBCD	Hap6, Hap9, Hap10, Hap15, Hap18, Hap22, Hap23, Hap24*, Hap27*, Hap28*, Hap29*, Hap30, Hap31
HBZJK	Hap1, Hap3*, Hap9, Hap30, Hap37, Hap39
SXLF	Hap9, Hap13, Hap23, Hap30, Hap33, Hap39
JLBS	Hap18
HNXX	Hap5*, Hap6, Hap11, Hap19, Hap20, Hap25*, Hap31, Hap32*
HNLY	Hap9, Hap13, Hap26, Hap30, Hap33
HLJDX	Hap12*, Hap13, Hap15, Hap16*, Hap17*, Hap30, Hap35*, Hap36*
SXXA	Hap4*, Hap9, Hap21, Hap37
SXSL	Hap9, Hap13, Hap14*, Hap26
SDZB	Hap9, Hap10, Hap15, Hap19, Hap31
BJMY	Hap9

*UGG-trnL* 可作为苍术属植物潜在的特异 DNA 条形码, 但目前没有相关的实验验证<sup>[21]</sup>。本研究通过比较不同地区的 4 份北苍术样品, 发现并验证了 *ccsA* 和 *trnC-GCA\_petN* 是北苍术潜在的种内鉴定的特异 DNA 条形码, 进一步选择这两个片段对来自 9 省 14 产区 256 份样品进行种质资源鉴定。结果表明, *ccsA* 和 *trnC-GCA\_petN* 分别形成 16 和 22 个单倍型, 两者联合分析发现 39 个单倍型; 河北承德 (HBCD)、河北张家口 (HBZJK) 等地有特有的单倍型, 可用于北苍术种质资源中产地鉴别的分子标识; 根据河北、内蒙古、黑龙



**Figure 7** The 39 haplotypes accounted for the distribution

**Table 8** Genetic diversity parameters of *Atractylodes chinensis* inferred from chloroplast DNA. S: Number of segregating sites; h: Number of haplotypes;  $H_d$ : Haplotype diversity

Population	S	h	$H_d$	$P_i \times 10^{-3}$
NZL	8	9	0.735	1.63
NHL	8	5	0.808	2.49
NC	3	3	0.800	1.38
HBCD	8	13	0.919	1.59
HBZJK	6	6	0.818	1.93
SXLF	7	6	0.775	1.93
JLBS	0	1	0.000	0.00
HNXX	11	7	0.911	2.30
HNLY	5	5	0.805	1.66
HLJDX	9	8	0.956	1.81
SXXA	3	4	0.857	1.26
SXSL	4	4	0.972	1.38
SDZB	5	5	0.861	1.38
BJMY	0	1	0.000	0.00
Total	24	39	0.896	2.22

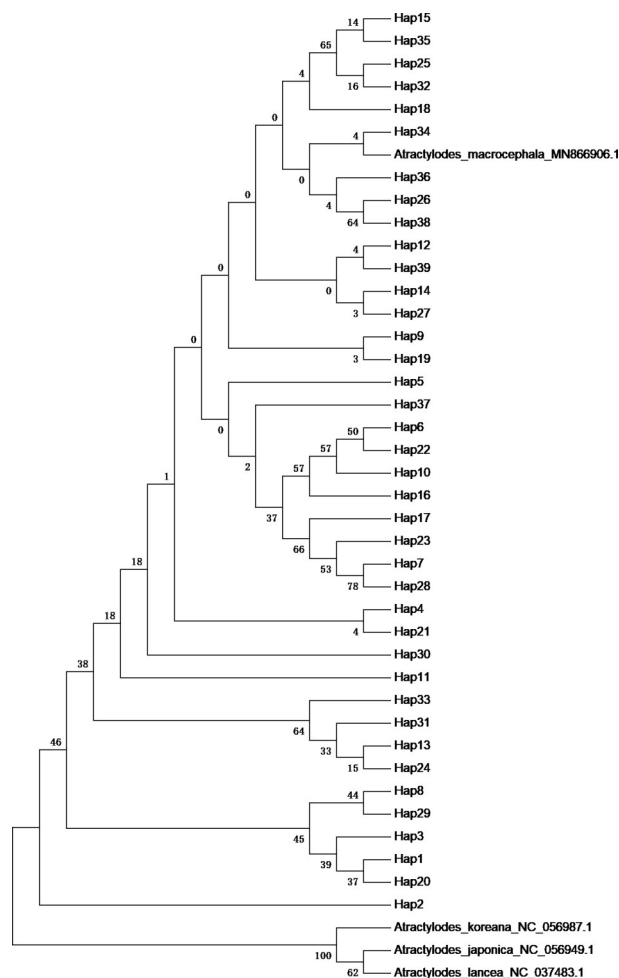


Figure 8 Neighbor-joining algorithm phylogenetic tree of 39 haplotypes and others

江等地区不同的地理环境, 区分出不同地理类型特有的单倍型, 可引种其他地理环境相似地区进行栽培。通过 mVISTA 及 Pi 对北苍术、白术、茅苍术、关苍术、朝鲜苍术进行分析, 也构建了系统发育进化树进行分析, 结果表明 *ccsA* 和 *trnC-GCA\_petN* 只能区分苍术属内的部分物种, 没有办法完全区分苍术属中的所有植物, 故认为 *ccsA* 和 *trnC-GCA\_petN* 具有一定的种间鉴别能力。经上述的叶绿体基因组分析及实验验证, *ccsA* 和 *trnC-GCA\_petN* 可用于北苍术种内资源的鉴别, 为后续研究北苍术产地鉴别、种质研究奠定基础。

Jiang 等<sup>[22]</sup>利用 ISSR 实验对 22 个不同来源的北苍术种子个体进行研究, 表明北苍术种质资源的遗传多样性非常丰富。物种的遗传多样性是常用于衡量其进化潜力的一个重要指标, 而核苷酸多样性和单倍型多样性又是度量物种遗传多样性的两个指标。单倍型多样性以 0.5 为临界<sup>[23]</sup>, 越高多样性越好。本研究基于特异性 DNA 片段 *ccsA* 和 *trnC-GCA\_petN* 联合分析对来自 9 省 14 产区北苍术进行遗传多样性分析, 结果表明

北苍术物种水平上有着较高水平的单倍型多样 ( $H_d = 0.896$ ) 和核苷酸多样性 ( $P_i = 2.22$ ), 庭藤复合群  $H_d = 0.778$ <sup>[24]</sup>, 羌活  $H_d = 0.873$ <sup>[25]</sup>, 遗传多样性丰富; 北苍术各群居间有较大的基因流 ( $N_m = 14.07$ ), 基因流有助于提高植物群体的遗传多样性水平防止种群分化<sup>[26]</sup>。总体遗传距离为 0.000 00~0.004 88, 总体的遗传距离比较小; 同时进化树结果表明单倍型聚为两支, 这些结果表明各群居之间存在较大的基因交流, 通过基因流新的北苍术个体或稀有基因被带到不同群居中增加了群居内的遗传变异, 丰富了北苍术遗传多样性。这些结果也为北苍术遗传育种、种质资源评估和品种鉴定提供了理论基础。

**作者贡献:** 第一作者尹光耀负责论文设计、实验、数据分析及论文撰写; 通讯作者魏胜利、王晓晖负责论文设计、数据分析和论文指导; 袁林参与全国实验样品的收集及实验指导; 王馨、张志飞、陈颖、刘珊瑚参与实验及数据处理; 满金辉、石玥、黄钰莹、张晓芹负责参与数据分析指导。

**利益冲突:** 无相关利益冲突。

## References

- [1] Guo FQ, Huang LF, Zhou SY, et al. Comparison of the volatile compounds of *Atractylodes* medicinal plants by headspace solid-phase microextraction-gas chromatography-mass spectrometry [J]. *Analyt Chim Acta*, 2006, 570: 73-78.
- [2] Zhang MF, Shen YQ. Advances in studies on anti-inflammation, antitumor, and immunoregulation of *Atractylodis Rhizoma* [J]. *Drug Eval Res (药物评价研究)*, 2016, 39: 885-886.
- [3] Kesara NB, Tullayakorn P, Juntra K. Research and development of *Atractylodes lancea* (Thunb) DC. as a promising candidate for cholangiocarcinoma chemotherapeutics [J]. *Evid Based Complement Alternat Med*, 2017, 2017: 5929234.
- [4] Muhammad JH, Chou JY, Li SM. An ethanol extract of the rhizome of *Atractylodes chinensis* exerts anti-gastritis activities and inhibits Akt/NF- $\kappa$ B signaling [J]. *Ethnopharmacol*, 2019, 228: 18-25.
- [5] Li H, Zhang L. Study on the molecular mechanism of *Atractylodes lancea* in treating corona virus disease 2019 (COVID-19) based on network pharmacology [J]. *J Chin Med Mater (中药材)*, 2020, 43: 2610-2615.
- [6] Yu YP. Study on the Prescription of Rhizoma *Atractylodis* and the Application of Chronic Gastritis with Depression and Anxiety (苍术方药的研究及苍术在慢性胃炎伴抑郁焦虑状态的应用) [D]. Nanning: Guangxi University of Chinese Medicine, 2017.
- [7] Cheng H, Li J, Zhang H, et al. The complete chloroplast genome sequence of strawberry (*Fragaria × ananassa* Duch.) and comparison with related species of Rosaceae [J]. *PeerJ*, 2017, 5: e3919.

- [8] Long BH, Jiang XH, Song R, et al. Application of DNA barcodes in identification and genetic diversity analysis of medicinal plants of the genus *Polygonatum* [J]. *Plant Sci J (植物科学学报)*, 2022, 40: 533-543.
- [9] Zhang XQ, Liu CS, Yan XL, et al. Sequence analysis and identification of a chloroplast *matK* gene in Rhei Rhizoma from different botanical origins [J]. *Acta Pharm Sin (药学报)*, 2013, 48: 1722-1728.
- [10] Dong BR, Zhao ZL, Ni LH, et al. Molecular markers based upon whole chloroplast genomes and identifying alpine *Gentiana waltonii* and *G. thassica* (Gentianaceae) [J]. *Acta Pharm Sin (药学报)*, 2021, 56: 2584-2591.
- [11] Wen YF, Han WJ, Wu S. Plant genetic diversity and its influencing factors [J]. *J Cent South Univ Forest Technol (中南林业科技大学学报)*, 2010, 30: 80-87.
- [12] Li JL, Wang S, Yu J, et al. A modified CTAB protocol for plant DNA extraction [J]. *Chin Bull Bot*, 2013, 48: 72-78.
- [13] Kurtz S, Choudhuri JV, Ohlebusch E, et al. REPuter: the manifold applications of repeat analysis on a genomic scale [J]. *Nucleic Acids Res*, 2001, 29: 4633-4642.
- [14] Benson G. Tandem repeats finder: a program to analyze DNA sequences [J]. *Nucleic Acids Res*, 1999, 27: 573-580.
- [15] Beier S, Thiel T, Munch T, et al. MISA-web: a web server for microsatellite prediction [J]. *Bioinformatics*, 2017, 33: 2583-2585.
- [16] Frazer KA, Pachter L, Poliakov A, et al. VISTA: computational tools for comparative genomics [J]. *Nucleic Acids Res*, 2004, 32: W273-W279.
- [17] Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data [J]. *Bioinformatics*, 2009, 25: 1451-1452.
- [18] Zhang MY, Zhang YQ, Li YM, et al. Complete plastid genomes of *Bupleurum chinense* DC. and *B. boissieuanum* H. Wolff, with comparative and phylogenetic analyses of medicinal *Bupleurum* species [J]. *Acta Pharm Sin (药学报)*, 2021, 56: 618-629.
- [19] Li RJ, Wu LW, Xin TY, et al. Analysis of chloroplast genomes and development of specific DNA barcodes for identifying the original species of Rhei Radix et Rhizoma [J]. *Acta Pharm Sin (药学报)*, 2022, 57: 1495-1505.
- [20] Wang XY, Chen XC, Liao BS, et al. Identification of Amomi Fructus Rotundus based on DNA barcoding [J]. *Mod Chin Med (中国现代中药)*, 2014, 16: 888-894.
- [21] Wang Y, Wang S, Liu Y, et al. Chloroplast genome variation and phylogenetic relationships of *Atractylodes* species [J]. *BMC Genomics*, 2021, 22: 103.
- [22] Jiang YX, Xiao CP, Sun J, et al. Genetic relationship analysis of *Atractylodes chinensis* germplasm resources based on ISSR markers [J]. *Seed (种子)*, 2021, 40: 32-38.
- [23] Grant WS, Bowen BW. Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation [J]. *J Hered*, 1998, 89: 415-426.
- [24] Peng CQ, Zhao XL. Study on the population genetics of *Indigofera decora* complex based on cpDNA and nrDNA ITS sequences [J]. *Plant Sci J (植物科学学报)*, 2015, 33: 425-437.
- [25] Yang LC, Liu HC, Zhou XL, et al. Polymorphism analysis of endangered *Notopterygium incisum* and endemic species based on cpDNA *trnT-trnL* sequences [J]. *Chin Tradit Herb Drugs (中草药)*, 2015, 46: 3390-3395.
- [26] Slatkin M. Gene flow and the geographic structure of natural populations [J]. *Science*, 1987, 236: 787-792.