

· 专论 ·

中药材 DNA 条形码分子鉴定指导原则

陈士林^{1,2*}, 姚辉¹, 韩建萍¹, 辛天怡¹, 庞晓慧¹, 石林春¹,
罗焜¹, 宋经元¹, 侯典云¹, 石上梅³, 钱忠直³

(1. 中国医学科学院北京协和医学院药用植物研究所濒危药材繁育国家工程实验室, 北京 100193;
2. 中国中医科学院中药研究所, 北京 100700; 3. 国家药典委员会, 北京 100061)

[摘要] 随着中药材 DNA 条形码分子鉴定方法研究的深入与普及, 国家药典委员会讨论通过在《中国药典》增补本中列入中药材 DNA 条形码分子鉴定指导原则, 该指导原则通过对大样本量中药材进行 DNA 条形码分子鉴定研究, 建立以 ITS2 为核心, *psbA-trnH* 为辅的植物类药材 DNA 条形码鉴定体系和以 COI 为主、ITS2 为辅的动物类药材 DNA 条形码鉴定体系。该文介绍了中药材 DNA 条形码分子鉴定指导原则及其起草说明, 并对该原则在中药材鉴定方面的应用前景进行展望。

[关键词] 中国药典; 中药材; DNA 条形码; 分子鉴定; 指导原则

历代本草记载的差异, 以及不同医学流派在传承过程中产生的异化使得中药存在同名异物、同物异名等现象; 部分药材由于历史的沿革, 品种产生变迁, 形成多基原药材同用现象; 此外, 因正品药材短缺难以满足市场需求, 民间常用其他类似品种取而代之, 代用品、习用品的随意使用使得中药材品种复杂混乱的情势加剧。由于传统的性状鉴定、显微鉴定和理化鉴定在中药材的物种鉴定上有着一定的局限性, 为了保证中药材临床应用准确、安全、有效, 有必要在现有鉴定方法的基础上增加中药材 DNA 条形码分子鉴定。随着中药材 DNA 条形码分子鉴定方法研究的深入与普及, 国家药典委员会讨论通过在《中国药典》增补本中列入中药材 DNA 条形码分子鉴定指导原则, 该指导原则通过对大样本量中药材进行 DNA 条形码分子鉴定研究, 建立以 ITS2 为核心, *psbA-trnH* 为辅的植物类药材 DNA 条形码鉴定体系和以 COI 为主、ITS2 为辅的动物类药材 DNA 条形码鉴定体系。本文介绍了中药材 DNA 条形码分子鉴定指导原则及其起草说明, 并对该原则在中药材鉴定方面的应用前景进行展望, 为中药材准确、可靠鉴定及临床用药安全提供新的技术手段。

1 中药材 DNA 条形码分子鉴定指导原则

1.1 定义及原理

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 DNA 序列是由腺嘌呤(A)、鸟嘌呤(G)、胞嘧啶(C)、胸腺嘧啶(T)4 种碱基以不同顺序排列组成, 因此一定长度 DNA 序列能够区分不同物种。

中药材 DNA 条形码分子鉴定是以 ITS2 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 为主体序列, *psbA-trnH* 为辅助序列, 动物类中药材采用 COI 为主体序列, ITS2 为辅助序列, 符合中药材鉴定简单、精确的特点, 有明确的判断标准, 能够实现中药材及其基原物种的准确鉴定。本指导原则用于规范中药材 DNA 条形码分子鉴定法, 为其应用提供指导。

该鉴定方法主要适用于中药材(包括药材、药材粉末及部分药材饮片)及基原物种的鉴定。

1.2 方法流程

中药材 DNA 条形码分子鉴定法主要包括供试品处理、DNA 提取、PCR 扩增、测序、序列拼接及结果判定, 以下内容详细说明各流程中的主要原理及注意事项。

1.2.1 供试品处理 除特殊标明外, 药材使用 75% 乙醇擦洗表面后晾干, 称取 10 ~ 100 mg 备用。

1.2.2 DNA 提取 DNA 的提取包括破碎细胞壁、释放 DNA, DNA 的分离和纯化, DNA 的浓缩、沉淀与洗涤等基本步骤, 目前常用试剂盒法, 包括植物基因组 DNA 提取试剂盒和动物组织/细胞基因组 DNA 提取试剂盒。

由于植物类中药材种类繁多, 可根据所研究中药材的具体情况对提取方法加以改进。植物细胞内含大量次生代谢产物, 如多糖、多酚等, 这些物质在提取 DNA 的过程中与 DNA 共沉淀, 形成黏稠的胶状物难以溶解或产生褐变, 严重影响 DNA 提取的产量与质量, 以及后续的 PCR 扩增实验。 β -巯基乙醇是抗氧化剂, 在提取 DNA 过程中加入 β -巯基乙

[稿件编号] 20130105018

[基金项目] 国家自然科学基金项目(81130069); 国家高技术研究发展计划(863)项目(2012AA021602); 教育部长江学者和创新团队发展计划项目(IR1150)

[通信作者] * 陈士林, Tel: (010)57833199, Fax: (010)57833028, E-mail: slchen@implad.ac.cn

醇,可以抑制氧化反应,避免褐化。PVP(聚乙烯吡咯烷酮)是酚的络合物,能与多酚形成一种不溶的络合物,有效去除多酚,减少DNA提取过程中多酚的污染;同时它也能和多糖结合,有效去除多糖。因此将PVP和 β -巯基乙醇配合使用,能够有效地防止DNA提取过程中多酚及多糖的污染。EDTA(乙二胺四乙酸)螯合 Mg^{2+} 或 Mn^{2+} ,抑制DNase(DNA酶)活性;CTAB(十六烷基三甲基溴化铵)是一种阳离子表面活性剂,可溶解细胞膜,与DNA形成复合物溶于高盐溶液,降低溶液盐浓度至一定程度,则从溶液中沉淀,经离心即可将CTAB与DNA复合物同蛋白质、多糖类物质分开。三羟甲基氨基甲烷(Tris-HCl)(pH 8.0)提供一个缓冲环境,防止DNA被破坏。

根、根茎、茎木类、皮类 通常根和根茎组织中多酚、多糖含量高,在研磨时多酚极易氧化成醌类,在纯化过程中很难去除,使DNA带有一定颜色,影响后续的PCR反应,所以在提取根及根茎类药材DNA时一定要注意多糖、多酚的去除。提取根及根茎类药材DNA时水浴时间一般为90 min,对于质地坚硬的根及根茎类和茎木类药材,可以采用延长水浴时间,如56℃水浴过夜,使得DNA充分释放到缓冲溶液中。此外,根茎类药材由于富含纤维和贮藏物质,需较多样品量才能提取到足量DNA,可用大体积离心管(5 mL或15 mL)抽提。皮类中药材组织中富含薄壁组织和纤维等,加液氮不易研磨成细粉,需适当增加样品量,同时应增加 β -巯基乙醇和PVP的使用量。

叶、花、全草类 该类药材采用试剂盒一般都能成功提取其DNA,对于保存时间较长的叶、花、全草类药材可适当增加水浴时间,同时适当降低水浴温度,如56℃水浴过夜等。

果实、种子类 果实及种子类中药材中多富含油脂,研磨时易被氧化,且易粘着在研钵壁上,损失较大,提取时需增加样品量。另外,对研磨后的材料可用丙酮浸提,去除脂溶性酚类化合物。

动物药材 常用基于SDS(十二烷基硫酸钠)原理的试剂盒法提取DNA。SDS是一种阴离子表面活性剂,在55~65℃条件下能裂解细胞,释放出核酸。对肌肉类动物药材如海龙、蛇类、蛤蚧等,需进行紫外杀菌处理,并且需要充分磨碎;含有脂类较多的动物内脏器官如蛤蟆油,先用不含蛋白酶K和SDS的缓冲液浸泡药材,然后在试剂盒消化缓冲液中增加SDS含量,有利于脱去脂类;骨甲类药材如龟甲、鳖甲和鹿茸等,由于DNA含量较低,样品量要适当增大,也可用大体积离心管(5 mL或15 mL)抽提。

1.2.3 PCR扩增 植物类中药材及其基原物种扩增ITS2或 $psbA-trnH$ 序列,动物类中药材及其基原物种扩增COI序列,通用引物及扩增条件如下,具体如有改变见各药材项下。

ITS2序列扩增正向引物ITS2F:5'-ATGCGATACTTGGTGTGAAT-3';反向引物ITS3R:5'-GACGCTTCTCCAGACTACAAT-3'。 $psbA-trnH$ 序列扩增正向引物psbAF:5'-GTTATG-

CATGAACGTAATGCTC-3';反向引物trnHR:5'-CGCGCATG-GTGATTCCACAATCC-3'。COI序列扩增正向引物HCO2198:5'-TAAACTTTCAGGGTGACCAAAAAATCA-3';反向引物LCO1490:5'-GGTCAACAAATCATAAAGATATTGG-3'。

PCR反应体系以25 μ L为参照,包括1 \times PCR缓冲液(不含 $MgCl_2$),2.0 mmol \cdot L⁻¹ $MgCl_2$,0.2 mmol \cdot L⁻¹ dNTPs,0.1 mmol \cdot L⁻¹引物对,模板DNA,1.0 U *Taq* DNA聚合酶,加灭菌双蒸水至25 μ L。设置未加模板DNA的PCR反应为阴性对照。

ITS2序列扩增程序:94℃ 5 min;94℃ 30 s,56℃ 30 s,72℃ 45 s,40个循环;72℃ 10 min。 $psbA-trnH$ 序列扩增程序:94℃ 5 min;94℃ 1 min,55℃ 1 min,72℃ 1.5 min,30个循环;72℃ 7 min。COI序列扩增程序:94℃ 1 min;94℃ 1 min,45℃ 1.5 min,72℃ 1.5 min,5个循环;94℃ 1 min,50℃ 1.5 min,72℃ 1 min,35个循环;72℃ 5 min。

1.2.4 PCR产物检测 采取琼脂糖凝胶电泳方法检测PCR产物。电泳后,PCR产物应在相应的DNA条形码序列长度位置出现一条目的条带,阴性对照应无条带。

1.2.5 测序 有PCR扩增条带的样品送测序公司进行DNA序列测定。使用DNA测序仪对目的条带进行双向测序,PCR扩增引物作为测序引物,测序原理同Sanger测序法。

1.2.6 中药材DNA条形码序列获得 主要包括序列拼接和序列质量与方向2个方面的内容。对双向测序峰图应用专业软件进行序列拼接,去除引物区,获得相应的DNA序列。为确保DNA条形码序列的可靠性,需对测序质量进行评估,去除测序结果两端的低质量序列。序列方向应与PCR扩增正向引物方向一致。

1.2.7 结果判定 将获得的序列在中药材DNA条形码鉴定系统(<http://www.tembarcode.cn>)或GenBank数据库中应用BLAST(basic local alignment search tool)方法进行结果判定,结果中相似性最高的序列对应物种为查询序列最接近的物种。如果植物类药材ITS2序列比对后最接近的物种为真菌,则需采用 $psbA-trnH$ 序列重复上述方法流程。中药材DNA条形码鉴定数据库系统及GenBank数据库BLAST鉴定系统操作流程见下。

登录中药材DNA条形码鉴定数据库系统网站,在主页面点击“物种鉴定”。根据需鉴定物种的种类,选择对应的鉴定数据库,如点击植物类药材鉴定(ITS/ITS2)。将需要鉴定的物种序列复制后粘贴到“植物类药材鉴定(ITS/ITS2)”的序列输入栏,点击“提交”按钮,进行BLAST鉴定。在BLAST比对结果中,在“序列比对信息”栏查看相关比对信息,在“物种鉴定结果”栏显示与所查询序列最接近的物种。

登录GenBank数据库BLAST鉴定系统(<http://blast.ncbi.nlm.nih.gov/Blast.cgi>),在Basic BLAST中选择nucleotide blast,在Enter Query Sequence中粘贴需要鉴定的序列(Query)(建议用fasta格式)。在Choose Search Set中选

others(nr etc.)数据库,点击左下角 BLAST。在 BLAST 结果中查看序列相似性最高(max ident)的物种,一般为与查询序列最接近的物种。

1.3 中药材 DNA 条形码数据库构建原则

建立中药材 DNA 条形码数据库时,应在药材原产地及主产地采集供试品,采用经典分类方法确定其基原,采集样本尽可能覆盖其分布区,每个物种采集来自不同产地样本 20 份以上。根据具体药材测定 ITS2 或 *psbA-trnH* 或 COI 等 DNA 条形码序列,分析确定该药材 DNA 条形码序列种内变异大小。对实验获得或 GenBank 中的中药材 DNA 条形码序列必须经过核准后方可录入中药材 DNA 条形码数据库,同时需定期对数据库进行更新。

1.4 方法学验证

1.4.1 影响因素考察 考察 DNA 条形码分子鉴定法的影响因素,包括 DNA 提取(样品量、水浴温度和水浴时间)、PCR 条件(变性时间、退火温度与时间及延伸时间)及产物纯化(考察不同纯化试剂盒),保证实验方法的准确性。

1.4.2 精密度考察 主要分为重复性考察、中间精密度考察及重现性考察。

重复性,至少用 3 批供试品,每批 3 次或同批供试品进行 6 次测定试验后对结果进行评价。实验结果判定应基本一致。

中间精密度,考察实验室内部条件改变(如不同人员、不同仪器、不同工作日和实验时间)对测定结果的影响,至少应对同实验室不同操作人员的结果进行考察。

重现性,实验结果在 3 家以上实验室能够重现,相同样品在不同实验室获得 DNA 条形码序列应相同。

1.4.3 方法适用性考察 采用 DNA 条形码分子鉴定法对 20 批次以上药材或基原物种进行测定,积累数据,确定种内序列变异大小,保证该测定方法的适用性。

1.4.4 基原物种对比验证 以分类学家确认的基原物种叶片为对象,采用该方法获得 DNA 条形码数据,与相应药材产生的 DNA 条形码数据进行对比,避免内生真菌等污染,保证结果准确性。

1.5 注意事项

实验场所应具备分子生物学实验室的基本条件。

为防止外源真菌污染,实验前须将实验用具进行高压灭菌,并用 75% 乙醇擦洗药材表面。有些药材本身含有内生真菌,如果内生真菌存在于药材的外围组织,则选用内部组织进行实验。如果真菌遍布整个药材,植物类药材需选用 *psbA-trnH* 条形码(真菌内不含有该基因片段),不能选用 ITS2 序列。为进一步确保实验结果不被真菌污染,研究者可参考 GenBank 数据库辨别是否为真菌序列。

中药材 DNA 条形码分子鉴定法不用于确定药用部位,暂不用于混合物及炮制品的鉴定。DNA 条形码分子鉴定是利用通用 DNA 序列来进行物种鉴定,因此该方法可鉴定药

材的基原物种,而不能确定药用部位。

种内阈值的确定。同一物种的不同样品间存在一定的变异范围,即种内变异阈值。不同物种,不同条形码序列均会影响种内变异范围。各基原物种的种内变异范围(种内遗传距离阈值)会在药材品种中具体明确。

2 中药材 DNA 条形码分子鉴定指导原则起草说明

中药材 DNA 条形码分子鉴定指导原则在大量实验样本基础上形成,实验涉及的样本量如下。

在 2 373 份药材样品中进行了 DNA 提取和 DNA 条形码序列 ITS/ITS2 及 *psbA-trnH* 的 PCR 扩增,其中包括根及根茎类药材 381 个样品(来自 26 科 47 属 58 种),全草类药材 1 679 个样品(来自 142 科 482 属 912 种),花类药材 70 个样品(来自 10 科 16 属 22 种),茎木类药材 52 个样品(来自 16 科 20 属 24 种),果实类药材 90 个样品(来自 16 科 20 属 33 种),种子类药材 18 个样品(来自 8 科 11 属 11 种),皮类药材 20 个样品(来自 6 科 6 属 6 种),真菌类药材 63 个样品(来自 3 科 3 属 4 种)。

中药材 DNA 条形码分子鉴定指导原则用于规范中药材 DNA 条形码分子鉴定法,为其应用提供指导。中药材 DNA 条形码分子鉴定法主要包括核心条形码序列选择、DNA 提取、PCR 扩增、电泳、测序、序列拼接及结果判定等步骤。中药材 DNA 条形码分子鉴定指导原则起草说明具体如下。

2.1 中药材 DNA 条形码分子鉴定核心序列选择依据

2.1.1 植物类中药材 ITS2 和 *psbA-trnH* 条形码的选择依据

DNA 条形码研究的首要任务是确定通用 DNA 条形码序列。在植物界,研究者主要从叶绿体基因组和核基因组中寻找理想的 DNA 条形码,曾提出诸多候选条形码序列或组合^[1-9]。2009 年,国际条形码协会植物工作组对来自 550 个物种 907 个样品的 7 个序列(*rbcL*, *matK*, *rpoC1*, *ropB*, *psbA-trnH*, *psbK-psbI*, *atpF-atpH*)进行了分析比较,建议将 *rbcL* + *matK* 组合作为植物通用条形码^[10]。然而,植物工作组认为该组合还远不够完美,寻找植物 DNA 条形码的工作还没有结束^[11]。同年,在墨西哥召开的第三届国际条形码大会上,与会代表一致认为应对 ITS/ITS2 和 *psbA-trnH* 序列进行进一步评估。

2010 年,陈士林等^[12]分析比较了 7 个候选 DNA 条形码(*psbA-trnH*, *matK*, *rbcL*, *rpoC1*, *ycf5*, ITS2, ITS)。研究对象为药用植物及其密切相关物种。研究结果表明 ITS2 表现突出,在物种水平的鉴定效率高达 92.7%。因此,陈士林等建议将 ITS2 作为药用植物标准 DNA 条形码。鉴于研究中 *psbA-trnH* 仅次于 ITS2 序列,也具有较好鉴定能力,因而推荐其作为 ITS2 的辅助条形码。Yao 等^[13]基于大样本量分析证实 ITS2 序列可以作为植物物种鉴别的通用条形码,并以 ITS2 序列为基础初步建立了药用植物 DNA 条形码数据库网络查询系统(<http://its2-plantidit.dnsalias.org/>)。2011 年,中国植物条形码工作组(Chinese Plant BOL Group)对来自 42 目 75

科141属1757物种的6286样本的 *rbcL*, *matK*, *psbA-trnH*, ITS序列进行研究,其结果进一步验证了ITS2的鉴定能力,建议ITS/ITS2应成为种子植物的核心条形码,当ITS难以扩增和测序时,ITS2可以有效地弥补该缺陷^[14]。陈士林等^[15]提出建立以ITS2为核心、*psbA-trnH*为补充序列的植物类药材DNA条形码鉴定体系。

2.1.2 动物类中药材 COI 和 ITS2 条形码的选择依据 在动物界,Herbert等于2003年首次提出将一段长度约为650 bp的COI基因序列作为动物条形码鉴定的基础片段^[16]。同年,Herbert等^[17]对11门13320个物种的线粒体细胞色素c氧化酶亚基I(cytochrome c oxidase subunit I, COI)基因序列进行比较分析,结果表明所研究物种的种间遗传距离在0.0%~53.7%,物种间遗传距离平均可达到11.3%,79%的物种间遗传距离均大于8%。以上研究结果进一步支撑了前期的结论。随后,COI基因特定片段的鉴定能力在鸟类、鱼类、节肢动物、哺乳动物等具体动物类群的鉴定研究中均获得了印证,因此研究者一致建议将COI序列作为动物的通用条形码^[18-27]。

2010年,陈士林等^[12]推荐将ITS2作为药用植物通用条形码,同时研究表明ITS2序列在动物鉴定中也有较好表现^[28-30]。为进一步评估ITS2序列对动物物种的鉴定能力,Yao等^[13]对GenBank中12221条动物ITS2序列进行了分析,结果表明ITS2序列在物种水平和属水平的鉴定成功率分别为91.7%,99.7%。此外,研究表明ITS2序列对动物密切相关物种也具有较高的鉴定能力,COI序列能够成功鉴定大多数动物类群,但刺胞动物^[17]、West Palaearctic *Pandasyopthalmus*^[31]等动物类群的COI基因序列变异性小或不变。而且线粒体DNA是母系遗传,应用其进行物种鉴定时应考虑核DNA、形态或生态特征^[32]等方面的数据。Yao等基于较大样本量的研究表明ITS2序列对动物物种的鉴定成功率高达91.7%,且在刺胞动物中的鉴定成功率大于77%,鉴于ITS2对动物具有较高的鉴定能力,Yao等建议将其作为COI序列的辅助序列对动物物种进行鉴定^[13]。

基于大量前期研究成果^[12,33-50],中药材DNA条形码分子鉴定法选用ITS2和*psbA-trnH*作为植物类药材的核心条形码,COI和ITS2作为动物类中药材的核心条形码。

2.2 中药材 DNA 提取方法的确定

2.2.1 植物基因组 DNA 提取方法的确定依据 植物基因组DNA提取方法较多,常用基于CTAB原理的试剂盒法和改良的CTAB法^[51]。应用不同公司植物基因组DNA提取试剂盒等提取中药材基因组DNA,结果表明试剂盒法适用于中药材基因组DNA的提取。由于中药材所含成分复杂,可针对不同入药部位改进DNA提取试剂盒方法^[52]。

常用液氮研磨和球磨仪(钢珠法)法来粉碎样品,液氮研磨的优点是保持低温状态,防止DNA降解,缺点是耗时且在样品种类较多时,容易产生交叉污染。球磨仪的优点是可以

同时粉碎多份样品,有效避免样品间的污染。

针对根及根茎、花、果实、种子和皮类等药材质地,确定各自合适的样品量。通过对2373份药材样品的DNA提取实验,叶类、花类药材取样量约10~20 mg,根类、果实类、种子类、皮类药材约20~40 mg。某些药材需要增加样品量,如沉香药材DNA提取时取样量为80 mg。因此,通过对大量样品DNA提取的研究,确定中药材DNA条形码鉴定法取样量为10~100 mg。

通过对不同入药部位实验不同水浴时间梯度(20,30,40,50,60 min,3 h及水浴过夜),发现叶类药材水浴时间一般在20~40 min,根类药材在60 min~3 h,一些质地坚硬的药材可56℃水浴过夜。

2.2.2 动物类药材 DNA 提取方法的确定依据 动物基因组DNA的提取采用血液/细胞、组织基因组DNA提取试剂盒。采用天根生化科技(北京)有限公司的血液/细胞、组织基因组DNA提取试剂盒(基于SDS法原理)提取动物药材的DNA,结果表明试剂盒法适用于动物基因组DNA的提取,根据动物药材药用部位的不同(肌肉、角甲、壳类),选用不同的DNA提取试剂盒。

动物类药材DNA提取前,应针对不同取样部位对样品进行不同的前期处理。肌肉类动物药材需进行紫外杀菌处理并充分捣碎;骨甲类药材由于DNA含量稍少,应适量增加取样量,并充分研磨粉碎;含有脂类较多的动物内脏器官可先用不含蛋白酶K和SDS的缓冲液浸泡药材,并试剂盒消化缓冲液中适量增加SDS,以利于脱去脂类;分泌物类动物药材(如胆汁),在消化前同样需进行必要的处理。如干蛇胆用双蒸水浸泡6 h,其间更换水数次,使其软化并洗去表面污染物;乙醇浸泡蛇胆,用流水浸泡1 d,以除去乙醇及表面污染物^[53]。目前多选用试剂盒法提取DNA,方法相对简洁、易控,而且在大多数动物药材中都获得了较为理想的结果。

2.3 PCR 条件及反应程序的确定依据

2.3.1 ITS2 序列最佳退火温度 采用梯度PCR仪对ITS2序列进行扩增,温度梯度设置为50~61℃,筛选得到56℃为最佳退火温度。优化的ITS2序列扩增程序:94℃预变性5 min;94℃变性30 s,56℃退火30 s,72℃延伸45 s,40个循环;72℃延伸10 min。

2.3.2 psbA-trnH 序列最佳退火温度 *psbA-trnH*的反应程序依照文献^[6,8]。*psbA-trnH*序列扩增程序:94℃预变性5 min;94℃变性1 min,55℃退火1 min,72℃延伸1.5 min,30个循环;72℃延伸7 min。

2.3.3 COI 序列最佳退火温度 COI的反应程序依照文献^[17-18]。COI序列扩增程序:94℃预变性1 min;94℃变性1 min,45℃退火1.5 min,72℃延伸1.5 min,5个循环;94℃变性1 min,50℃退火1.5 min,72℃延伸1 min,35个循环;72℃延伸5 min。该方法适合于大多数动物的PCR扩增,针对特殊类群,引物及反应条件可参见文献^[54-57]。

2.4 确定可靠的DNA条形码鉴定数据库

DNA条形码数据库的可靠性是中药材DNA条形码鉴定的关键。目前已有的公共数据库,如GenBank,EMBL等,由世界各地的研究者进行独立提交,缺乏相互之间的验证,数据质量参差不齐,中药材DNA序列数据的系统性和代表性尚显不足。

为保证中药材DNA条形码数据库的可靠性,首先需要在基原上保证物种鉴定的可靠性,然后采用严格的序列校对机制确保获得序列和基原样品的一致性,最后规范管理数据库,确保数据库的安全维护和有序增减。

2.4.1 获得可靠的基原样品 为保证实验结果具有代表性,实验样品需采集自药材的原产地和其分布区。通过调查和文献检索了解药材的产地和分布信息,每个产地需采集3份以上的样品,每种药材采集的样品总数原则上需在20份以上,并且至少覆盖该药材80%以上的产地和分布区。

所有采集的药材样品均需先请国内外专科专属的形态学鉴定专家进行鉴定以确保基原的准确性,同时应保存好凭证标本和凭证照片等信息以备复查。

采集的药材样品及其提取的DNA样品进行规范化管理,以避免真菌污染和样品交叉污染。采集的药材样品经专家鉴定后,根据样品编码规则对采集样品进行标准编码,并进行专门存放。样品DNA经专人按标准流程提取,统一编码后存放于专门的-20℃冰箱中。对存放的药材样品和药材DNA样品需有专人进行定期检查和维护。

2.4.2 确定序列质量判定标准 Q 值,即quality value,是评价测序过程中碱基可靠性的一个重要参数,其计算公式为 $Q = -10\lg Pe$ (Q 指碱基的测序质量值, Pe 指碱基的出错概率),一般情况下,单向测序的碱基质量需要大于 $Q20$ (99%的可靠性),经双向测序拼接后的碱基质量需要大于 $Q30$ (99.9%的可靠性)。

序列质量判断标准采用国际DNA条形码协会(CBOL)通用的标准^[10],即以20bp的窗口分别从序列5'端和3'端进行滑动,如果窗口内有多余2个碱基的 Q 小于20,则删除一个碱基,窗口继续滑动一个碱基,如果窗口内碱基 Q 小于20的数目小于或等于2个,窗口停止滑动。测序结果的剩余部分需大于150bp,且平均 Q 大于等于30。

2.4.3 保证序列和基原样品的一致性 除实验过程规范操作外,采用严格校对机制,即BLAST分析防错、系统树分析防错、barcoding gap检验防错,确保实验获得的序列和基原样品的一致性,避免在实验过程中由于操作不当引入的真菌污染和样品交叉污染。

BLAST分析防错:基于局部序列比对,对数据库进行快速搜索,其特点是仅搜索序列之间高度相似的区域,可以兼顾搜索的精确性和搜索的速度,是目前应用最广泛的序列相似性搜索工具之一。利用BLAST工具,在一定的 E 值下(通常是 $E-20$),在GenBank公共数据库或者本地数据库进行序

列相似性搜索,确保查询序列和基原样品的一致性。

系统树分析防错:收集查询序列同物种序列和同属其他物种序列,以及2~3条与查询序列同科但不同属的物种的序列。利用MEGA5软件(版本V5.1Beta2),采用K-2P遗传距离,构建这些序列的NJ(neighbor-joining)系统发育树,查询序列应与同物种序列聚为一支。

barcoding gap检验防错:收集与查询序列同物种和同属其他物种的序列,分别计算查询序列与其他序列的K2P遗传距离,查询序列与同物种其他样品序列的最大遗传距离应该不大于该序列与同属其他物种样品之间的遗传距离。

3 展望

部分中药材品种混乱给中药安全使用造成极大隐患,而传统鉴定方法的局限影响我国中药安全有效使用。中药鉴定技术方法贯穿于中药材种植、加工、生产等各个环节,随着中药产业的发展,传统鉴定方法已经难以满足医院、药店、企业、海关等各行业快速、标准的鉴定需求,建立新的鉴定方法十分迫切^[58-64]。DNA条形码分子鉴定技术为中药材基原物种的鉴定提供了强有力的科技支撑。构建中药材DNA条形码分子鉴定数据库,可以实现对中药材基原物种、粉末以及细胞、组织等材料来源的准确快速鉴定,可以满足不同行业不同科研背景工作者对中药材鉴定的要求。将中药材DNA条形码分子鉴定指导原则纳入《中国药典》,能够为中药材DNA条形码分子鉴定提供依据和参考,使其在中药材流通领域、中药生产企业的原料监管、医院饮片的真伪鉴定、海关中药材进出口鉴定与管理、出入境检验检疫等方面发挥强有力的作用。

【参考文献】

- [1] Newmaster S G, Fazekas A J, Ragupathy S. DNA barcoding in land plants: evaluation of *rbcL* in a multigene tiered approach [J]. *Can J Bot*, 2006, 84: 335.
- [2] Kress W J, Wurdack K J, Zimmer E A, et al. Use of DNA barcodes to identify flowering plants [J]. *Proc Natl Acad Sci USA*, 2005, 102: 8369.
- [3] Lahaye R, van der Bank M, Bogarin D, et al. DNA barcoding the floras of biodiversity hotspots [J]. *Proc Nat Acad Sci USA*, 2008, 105: 2923.
- [4] Lahaye R, Savolainen V, Duthoit S, et al. A test of *psbK-psbI* and *atpF-atpH* as potential plant DNA barcodes using the flora of the Kruger National Park (South Africa) as a model system [J]. *Nature Precedings*, 2008, hdl.handle.net/10101/npre.2008.1896.1.
- [5] Newmaster S G, Fazekas A J, Steeves R A D, et al. Testing candidate plant barcode regions in the Myristicaceae [J]. *Mol Ecol Resour*, 2008, 8: 480.
- [6] Sasser C, Little D P, Stevenson D W, et al. DNA barcoding in the cycadales: testing the potential of proposed barcoding markers for species identification of cycads [J]. *PLoS ONE*, 2007, 2: e1154.

- [7] Presting G G. Identification of conserved regions in the plastid genome: implications for DNA barcoding and biological function [J]. *Can J Bot*, 2006, 84: 1434.
- [8] Kress W J, Erickson D L. A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region [J]. *PLoS ONE*, 2007, 2: e508.
- [9] Chase M W, Cowan R S, Hollingsworth P M, et al. A proposal for a standardised protocol to barcode all land plants [J]. *Taxon*, 2007, 56: 295.
- [10] CBOL Plant Working Group. A DNA barcode for land plants [J]. *Proc Natl Acad Sci USA*, 2009, 106: 12794.
- [11] Thomas C. Plant bar code soon to become reality [J]. *Science*, 2009, 325: 526.
- [12] Chen S L, Yao H, Han J P, et al. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species [J]. *PLoS ONE*, 2010, 5: e8613.
- [13] Yao H, Song J Y, Liu C, et al. Use of ITS2 region as the universal DNA barcode for plants and animals [J]. *PLoS ONE*, 2010, 5 (10): e13102.
- [14] China Plant BOL Group. Comparative analysis of a large dataset indicates that ITS should be incorporated into the core barcode for seed plants [J]. *Proc Natl Acad Sci USA*, 2011, 108: 19641.
- [15] 陈士林, 庞晓慧, 姚辉, 等. 中药 DNA 条形码鉴定体系及研究方向 [J]. *世界科学技术——中医药现代化*, 2011, 13 (5): 747.
- [16] Hebert P D N, Cywinska A, Ball S L, et al. Biological identifications through DNA barcodes [J]. *Proc R Soc Lond B*, 2003, 270: 313.
- [17] Hebert P D N, Ratnasingham S, Dewaard J R. Barcoding animal life: cytochrome *c* oxidase subunit I divergences among closely related species [J]. *Proc R Soc Lond B (Suppl)*, 2003, 270: S96.
- [18] Hebert P D N, Stoeckle M, Zemplak T, et al. Identification of birds through DNA barcodes [J]. *PLoS Biol*, 2004, 2: 1657.
- [19] Hebert P D N, Penton E H, Burns J M, et al. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator* [J]. *Proc Natl Acad Sci USA*, 2004, 101: 14812.
- [20] Smith M A, Fisher B L, Hebert P D N. DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar [J]. *Philos Trans R Soc Lond B Biol Sci* 2005, 360: 1825.
- [21] Hajibabaei M, Janzen D H, Burns J M, et al. DNA barcodes distinguish species of tropical Lepidoptera [J]. *Proc Natl Acad Sci USA* 2006, 103: 968.
- [22] Yoo H S, Eah J Y, Kim J S, et al. DNA barcoding Korean birds [J]. *Mol Cells*, 2006, 22: 323.
- [23] Elias M, Hill R I, Willmott K R, et al. Limited performance of DNA barcoding in a diverse community of tropical butterflies [J]. *Proc Biol Sci*, 2007, 274: 2881.
- [24] Kerr K C R, Stoeckle M Y, Dove C J, et al. Comprehensive DNA barcode coverage of north American birds [J]. *Mol Ecol Notes*, 2007, 7: 535.
- [25] Tavares E S, Baker A J. Single mitochondrial gene barcodes reliably identify sister-species in diverse clades of birds [J]. *BMC Evol Biol*, 2008, 8: 81.
- [26] Ward R D, Zemplak T S, Innes B H, et al. DNA barcoding Australia's fish species [J]. *Philos Trans R Soc Lond B Biol Sci*, 2005, 360: 1847.
- [27] Savolainen V, Cowan R S, Vogler A P, et al. Towards writing the encyclopedia of life: an introduction to DNA barcoding [J]. *Philos Trans R Soc Lond B Biol Sci*, 2005, 360: 1805.
- [28] Ben-David T, Melamed S, Gerson U, et al. ITS2 sequences as barcodes for identifying and analyzing spider mites (Acari: Tetranychidae) [J]. *Exp Appl Acarol*, 2007, 41: 169.
- [29] Li Y W, Zhou X, Feng G, et al. COI and ITS2 sequences delimit species, reveal cryptic taxa and host specificity of fig-associated *Sycophila* (Hymenoptera, Eurytomidae) [J]. *Mol Ecol Resour*, 2010, 10: 31.
- [30] Prasad P K, Tandon V, Biswal D K, et al. Phylogenetic reconstruction using secondary structures and sequence motifs of ITS2 rDNA of *Paragonimus westermani* (Kerbert, 1878) Braun, 1899 (Digenea: Paragonimidae) and related species [J]. *BMC Genomics*, 2009, 10 (Suppl 3): S25.
- [31] Rojo S, Stahls G, Perez-Banon C, et al. Testing molecular barcodes: invariant mitochondrial DNA sequences vs the larval and adult morphology of West Palaearctic *Pandasyopthalmus* species (Diptera: Syrphidae: Paragini) [J]. *Eur J Entomol*, 2006, 103: 443.
- [32] Rubinoff D, Cameron S, Will K. A genomic perspective on the shortcomings of mitochondrial DNA for "barcoding" identification [J]. *J Hered*, 2006, 97: 581.
- [33] 罗焜, 陈士林, 陈科力, 等. 基于芸香科的植物通用 DNA 条形码研究 [J]. *中国科学: 生命科学*, 2010, 40(4): 342.
- [34] Pang X H, Song J Y, Zhu Y J, et al. Applying plant DNA barcodes for Rosaceae species identification [J]. *Cladistics*, 2011, 27: 165.
- [35] Pang X H, Song J Y, Zhu Y J, et al. Using DNA barcoding to identify species within Euphorbiaceae [J]. *Planta Med*, 2010, 76: 1784.
- [36] Gao T, Yao H, Song J Y, et al. Identification of medicinal plants in the family Fabaceae using a potential DNA barcode ITS2 [J]. *J Ethnopharmacol*, 2010, 130: 116.
- [37] Gao T, Yao H, Song J Y, et al. Evaluating the feasibility of using candidate DNA barcodes in discriminating species of the large Asteraceae family [J]. *BMC Evol Biol*, 2010, 10: 324.
- [38] 朱英杰, 陈士林, 姚辉, 等. 重楼属药用植物 DNA 条形码鉴定研究 [J]. *药学报*, 2010, 45(3): 376.
- [39] Yao H, Song J Y, Ma X Y, et al. Identification of *Dendrobium* species by a candidate DNA barcode sequence: the chloroplast *ps-*

- ba-trnH* intergenic region[J]. *Planta Med*, 2009, 75: 667.
- [40] Han J P, Song J Y, Liu C, et al. Identification of *Cistanche* species (Orobanchaceae) based on sequences of the plastid *psbA-trnH* intergenic region[J]. *Acta Pharm Sin*, 2010, 45: 126.
- [41] Sun Z Y, Gao T, Yao H, et al. Identification of *Lonicera japonica* and its related species using the DNA barcoding method[J]. *Planta Med*, 2011, 77: 301.
- [42] Ma X Y, Xie C X, Liu C, et al. Species identification of medicinal pteridophytes by a DNA barcode marker, the chloroplast *psbA-trnH* intergenic region[J]. *Biol Pharm Bull*, 2010, 33: 1919.
- [43] Song J Y, Yao H, Li Y, et al. Authentication of the family Polygonaceae in Chinese pharmacopoeia by DNA barcoding technique[J]. *J Ethnopharmacol*, 2009, 124:434.
- [44] 辛天怡, 姚辉, 罗焜, 等. 羌活药材 ITS/ITS2 条形码鉴定及其稳定性与准确性研究[J]. *药学报*, 2012, 47(8): 1098.
- [45] 罗焜, 马培, 姚辉, 等. 多基原药材秦艽 ITS2 条形码鉴定研究[J]. *药学报*, 2012, 47(12): 1710.
- [46] Song J Y, Shi L C, Li D Z, et al. Extensive pyrosequencing reveals frequent intra-genomic variations of internal transcribed spacer regions of nuclear ribosomal DNA[J]. *PLoS ONE*, 2012, 7:e43971.
- [47] 韩建萍, 宋经元, 刘昶, 等. 基于叶绿体 *psbA-trnH* 基因间区序列鉴定肉苁蓉属植物[J]. *药学报*, 2010, 45(1): 126.
- [48] 陈士林, 宋经元, 姚辉, 等. 药用植物 DNA 条形码鉴定策略及关键技术分析[J]. *中国天然药物*, 2009, 7(5):322.
- [49] 陈士林, 姚辉, 宋经元, 等. 基于 DNA barcoding(条形码)技术的中药材鉴定[J]. *世界科学技术——中医药现代化*, 2007, 9(3):7.
- [50] Pang X, Liu C, Shi L, et al. Utility of the *trnH-psbA* intergenic spacer region and its combinations as plant DNA barcodes: a meta-analysis [J]. *PLoS ONE*, 2012, 7(11): e48833.
- [51] Chiou S J, Yen J H, Fang C L, et al. Authentication of medicinal herbs using PCR-amplified ITS2 with specific primers[J]. *Planta Med*, 2007, 73: 1421.
- [52] 罗焜, 马培, 姚辉, 等. 中药 DNA 条形码鉴定中的 DNA 提取方法研究[J]. *世界科学技术——中医药现代化*, 2012, 14(2): 1433.
- [53] 曹琳, 刘忠权, 郝明, 等. 不同种类中药材的 DNA 提取方法[J]. *中国现代应用药学*, 2004, 21(6): 465.
- [54] Folmer O, Black M, Hoeh W, et al. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan vertebrates[J]. *Mol Mar Biol Biotechnol*, 1994, 3: 294.
- [55] Hajibabaei M, Janzen D H, Burns J M, et al. DNA barcodes distinguish species of tropical Lepidoptera[J]. *Proc Nat Acad Sci USA*, 2006, 103: 968.
- [56] Ivanova N V, Zemlak T S, Hanner R H, et al. Universal primer cocktails for fish DNA barcoding[J]. *Mol Ecol Notes*, 2007, 7: 544.
- [57] Messing J. New M13 vectors for cloning[J]. *Methods Enzymol*, 1983, 101: 20.
- [58] 陈士林, 郭宝林, 张贵君, 等. 中药鉴定学新技术新方法研究进展[J]. *中国中药杂志*, 2012, 37(8):1043.
- [59] 韩建萍, 宋经元, 姚辉, 等. 中药材 DNA 条形码鉴定的基因序列比较[J]. *中国中药杂志*, 2012, 37(8):1056.
- [60] 赵中振, 梁之桃. 近红外光谱技术在中药鉴定中的应用与优势[J]. *中国中药杂志*, 2012, 37(8):1062.
- [61] 胡咏川, 田晓鑫, 刘蕾, 等. 近红外光谱技术鉴定中药的进展[J]. *中国中药杂志*, 2012, 37(8):1066.
- [62] 黄家乐, 邵鹏柱. 生物信息学在中药材分子鉴别中的应用[J]. *中国中药杂志*, 2012, 37(8):1072.
- [63] 龙芳, 李会军, 李萍. 新技术和新方法在中药性状与显微鉴别中的应用[J]. *中国中药杂志*, 2012, 37(8):1076.
- [64] 刘江, 陈兴福, 邹元峰. 基于中药指纹图谱多维信息的化学模式识别研究进展[J]. *中国中药杂志*, 2012, 37(8):1081.

Principles for molecular identification of traditional Chinese materia medica using DNA barcoding

CHEN Shi-lin^{1,2*}, YAO Hui¹, HAN Jian-ping¹, XIN Tian-yi¹, PANG Xiao-hui¹,

SHI Lin-chun¹, LUO Kun¹, SONG Jing-yuan¹, HOU Dian-yun¹, SHI Shang-mei³, QIAN Zhong-zhi³

(1. *The National Engineering Laboratory for Breeding of Endangered Medicinal Materials, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100193, China;*

2. *Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China;*

3. *Chinese Pharmacopoeia Commission, Beijing 100061, China*)

[**Abstract**] Since the research of molecular identification of Chinese Materia Medica (CMM) using DNA barcode is rapidly developing and popularizing, the principle of this method is approved to be listed in the Supplement of the Pharmacopoeia of the People's Republic of China. Based on the study on comprehensive samples, the DNA barcoding systems have been established to identify CMM, i. e. ITS2 as a core barcode and *psbA-trnH* as a complementary locus for identification of planta medica, and COI as a core barcode and ITS2 as a complementary locus for identification of animal medica. This article introduced the principle of molecular identification of CMM using DNA barcoding and its drafting instructions. Furthermore, its application perspective was discussed.

[**Key words**] Chinese pharmacopoeia; traditional Chinese materia medica; DNA barcoding; molecular identification; principle

doi:10.4268/cjcmm20130201

[责任编辑 曹阳阳]