

· 技术与应用 ·

利用人体携带植物基因信息推断居住地

刘萌妍¹, 刘艳磊², 吴平², 陈庆¹, 周世良²

(1. 北京市公安司法鉴定中心, 北京 100192; 2. 中国科学院植物研究所 系统与进化植物学国家重点实验室, 北京 100093)

摘要: **目的** 利用从未知名死者肺组织中获取的植物花粉基因序列鉴定植物种类, 再根据植物的分布区域推断死者可能长期生活的地区, 缩小侦查范围, 为案件侦破提供线索。 **方法** 提取死者肺组织, 采用mCTAB法提取总DNA, 利用专一扩增 *matK* 和 *rbcL* 的引物获取这两个植物DNA条形码的基因片段并测定基因序列, 经生物信息学分析得到目的基因片段的DNA序列, 将这些序列与参考序列数据合并, 进行系统发育分析, 确定所获取的DNA序列所属的物种, 根据物种的分布信息推断死者生前可能长期生活的地区。 **结果** 该死者肺组织中存在27科31属32种植物的基因片段, 其中有一定指示作用的9个属的植物大多为海南、广东、广西和云南的特有植物, 表明死者生前可能长期在这些地区居住。经调查, 死者生前于广西南部某县长期生活, 与上述研究结果相符。 **结论** 利用人体肺组织内所携带植物的基因信息推断居住地的方法, 可以辅助推断死者生前的长期居住地, 为无名尸案件的尸源查找提供了新的思路。

关键词: 法医遗传学; 植物学; DNA, 植物; 地理分布; 系统发育树; 尸源鉴定

中图分类号: DF795.2 文献标志码: B doi: 10.12116/j.issn.1004-5619.2019.06.012

文章编号: 1004-5619(2019)06-0710-06

Determination of Place of Residence Using the Gene Information of Plants Carried by the Human BodyLIU Meng-yan¹, LIU Yan-lei², WU Ping², CHEN Qing¹, ZHOU Shi-liang²

(1. Forensic Judicial Appraisal Center of Beijing Public Security Bureau, Beijing 100192, China; 2. State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China)

Abstract: **Objective** To identify the plant species using the DNA sequence of plant pollen from lung tissues of a unidentified body, infer the possible long-term places of residence of the deceased according to the distribution area of the pollen in the lung tissues, therefore narrow the scope of criminal investigation and provide clues for case solving. **Methods** Lung tissues were extracted from the deceased, total DNA was extracted by the mCTAB method. Gene fragments of the two plant DNA barcodes, *matK* and *rbcL*, were acquired using specific primers for amplification, then sequenced. The DNA sequences of target gene fragments were acquired through bioinformatics analysis. The sequences were combined with reference sequence data. Phylogenetic analysis was made to identify the species that the DNA sequences belonged to. The places where the deceased could have lived for a long time were inferred, according to the distribution information of plant species. **Results** Gene fragments of 32 plant species which belonged to 31 genera of 27 families were in the lung tissues of the deceased. Among them, plants of 9 genera that had certain indicative function were mainly endemic plants from Hainan, Guangdong, Guangxi and Yunnan. These results showed that the deceased may have stayed in these areas for a long time before death. After further investigation, the victim was confirmed to have come from a county in southern Guangxi, which was in accordance with the research results. **Conclusion** The method of using gene information of plants from lung tissues of human bodies to infer places of residence can assist inference of the places where the deceased could have lived for a long time. The present study may also provide new ideas for locating sources of the corpses in cases with unidentified victims.

Keywords: forensic genetics; botany; DNA, plant; geographical distribution; phylogenetic tree; identification of corpse source

作者简介: 刘萌妍(1986—), 女, 主检法医师, 主要从事法医病理学研究; E-mail: namilmy@sina.com

通信作者: 陈庆, 男, 主任法医师, 主要从事法医病理学研究; E-mail: fayibingli@163.com

通信作者: 周世良, 男, 研究员, 博士研究生导师, 主要从事植物DNA库建设与物种分子鉴定研究; E-mail: slzhou@ibcas.ac.cn

无名尸体尸源信息的破解对于案件性质的判断有着至关重要的作用。确认尸源信息一直是刑侦领域的难点。常用的尸源查找手段主要为“内源性”信息技术手段,包括通过法医人类学判断死者体貌信息、颅像复原、Y系亲缘库比对等。随着流动性人口的增加,“内源性”信息技术手段已无法完全解决尸源查找这一难点。

孢粉学证据作为一种“外源性”信息能够反映一定的死者生活信息,但以往孢粉学证据基本采自无名尸体的附属物、体表、消化系统和上呼吸道^[1]。本研究尝试通过提取死者肺组织中的植物遗传信息这一“外源性”信息,利用目前数据库信息较为丰富的植物通用 *matK* 和 *rbcL* 基因作为 DNA 条形码进行研究,根据植物所属地推断死者可能长期生活的地区,为案件调查提供一种新的技术手段,缩小侦查范围,提供有效侦查方向。

1 材料与方法

1.1 样本

本研究样本来源于一具无身份信息的男性尸体。尸体在北方某地被发现时已呈腐败状态,肺组织呈中度腐败。切取双肺各叶边缘组织 10 块,每块质量约为 1 g。由于肺组织中花粉分布不均匀,为了降低取样偏差,本实验将 10 份样本进行混合后提取 DNA。本研究相关实验样本及方法已通过北京伦理学会的审查。

1.2 方法

1.2.1 DNA 提取与叶绿体基因扩增

将获得的肺组织用液氮冷冻后磨碎,采用 mCTAB 法(改良 CTAB 法)提取总 DNA^[2]。利用 Wizard® DNA Clean-Up System(美国 Promega 公司)对得到的 DNA 进行纯化,采用专一扩增植物叶绿体基因 *matK* 的引物(F: 5'-CATTATGTTAGATATCAAGGAAA-3'; R: 5'-GCTRTRATAATGAGAAAGATTTCTGC-3')和 *rbcL* 的引物(F: 5'-AGACCTWTTTGAAGAAGGTTTCWGT-3'; R: 5'-CATGTACCTGCAGTAGCATTCAGT-3')扩增 *matK* 和 *rbcL* 基因片段^[3-4]。

PCR 扩增实验采用 10 μ L 反应体系,包括: 10 \times PCR 缓冲液(含 Mg^{2+}) 1 μ L, 2 mmol/L dNTP 混合液 1 μ L, 5 μ mol/L 正向引物 0.5 μ L, 5 μ mol/L 反向引物 0.5 μ L, 2.5 U/ μ L *Taq* DNA 聚合酶 0.1 μ L, 模板 DNA 1 μ L, ddH₂O 5.9 μ L。PCR 扩增反应在 Mastercycler pro S PCR 仪(德国 Eppendorf 公司)上进行。扩增程序: 94 $^{\circ}$ C 预变性 2 min; 94 $^{\circ}$ C 变性 30 s, 52 $^{\circ}$ C 退火 30 s, 72 $^{\circ}$ C 延伸 40 s, 35 个循环; 72 $^{\circ}$ C 延伸 5 min; 4 $^{\circ}$ C 保存。PCR 扩增产物用 1.2% 琼脂糖凝胶电泳(JY04S-3D

型,北京君意东方电泳设备有限公司)检测,观察并拍照。胶检时加入 1 μ L PCR 产物和 4 μ L 1 \times 溴酚蓝缓冲液。利用 Wizard® DNA Clean-Up System 对所得到的 PCR 产物进行纯化,纯化后的产物用于后续高通量测序实验。

1.2.2 高通量测序

经过 PCR 扩增实验获得的 *matK* 和 *rbcL* 基因片段为不同物种该基因片段的混合物,常规的 Sanger 测序无法实现多样品同一基因混合物的测序。为了测定这些基因片段,本研究采用 Ion S5™ 系统(美国 Thermo Fisher Scientific 公司)高通量测序平台。基因片段混合物经纯化后,由北京泛生子基因科技有限公司利用 Ion Plus Fragment Library Kit(美国 Thermo Fisher Scientific 公司)标准流程^[5]测序文库并利用 Ion S5™ 系统标准参数进行测序。

1.2.3 *matK* 和 *rbcL* 基因序列获取

在所有种子植物中,每科选取一条序列作为代表,构建高质量的参考数据库 *matK*-db 和 *rbcL*-db。使用 Divide 程序(<https://github.com/wpwupingwp/divide>),根据引物序列将 Ion S5™ 系统测序数据拆分为 *matK* 和 *rbcL* 基因数据,并以 97% 相似性为阈值,将各基因数据进行聚类,生成多条共有序列(consensus sequence),舍去丰度为 1 的序列(singleton)。使用 Filter 程序(<https://github.com/wpwupingwp/filter>),以 *matK*-db 和 *rbcL*-db 作为参考,过滤非特异性扩增序列,得到候选序列 *matK*-contigs 和 *rbcL*-contigs 后,将其与参考数据库进行合并,使用 MAFFT 程序进行多序列比对^[6]。得到的比对文件,分别调用 IQTREE 软件^[7],选用 GTR+F+R4 模型构建最大似然法系统发育树,并进行 1 000 次超速自举(ultrafast bootstrap)检验^[8]。根据候选序列在系统发育树中的位置、姊妹群、最近共同祖先和支长以及可信度等信息,筛选得到序列 *rbcL*-ML 和 *matK*-ML。使用 Usearch 程序^[9],以 *rbcL*-ML 和 *matK*-ML 为模板,从原始的 Ion S5™ 系统测序数据中调取与这些序列相似性大于 97% 的原始序列,使用 MAFFT 程序进行全局比对,手工删除长度异常的序列及空位(gap),调用 VSEARCH 程序^[10],以相似性 95% 为阈值,生成最终的结果序列 *rbcL*-result 和 *matK*-result。以 GenBank 数据库中 nr 数据集为参考进行在线 BLAST^[11]比对,确定这些序列可能的科和属。

1.3 系统发育分析

根据 BLAST 比对结果,将同属物种所有的 *matK* 或 *rbcL* 基因序列下载,与核准的序列合并,用 MAFFT 程序进行比对,并使用 BioEdit 软件(<https://bioedit.software.informer.com>)进行手工校正,得到数据矩

阵。将这些数据矩阵通过PAUP软件^[12]构建最大简约树,确定与核准序列具有最近亲缘关系的物种。查询这些物种的分布区,推断尸体可能长期居住的地区。

2 结 果

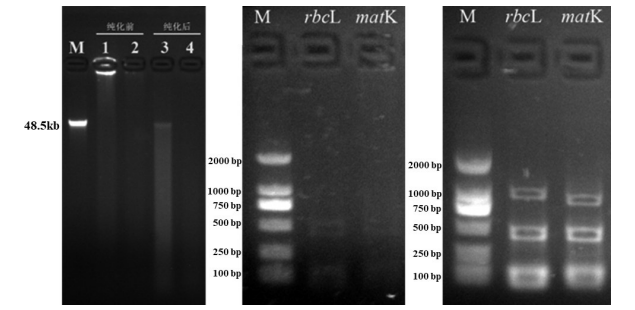
2.1 DNA 提取及PCR 扩增

从有一定程度腐败的肺组织中成功提取到DNA(图1A)。其中,样品2的DNA纯化前后未见明显改善,但在后续PCR扩增实验中获得明显扩增条带。纯化前后DNA的扩增结果见图1B~C。

2.2 高通量数据

从Ion S5™系统测序共获得1 957 440条序列,数据量可以保证后续数据分析。使用Divide数据拆分程序,共得到83 144条*rbcL*基因的序列,60 665条*matK*基因的序列。经过聚类及Filter程序过滤后得

到*rbcL*基因一致性序列3 355条,*matK*基因一致性序列2 721条。与参考序列合并、构建系统发育树后,舍去丰度过低的数据,从死者的肺组织中得到的种子植物共计27科31属32种,见表1。



A:从腐败肺组织中提取的DNA;B:未纯化DNA的扩增结果;C:纯化DNA的扩增结果。A图中M为Marker,1、2为未纯化的DNA,3、4为纯化的DNA。

图1 DNA提取及PCR扩增结果

表1 从肺组织中检测到的植物种类

序号	科	属	最接近的种
1	棕榈科(Arecaceae)	刺葵属(<i>Phoenix</i>)	加拿利海枣(<i>canariensis</i>)
2	马兜铃科(Aristolochiaceae)	马兜铃属(<i>Aristolochia</i>)	马兜铃属一种
3	菊科(Asteraceae)	蒿属(<i>Artemisia</i>)	艾(<i>vulgaris</i>)
4	桦木科(Betulaceae)	桦木属(<i>Betula</i>)	桦木属一种
5	十字花科(Brassicaceae)	鼠耳芥属(<i>Arabidopsis</i>)	鼠耳芥属一种
6	红厚壳科(Calophyllaceae)	铁力木属(<i>Mesua</i>)	铁力木(<i>ferrea</i>)
7	桔梗科(Campanulaceae)	同钟花属(<i>Homocodon</i>)	同钟花(<i>brevipes</i>)
8	大麻科(Cannabaceae)	大麻属(<i>Cannabis</i>)	大麻(<i>sativa</i>)
9	金鱼藻科(Ceratophyllaceae)	金鱼藻属(<i>Ceratophyllum</i>)	金鱼藻(<i>demersum</i>)
10	葫芦科(Cucurbitaceae)	西瓜属(<i>Citrullus</i>)	西瓜(<i>lanatus</i>)
11	柏科(Cupressaceae)	侧柏属(<i>Platycladus</i>)	侧柏(<i>orientalis</i>)
12	龙脑香科(Dipterocarpoideae)	青梅属(<i>Vatica</i>)	青梅(<i>mangachapoi</i>)
13	豆科(Fabaceae)	补骨脂属(<i>Psoralea</i>)	胡韭子(<i>corylifolia</i>)
14	买麻藤科(Gentaceae)	买麻藤属(<i>Gnetum</i>)	买麻藤属一种
15	牻牛儿苗科(Geraniaceae)	老鹳草属(<i>Geranium</i>)	老鹳草属一种
16	锦葵科(Malvaceae)	苘麻属(<i>Abutilon</i>)	苘麻属一种
17	锦葵科(Malvaceae)	银叶树属(<i>Heritiera</i>)	银叶树属一种
18	桑科(Moraceae)	波罗蜜属(<i>Artocarpus</i>)	二色波罗蜜(<i>styracifolius</i>)
19	桑科(Moraceae)	构属(<i>Broussonetia</i>)	构树(<i>papyrifera</i>)
20	罂粟科(Papaveraceae)	罂粟属(<i>Papaver</i>)	罂粟(<i>somniferum</i>)
21	松科(Pinaceae)	松属(<i>Pinus</i>)	松属一种
22	禾本科(Poaceae)	稻属(<i>Oryza</i>)	水稻(<i>sativa</i>)
23	禾本科(Poaceae)	小麦属(<i>Triticum</i>)	小麦属一种
24	毛茛科(Ranunculaceae)	黄连属(<i>Coptis</i>)	三角叶黄连(<i>deltoidea</i>)
25	毛茛科(Ranunculaceae)	黑种草属(<i>Nigella</i>)	黑种草(<i>arvensis</i>)
26	毛茛科(Ranunculaceae)	黑种草属(<i>Nigella</i>)	腺毛黑种草(<i>damascena</i>)
27	蔷薇科(Rosaceae)	绣线菊属(<i>Spiraea</i>)	绣线菊属一种
28	芸香科(Rutaceae)	柑橘属(<i>Citrus</i>)	桔(<i>sinensis</i>)
29	杨柳科(Salicaceae)	柳属(<i>Salix</i>)	柳属一种
30	无患子科(Sapindaceae)	龙眼属(<i>Dimocarpus</i>)	龙眼(<i>longan</i>)
31	茄科(Solanaceae)	枸杞属(<i>Lycium</i>)	枸杞属一种
32	香蒲科(Typhaceae)	香蒲属(<i>Typha</i>)	宽叶香蒲(<i>latifolia</i>)

2.3 比对结果

将上述序列与美国国立生物技术信息中心(National Center for Biotechnology Information, NCBI)公共数据库进行初步对比分析,筛选出与中国物种关系密切的物种。

在31个属中有9个属的物种有一定程度的指示作用,分别是波罗蜜属(*Artocarpus*)、黄连属(*Coptis*)、龙眼属(*Dimocarpus*)、买麻藤属(*Gnetum*)、银叶树属(*Heritiera*)、同钟花属(*Homocodon*)、铁力木属(*Mesua*)、刺葵属(*Phoenix*)和青梅属(*Vatica*)。

在波罗蜜属(*Artocarpus*)中,经与NCBI公共数据库中该属47种186条*rbcL*基因序列进行初步对比分析,检测到的序列与3个中国产物种关系密切。这3个种的8条代表性*rbcL*序列的系统发育分析结果表明,检测到的序列与光叶桂木(*A. nitidus*)聚为一支,最可能属于光叶桂木(*A. nitidus*,图2A)。

在黄连属(*Coptis*)8个物种有*matK*基因序列,五裂黄连(*C. quinquesecta*)和五叶黄连(*C. quinquefolia*)数据缺失。经与8个物种的12条*matK*序列的系统发育分析,检测到的序列与峨眉黄连(*C. omeiensis*)和三角叶黄连(*C. deltoidea*)聚为一支,关系最近(图2B)。

在龙眼属(*Dimocarpus*)中除了滇龙眼(*D. yunnanensis*)外,其余中国产物种均有*rbcL*序列。系统发育分析表明,检测到的序列属于龙眼(*D. longan*,图2C)。

在买麻藤属(*Gnetum*)中,经与19种38条*rbcL*序列的系统发育分析,确定检测到的序列与中国产海南买麻藤(*G. hainanense*)、罗浮买麻藤(*G. luofuense*)、买麻藤(*G. montanum*)和小叶买麻藤(*G. parvifolium*)关系最密切(图2D)。

在银叶树属(*Heritiera*)中,所有物种均有*rbcL*参考序列。然而,*rbcL*不能有效区分这3个物种(图2E)。

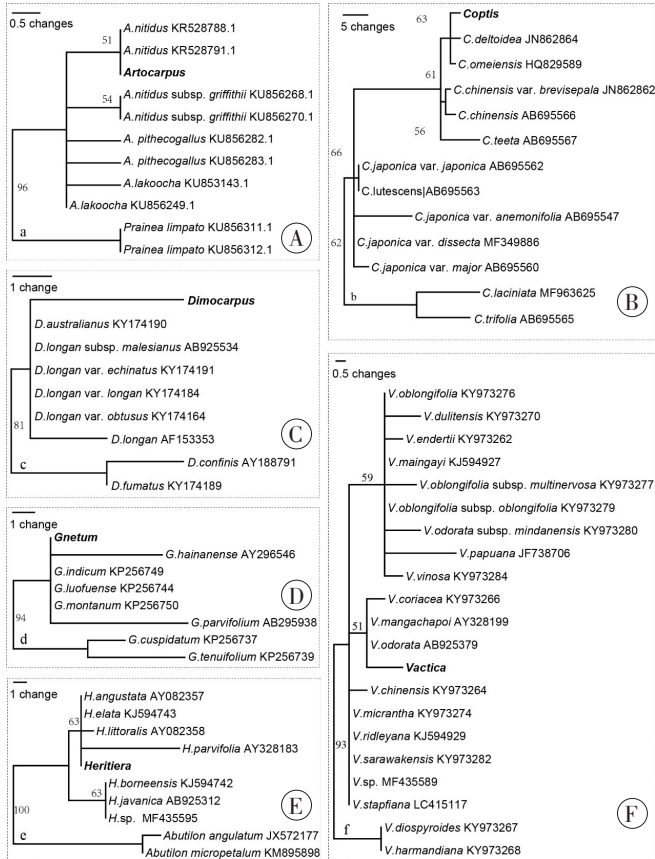
同钟花属(*Homocodon*)仅2种。

铁力木属(*Mesua*)仅1种。

刺葵属(*Phoenix*)中有2个中国产物种有*matK*参考序列,但该基因没有分辨率。

在青梅属(*Vatica*)中,系统发育分析结果表明,检测到的序列与青梅(*V. mangachapoi*)和婆罗洲青梅(*V. coriacea*,马来西亚的婆罗洲)关系密切(图2F)。

依据《中国植物志》(<http://frps.iplant.cn>)对9个属的物种进行生长地域的检索,得到相对相关度高的地域,详见表2。



A: 基于*rbcL*基因的波罗蜜属系统发育关系; B: 基于*matK*基因的黄连属系统发育关系; C: 基于*rbcL*基因的龙眼属系统发育关系; D: 基于*rbcL*基因的买麻藤属系统发育关系; E: 基于*rbcL*基因的银叶树属系统发育关系; F: 基于*matK*基因的青梅属系统发育关系。所有系统发育树均为严格一致树,分支上的数字为bootstrap支持率; 黑体全拼属名代表检测出的DNA序列,其他为参照序列。

图2 从肺组织中检测到的属于地区特有植物的基因片段及其系统发育树

表2 9个具有一定指示作用植物属的基本情况

名称	世界种数	中国产种数	相对相关度高的地域
波罗蜜属(<i>Artocarpus</i>)	50	14	光叶桂木分布于广东、广西、海南、湖南和云南等地
黄连属(<i>Copti</i>)	15	6	峨眉黄连:河南和四川;三角叶黄连:四川
龙眼属(<i>Dimocarpus</i>)	7	4	龙眼产于广东、广西、海南、云南,在福建等南方广泛栽培
买麻藤属(<i>Gnetum</i>)	55	9	分布于广东、广西、海南、云南,个别可北延至江西、福建、湖南、贵州等地
银叶树属(<i>Heritiera</i>)	35	3	长柄银叶树分布于海南和云南,蝴蝶树为海南特产
同钟花属(<i>Homocodon</i>)	2	2	同钟花分布于贵州、四川和云南,长梗同钟花特产于四川泸定县的海螺沟
铁力木属(<i>Mesua</i>)	1	1	通常在广东、广西、云南等地栽培
刺葵属(<i>Phoenix</i>)	14	3	海枣在福建、广东、广西、云南有栽培;刺葵产于福建、广东、广西、海南、台湾和云南;江边刺葵特产于云南澜沧江和怒江流域低海拔地区
青梅属(<i>Vatica</i>)	65	4	青梅特产于海南,分布于广西和云南的属于广西青梅,分布于西藏东南部的属于西藏青梅

2.4 调查结果

经过实地调查走访,死者生前于广西南部某县长期生活过。

3 讨 论

肺组织长期与外界环境接触,当花粉进入肺泡后刺激机体产生应激反应,从而将花粉颗粒包裹、激化,随着细胞死亡而钙化,永久性沉积于肺组织中。花粉自身拥有两层花粉壁,因此具有很强的保护遗传物质的能力。上述两个条件使其遗传物质可以在人体内长久保存,且不易遭到破坏,甚至在腐败组织中都可以提取到植物遗传物质。植物的突出特点之一是固着生长,某些植物对环境要求高,形成了特异性地理分布区域,因此植物具有天然的地理位置指示作用。在上述理论基础上,本研究进行了采用人体所携带植物的基因信息判断居住地的实验。

本研究在死者肺组织内共取得27科31属32种植物,31个属中,物种数目差异很大,有超过15个属都难以鉴定到种。

一些序列如艾(*Artemisia vulgaris*)、构树(*Broussonetia papyrifera*)、金鱼藻(*Ceratophyllum demersum*)、西瓜(*Citrullus lanatus*)、桔(*Citrus sinensis*)、水稻(*Oryza sativa*)、侧柏(*Platycladus orientalis*)、宽叶香蒲(*Typha latifolia*)等,虽然能鉴定到种,但由于物种分布区域广泛,地区指示作用有限。有的植物,如大麻(*Cannabis sativa*)和罂粟(*Papaver somniferum*),流通普遍,没有指示作用。

最终,本研究筛选出9个属具有一定指示作用的物种,并分析这9个属物种的分布数据,主要分布在广东、广西、福建、云南、海南、四川、湖南等省,其中以广东、广西、海南和云南特有植物出现率最高,提示死者生前可能在长江以南流域长期生活,尤其是广东、广西、云南、海南等地可能性更高。

调查结果显示,死者生前确实于广西南部某县长期生活过,与分析结果相符,表明利用人体肺组织内所携带植物的基因信息推断居住地的方法可以辅助推断死者生前的长期居住地,为无名尸案件的尸源查找提供新的法医学鉴定思路。本研究采用的是DNA条形码技术,该技术由HEBERT等^[13]提出,是利用一小段易于扩增的DNA片段区分、鉴别生物物种的技术。不同于传统的形态学鉴别物种的方式,DNA条形码技术更加高效、准确、可重复。

DNA条形码的选择对于种属鉴定也至关重要。通常评价DNA条形码有以下几个条件:(1)DNA片段短小且易于扩增,尤其适用于腐败的检材。(2)目标DNA片段可以作为标准在绝大多数植物分类中应用。(3)目标DNA片段在种间变异度大,具有种间高分辨率;在种内变异度小,具有相对的保守性;含有系统进化信息,可以显示在分类系统中的位置。(4)该目标DNA片段应有高度保守区域以方便引物设计^[14]。

本研究选择了*rbcL*和*matK*这两个DNA条形码组合提取植物DNA。*rbcL*和*matK*是第三届国际DNA条形码会议上被推荐作为植物DNA标准条形码的核心标记,具备以下几项优势^[15]:(1)*rbcL*和*matK*均是叶绿体基因组的蛋白编码基因,在PCR扩增时可以直接筛选出植物基因,避免人类、真菌等其他非植物真核生物的基因干扰。(2)本研究所用*rbcL*长度为402bp,*matK*长度为349bp,片段短小,即使是降解检材亦可扩增。(3)*rbcL*的属内变异小,相对保守,可以很好地显示在属水平的系统进化位置;*matK*在叶绿体基因组中具有较快的进化速率,因此具有一定的种间差异,具备相对较高的种间分辨率。(4)由于*rbcL*和*matK*为植物通用条形码,研究者应用较多,通用引物相对完善,数据库充足。但是这两个DNA条形码并不完美,对部分植物的分辨能力往往仅到科属水平,并不能全面达到种的水平,因此未来需要开发分辨率更高

的基因标记。

内部转录间隔区(internal transcribed spacer, ITS),即5.8S rDNA和28S rDNA基因间隔序列,是目前系统发育构建中最受欢迎的一个DNA条形码,也是中国植物条形码协会进一步建议补充的DNA条形码。不同于叶绿体基因组,ITS属于核基因组,且高度重复,其进化速率比*matK*更快,同时具有拷贝间序列随时间推移趋于一致的特点,因此其分辨率相对*rbcL*和*matK*更高,可以准确追溯种间关系,即将大部分植物区分至种^[16]。本次实验的预实验也采用了ITS序列,但ITS序列在真核生物中广泛存在,本实验中提取到的DNA绝大部分是人类的,会对扩增结果产生干扰。如何在扩增此DNA条形码时减少人类ITS序列的扩增是目前正在探索的问题。

值得注意的是,花粉的传粉方式大致分为风媒传粉和虫媒传粉。约87.5%的被子植物由昆虫传粉^[17],这意味着虫媒花粉传播范围小,从而形成具有不同群落结构的花粉谱。相较于风媒花粉,虫媒花粉可能更具区域特异性,提示携带者在该区域停留时间更长。风媒传粉和虫媒传粉植物应在未来的研究中列入数据分析参考系。

DNA琼脂糖凝胶电泳结果显示,纯化后扩增效果明显提高,得到了与目的片段长度相符的扩增片段。虽然样品2纯化后未见DNA条带,但PCR产物的电泳图显示获得明显扩增条带。经后续高通量数据分析,确实发现植物DNA信息,证明纯化后样品2中确实有微量的植物DNA被提取出来。

在PCR的3个条带中,较长的片段可能为低等植物相应基因片段,较短的为引物二聚体。本研究的重点是种子植物,这些植物的分布有非常明确的地理分布记载,所以片段长度在250~500bp的部分才会被用于后续分析。

本研究提供了一种利用人体肺组织内所携带植物的基因信息推断居住地的方法,可以辅助推断死者生前的长期居住地,为无名尸案件的尸源查找提供了新的法医学思路。本研究原本希望能够根据植物成分含量比例推测具体的来源地域,但目前通过reads数并不能完全衡量各物种的比例关系,对植物所属地的推断不够理想。若联合其他尸源溯源手段会有效提高案件侦破率,希望未来该技术能够成为刑事法医学的一种常用尸源溯源手段,为案件侦破提供有效方向。

参考文献:

- [1] HALL D W, BYRD J H. 法医植物学实践指南[M]. 吕宙,译. 西安:西安交通大学出版社,2017:134-150.

- [2] 李金璐,王硕,于婧,等. 一种改良的植物DNA提取方法[J]. 植物学报,2013,48(1):72-78.
- [3] YU J, XUE J H, ZHOU S L. New universal *matK* primers for DNA barcoding angiosperm[J]. J Syst Evol,2011,49(3):176-181.
- [4] DONG W, CHENG T, LI C, et al. Discriminating plants using the DNA barcode *rbcLb*: An appraisal based on a large data set[J]. Mol Ecol Resour,2014,14(2):336-343.
- [5] Prepare amplicon libraries without fragmentation using the Ion Plus Fragment Library Kit[Z/OL]. [2018-10-25]. https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0006846_PrepAmpliconLibr_using_Ion-PlusFragLibraryKit_UB.pdf.
- [6] KATO H K, STANDLEY D M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability[J]. Mol Biol Evol,2013,30(4):772-780.
- [7] NGUYEN L T, SCHMIDT H A, von HAESLER A, et al. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies[J]. Mol Biol Evol,2015,32(1):268-274.
- [8] MINH B Q, NGUYEN M A, von HAESLER A. Ultrafast approximation for phylogenetic bootstrap[J]. Mol Biol Evol,2013,30(5):1188-1195.
- [9] EDGAR R C. Search and clustering orders of magnitude faster than BLAST[J]. Bioinformatics,2010,26(19):2460-2461.
- [10] ROGNES T, FLOURI T, NICHOLS B, et al. VSEARCH: A versatile open source tool for metagenomics[J]. PeerJ,2016,4:e2584.
- [11] ALTSCHUL S F, GISH W, MILLER W, et al. Basic local alignment search tool (BLAST) [J]. J Mol Biol,1990,215(3):403-410.
- [12] RÉDEI G P. Encyclopedia of genetics, genomics, proteomics and informatics[M]. 3rd ed. Dordrecht: Springer,2008.
- [13] HEBERT P D, CYWINSKA A, BALL S L, et al. Biological identifications through DNA barcodes[J]. Proc Biol Sci,2003,270(1512):313-321.
- [14] 庞晓慧. 植物DNA条形码序列筛选与鉴定研究—以裸子植物、蔷薇科和大戟科为例[D]. 北京:中国协和医科大学,2010.
- [15] 郭义红,孙威江,林伟东,等. 植物DNA条形码鉴定研究进展[J]. 江苏农业科学,2016,44(7):19-21.
- [16] CHENG T, XU C, LEI L, et al. Barcoding the kingdom Plantae: New PCR primers for ITS regions of plants with improved universality and specificity[J]. Mol Ecol Resour,2016,16(1):138-149.
- [17] OLLERTON J, WINFREE R, TARRANT S. How many flowering plants are pollinated by animals?[J]. Oikos,2011,120(3):321-326.

(收稿日期:2019-01-07)

(本文编辑:张建华)