

DOI: 10.13718/j.cnki.xdzk.2017.07.013

# RAD-seq 技术在植物物证个体认定中的应用<sup>①</sup>

庄艳<sup>1</sup>, 孔庆波<sup>1</sup>, 陈新平<sup>1</sup>, 陆海新<sup>2</sup>,  
王晓丽<sup>2</sup>, 王崇志<sup>2</sup>, 孙海陆<sup>2</sup>

1. 重庆警察学院 科研处, 重庆 401331; 2. 深圳华大基因, 广东 深圳 518000

**摘要:** 在犯罪案件现场发现与嫌疑人(或受害人)相关联的植物物证, 鉴定它们来源和进行个体识别, 可为案件的调查取证提供有价值的证据或线索. 该研究采用一种 RAD-seq 技术对 18 个桂花样品(15 个已知样本、3 个盲测样本)进行 DNA 建库和高通量测序, 序列多态性分析, 评估其遗传多样性, 并用基于 SNP 和基于 k-mer 频谱的 2 种比较分析方法探讨了 RAD-seq 数据用于个体识别的科学性和对植物物证进行同一性认定的可行性. 结果表明, 2 种方法都能成功识别 3 个盲测样本. 相比较而言, 基于 k-mer 频谱的方法在稳定性和自动化程度上要优于基于 SNP 亲缘关系树的分析方法, 更易推广到公安物证鉴别应用中. RAD-seq 技术结合该文的分析方法能够识别出未知植物物证样品 A、B 和 C 分别对应于植物物证桂花 1、4 和 11, 实现了对案件中关联植物物证进行精确检验鉴定的方法和一个小型鉴定实例, 但其识别能力有待扩展到更大样本集和更多物种中.

**关键词:** 植物物证; RAD-seq; 单核苷酸多态性; 个体识别; 桂花

**中图分类号:** Q819

**文献标志码:** A

**文章编号:** 1673-9868(2017)07-0084-07

随着科技的进步, 一些犯罪案件的科技含量也在提高, 手段也在不断多样化, 并且常规物证被有意识地销毁, 往往给侦破工作带来难度, 因此微量物证尤其是生物检材越来越引起物证工作者的重视. 2011 年公安部关于《公安机关侦办命案工作规范》中要求, 外围现场和关联现场要受到重视, 尽可能多地发现并提取有用价值的物证, 尤其是微量物证和生物检材. 植物是犯罪现场中一种重要的环境和过程证据, 而且很多案件中常常出现植物类物证, 这些植物类物证大多都无法用形态学检验方法来鉴别, 需要借助植物 DNA 技术手段对植物物证的来源和种类进行鉴定. 这对于查找破案线索、划定侦查范围和方向提供了有价值的证据, 也为法庭诉讼产生重要的指导和支撑作用<sup>[1-2]</sup>. 近年来, CBOL(生物条形码协会)植物工程组对约 550 个物种进行通用引物筛选, 试图用一个或组合的 DNA 条形码来鉴别植物物种, 这些研究成果已在鉴别违禁植物、毒品、出入境物种资源等方面得到应用<sup>[3]</sup>. 植物物证个体识别方面的研究报道不多, 虽然也有用传统 DNA 分子标记技术进行 DNA 多态性比对分析, 进行同一认定的成功案例, 但该方法难以实现自动化, 检测方法相对缓慢而且费力. 目前, 利用 RAD-seq 技术来进行植物物证的个体识别还未见报道.

简化基因组测序是在第二代测序技术基础上发展起来的, 是利用酶切技术、序列捕获芯片技术或其他

① 收稿日期: 2016-12-05

基金项目: 重庆市基础与前沿研究计划项目(cstc2014jcyjA00012); 重庆市教委科学技术项目(KJ1501502); 重庆市科学技术委员会青年科技人才培养计划项目(cstc2014kjrc-qnc70003); 重庆市高校创新团队建设计划项目(CXTDX201601041).

作者简介: 庄艳(1980-), 女, 新疆昌吉人, 博士, 讲师, 主要从事植物物证鉴定技术研究.

实验手段降低物种基因组复杂程度, 针对基因组特定区域进行测序, 进而反映部分基因组序列结构信息的综合实验技术<sup>[4]</sup>, 目前运用较广泛的方法是限制性酶切相关位点的 DNA 测序(Restriction Association site DNA sequencing, RAD-seq)<sup>[5]</sup>. RAD-seq 所得的全基因组范围特异酶切位点附近的小片段 DNA 标签, 能较好地代表整个基因组的序列特征, 从而能够在大多数生物中获得成千上万的单核苷酸多态性(Single Nucleotide Polymorphism, SNP)标记. RAD-seq 技术操作简单, 不受参考基因组限制, 并可简化复杂基因组, 目前已广泛应用于分子育种、系统进化、种质资源等领域<sup>[6-8]</sup>, 但在植物物证个体识别方面, RAD-seq 应用于法庭科学中具体的实验实施和分析方法还需要摸索细化.

桂花 *Osmanthus* sp. 在重庆地区分布很广, 在公园、路边、居民区随处可见, 桂花可以年年开花结果, 且花期较长, 与我们的日常生活紧密联系, 如果涉案出现的可能性较大. 本文的研究对象选择没有可参考基因组的桂花作为植物物证个体识别的样本, 利用简化基因组测序技术来降低基因组测序和分析的复杂度, 进而构建植物物证桂花样品的个体识别数据库, 试图找到能识别植物物证样品桂花同一个体的方法, 为后续利用 RAD-seq 技术进行植物类物证的个体鉴别提供可靠的方法和支持.

## 1 植物材料

植物物证桂花样品全部于 2015 年 10 月采自西南大学校园内, 共随机采集 15 棵桂花树上的叶片, 每个样品随机采集 2~4 片叶片, 分别装入自封袋中, 按 1~15 顺序号标注, 置于 4 °C 冰箱备用. .

## 2 DNA 提取及分析

提取叶片样品基因组总 DNA 进行分析. 把随机采集的 15 棵桂花树叶片, 按序号随机从中取叶片 0.5 克, 在 CTAB(Hexadecyltrimethyl Ammonium Bromide)法<sup>[9]</sup>的基础上略做改进以提取植物总 DNA. 采用 NanoDroP 2000 测定 A260/A280 比值评价其 DNA 浓度及纯度.

## 3 RAD-seq 检测

选择 EcoR1 六碱基限制性内切酶对基因组 DNA 样品进行酶切, 保证产生的 RAD 标记能够在基因组上均有分布, 同时获得的 RAD 标记数量能够达到实验所需的饱和度. 将 PCR 后产物进行 DNA 片段回收, 并对最终构建完成的文库进行检测. 测序仪器为 IlluminaHiSeq4000. 对所测得的原始数据进行过滤, 过滤标准为: 所有 SNP 位点总深度须大于等于 4, 如果 SNP 为杂合型则次好碱基深度须大于等于 2. 过滤后的序列根据 index 序列划分到具体个体, 便于后续分析.

## 4 结果与分析

### 4.1 测序基本信息分析

二代高通量测序获得 15 个已知桂花样本及 3 个未知植物物证样品桂花的原始 DNA 序列, 本文依据如下标准对原始数据进行过滤: ① 仅使用含有 Hind III 酶切识别位点 Read1 序列; ② 利用 Q30 标准对序列质量进行评估; ③ 所得的序列的前 50 bp 不存在不确定碱基; ④ 整条序列中不确定碱基不多于 3 个, 过滤后 reads 数目比例如图 1 所示. 从数据量的整体上看, 数据有效率高达 97.28%, 经过滤后共获得 6,638 bp, 587,602 bp 的 Clean data, 数据量统计如表 1 所示.

### 4.2 SNP 的查找检测和建库

利用 RAD 技术对过滤后的数据进行聚类 and SNP 查找, 并对 15 个样品通过聚类获得的初始 SNP 进行初步过滤, 得到 SNP 数量为 1197085 个(表 2).

表 1 过滤数据量统计表

样本	过滤后的 reads 数	过滤后的碱基数	样本	过滤后的 reads 数	过滤后的碱基数
<i>Osmanthus</i> sp. 1	1, 930, 652	362, 962, 576	<i>Osmanthus</i> sp. 11	2, 013, 516	374, 513, 976
<i>Osmanthus</i> sp. 2	1, 644, 692	312, 491, 480	<i>Osmanthus</i> sp. 12	1, 265, 587	232, 868, 008
<i>Osmanthus</i> sp. 3	1, 656, 527	308, 114, 022	<i>Osmanthus</i> sp. 13	2, 002, 268	376, 426, 384
<i>Osmanthus</i> sp. 4	1, 606, 295	295, 558, 280	<i>Osmanthus</i> sp. 14	1, 960, 908	372, 572, 520
<i>Osmanthus</i> sp. 5	2, 758, 437	518, 586, 156	<i>Osmanthus</i> sp. 15	2, 260, 074	418, 308, 332
<i>Osmanthus</i> sp. 6	2, 791, 362	530, 358, 780	<i>Osmanthus</i> sp. A	1, 895, 016	356, 263, 008
<i>Osmanthus</i> sp. 7	3, 662, 058	681, 142, 788	<i>Osmanthus</i> sp. B	1, 960, 789	372, 549, 910
<i>Osmanthus</i> sp. 8	353, 502	65, 044, 368	<i>Osmanthus</i> sp. C	2, 354, 668	436, 015, 002
<i>Osmanthus</i> sp. 9	1, 748, 449	328, 708, 412	总 计	35, 423, 240	6, 638, 587, 602
<i>Osmanthus</i> sp. 10	1, 558, 440	296, 103, 600			

表 2 经过滤所得的 SNP 信息

样本	总 SNP	纯合的 SNP	纯合率/%	杂合 SNP	杂合率/%
<i>Osmanthus</i> sp. 1	93, 463	75, 897	81.21	17, 566	18.79
<i>Osmanthus</i> sp. 2	77, 768	61, 304	78.83	16, 464	21.17
<i>Osmanthus</i> sp. 3	80, 994	62, 544	77.22	18, 450	22.78
<i>Osmanthus</i> sp. 4	39, 629	6, 529	16.48	33, 100	83.52
<i>Osmanthus</i> sp. 5	83, 420	66, 103	79.24	17, 317	20.76
<i>Osmanthus</i> sp. 6	63, 774	52, 570	82.43	11, 204	17.57
<i>Osmanthus</i> sp. 7	68, 167	55, 654	81.64	12, 513	18.36
<i>Osmanthus</i> sp. 8	3, 373	2, 835	84.05	538	15.95
<i>Osmanthus</i> sp. 9	158, 467	129, 413	81.67	29, 054	18.33
<i>Osmanthus</i> sp. 10	127, 311	101, 440	79.68	25, 871	20.32
<i>Osmanthus</i> sp. 11	125, 369	102, 350	81.64	23, 019	18.36
<i>Osmanthus</i> sp. 12	61, 719	50, 821	82.34	10, 898	17.66
<i>Osmanthus</i> sp. 13	69, 229	56, 272	81.28	12, 957	18.72
<i>Osmanthus</i> sp. 14	65, 116	53, 650	82.39	11, 466	17.61
<i>Osmanthus</i> sp. 15	79, 286	64, 002	80.72	15, 284	19.28

McCarroll 等<sup>[10]</sup>认为群体遗传上的差异主要是通过对其群体上 SNP 位点信息的分析, 来进行群体间遗传多样性的分析. 植物物证桂花样品 SNP 数量从 3373 到 158 467 不等, 样品平均 SNP 数量为 79 805 个. SNP 数量最多的是植物物证桂花样品 9, 最少的是植物物证桂花样品 8. 每个个体的 SNP 差异性都反映了个体的多态性. 杂合度最高个体是植物物证桂花样品 4, 高达 83.52%.

#### 4.3 系统进化树分析

通过 RAD 技术获得 18 个(15 个植物物证已知样品和 3 个未知植物物证样品)桂花 SNP 数据, 然后对 SNP 进行筛选(筛选原则: 纯合 SNP 深度至少为 4, 杂合 SNP 次好深度至少为 2), 最后依据至少在 18 个植物物证样品都存在的 SNP 位点(118)和至少在 17 个植物物证桂花样品都存在的 SNP 位点(797)总信息, 计算样品的遗传距离, 利用 MEGA6 构建系统发育树(图 1).

从图 1 中可以得出, 未知的 A 植物物证样品与桂花 1 样品最为接近, 未知的 B 植物物证样品与桂花 4

样品最为接近,未知的 C 植物物证样品与桂花 11 样品最为接近,从而可初步估计 A,B,C 3 个植物物证样品分别对应桂花 1、桂花 4 和桂花 11 这 3 个样品. 而样品 12 和样品 15 在很接近的同一分支上,笔者推测:① 它们存在很近的亲缘关系;② 所在分支取样稀疏、从而相对位置接近.

#### 4.4 建库比对分析法

为进一步解析并验证系统发育树分析结果,本文进一步通过建立 SNP 数据库和未知植物物证样品的 SNP 位点比较来找出具有区分效力的 SNP 子集. 通过 RAD 技术获得 15 个桂花样品的 SNP 数据集,并对此数据集进行筛选,获得 1 182 个 SNP 位点,然后用这些数据集建库(筛选原则:纯合 SNP 深度至少为 4,杂合 SNP 次好深度至少为 2),将至少在 14 个植物物证样本都存在的 SNP 位点挑选出来建立植物物证桂花样品的 SNP 数据库.

通过 RAD 技术分别获得 3 个未知样品的 SNP 数据集,和以上原则一样进行筛选,然后将得到的 SNP 位点信息与 SNP 库做比对,最终得出统计结果(表 3).

表 3 未知植物物证样品 A,B,C 的 SNP 位点信息与桂花数据库比对的结果

样本	A		B		C	
	符合数	符合度	符合数	符合度	符合数	符合度
<i>Osmanthus</i> sp. 1	148	12.52%	86	7.28%	25	2.12%
<i>Osmanthus</i> sp. 2	116	9.81%	94	7.95%	37	3.13%
<i>Osmanthus</i> sp. 3	120	10.15%	87	7.36%	39	3.30%
<i>Osmanthus</i> sp. 4	107	9.05%	116	9.81%	41	3.47%
<i>Osmanthus</i> sp. 5	116	9.81%	86	7.28%	41	3.47%
<i>Osmanthus</i> sp. 6	121	10.24%	98	8.29%	34	2.88%
<i>Osmanthus</i> sp. 7	118	9.98%	96	8.12%	33	2.79%
<i>Osmanthus</i> sp. 8	12	1.02%	7	0.59%	6	0.51%
<i>Osmanthus</i> sp. 9	118	9.98%	85	7.19%	38	3.21%
<i>Osmanthus</i> sp. 10	123	10.41%	90	7.61%	34	2.88%
<i>Osmanthus</i> sp. 11	34	2.88%	38	3.21%	112	9.48%
<i>Osmanthus</i> sp. 12	89	7.53%	64	5.41%	62	5.25%
<i>Osmanthus</i> sp. 13	87	7.36%	76	6.43%	62	5.25%
<i>Osmanthus</i> sp. 14	79	6.68%	68	5.75%	66	5.58%
<i>Osmanthus</i> sp. 15	88	7.45%	66	5.58%	63	5.33%

从表 3 中可以看出,未知植物物证样品 A 的 SNP 位点信息与桂花数据库比对后,桂花 1 与未知植物物证样品 A 的比对符合度最高,有 148 个位点符合,符合率达到 12.52%;未知植物物证样品 B 的 SNP 位点信息与桂花数据库比对后,桂花 4 与未知植物物证样品 B 的比对符合度最高,有 116 个位点符合,符合率达到 9.81%;未知植物物证样品 C 的 SNP 位点信息与桂花数据库比对后,桂花 11 与未知植物物证样品 C 的比对符合度最高,有 112 个位点符合,符合率达到 9.48%. 从而推测出 A,B,C 这 3 个未知样品分别是桂花 1、桂花 4、桂花 11,与实际情况吻合.

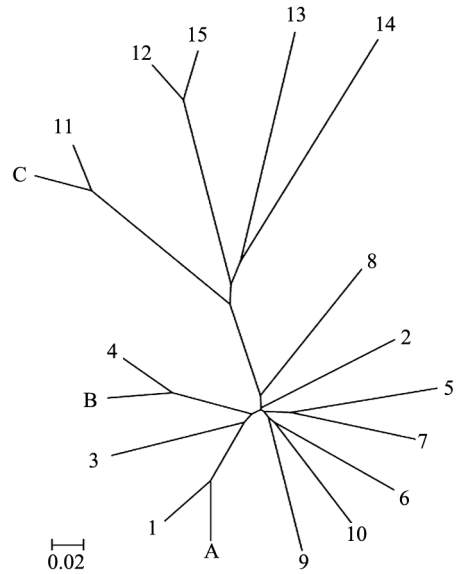


图 1 18 个桂花物证样品的系统发育树

#### 4.5 用 jellyfish 算法绘制聚类图分析

RAD-seq 测序的片段是有 RAD 酶切后的产物, 过滤后的 clean data 两端含有相同的酶切位点. 先切除 reads 两端的酶切序列 (reads1 切除 5 bp, reads2 切除 15 bp), 从而避免酶切序列造成 kmer 偏好性. 然后以 9 bp (kmer) 为滑动窗口, 用 jellyfish 计算所有 reads kmer 出现的频数<sup>[11]</sup>, 考虑到样品测序量的不同, 将频数转化成频率. 根据所得的频率表用欧式距离公式计算样品间的差异度. 距离越大, 表示 2 个样品的遗传相似度越远. 最后利用差异度矩阵通过离差平方和的方式将样品聚类, 2 个物种的差异度越小, 在聚类图中的位置越近. 欧式距离算法特性是高频 kmer 在计算样品差异度中起主导作用, 但是某些序列的特性 (如 *poloA*) 和 PCR 扩增过程等因素会造成某些 kmer 的频数急剧增大, 干扰物种鉴别结果. 因此, 在计算差异度之前, 我们过滤掉高频 kmer (大于 0.005%).

从图 2 中可以看出, 盲测的 3 个样品都能很好地和对照组聚类. A 样品与样本桂花 1 聚类后距离最近, B 样品与样本桂花 4 聚类后距离最近, C 样品与样本桂花 11 聚类后距离最近. 而且 3 组盲测样品聚类的顺序都先于其他样品, 说明 3 组样品的遗传距离是最近的, 聚类结果能很好地和实验设计吻合.

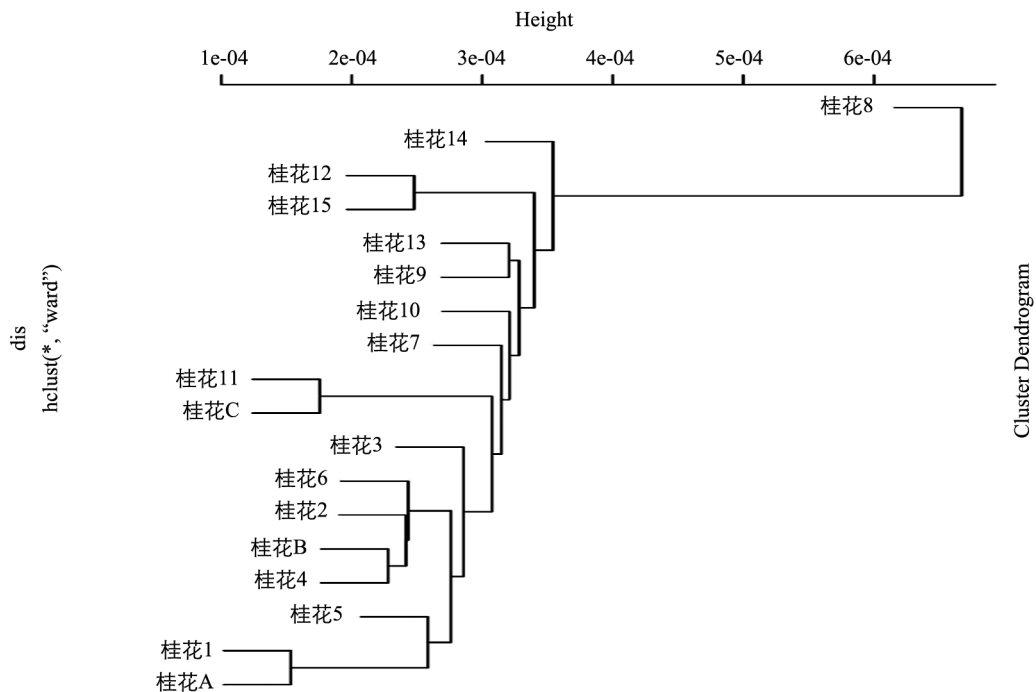


图 2 18 个桂花植物物证样品 kmer 频数聚类图

## 5 结果与讨论

本文建库提取的植物桂花总 DNA 选取的材料是新鲜叶片, 是一种理想状态, 虽然本文首次尝试利用 RAD-seq 技术来识别植物物证桂花个体, 并通过 3 种分析方法都可以成功识别未知的植物物证桂花个体, 鉴定结论成立, 存在同一性联系, 但是在实际案件中植物材料不是干枯残缺就是量小且已破碎、腐败, 肉眼只能鉴别是否植物检材的粗浅辨识, 因此微量残败植物物证 DNA 提取方法的成熟建立也是满足下游实验的重要工作. 本文作者也对植物基因组 DNA 提取方法进行了比较, 发现改良的 CTAB 法比一般的植物 DNA 提取试剂盒获得总 DNA 的产量多, 且纯度高, 但是此方法操作步骤复杂, 耗时长. 在今后的工作中, 要针对不同植物物证样本的状态和部位来建立合适的提取方法, 满足实战需要.

二代测序技术在法庭科学领域中应用发展比较滞后, 尚处于刚刚起步阶段<sup>[12-13]</sup>. 本文从 3 种分析方法都得到了相同的结果, 即未知植物桂花样品 A, B, C 对应库中个体桂花 1, 4, 11. 在建库比对分析法的严格



与宽松 2 种条件下, 聚类图分析法显示 3 个植物未知样品都与桂花 8 的匹配度最低. 这说明虽然分析方法的角度不同, 但是数据内在的关联性被很好地反映了出来, 进一步说明了数据分析的可靠性. 相比较而言, 基于 k-mer 频谱的方法采用了 jellyfish 软件、欧式距离等数学手段对数据进行处理和计算, 减少了数据分析的误差, 更易于推广到公安植物物证鉴别应用中.

从分析流程来看, 还不能对个体鉴定是否存在阈值和是否有物种特异性做出评判. 比如在建库比对分析中, 未知植物样品 A 与桂花 1 有 148 条 SNP 位点信息相吻合, 符合度达到了 12.5%, 这 12.5% 的吻合度是否可以对未知样品 A 与桂花 1 的同一认定做出定性的判断, 在今后不同物种个体识别的研究中, 是否可以找到一个最低阈值来做定性判断标准, 基于 k-mer 频谱的方法给出了通过聚类距离远近来进行同一认定, 但这个距离近到多少才认为是同一个个体也没有明确指出. 这些都有待于下一步的研究, 这项技术要被普及应用到法庭科学, 还需要国际学者的广泛参与, 建立公开明了的工作原理, 优化完善技术流程, 保证数据风险可控及价格稳定等, 才可能被法庭和立法机构接受.

## 6 结 论

RAD-seq 技术结合该文的分析方法能够识别出未知植物物证样品 A, B 和 C 分别对应于桂花 1, 4 和 11, 实现对案件中关联植物物证进行精确检验鉴定的方法和一个小型鉴定实例, 但其识别能力有待扩展到更大样本集和更多物种中.

### 参考文献:

- [1] HEBERT P D N, CYWINSKAA, BALL S L, et al. Biological Identifications Through DNA Barcodes [J]. Proceedings of the Royal Society B: Biological Sciences, 2003, 270(1512): 313–321.
- [2] 王彦学. 植物 DNA 检验技术在命案现场重建中的应用策略 [J]. 中国司法鉴定, 2014, 74(3): 53–57.
- [3] 付 涛, 王志龙, 钱萍仙, 等. 高等植物 DNA 条形码最新研究进展及其应用 [J]. 核农学报, 2016, 30(5): 0887–0896.
- [4] ALTSHULER D, POLLARA V J, COWLES C R, et al. An SNP Map of the Human Genome Generated by Reduced Representation Shotgun Sequencing [J]. Nature, 2000, 407(6803): 513–516.
- [5] 王洋坤, 胡 艳, 张天真. RAD-seq 技术在基因组研究中的现状及展望 [J]. 遗传, 2014, 36(1): 41–49.
- [6] POLAND J A, BROWN P J, SORRELLS M E, et al. Development of High-Density Genetic Maps for Barley and Wheat Using a Novel Two-Enzyme Genotyping-by-Sequencing Approach [J]. PLoS ONE, 2012, 7(2): e32253.
- [7] BARCHI L, LANTERI S, PORTIS E, et al. Identification of SNP and SSR Markers in Eggplant Using RAD Tag Sequencing [J]. BMC Genomics, 2011, 12(1): 304.
- [8] DAVIDE S, ALBERTO A, EZIOP, et al. Tirone M Knapp SJ Lanteri S RAD Tag Sequencing as a Source of SNP Markers in *Cynara cardunculus* L. [J]. BMC Genomics, 2012, 13(1): 3.
- [9] DOYLE J J, DOYLE J L. Arapid. DNA Isolation Procedure for Small Quantities of Fresh Leaf Tissue [J]. Phytochemical Bulletin, 1987, 19(1): 11–15.
- [10] MCCARROLL S A, KURUVILLA F G, KORN J. Integrated Detection and Population-Genetic Analysis of SNPs and Copy Number Variation [J]. Nature Genetics, 2008, 40(1): 1166–1174.
- [11] MARCAIS G, KINGSFORD C. A Fast Lock-Free Approach for Efficient Parallel Counting of Occurrences of k-mers [J]. Bioinformatics, 2011, 27(6): 764–770.
- [12] 王 乐, 叶 健, 白 雪, 等. 二代测序技术及其在法医遗传学中的应用 [J]. 刑事技术, 2015 40(5): 353–358.
- [13] GETTINGS K B, KIESLER K M, VALLONE P M. Performance of a Next Generation Sequencing SNP Assay on Degraded DNA [J]. Forensic Science International Genetics, 2015, 19(1): 1–9.

## Application of RAD-seq Technology in Individual Identification of Plant Material Evidence

ZHUANG Yan<sup>1</sup>, KONG Qing-bo<sup>1</sup>, CHEN Xin-ping<sup>1</sup>, LU Hai-xin<sup>2</sup>,  
WANG Xiao-li<sup>2</sup>, WANG Chong-zhi<sup>2</sup>, SUN Hai-lu<sup>2</sup>

1. Scientific Research Office, Chongqing Police College, Chongqing 401331, China;

2. Shenzhen BGI, Shenzhen Guangzhou 518000, China

**Abstract:** Plant evidence related to the suspect/victim found at the crime scene can provide valuable evidence or clues for investigation and evidence collection of the case. In a study reported in this paper, an RAD-seq (restriction-site associated DNA sequencing) technique was used to build a database and make high-throughput sequencing for 18 samples of *Osmanthus* (15 known samples and 3 blind samples), analyze their sequence polymorphism and evaluate their genetic diversity. Two kinds of comparative analysis method, one based on SNP (single nucleotide polymorphism) and the other on k-mer spectrum, were employed to discuss the scientific nature of RAD-seq data for individual identification and the feasibility of identity identification of plant material evidence. The results showed that both methods could successfully identify the three blind samples. Comparatively, the k-mer spectrum-based method was superior to the analysis method based on the SNP genetic relationship tree in stability and automation degree. It was easier to apply to the public security material evidence of identification. In conclusion, the RAD-seq technology used in combination with the analysis method successfully identified the three unknown plant material evidence samples A, B and C, and they corresponded to the sweet scented osmanthus 1, 4 and 11, respectively. It has provided a method for accurate examination and identification of related plant material evidence in a case and a small identification example, but its ability of identification is to be extended to a larger set of samples and more plant species.

**Key words:** plant evidence; RAD-seq; SNP; individual identification; *Osmanthus* sp.

责任编辑 夏 娟

