

实时直接分析质谱技术在木材识别研究中的应用

张毛毛¹, 蒋劲东², 刘波¹, 郭娟¹, 焦立超¹, 殷亚方¹

(1. 中国林科院木材工业研究所, 北京 100091; 2. 浙江省木雕红木家具产品质量检验中心, 浙江东阳 322100)

摘要: 实时直接分析质谱 (Direct Analysis in Real Time Mass Spectrometry, 简称 DART-MS) 是一种新型常温常压离子化质谱技术, 可直接对固体样品进行检测, 操作方便、准确度高、稳定性好、分析速度快, 目前已开始应用于木材识别领域研究。文中介绍 DART-MS 技术的基本原理以及在木材识别研究中的应用, 比较分析了 DART-MS 数据处理和分析方法, 并建议采用 DART-MS 技术实现木材树种的科学精确识别。

关键词: 木材; 树种鉴定; 实时直接分析质谱; 化学指纹图谱; 无监督分类法; 有监督分类法

中图分类号: S781 文献标识码: A 文章编号: 1001-8654 (2019) 01-0029-05

Application of Direct Analysis in Real Time Mass Spectrometry (DART-MS) for Wood Identification

ZHANG Mao-mao¹, JIANG Jin-dong², LIU Bo¹, GUO Juan¹, JIAO Li-chao¹, YIN Ya-fang¹

(1. Research Institute of Wood Industry, Chinese Academy of Forestry, Beijing 100091, China;

2. Zhejiang Hongmu Carved Furniture Product Quality Inspection Center, Dongyang 322100, Zhejiang, China)

Abstract: Direct analysis in real time mass spectrometry (DART-MS) is a new atmospheric ionization technology under ambient temperature. It can be used in chemical investigation for solid samples in an easy, accurate, stable and direct way. Recently, DART-MS was introduced into wood forensic identification. In this paper, working principles of DART-MS and its applications in wood identification, data processing and statistical methods were reviewed to provide basic information for more scientific and precise wood identification.

Key words: wood; species identification; DART-MS; chemical fingerprint spectrum; unsupervised classification; supervised classification

木材是一种天然的高分子复合材料, 具有独特的自然属性。不同的树种, 其木材具有不同的构造特征与加工利用性质。正确地认识与鉴别木材, 是

实现合理有效地保护、管理和利用木材资源的重要前提, 也是规避和化解木材进出口贸易风险的技术基础。此外, 木材识别技术在木质文物的保护和修复、木制品市场的健康发展等方面, 也具有重要的实际意义。

传统的木材识别技术建立在木材解剖学基础上, 主要通过木材宏观和微观特征相结合的方法实现, 但操作过程复杂、耗时较长, 对专业知识要求高, 且一般只能识别到“属”或“类”, 同时也无法实现木材产地的识别^[1]。最新发展的木材识别技术

收稿日期: 2018-05-06; 修改日期: 2018-12-21

基金项目: 中国林业科学研究院林业新技术研究所基本科研业务费专项资金“紫檀属木材荧光识别技术研究”(CAFINT2015C06); 国家“万人计划”特殊支持项目“木材 DNA 条形码与化学识别新技术研究”(W02020331)。

作者简介: 张毛毛 (1990—), 女, 中国林科院木材工业研究所博士研究生。

责任作者: 殷亚方, 男, 中国林科院木材工业研究所研究员。

中, DNA 条形码技术, 是通过提取和分析木质部细胞中特异性的 DNA 片段, 实现木材“种”的鉴别, 但也因为测试周期长、成本高, DNA 提取难度大, 技术不易普及^[2]。稳定同位素技术可以对木材进行产地追踪, 但对样品来源要求高, 识别精度低(一般大于 250 KM), 且需要建立通用的分析数据库^[3]。因此, 亟待开发新的木材识别技术, 对实现木材树种的快速准确识别, 意义重大。

化学指纹图谱(Chemical Fingerprint)是木材识别研究领域的重要应用技术。木材的化学成分主要包括纤维素、半纤维素、木质素和抽提物, 不同树种之间木材化学成分差异很大。但由于受遗传因素影响, 同种木材的化学成分通常比较相似, 难以细分。

近年来, 基于化学指纹图谱技术, 国内外学者相继采用了包括色谱质谱、核磁共振、光谱等在内的化学分析方法, 应用在木材识别研究领域。其中, 实时直接分析质谱(Direct Analysis in Real Time Mass Spectrometry, 简称 DART-MS)是最近快速发展起来的一种新技术, 由实时直接分析离子源与不同类型质谱仪串联而成, 不需要复杂的样品前处理、可以直接对固体样品进行检测, 具有操作方便、准确度高、稳定性好、分析速度快等优点, 已经在食品、药品、材料以及环境检测领域得到广泛应用^[4-7]。

DART-MS 技术在木材识别方面的应用研究国外已有报道^[8-11], 然而在我国木材科学研究领域至今未见报道。笔者通过归纳近年来 DART-MS 在木材识别领域方面的研究进展, 旨在突破传统方法的局限性, 为木材识别技术研究的进一步发展, 提供新的思路。

1 DART-MS 技术的应用

1.1 应用原理及范围

实时直接分析(DART), 是 Robert Cody 和 Jim Laramée 等在 2005 年发明的常温常压离子化质谱技术^[12], 原理是在常温常压下, 载气(如 He 或 N₂)经放电产生的激发态原子, 解吸并离子化样品中的化合物, 进而以质谱或串联质谱检测, 实现对样品的实时直接分析。该技术对样品的适用性强, 可检测液体、气体、固体样品, 摒弃了传统质谱复杂的样品前处理过程, 具有快速实时反应机制, 可实现高通量检测等特点^[12-14]。迄今为止, DART 已成

功与各种质谱仪实现串联, 包括飞行时间质谱、傅立叶变换质谱、离子阱质谱、三重四级杆质谱等。

自 2005 年 DART 商业化设备问世以来, DART-MS 技术发展迅速, 可直接用于分析固体样品, 因此在中药质量控制领域得到了广泛应用, 也被引入医学临床分析领域, 及成为法检领域物证分析重要工具, 可快速筛查有毒化学品、区分纸张和油墨、检测爆炸物等。

同时, DART-MS 技术在食品安全检测^[15]、环境污染检测领域^[6]以及在包装、纺织、建筑等材料的有毒有害物质检测^[16-17]方面, 也取得了突破性进展。

1.2 在木材识别研究领域的应用

相比其他领域, DART-MS 在木材识别领域的应用起步较晚, 相关研究报道主要来自于 DART 的发明人 Cody, 以及美国鱼类及野生动物管理局的研究团队。

2012 年, DART-MS 被引入到沉香产品中的 2-(2-苯基乙基)色酮研究中, 确定了鉴定沉香的特征离子 319.118 m/z 或 349.129 m/z, 为沉香的鉴定提出了新方法^[18]。DART-MS 还被应用于针叶材的鉴别, 通过研究窄叶南洋杉(*Araucaria angustifolia*)、异叶南洋杉(*A. heterophylla*)、新西兰贝壳杉(*Agathis australis*)和瓦勒迈杉(*Wollemia nobilis*)的心材和边材, 证实 DART-MS 有助于帮助区分南洋杉科(*Araucariaceae*)中不同树种的木材^[19]。

此外, 非法的木材贸易和跨国运输往往会掩盖其地理来源, DART-MS 技术除了用于木材解剖特征相似树种的鉴别之外, 在确定木材地理来源方面也发挥着作用。研究表明, 来自两个不同地理位置的北美黄杉(*Pseudotsuga menziesii*)木材, 其 DART-MS 质谱图中呈现 32 种与地理位置显著相关的化合物。虽然由于可选木材样本产地较少, 对距离 100 km 内两个产地的识别正确率仅为 70%~76%, 但未来通过扩大更多产地样本的研究, 将有助于提高木材识别精度^[20]。

已有研究表明 DART-MS 在木材识别方面具有很大潜力, 但离子模式、样品状态、树种等因素还是会影响 DART-MS 对木材树种的识别效果。

1) 离子模式: DART-MS 在正负离子模式下得到的结果会有所差异, 进而影响对木材树种的识别

效果。如正离子模式下,美洲白栎(*Quercus alba*)和红栎(*Q. rubra*)木材之间差异不显著,但在负离子模式下,两个树种差异显著,经进一步线性判别分析后,未知样品均被准确识别,证实 DART-MS 负离子模式鉴别美洲白栎和红栎木材的可行性,同时揭示了离子模式对识别效果的显著影响^[13]。

2) 样品状态:在针对濒危野生动植物种国际贸易公约(CITES)管制的黄檀属(*Dalbergia*)木材的识别研究中,通过 DART-MS 方法,可以有效区分巴西黑黄檀(*Dalbergia nigra*)和亚马逊黄檀(*D. spruceana*),木材甲醇提取物的分析结果与干燥木材薄片的结果一致^[10],说明 DART-MS 对不同形态的木材样品具有一定的适用性。

同时,现有应用 DART-MS 的木材识别研究中,大部分木材样品都为自然状态,未经过加热或粉碎处理。尽管有研究推测,当质谱中能够有效保留鉴别物种的特征标记物时,样品处理条件对木材树种鉴别能力可能影响不大^[21]。但由于干燥或其他加热处理所引起的木材化学组分的变化,将会直接影响其 DART-MS 质谱图的主要特征,从而最终影响木材树种的识别精度。

3) 树种:通过对 8 种黄檀属树种以及 6 种非黄檀属树种心材的 DART-MS 谱图进行分析,发现巴西黑黄檀和亚马逊黄檀之间,以及伯利兹黄檀(*D. stevensonii*)和危地马拉黄檀(*D. tucurensis*)之间较易区分,而微凹黄檀(*D. retusa*)和中美洲黄檀(*D. granadillo*)则无法区分。从而推断,两者可能属于同一物种^[9, 11]。虽然微凹黄檀和中美洲黄檀是否同种异名仍有待考证,但树种的正确定名会直接影响 DART-MS 对木材的识别精度。因此,在利用 DART-MS 建立木材指纹图谱数据库时,应采用经正确定名、来源可靠的木材标本。

2 DART-MS 数据处理和分析方法

DART-MS 由实时直接分析离子源与不同类型质谱仪串联而成,最终得到不同类型的质谱数据。这些数据一般具有数量多、数据复杂、各变量之间具有相关性等特点,无法通过简单的分析方法提取有效信息。因此,需从海量数据中挖掘出变量之间的相互关系,对于后续的筛选特征标记物,具有重要

意义。数据的分析过程主要包括:数据预处理、无监督分类以及有监督分类等。

2.1 数据预处理

通过质谱仪得到的数据量比较大,但噪声较多,需进行质谱数据峰提取、峰对齐、填补空缺峰以及归一化之后,才能进行统计分析,因此,前期的数据预处理必不可少。常用的质谱预处理软件包括 ChromaTOF、SpectConnect、MetAlign、MarkerLynx、MassHunter、MassLynx、MarkerView、XCMS^[22]、Compound Discoverer、Progenesis QI^[23-24]、MZmine^[25]、MetaboAnalyst^[26]和 TSSPro3^[10-11, 18-20]。其中,大部分软件需与特定仪器配套使用,具有一定局限性。

TSSPro3 由 DART 发明公司开发,广泛用于 DART-MS 的木材分类识别研究中,前述黄檀属、沉香、南洋杉科的树种识别研究中,均使用了 TSSPro3 软件。但该软件价格昂贵,目前尚未被引入我国。

XCMS、MetAlign、MZmine 以及 MetaboAnalyst 为公开的免费软件,已在质谱数据的预处理方面得到广泛应用。同时,XCMS 是基于 R 语言开发的软件包,源代码公开并且可根据用户需求修改^[27],可满足国内 DART-MS 用户的数据预处理需求。

2.2 无监督分类法

无监督分类法,是指在样品分组信息未知的前提下,通过数据降维和分析,实现样品间潜在关系的可视化^[22]。主要方法包括:主成分分析(PCA)和聚类分析(HCA),其中,PCA 最为常见^[28]。

经 PCA 降维之后,可以直观看到样本之间、样本与变量之间的相关关系,这种方法在质谱、核磁共振波谱、光谱等技术的数据分析中应用十分广泛。PCA 法已分别用于野生沉香与人工诱导沉香、南洋杉科以及黄檀属木材识别研究的质谱分析中^[8-9, 29]。

此外,PCA 法也在不断发展改进。核主成分(Kernel PCA)方法的提出,可实现非线性数据的分析,具有更好的特征提取效果^[30]。随后提出的稳态 PCA(Robust PCA)的方法,通过投射原理和低维空间估算,改进了 PCA 的分类效果,可更好地找出异常点^[31]。

2.3 有监督分类法

有监督分类法,是指在样品分类信息已知的前

提下,寻找不同组别之间的变量差异^[22]。常用的方法包括:线性判别(LDA)、最近邻算法(KNN)、神经网络(NN)、随机森林以及偏最小二乘判别分析(PLS-DA)等,其中以PLS-DA法最为常见。

PLS-DA的原理是先对数据进行PCA分析,然后线性回归建立校正模型,再对预测集样本进行预测^[32]。在PLS分析模型中,忽略了矩阵X和Y的某些正交信息,为了提高PLS-DA的分类性能,正交偏最小二乘法(OPLS)被提出并得到发展^[33],同时其分类判别的能力,已经在植物代谢组学以及野生沉香和人工诱导沉香的分类研究中得到证实^[34-35]。

LDA法也被广泛用于基于DART-MS技术的植物分类研究中^[10]。通过随机森林法建立木材年份、树种来源、年份和来源的识别模型,预测样品来源的正确率达到70%~76%^[20]。神经网络方法适合大量数据的分析,能有效解决样品和相关参数之间的非线性问题。

不同算法具有不同的适用性,通过选用多种算法进行综合比较分析,最终根据判别模型及其精度来确定最优算法。2016年,研究人员利用LDA、KNN和NN三种算法,对肉桂样品的DART-MS数据进行了分析和鉴别,三种模型分类正确率依次为100%、92%和92%^[36]。这些研究为选择科学、可靠的木材样品质谱数据分析方法,奠定了重要基础。

3 未来研究重点及前景展望

DART-MS方法作为一种新的化学指纹图谱技术,将为现有木材识别技术开辟新的途径。尽管该技术在木材树种识别方面具有较好的潜力及应用前景,但在木材识别领域的应用仍处于起步阶段。因此,今后的木材识别研究中,需加强DART-MS新技术的完善,优化质谱数据分析方法,挖掘可用于不同木材物种识别的特征化合物。

未来的研究重点:

1) 建立木材DART-MS指纹图谱数据库,作为适应绝大多数树种的化学指纹图谱参照标准,是该方法能否广泛和准确应用的技术关键;

2) 筛选和确定适用于木材树种识别的特征化合物;

3) 探求不同样本处理条件对DART-MS识别模

型的影响机制;

4) DART-MS数据处理和分析方法的选择和优化。

木材DART-MS信息数据库的建立,需全球范围内的共同合作与努力;木材DART-MS识别技术的开发和储备,将为木材贸易执法和监管提供技术支持。通过基于形态学的传统木材解剖方法与DART-MS技术相结合,可实现木材树种的准确识别,突破传统木材识别方法的局限性,提高木材化学法识别技术的精度,最终实现该技术的工业化应用。

参考文献:

- [1] 姜笑梅,殷亚方,刘波. 木材树种识别技术现状、发展与展望[J]. 木材工业, 2010, 24(4): 36-39.
- [2] Hartvig I, Czako M, Kjær E D, et al. The use of DNA barcoding in identification and conservation of rosewood (*Dalbergia* spp.)[J]. PLoS One, 2015, 10(9): e0138231.
- [3] Dormontt E E, Boner M, Braun B, et al. Forensic timber identification: It's time to integrate disciplines to combat illegal logging[J]. Biological Conservation, 2015, 191: 790-798.
- [4] Vaclavik L, Cajka T, Hrbek V, et al. Ambient mass spectrometry employing direct analysis in real time (DART) ion source for olive oil quality and authenticity assessment[J]. Analytica Chimica Acta, 2009, 645(1): 56-63.
- [5] Zeng S, Wang L, Chen T, et al. Direct analysis in real time mass spectrometry and multivariate data analysis: a novel approach to rapid identification of analytical markers for quality control of traditional Chinese medicine preparation[J]. Analytica chimica acta, 2012, 733: 38-47.
- [6] Grange A H. Semi-quantitative analysis of contaminants in soils by direct analysis in real time (DART) mass spectrometry[J]. Rapid Communications in Mass Spectrometry, 2013, 27(2): 305-318.
- [7] Haunschild M, Klampfl C W, Buchberger W, et al. Rapid identification of stabilisers in polypropylene using time-of-flight mass spectrometry and DART as ion source[J]. Analyst, 2010, 135(1): 80-85.
- [8] Espinoza E O, Lancaster C A, Kreitals N M, et al. Distinguishing wild from cultivated agarwood (*Aquilaria* spp.) using direct analysis in real time and time of - flight mass spectrometry[J]. Rapid Communications in Mass Spectrometry, 2014, 28(3): 281-289.
- [9] Espinoza E O, Wiemann M C, Barajas-Morales J, et al. Forensic analysis of CITES-protected *Dalbergia* timber from the Americas[J]. IAWA Journal, 2015, 36(3): 311-325.
- [10] Lancaster C, Espinoza E. Analysis of select *Dalbergia* and trade timber using direct analysis in real time and time-of-flight mass spectrometry for CITES enforcement[J]. Rapid Communications in Mass Spectrometry, 2012, 26(9): 1147-1156.
- [11] Musah R A, Espinoza E O, Cody R B, et al. A high throughput ambient mass spectrometric approach to species identification and classification from chemical fingerprint signatures[J]. Scientific reports, 2015, 5: 11520.
- [12] Cody R B, Laramée J A, Durst H D. Versatile new ion source for the analysis of materials in open air under ambient conditions[J]. Analytical chemistry, 2005, 77(8): 2297-2302.
- [13] Cody R B, Dane A J, Dawson-Andoh B, et al. Rapid classification of

- (责任编辑 劳万里、姜 征)

(上接第24页)

- (责任编辑 孟凡丹、向 琴)