网络出版时间:2026-03-13 DOI:10.13207/j.jnwafu.2026.09.006

柠檬桉种源遗传背景的近红外光谱分析

林彦[1],卢万鸿[1],翟江博[1],刘果[1],罗建中[1],张鹏[2]
(1 中国林业科学研究院 速生树木研究所,广东 湛江 524022;2 黄石市鹏程中学,湖北 黄石 435100)

[摘 要] 【目的】借助近红外光谱(near infrared spectroscopy,NIRs)信息分析柠檬桉(*Corymbia citriodora* Hook.)不同种源间的亲缘关系,探索NIRs用于其遗传信息分析的可行性和准确性。【方法】以国家桉树种质资源库保存的柠檬桉10个种源家系材料为研究对象,每个种源按照广泛性和代表性的取样原则共采集296个家系的叶样,每个家系仔细选择6~10片外形健康、生理条件一致的叶片,烘干、粉碎后装袋,采用便携式近红外光谱仪Phazir Rx采集叶片粉样的NIRs信息。NIRs数据经Savitzky-Golay平滑二阶导数预处理后,建立柠檬桉种源的主成分分析(PCA)模型。通过无监督的判别分析(SIMCA),估算10个种源两两之间的NIRs光谱距离。基于NIRs光谱距离构建各种源的层级聚类树。通过不同波长建模能力(modeling power)分析,解析柠檬桉种源间遗传差异的物质基础。【结果】NIRs能够有效量化并区分柠檬桉种源。各种源间的NIRs欧氏距离为0.016~0.024,并将10个种源明确地划分为澳大利亚原生和国内引种两大组,且原生种源的聚类关系与其地理分布高度吻合,为追溯国内引种种源的来源提供了直接证据。SIMCA进一步证实,即使是NIRs光谱距离相近的种源,其包含的296个家系样本也能被彼此模型清晰、准确分类。建模能力分析显示,1 708~1 785,1 880~1 935及2 292~2 370 nm 3个波段的建模能力尤为突出,其分别对应纤维素、半纤维素和C—H基团等关键化学成分的特征吸收峰,而种源间这些化合物或基团的差异很可能是其遗传分化的根源。【结论】NIRs分析能有效解析柠檬桉种源间的亲缘关系和地理分化信息,为厘清国内早期引种材料的来源问题提供了新的技术方案。

[关键词] 柠檬桉;遗传背景;近红外光谱(NIRs);光谱距离;层级聚类;判别分析
[中图分类号] S792.39;S126;O657.33 [文献标志码] A [文章编号] 1671-9387(2026)09-0059-08

Near infrared spectroscopy analysis of the genetic background of *Corymbia citriodora* provenances

LIN Yan[1],LU Wanhong[1],ZHAI Jiangbo[1],LIU Guo[1],LUO Jianzhong[1],ZHANG Peng[2]
(1 *Research Institute of Fast-growing Trees,Chinese Academy of Forestry,Zhanjiang,Guangdong* 524022,*China*;
2 *Huangshi Pengcheng Junior High School,Huangshi,Hubei* 435100,*China*)

Abstract:【Objective】This study aims to analyze the genetic relationships among different provenances of *Corymbia citriodora* using near-infrared (NIR) spectral data,and to explore the feasibility and accuracy of applying NIR for genetic analysis.【Method】The study utilized genetic materials from 10 provenances of *C. citriodora* preserved in the National Eucalypt Germplasm Repository. Based on the principles of representativeness and broad sampling,leaf samples from 296 families were collected. For each family,6—10 healthy leaves with consistent physiological conditions were carefully selected,dried,ground,and packed into plastic self sealing

[收稿日期] 2025-10-29 [修回日期] 2026-01-07
[基金项目] 广东省林业科技创新项目(2023KJCX012);国家重点研发计划子课题(2022YFD2200203-3);广东省自然科学基金项目(2021A1515010618)
[作者简介] 林彦(1977—),女,广西隆安人,高级工程师,硕士,主要从事桉树无性繁殖、杂种优势理论及新品种开发测试评价研究。E-mail:liny10000@126.com
[通信作者] 卢万鸿(1982—),男,陕西扶风人,副研究员,博士,硕士生导师,主要从事林木遗传育种研究。E-mail:luwanhong@caf.ac.cn

bags. NIRs of the powdered samples were acquired using a portable NIR spectrometer (Phazir Rx). After pretreating the original spectra with Savitzky-Golay second derivative smoothing, a principal component analysis (PCA) model was established. Spectral distances between each pair of the 10 provenances were estimated using Soft Independent Modeling of Class Analogy (SIMCA), an unsupervised discriminant analysis method. A hierarchical clustering tree of provenances was then constructed based on the NIRs distances. Furthermore, the analysis of modeling power at different wavelengths were conducted to elucidate the chemical basis of genetic variations among provenances.【Result】NIRs proved effective for quantifying and distinguishing *C. citriodora* provenances. The Euclidean spectral distances among provenances ranged from 0.016 to 0.024, clearly clustering the ten provenances into two main groups: native Australian and domestic introduced provenances. The clustering pattern of native provenances strongly aligned with their geographic distribution, providing direct evidence for tracing the origins of the introduced germplasms in China. SIMCA analysis further confirmed that the 296 family samples could still be accurately and distinctly classified by their respective models, even for closely related provenances with similar spectral distances. Modeling power analysis highlighted three spectral ranges with particularly high discriminatory power: 1 708—1 785, 18 80—1 935, and 2 292—2 370 nm. And these waves corresponded to the characteristic absorption peaks of key chemical compounds and functional groups, namely cellulose, hemicellulose, and C—H groups, respectively. The differences in these compounds or functional groups among provenances were likely be the inner reasons of genetic differentiation on *C. citriodora*.【Conclusion】The present study showed that NIR can effectively reveal the genetic relationships and geographic differentiation among provenances of *C. citriodora*, which provided a novel solution for clarifying genetic origins of early-introduced germplasm of this species in China.

Key words: *Corymbia citriodora* Hook.; genetic background; near-infrared spectroscopy (NIRs); spectral distance; hierarchical clustering; discriminant analysis

柠檬桉(*Corymbia citriodora* Hook.)是桃金娘科(Myrtaceae)伞房属(*Corymbia*)重要树种,因其速生、材质优良及富含高经济价值的精油,在全球亚热带地区被广泛引种栽培。柠檬桉具有丰富的遗传多样性,HEALEY等[1]的基因组学研究表明,病虫害和干旱等环境压力是塑造其基因组多样性的关键驱动力,这导致了不同地理种源在适应性和经济性状上存在显著差异。此外,柠檬桉桉叶油成分及其含量也因种植环境的差异而存在丰富变异。TOLBA等[2]研究表明,广西南宁的柠檬桉桉叶油主要成分为香茅醛(57.0%)、香茅醇(15.9%)和乙酸香茅酯(15.3%);HEALEY等[3]研究发现,阿尔及尔柠檬桉的桉叶油主要含有香茅醛(69.8%)、香茅醇(10.6%)和异蒲勒醇(4.7%)。这些成分在医药领域展现出巨大的应用潜力,ANSARI等[4]的研究证实其具有调节胰岛素分泌的生物活性。在工业应用上,DA COSTA等[5]研究了伞房属与桉属杂种无性系的制浆性能,明确其可作为优质纸浆材树种,并强调优良种源选择对提升两者杂交种的生产力具有决定性作用。因此,深入解析柠檬桉种源的遗传背景,对其高效遗传改良和产业化利用至关重要。

中国对柠檬桉的引种可追溯至1920s左右,当时对其仅限树种水平的适应性研究,未开展种源(家系)水平的遗传差异分析,对其种内遗传背景信息重视不够,甚至忽视[6]。因此,早期引种的柠檬桉种源存在档案信息丢失或混杂的问题,严重阻碍了现存种质资源的系统评估、优良亲本的组配及世代改良进程。准确鉴定并厘清这些材料的遗传背景,是当前柠檬桉育种工作面临的瓶颈问题。传统的种源鉴定方法主要依赖于形态学和表型性状的观测,该方法虽然简便,但性状极易受到立地环境与基因型相互作用的影响,区分精度有限,难以揭示深层的遗传关系。近年来,简单重复序列(simple sequence repeats, SSR)、单核苷酸多态性(single nucleotide polymorphism, SNP)等标记或测序技术已成为群体遗传基础分析的常规方法[7],其能直接从DNA层面提供精确的遗传信息,但其分析流程复杂、成本高昂且耗时较长,限制了其在大规模种质资源快速筛选和初步普查中的应用。因此,亟需一种兼具高通量、低成本和高精度的技术手段来弥补现有方法的不足。近红外光谱(near-infrared

spectroscopy，NIRs）分析技术为此提供了新的解决方案。该技术通过检测样品有机分子中—CH、—OH、—NH等主要化学键的振动倍频与合频吸收，快速获取其整体的化学组分指纹信息。其科学依据在于，基因型即遗传背景的差异决定了植物新陈代谢途径的差异，最终体现为叶片中纤维素、半纤维素、木质素、蛋白质等化学物质含量与状态的差异。NIRs能够灵敏地捕捉到这些由基因决定的细微化学变化[8-9]。因此，通过分析NIRs可以间接推断样品的遗传背景。相较于传统方法，NIRs分析具有快速、无损、成本低廉等突出优点，已在林木品质性状预测等领域展现巨大潜力。

本研究使用便携式近红外光谱仪，采集10个国内外柠檬桉种源的叶片光谱信息，并通过化学计量学方法系统分析不同种源间的亲缘关系。具体包括：验证NIRs光谱信息有效区分不同地理来源柠檬桉种源的能力；探索利用NIRs光谱数据为国内早期引种种源提供溯源线索的可能性；解析造成种源间光谱差异的内在物质基础。本研究旨在探索NIRs技术用于柠檬桉遗传信息分析的可行性与准确性，尝试为林木种质资源亲缘关系等遗传基础分析提供新方法。

1　材料与方法

1.1　材　料

本研究以团队收集保存于南方国家级林木种苗示范基地（广东湛江）的柠檬桉10个种源家系材料为对象，其中5个种源来自柠檬桉的原生地澳大利亚，以来源地划分的5个国内种源为早期引种的材料。2024年3月于桉树生长旺季，采集目标植株中上部枝条，收集生理成熟、健康叶片10～20片，置于纸质信封，并确保不同单株所采叶样生理成熟度一致。每个种源采集尽量多的家系材料，最终共采集了296个家系的叶样（表1），基本满足NIRs分析对材料广谱性和代表性的需求。

表1　柠檬桉种源信息及采样家系数量

Table 1　Provenance information and sampled family counts for *Corymbia citriodora*

种源 Provenances	经度/纬度 Longitude/Latitude	海拔/m Altitude	采样家系数 Number of families sampled
格伦戈登 Glen Gordon，SW Ravenshoe，N QLD（Gordon）	145°19.79′—24.47′ E/17°44.70′—46.24′ S	657～746	49
休恩登 N Hughenden，N QLD（Hughenden）	145°13.65′—15.67′ E/19°33.71′—50.54′ S	892～1 012	40
奎德 Quaid Rd，N QLD（Quaid）	145°26.77′—27.19′ E/16°45.67′—45.76′ S	485～524	29
银谷 Silver Valley Rd & Atherton herberton Rd，Herberton，N QLD（Silver）	145°18.48′—24.77′ E/17°21.04′—25.98′ S	900～968	33
马里巴 W Mareeba，N QLD（Mareeba）	145°17.32′—20.70′ E/16°59.91′—17°1.42′ S	517～620	54
云南元谋 Yuanmou，Yunnan（YNYM）	101°45′ E/25°50′ N	1 130	27
广西东门林场院区 Dongmen Forest Farm Campus，Guangxi（GXDM）	107°51′ E/22°20′ N	123	16
广西高峰林场延河分场 Yanhe Branch of Gaofeng Forest Farm，Guangxi（GXGF）	108°20′ E/22°57′ N	150	17
广东惠州博罗象头山 Xiangtou Mountain，Boluo，Huizhou，Guangdong（GDHZ）	114°24′ E/23°14′ N	300	15
华南农业大学 South China Agricultural University（SCAU）	113°21′ E/23°09′ N	18	16

1.2　光谱仪器

便携式近红外仪Phazir Rx（1624）（Polychromix，Thermo Scientific，USA）用于柠檬桉叶样的NIRs信息采集。该仪器的分辨率为12 nm，波长范围为1 600～2 400 nm，内置可编程的MEMS微衍射光栅，自带背景校正板，单次扫描可获得100维向量数据[9]。

1.3　方　法

1.3.1　光谱采集　柠檬桉每个家系选择6～10片外形健康、生理条件一致的叶片，置于Classic Line烘箱（BINDER，Germany），在60 ℃条件下，烘72 h至绝干（含水量0%）。用中药粉碎机将烘干后的叶样磨碎呈粉样，磨碎时间为1 min。粉样装于透明的自封口塑料袋，以备NIRs采集。将样品袋平放在PVC板上，用上述Phazir仪隔袋换点扫描5次，其均值代表该家系的NIRs信息[9-11]，10个种源共获得296条NIRs信息。

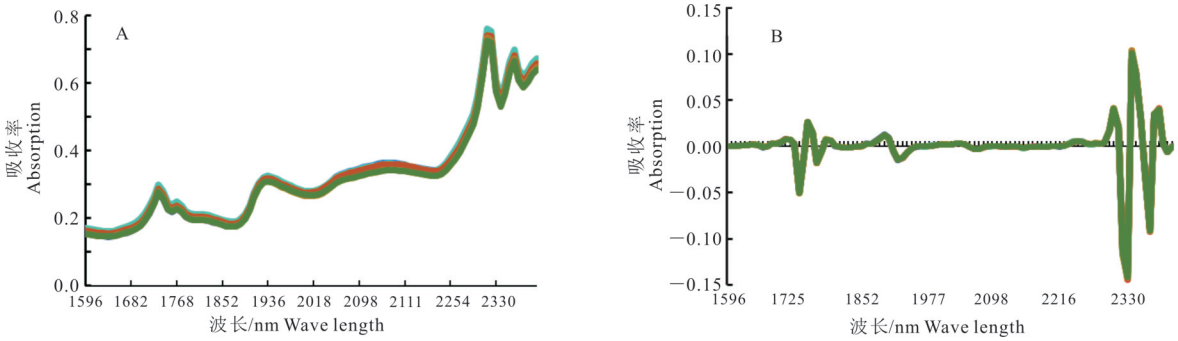
1.3.2　NIRs数据的预处理和分析　借鉴他人和作者之前的研究经验，本研究采用Savitzky-Golay平滑二阶导数方法对NIRs原始数据进行预处理，以消

除表面散射和基线漂移对数据精度的影响。经预处理的数据选择全交互式内部交叉验证算法,建立每个柠檬桉种源的PCA模型。通过簇类独立软模式(SIMCA)判别分析,估算10个种源两两之间的光谱距离[11-12]。该计算将生成一个10×10的欧氏距离矩阵,该矩阵具有对称性,即任意2个种源i和j之间的距离$D(i,j)$等于j到i的距离$D(j,i)$。基于NIRs欧氏距离按全链接算法对柠檬桉按种源构建层级聚类树,直观展示柠檬桉种源间的亲缘关系。基于柠檬桉各种源模型在不同波长处的建模能力(modeling power),分析柠檬桉各种源间的遗传差异信息。采用The Unscrambler x10.4(CAMO, Oslo, Norway)进行NIRs数据的预处理和分析,Excel 2016(Microsoft, USA)用于部分数据整理和作图[8]。

2 结果与分析

2.1 柠檬桉遗传背景分析的NIRs模型参数

为了实现对10个不同来源柠檬桉种源的准确识别,本研究为每个种源分别建立了独立的类别模型。在建模之前,对296条原始光谱数据进行Savitzky-Golay平滑二阶导数预处理,以消除样本物理状态、背景噪声和基线漂移带来的干扰,增强与化学成分相关的特征吸收信息[9,13-14]。如图1所示,柠檬桉种源NIRs原始光谱经预处理后,突出了种源样本遗传差异的特征波段,提示研究可重点关注1 700~1 800,1 900左右和2 300~2 400 nm处的波长及其对应化合物或基团信息。



A. 原始光谱;B. 二阶导数平滑处理后光谱。

A. Original spectrum; B. Spectrum after pretreatment with Savitzky-Golay second derivative smoothing.

图1 柠檬桉种源叶样NIR光谱曲线

Fig. 1 NIR spectra of leaf samples from the *Corymbia citriodora* provenances

NIRs定量预测和定性判别模型的关键参数之一是确定最佳主成分数量。过少的主成分无法充分捕捉该类别的有效信息,导致模型欠拟合;过多的主成分则会引入噪声,导致模型过拟合,降低其对新样本的预测能力。本研究采用留一法交叉验证(leave-one-out cross-validation)来确定每个种源模型的最佳主成分数[9-10]。

表2总结了10个柠檬桉种源PCA模型的关键性能参数。从中可以看出,所有模型的主成分数均为3~5个,这表明各种源内部的光谱共性特征可以用较少的变量(即主因子)高效概括。各模型的累计方差解释率均高达95%以上,说明所选主成分很好地解释了NIR光谱数据的绝大部分变异。同时,所有模型的累计交叉验证方差解释率均超过了0.70,这表明模型具有良好的稳健性和对新样本的预测能力,避免了过拟合[13-14],确保本研究后续分析的可靠性。

柠檬桉在适应不同环境如水分、温度、土壤时,其生长策略和细胞壁构建机制发生了系统性的微调。如不同种源在叶片结构、厚度、机械强度等方面可能存在差异,以适应不同的水分蒸腾和物理胁迫。NIR光谱对这些结构性大分子的状态变化极为敏感。因此,本研究揭示的"化学指纹"(chemical fingerprint)可能是植物长期环境适应性的"生理印记",其较易变的次生代谢物更能稳定地反映种源的遗传信息。

2.2 柠檬桉不同种源间的NIR光谱距离

本研究首先采用无监督的判别分析方法(SIMCA)[11-12,15]计算了10个柠檬桉种源两两之间的NIR光谱欧氏距离,结果如表3所示。从表3数据可以看出,各种源间的光谱距离绝对值均较小,在0.016~0.024之间窄幅波动。例如,'Silver'种源与'SCAU'种源的距离最小(0.016),而'GDHZ'种源与'GXGF'种源的距离最大(0.024)。以上数值上的微小差异似乎不足以支持"种源间存在显著差异"的结论。然而,必须强调的是,这里的欧氏距

离是在一个由数百个波长变量构成的高维光谱空间中计算得出的综合度量。在这个空间中,微小但系统性的偏移,比在少数几个变量上的巨大差异更能反映样本类别的不同本质。它代表了叶片整体“化学指纹”的综合性差异。因此,不能仅凭绝对距离来判断其区分潜力,而需要进一步的分析来揭示这些微小差异背后的结构性信息[16]。

表 2　10 个柠檬桉种源的 PCA 模型性能参数

Table 2　Performance parameters of the PCA models for the ten provenances

种源 Provenances	样本数 Number	最佳主成分数 Optimal principal component (PCs)	累计方差解释率/% Cumulative percentage of variance	累计交叉验证方差解释率 Cumulative cross validation variance explanation
Gordon	49	4	97.2	0.78
Hughenden	40	3	96.5	0.75
Quaid	29	3	95.8	0.71
Silver	33	4	98.1	0.82
Mareeba	54	5	97.6	0.79
YNYM	27	3	96.9	0.76
GXDM	16	3	95.5	0.72
GXGF	17	4	97.0	0.77
GDHZ	15	4	96.2	0.73
SCAU	16	3	97.8	0.81

表 3　柠檬桉不同种源间的 NIRs 欧氏距离

Table 3　NIRs Euclidean distances among different provenances of *Corymbia citriodora*

模型 Models	Gordon	Hughen	Quaid	Silver	Mareeba	GXDM	GXGF	SCAU	GDHZ	YNYM
Gordon		0.021	0.021	0.020	0.021	0.021	0.021	0.021	0.022	0.021
Hughen	0.021		0.022	0.021	0.021	0.022	0.021	0.022	0.022	0.021
Quaid	0.022	0.022		0.021	0.022	0.022	0.022	0.022	0.022	0.022
Silver	0.020	0.020	0.020		0.020	0.020	0.020	0.020	0.021	0.020
Mareeba	0.020	0.020	0.020	0.019		0.020	0.020	0.020	0.020	0.020
GXDM	0.021	0.022	0.022	0.021	0.021		0.021	0.022	0.022	0.021
GXGF	0.023	0.023	0.023	0.023	0.023	0.023		0.023	0.024	0.023
SCAU	0.017	0.017	0.017	0.016	0.017	0.017	0.017		0.017	0.017
GDHZ	0.019	0.020	0.020	0.019	0.020	0.020	0.019	0.020		0.020
YNYM	0.019	0.020	0.020	0.019	0.020	0.020	0.020	0.020	0.020	

注:表中数据为种源 i 和 j 间的 NIRs 欧氏距离,下三角以列种源为模型,上三角以行种源为模型。对角线代表自身距离,为 0,故留空。

Note: Values represent pairwise distances between provenances based on NIR spectra. The lower triangle uses the column provenance as the model, while the upper triangle uses the row provenances. The diagonal represents zero self-distance and is thus omitted.

为了更直观地揭示种源间的相对亲疏关系,对表 3 的原始距离数据进行了标准化处理(Z-score 标准化),将绝对距离转化为相对亲疏关系的度量(表 4)。标准化后的相对距离清晰地揭示了隐藏在原始数据背后的规律:(1)亲缘关系聚类清晰。以国内‘SCAU’种源为参照,其与澳大利亚的‘Silver’和‘Mareeba’种源在相对距离上表现为负值,即显著“更近”,而与国内的‘GXGF’种源、澳大利亚的‘Quaid’种源则表现为较大的正值,即显著“更远”;(2)潜在溯源线索。‘SCAU’种源与‘Silver’种源的相对距离最近,这为“华南农业大学引种材料可能来源于‘Silver’或其相似生态型地区”这一推测提供了初步的数据支持。

2.3　基于 NIR 光谱距离的柠檬桉种源层级聚类

作为无监督的学习方法,层级聚类通过全链接算法,对柠檬桉 10 个种源 NIR 光谱数据进行了聚类分析(图 2)。柠檬桉各种源的 NIRs 层级聚类结果显示,10 个种源清楚地分为两大类,即澳大利亚原生种源聚为一类,国内早期引种的种源聚为一类。可以看出,国外种源的聚类规律基本符合其地理相似性,国内种源的聚类结果则可能有助于早期引种信息的确认。

在国内种源中,‘GXGF’种源单独分类,‘YNYM’和‘GDHZ’、‘SCAU’和‘GXDM’种源分

别聚类,此结论有望为国内早期引种的柠檬桉遗传背景的确认提供参考。国外的原生种源间,'Silver'和'Gordon'种源聚为一类,可能表明两者地理距离更近;'Hughen'、'Quaid'和'Mareeba'种源聚类,其中后2个种源的纬度和海拔更接近(表1),'Hughen'种源的纬度及海拔与两者存在差异,但聚为一类,这可能说明3个种源间的地理距离相近,因此存在一定程度的花粉、种子或传粉媒介等某种形式的基因流。

表4 数据标准化处理后柠檬桉种源间的NIRs相对欧氏距离

Table 4 Relative Euclidean distances based on NIRs among provenances of *Corymbia citriodora* after data standardization

模型 Models	Gordon	Hughen	Quaid	Silver	Mareeba	GXDM	GXGF	SCAU	GDHZ	YNYM
Gordon		0.349	0.403	0.223	0.261	0.363	0.460	0.098	0.281	0.255
Hughen	0.500		0.598	0.438	0.426	0.562	0.782	0.408	0.422	0.444
Quaid	0.813	0.876		0.756	0.759	0.928	1.138	0.925	0.745	0.773
Silver	−0.241	−0.207	−0.167		−0.261	−0.213	−0.147	−0.730	−0.254	−0.286
Mareeba	−0.231	−0.359	−0.349	−0.312		−0.375	−0.244	−0.962	−0.432	−0.439
GXDM	0.462	0.545	0.627	0.414	0.431		0.880	0.378	0.404	0.420
GXGF	1.560	1.550	1.609	1.505	1.447	1.475		1.852	1.425	1.435
SCAU	−1.986	−1.987	−2.027	−2.042	−2.085	−2.017	−2.189		−2.128	−2.090
GDHZ	−0.457	−0.389	−0.338	−0.503	−0.497	−0.356	−0.381	−0.988		−0.511
YNYM	−0.419	−0.378	−0.355	−0.479	−0.481	−0.368	−0.300	−0.980	−0.463	

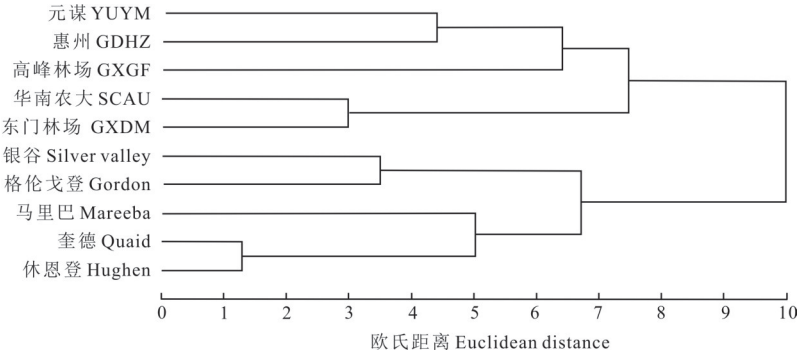


图2 基于叶样NIRs欧氏距离的柠檬桉种源层级聚类

Fig. 2 Hierarchical clustering of *Corymbia citriodora* provenances based on Euclidean distances of foliage NIRs

另外,澳大利亚原生种源组内部的聚类关系也提供了一些有趣的生态学线索。例如,'Hughen'种源在地理纬度和海拔上与同组的'Quaid'和'Mareeba'种源存在一定差异,但依然聚为一类。这可能暗示在原生环境下,尽管存在地理区隔,但种源间仍可能通过花粉、种子等途径发生一定程度的基因流,从而维持了表型相似性。研究表明环境压力是塑造柠檬桉基因组多样性的关键驱动力,相似的环境适应策略可能导致不同地理来源的种源表现出趋同的化学表型[12]。因此,NIR光谱聚类结果不仅反映了地理隔离,也可能间接揭示了柠檬桉种群进化过程中发生了某种形式的基因流或基因隔离事件。

2.4 柠檬桉国内种源间的判别分析

表型遗传变异和分子生物学的遗传亲缘关系分析,常用于解决遗传材料信息的错误和丢失问题。鉴于NIR光谱信息的灵敏性,本研究试图通过判别分析,探索其分析柠檬桉遗传背景的可行性。上述柠檬桉各种源间NIR光谱距离的分析显示,各种源距'SCAU'种源PCA模型间的距离最小,因此,就以'SCAU'种源作为参照,验证国内另外几个种源与其样本间的区分效果。图3是用'SCAU'家系样本建立的PCA模型分别对'GXDM'和'GDHZ'种源样本的判别散点图。从图3可以看出,尽管三者之间的NIR光谱距离差别不是很明显,但2个种源都能与'SCAU'种源清楚地区分开来。依同样的方法,'GXGF'、'YNYM'种源也可与'SCAU'种源清楚地区分,表明NIRs信息可以用于柠檬桉遗传背景的分析。为避免冗余,这里省去其他种源间的判别散点图。

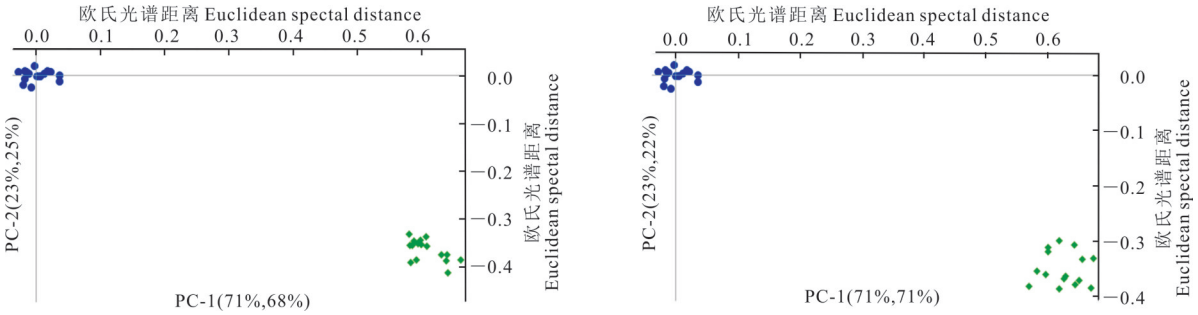


图 3　柠檬桉‘SCAU’种源模型（蓝色）对‘GXDM’（绿色，左）和‘GDHZ’（绿色，右）种源的判别

Fig. 3　Discrimination between provenance ‘SCAU’ (blue) and provenances ‘GXDM’ (green, left) or ‘GDHZ’ (green, right) using the PCA model for *Corymbia citriodora*

对于林木育种者来说，图 3 中还有一个信息值得解读，即‘SCAU’种源的家系样本相对集中，‘GXDM’种源样本次之，‘GDHZ’种源样本的因子得分最为分散。扣除取样误差因素，从遗传角度讲，这种离散程度则很有可能反映了种源内家系间的变异程度，集中表示遗传变异小，反之则遗传变异大。通过 PCA 散点图离散度来评估种源内遗传变异的方法，为林木育种实践提供了一种创新的快速评价工具。传统上，评估遗传多样性需要进行耗时费力的分子水平实验，而 NIRs 分析几乎可以“即时”地提供群体的表型多样性信息，极大地加速了育种决策过程，如 MIRANDA 等[17]关于桉树木材性状的表型分析就证明了这一点。

2.5　柠檬桉种源样本的建模能力

在 NIRs 分析中，建模能力（modeling power）是指特定波长对模型的贡献大小，其值越接近 1，表明该点波长对模型越重要。一般来说，建模能力小于 0.3 的波长对模型不重要。从遗传育种的关注点来说，希望借助对建模能力大于 0.3 的波长的分析，尝试探讨柠檬桉不同种源的遗传变异特征及种源间的差异。

图 4 是柠檬桉 10 个种源 NIRs 全波段（1 600～2 400 nm）建模能力的示意图。从图 4 可以看出，在 1 708～1 785，1 880～1 935 及 2 292～2 370 nm 波段处，10 个种源 NIRs 的建模能力都很强，而这几个波段对应的有机物或化学键可能正是柠檬桉遗传变异的关键物质。结合参考文献[13,18]可知，1 708～1 785 nm 波段为木质素、半纤维素、纤维素、—CH_2 等的特征吸收谱带；1 880～1 935 nm 波段为淀粉、羧基（—COOH）、羟基（—OH）、酰亚胺基（—CONH）、半纤维素和水等的特征吸收谱带；2 292～2 370 nm 波段主要为纤维素、半纤维素、—CH_2 等的特征吸收谱带。另外，‘Gordon’和‘Silver’种源在 1 604～1 630，2 082 和 2 160 nm 波段附近的 NIRs 建模能力较其他种源强，文献[18]显示，前 2 个波段都是纤维素的特征吸收谱带；2160 nm 波段是乙酰基的特征吸收谱带。可以说上述这些物质和化学键含量的差异或状态是柠檬桉种源差异的物质基础。同时，这些特征物质的功能基因有望在探究对比基因型的差异表达基因时提供参考。

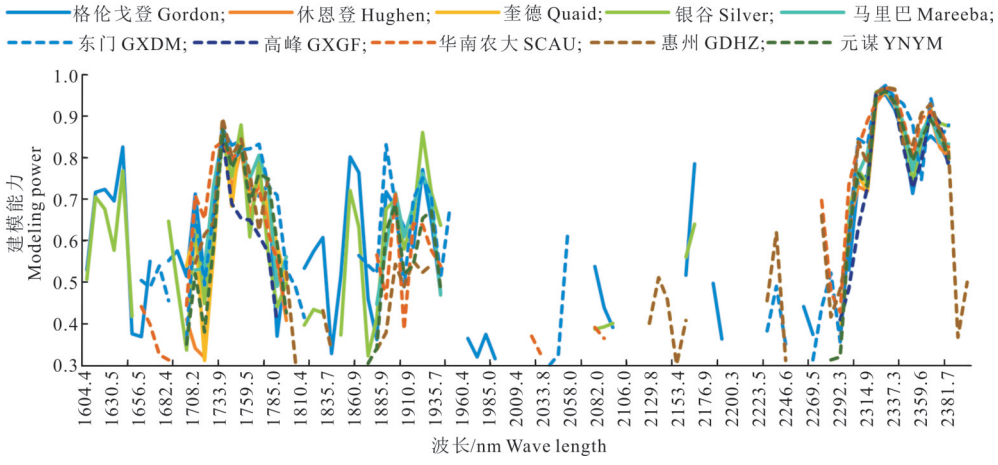


图 4　柠檬桉 10 个种源的 NIRs 全波段建模能力

Fig. 4　Modeling power of the full NIRs for ten provenances of *Corymbia citriodora*

3 讨 论

NIRs本质上是一种综合的表型信息,其与植物遗传信息的关系可以通过分析光谱距离与遗传距离的关系得到验证。如刘旭等[19]研究显示,玉米的遗传距离与其NIR光谱距离间的相关系数超过0.9。作者团队关于粗皮桉NIRs差异与其遗传差异关系的研究也表明,其种源间的遗传距离与光谱距离呈明显正相关[10]。本研究基于NIR光谱距离对柠檬桉种源进行分类,尤其是种源间的准确判别也间接证明NIRs信息能准确反映其遗传信息,但两者间具体的相关性有待后续基于SNP标记信息(已测试)进一步明确。

借助NIRs信息开展林木遗传基础信息研究的依据在于NIRs信息能准确反映其遗传信息,或与其目标性状间存在密切的关联。通过NIRs信息开展基因型定性判别和性状定量预测的研究屡见不鲜。作者前期将NIRs技术用于桉树耐旱性状[9]、木材化学成分[20]检测的研究,结果表明该技术具备理想的可靠性。LUAN等[21]关于湿地松的研究也表明,NIRs可准确用于不同坡向湿地松松脂含量的预测(R^2为0.82~0.85)。本研究则是基于柠檬桉种源NIRs数据,从光谱距离、种源区分与聚类,以及建模能力等定性判别角度证明了NIRs信息能真实反映其遗传信息,说明借助NIRs技术开展柠檬桉种源遗传背景分析是可行的。

4 结 论

在利用基因组学分析林木遗传基础的同时,本研究尝试借助NIRs灵敏、快速的分析优势,探索其用于柠檬桉种源区分的可行性和准确性。结果表明,虽然柠檬桉国内各种源与国外种源间光谱距离的差异微小,但基于NIRs的层级聚类显示,澳大利亚原生种源聚为一类,国内引种种源聚为一类,且国外种源的聚类基本与其地理距离一致;基于SIMCA的判别分析也成功地将各种源的家系样本进行了准确分类;聚类特例则说明种源间可能存在某种形式的基因流,同时一定程度上表明国内引种的柠檬桉种质未经过世代传递发生基因重组。对于柠檬桉遗传信息来说,不同波段建模能力的分析则可从化合物和化学键含量或状态的差异角度解析柠檬桉种源差异的内在物质基础。本研究结果证明,可以综合柠檬桉的NIR光谱信息与其生物学信息的分析结果,溯源国内早期引种柠檬桉的遗传背景信息。

[参考文献]

[1] 田玉红,刘雄民,周永红,等. 柠檬桉叶挥发性成分的提取及成分分析[J]. 色谱,2005,23(6):4.
TIAN Y H,LIU X M,ZHOU Y H,et al. Extraction and determination of volatile constituents in leaves of *Eucalyptus citriodora*[J]. Chinese Journal of Chromatography,2005,23(6):4

[2] TOLBA H,MOGHRANI H,BENELMOUFFOK A,et al. Essential oil of Algerian *Eucalyptus citriodora*:Chemical composition,antifungal activity[J]. Journal of Medical Mycology,2015,25(4):e128.

[3] HEALEY A L,SHEPHERD M,KING G J,et al. Pests,diseases,and aridity have shaped the genome of *Corymbia citriodora*[J]. Communications Biology,2021,4:537.

[4] ANSARI P,CHOUDHURY S T,ABDEL-WAHAB Y H A. Insulin secretory actions of ethanol extract of *Eucalyptus citriodora* leaf,including plasma DPP-IV and GLP-1 levels in high-fat-fed rats,as well as characterization of biologically effective phytoconstituents[J]. Metabolites,2022,12(8):757.

[5] DA COSTA M M,NOGUEIRA T A P C,DE CARVALHO BITTENCOURT R,et al. Assessment of industrial performance for market pulp production between *Eucalyptus* and *Corymbia* hybrids clones[J]. Revista Árvore,2022,46:e4628.

[6] MICHAEL M. 中山大学与桉树[DB/OL]. (2025-10-15)[2025-10-29]. https://mp.weixin.qq.com/s/KfMGWz1BnhXzDCndCnOYSg.

[7] 刘思汝,林彦,刘晓华,等. 柠檬桉群体遗传多样性和遗传结构的SSR分析[J]. 分子植物育种,2016,14(7):1923-1929.
LIU S R,LIN Y,LIU X H,et al. Genetic diversity and genetic structure of *Corymbia citriodora* based on SSR analysis[J]. Molecular Plant Breeding,2016,14(7):1923-1929.

[8] HAYES R A,NAHRUNG H F,LEE D J. Consequences of *Corymbia* (Myrtaceae) hybridisation on leaf-oil profiles[J]. Australian Journal of Botany,2013,61(1):52-59.

[9] 卢万鸿,杨桂丽,林彦,等. 桉树叶片水分状态的近红外光谱检测[J]. 林业科学,2017,53(5):16-22.
LU W H,YANG G L,LIN Y,et al. Assessing leaf water status of *Eucalyptus* using NIRS[J]. Scientia Silvae Sinicae,2017,53(5):16-22.

[10] 王楚彪,杨艳,白卫国,等. 粗皮桉近红外光谱差异与其遗传差异间的关系[J]. 光谱学与光谱分析,2021,41(11):3399-3404.
WANG C B,YANG Y,BAI W G,et al. The relationship between genetic variations and NIRs differences of *Eucalyptus pellita* provenances[J]. Spectroscopy and Spectral Analysis,2021,41(11):3399-3404.

[11] LU W H,ARNOLD R J,WANG C B,et al. Defining *Eucalyptus urophylla* with its hybrid and the rules of genetic recombination using near infrared spectroscopy[J]. Journal of Near Infrared Spectroscopy,2021,29(4):236-244.

(下转第77页)