

DOI: 10.3724/SP.J.1006.2026.54108

紫肉甘薯品质性状的近红外光谱预测模型构建与优化

蒋嘉卉^{1,2} 江炳志¹ 刘冠明² 王章英^{1,*} 唐朝臣^{1,*}

¹ 广东省农业科学院作物研究所 / 广东省农作物遗传改良重点实验室, 广东广州 510640; ² 仲恺农业工程学院农业与生物学院, 广东广州 510225

摘要: 紫肉甘薯在品质评价上缺乏精确高效的方法, 已成为其种质资源精准鉴定与遗传改良的关键瓶颈。为此, 本研究基于近红外光谱(NIRS)技术, 针对紫肉甘薯的主要品质性状(包括总淀粉、粗蛋白、还原糖、总黄酮、总酚和花青素), 开展了高通量预测模型的构建与优化。本研究选取 150 份代表性样品, 结合双重优化策略与机器学习算法, 成功建立并优化了 6 个高性能预测模型。结果显示, 校正集决定系数(R^2_c)为 0.936~0.992, 交叉验证决定系数(R^2_{cv})为 0.918~0.987, 验证集决定系数(R^2_v)为 0.903~0.987, 预测偏差比(RPD)为 6.55~19.80, 范围误差比(RER)为 21.9~63.4, 这些结果表明, 所构建的模型具有良好的稳定性与可靠的预测能力。这些预测模型为紫肉甘薯品质的高通量评价提供了高效、可行的技术路径, 能够促进营养成分的定量评估和优质种质的高效筛选, 从而为提升紫肉甘薯品质和推动产业可持续发展提供核心驱动力。

关键词: 紫肉甘薯; 近红外光谱; 品质性状; 预测模型; 机器学习算法

Establishment and optimization of near-infrared spectroscopy models for quality traits of purple-fleshed sweet potato

Jiang Jia-Hui^{1,2}, Jiang Bing-Zhi¹, Liu Guan-Ming², Wang Zhang-Ying^{1,*}, and Tang Chao-Chen^{1,*}

¹ Crop Research Institute, Guangdong Academy of Agricultural Sciences / Guangdong Key Laboratory of Crop Genetics and Improvement, Guangzhou 510640, Guangdong, China; ² College of Agriculture & Biology, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, Guangdong, China

Abstract: The lack of precise and efficient methods for evaluating the quality of purple-fleshed sweet potato has become a major bottleneck in accurate germplasm identification and genetic improvement. To address this challenge, this study employed near-infrared spectroscopy (NIRS) to construct and optimize high-throughput predictive models for key quality traits of purple-fleshed sweet potato, including total starch, crude protein, reducing sugar, total flavonoid content, total phenolic content and total anthocyanin content. A total of 150 representative samples were selected, and six high-performance predictive models were successfully developed and optimized using a dual optimization strategy combined with machine learning algorithms. The models achieved high accuracy, with coefficients of determination (R^2_c) for calibration ranging from 0.936 to 0.992, cross-validation (R^2_{cv}) from 0.918 to 0.987, and external validation (R^2_v) from 0.903 to 0.987. The ratio of prediction deviation (RPD) ranged from 6.55 to 19.80, and the range error ratio (RER) ranged from 21.9 to 63.4, indicating strong model stability and prediction capability. The predictive models developed in this study offer an efficient and practical approach for the high-throughput analysis of purple-fleshed sweet potato quality. They enable quantitative evaluation of nutritional components and facilitate the screening of superior germplasm, thereby providing important technical support for quality improvement and the sustainable development of the purple-fleshed sweet potato industry.

Keywords: purple-fleshed sweet potato; near-infrared spectroscopy; quality trait; predictive model; machine learning algorithm

本研究由广东省基础与应用基础研究基金项目(2024A1515012158), 财政部和农业农村部国家现代农业产业技术体系建设专项(CARS-10)和广州市科技计划项目(2025B03J0007)资助。

This study was supported by the Guangdong Basic and Applied Basic Research Fund (2024A1515012158), the China Agriculture Research System of MOF and MARA (CARS-10), and the Guangzhou Science and Technology Project (2025B03J0007).

* 通信作者(Corresponding authors): 唐朝臣, E-mail: tangchaochen1988@163.com; 王章英, E-mail: wangzhangying@gdaas.cn

第一作者联系方式: E-mail: JanetChiang1998@163.com

Received (收稿日期): 2025-09-08; Accepted (接受日期): 2026-01-22; Published online (网络出版日期): 2026-02-04.

URL: <https://link.cnki.net/urlid/11.1809.S.20260204.1731.017>

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

紫肉甘薯 [*Ipomoea batatas* (L.) Lam.] 属于旋花科一年生草本植物, 其块根肉色呈紫色至深紫色, 富含花青素、淀粉、蛋白质、多酚等多种营养成分, 兼具粮食、工业和药用价值, 是具有重要开发潜力的特色作物^[1-2]。随着营养导向型育种理念的不断发展, 甘薯育种已逐渐由单纯追求高产向兼顾营养与品质提升转变^[3]。紫肉甘薯因其独特的营养成分和功能特性, 已成为甘薯育种研究的重要方向之一^[4]。

近年来, 紫肉甘薯在种质资源创新和栽培管理方面取得了显著进展, 但其主要品质性状的高效评价体系尚不完善^[5]。传统检测方法在测定总淀粉、粗蛋白、还原糖、总黄酮、总酚和花青素等指标时, 普遍存在检测周期长、成本高、样品损耗大及难以实现多指标同步分析等不足, 严重制约了紫肉甘薯的优质种质筛选与遗传改良^[6]。因此, 建立一种高通量、精准的品质检测方法, 对于实现紫肉甘薯品质的标准化评价和精准育种具有重要意义^[7-8]。

近红外光谱(near-infrared spectroscopy, NIRS)技术是一种基于含氢基团(C-H、N-H、O-H)振动吸收特征的非破坏性快速检测方法, 可在 780~2526 nm 波段内获取样品的丰富化学信息^[9-10]。该技术已广泛应用于农产品品质检测, 并因其能够实现多指标并行分析而备受关注。在 NIRS 建模过程中, 特征波长的有效筛选及化学计量学算法的优化协同, 是提升模型预测精度与稳健性的关键^[11]。截至目前, NIRS 结合先进算法已成功应用于水稻^[12]、小麦^[13]、高粱^[14]和谷子^[15]等作物的品质分析, 显著提高了检测效率与准确性。

在甘薯研究中, NIRS 技术同样展现出广泛的应用潜力。早期研究多集中于单一品质性状建模, 例如, 针对块根淀粉含量^[16-18]、蛋白质含量^[19]及糖类成分^[20-21]的预测模型, 验证集决定系数(coefficient determination of validation, R^2_v)达到 0.734~0.980; 部分研究还构建了紫肉甘薯花青素含量的预测模型, R^2_v 值在 0.847 至 0.888 之间^[22-23]。随着技术的进步, 研究逐渐转向多指标的同步分析。Lebot 等^[24]构建了淀粉、糖类和蛋白质 3 项指标的预测模型, R^2_v 为 0.710~0.870, 预测偏差比(ratio of prediction to deviation, RPD)为 2.11~2.93; Tang 等^[8]建立了总黄酮和总酚等 8 项指标的定量分析模型($R^2_v = 0.890\sim0.960$, RPD = 6.33~11.40)。Tang 等^[25]将机器学习算法与 NIRS 深度结合, 实现了不同处理状态下的 4 种糖分高通量预测($R^2_v = 0.898\sim0.971$, RPD = 5.83~10.30)。

尽管如此, 目前针对紫肉甘薯的多品质性状高通量预测体系仍存在研究空白, 特别是在特征波长筛选优化、模型稳健性提升和适应性拓展等方面, 仍需深入探索。这些不足在一定程度上限制了紫肉甘薯育种效率及品质评价的标准化进程。

基于此, 本研究以 150 份具有代表性的紫肉甘薯样品为对象, 围绕总淀粉、粗蛋白、还原糖、总黄酮、总酚和花青素 6 项核心品质性状, 构建并优化近红外高通量预测模型。通过整合先进化学计量学方法, 优化特征波长筛选与建模策略, 提升模型精度与稳健性。本研究旨在为紫肉甘薯多指标的同步检测提供可行的技术路径, 并为其种质资源的高效评价和遗传改良提供理论支撑与技术保障。

1 材料与方法

1.1 样品采集与制备

为了构建能够充分覆盖目标性状变异范围并适用于高稳健性 NIRS 模型开发的样品集, 本研究选取了 150 份具有显著基因型和表型差异的紫肉甘薯样品(附表 1)。所有样品均采自广东省农业科学院白云试验基地(23°23'N, 113°26'E), 来源于 2020 年至 2024 年间开展的甘薯种质资源鉴定、优新品种筛选及栽培试验。所选样品在块根肉色上展现了丰富的表型多样性, 包括深紫型、紫色型、浅紫型、紫红型及紫斑型等典型类型。

样品在种植 120~150 d 后人工采收, 每份样品随机选取 5 个形态完整、大小适中的新鲜块根(100~250 g), 并立即送实验室处理。块根经流动自来水冲洗、吸水纸擦干后切丝。前期研究表明, 热风干燥(如 75℃)易导致生物活性成分降解以及糖分含量降低^[26], 因此本研究采用冷冻干燥法。具体操作如下: 使用台式冷冻干燥机(Christ Beta 2-8 LD plus, 德国)在 -50℃ 条件下冻干 72 h, 随后用液氮研磨机(A10 basic, IKA, 德国)粉碎, 并过 40~80 目筛网(孔径 0.45~0.20 mm)以获取均一粉末, 用于后续化学分析与光谱采集。

1.2 化学成分测定

采用分光光度法(AOAC 方法 996.11)^[27]测定总淀粉含量, 使用 Megazyme 检测试剂盒(K-TSTA-100A, Wicklow, Ireland)在 510 nm 波长下测定吸光度。采用凯氏定氮法(Kjeldahl method)测定粗蛋白含量^[28], 其中, 氮-蛋白质换算系数为 6.25。操作遵循农业行业标准 NY/T 2017-2011《植物中氮的测定》。

采用 3,5-二硝基水杨酸比色法测定还原糖含量^[29], 于 540 nm 波长下测定吸光度。采用硝酸铝盐比色法测定总黄酮含量^[30], 0.1 g 样品经 10 mL 乙醇超声提取(1844×g 离心 10 min), 取上清液与亚硝酸钠-硝酸铝-氢氧化钠溶液梯度显色, 于 510 nm 波长下测定吸光度, 以芦丁当量(mg g⁻¹ 干重)表示(标准曲线 $R^2 = 0.999$)。采用 Folin-Ciocalteu 比色法测定总酚含量^[31], 0.1 g 样品经 10 mL 乙醇超声提取(1180×g 离心 10 min), 取上清液与 Folin-Ciocalteu 试剂-碳酸钠溶液显色, 避光条件下反应后于 760 nm 波长下测定吸光度, 以没食子酸当量(mg g⁻¹ 干重)表示(标准曲线 $R^2 = 0.997$)。采用 pH 差值法测定花青素含量^[32], 1.0 g 样品经 60 mL 盐酸乙醇溶液超声提取(1844×g 离心 10 min), 用 pH 1.0 和 4.5 缓冲液稀释, 避光平衡后分别于 530 nm 和 700 nm 波长下测定吸光度, 以矢车菊素-3-葡萄糖苷当量(mg g⁻¹ 干重)表示。

所有化学分析试验均进行 3 次生物学重复, 结果以干基表示。所有特定波长的吸光度测定均使用分光光度计(DU800, Beckman Coulter)完成。

1.3 光谱数据采集与预处理

使用 FT-9700 NIR 光谱仪(PerkinElmer, 美国)采集光谱数据。具体操作如下: 称取 5 g 样品置于样品杯中, 在 25℃ 恒温条件下以反射模式进行扫描。光谱范围覆盖 4000~10000 cm⁻¹, 分辨率设定为 8 cm⁻¹, 每个样品扫描 3 次, 每次累积 64 次扫描, 最终取 3 次扫描的平均光谱用于后续分析。光谱数据以 log (1/R) 记录, 其中 R 为样品的反射率。

使用 Unscrambler v10.4 软件(CAMO, 美国)对原始光谱预数据进行主成分分析(principal component analysis, PCA), 以剔除异常值。结合前期预建模结果^[8], 采用组合式预处理策略以优化模型性能, 包括标准正态变换(standard normal variate, SNV)校正散射效应, 以及 Savitzky-Golay (SG)平滑和二阶导数(2nd derivative)处理, 以增强特征峰的分辨率并稳定光谱基线^[33]。使用 Chem Data Solution v3.5 软件(Dalian ChemData Solution Technology Co. Ltd., 中国大连)进行上述预处理。

1.4 模型建立与性能评价

本研究采用双重优化策略结合机器学习算法, 通过系统比较与筛选, 确定最优的样本集划分方法以及基于机器学习的关键特征变量筛选算法, 从而构建预测模型。

在样本集划分方面, 比较了随机划分、光谱-理

化值共生距离(sample set partitioning based on joint X-Y distances, SPXY)算法和 Kennard-Stone (KS)算法 3 种策略, 结果表明, KS 算法具有最佳性能^[33-34]。按照 4 : 1 的比例将样本划分为校正集与验证集, 其中, 校正集用于建立模型, 并通过留一交叉验证法(leave-one-out cross-validation, LOOCV)进行内部交叉验证, 同时利用验证集中的独立样本进行外部验证。

为了去除冗余信息并提高预测精度, 本研究比较了 6 种特征筛选方法, 包括选择比率法(selectivity ratio, SR)、变量重要性投影(variable importance for projection, VIP)、无信息变量消除(uninformative variable elimination, UVE)、蒙特卡洛无信息变量消除(Monte Carlo uninformative variable elimination, MC-UVE)、竞争性自适应权重(competitive adaptive reweighted sampling, CARS)和随机蛙跳(random frog, RF)。最终, 采用 CARS 与 RF 这 2 种机器学习算法从全光谱中提取关键波长变量。CARS 算法基于“适者生存”原则, 通过竞争性自适应权重筛选出最优变量子集; RF 算法则是一种可逆跳跃马尔可夫链蒙特卡洛方法, 通过随机迭代跳跃增强变量选择的鲁棒性^[35-36]。利用筛选出的变量, 最终耦合 PLSR 算法建立预测模型^[33-34]。

模型性能的综合评价指标包括: 校正集决定系数(coefficient determination of calibration, R^2_c)、交叉验证决定系数(coefficient determination of cross-validation, R^2_{cv})、验证集决定系数(R^2_v)、校正标准偏差(root mean square error of calibration, RMSEC)、交叉验证标准偏差(root mean square error of cross-validation, RMSECV)和预测标准偏差(root mean square error of prediction, RMSEP)。此外, 通过 RPD 和范围误差比(range error ratio, RER)评估模型在种质评估与育种中的实用性^[33,37]。在实际应用中, 当 R^2_c 、 R^2_{cv} 和 R^2_v 值均高于 0.90 且 RPD 值大于 3 时, 模型被视为优秀模型。具体公式为:

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (1)$$

式中, y_i 是实际值, $\hat{y}_i$ 是预测值, $\bar{y}$ 是实际值的平均值。

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (2)$$

式中, n 是样本数量。

$$SEP = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (3)$$

$$RPD = \frac{SD}{SEP} \quad (4)$$

式中, SD 是标准差, SEP 是预测标准误差。

$$RER = \frac{\text{Max} - \text{Min}}{SEP} \quad (5)$$

式中, Max 是实际值的最大值, Min 是实际值的最小值。

1.5 统计分析

采用 Excel 2024 计算最大值、最小值、平均值、标准差、变异系数(coefficient of variation, CV)和遗传多样性指数(Shannon-Weiner diversity index, H')。使用 SPSS 29.0 软件(IBM SPSS Inc.)进行 t 检验, 以 $P < 0.05$ 为显著性水平, 评估校正集与验证集在品质性状上的差异。

2 结果与分析

2.1 样品的近红外光谱特征

为了探讨紫肉甘薯样品的光谱特征与品质性状之间的关系, 本研究使用近红外光谱仪对样品进行扫描, 并绘制光谱曲线图(图 1-a)。横轴为光谱波长(cm^{-1}), 纵轴为吸光度。结果表明, 所有样品的光谱曲线均处于预期范围内, 但吸光度存在明显差异, 反映出紫肉甘薯样品在光谱特征上的显著差异性, 间接验证了其理化性质的多样性。

在 $4000\sim 7000\text{ cm}^{-1}$ 范围内观察到多个明显的吸

收峰, 特别是在约 4016 、 $4320\sim 4376$ 、 4760 、 5176 、 5624 、 6408 和 6856 cm^{-1} 等波段, 这些吸收带表现出较强的异质性。这些特征峰主要源于 C-H、C-C、N-H、O-H 等振动模式的倍频与合频吸收的重叠。这表明, 近红外光谱对紫肉甘薯的主要品质性状具有较高的敏感性和分辨能力, 能够有效反映其理化成分的差异。

为了进一步提取数据的主要信息并降低数据维度, 本研究采用主成分分析(PCA)对原始光谱进行处理(图 1-b)。结果显示, 校正集与验证集的样品在二维主成分空间内均匀分布, 未观察到明显的聚类或分离现象。PC1 和 PC2 的累计方差贡献率为 94%, 说明绝大部分光谱信息已被有效提取, 且样品之间的光谱差异在主成分空间中得到了良好的展现。这为后续建立具有普适性和较高预测能力的光谱模型提供了坚实的基础。

2.2 样品化学成分多样化

为了分析紫肉甘薯的品质性状分布特征, 本研究绘制了 150 份代表性样品的 6 种主要化学成分的箱线图(图 2)。结果显示, 各项品质性状之间存在显著差异, 且整体接近正态分布, 反映出所选样品在化学组成上具有较高的遗传多样性。

具体而言, 总淀粉含量范围为 $34.8\sim 59.0\text{ g }100\text{ g}^{-1}$, 粗蛋白含量为 $2.13\sim 9.50\text{ g }100\text{ g}^{-1}$, 还原糖含量为 $0.20\sim 6.99\text{ g }100\text{ g}^{-1}$, 总黄酮含量为 $0.72\sim 42.30\text{ mg g}^{-1}$, 总酚含量为 $0.37\sim 25.10\text{ mg g}^{-1}$, 花青素含量为 $0.00\sim$

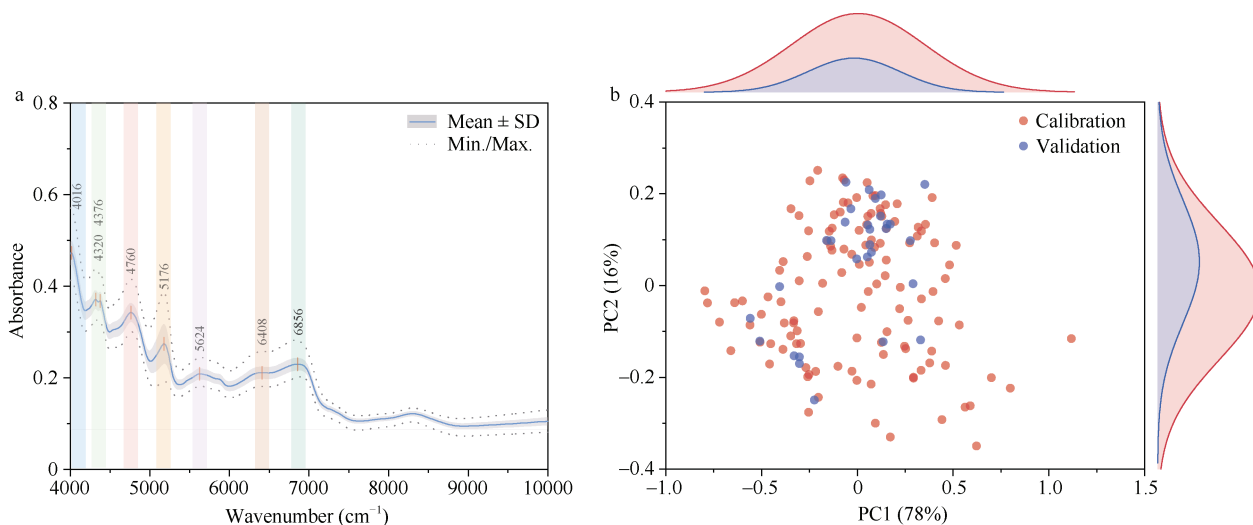


图1 紫肉甘薯样品的近红外光谱变化(a)和 PCA 分数(b)

Fig. 1 Variability of near-infrared spectra (a) and PCA scores (b) in purple-fleshed sweet potato samples

Mean $\pm$ SD: 平均值 $\pm$ 标准差; Min./Max.: 最小值/最大值。

Mean $\pm$ SD: mean $\pm$ standard deviation; Min./Max.: minimum/maximum.

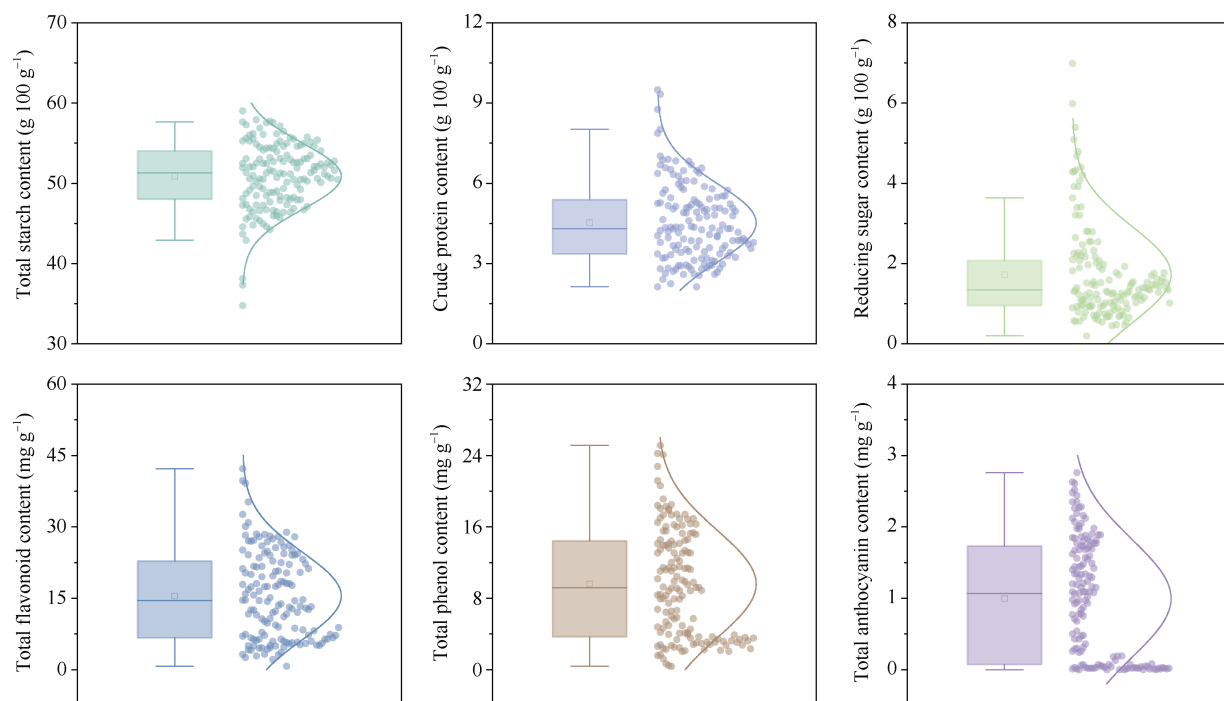


图 2 紫肉甘薯样品化学成分的多样性

Fig. 2 Variations in chemical composition in purple-fleshed sweet potato samples

2.63 mg g⁻¹。各成分的平均含量分别为 50.8 g 100 g⁻¹、4.51 g 100 g⁻¹、1.78 g 100 g⁻¹、15.3 mg g⁻¹、9.35 mg g⁻¹和 0.975 mg g⁻¹。

变异系数(CV)分析表明,花青素(86.60%)、还原糖(69.66%)、总酚(66.52%)和总黄酮(62.33%)的变异度较高,而总淀粉(8.43%)和粗蛋白(33.48%)相对较低;整体变异系数平均为 54.50%。Shannon-Weiner 多样性指数(H')依次为总淀粉(5.01)、粗蛋白(4.96)、总黄酮(4.83)、还原糖(4.81)、总酚(4.80)、花青素(4.59),平均值为 4.83。以上结果表明,紫肉甘薯样品在主要品质性状上表现出广泛的表型多样性,为 NIRS 建模提供了有力的数据支撑。

2.3 样本子集的划分

为了筛选具有代表性的样本并提高模型的稳健性,本研究采用 KS 算法对样本进行子集划分。统计分析结果表明,校正集与验证集在各项品质性状上的差异均未达到显著性水平($P > 0.05$),说明划分结果具有良好的均衡性和代表性(表 1)。2 组数据在最大值、最小值、平均值、标准差及 CV 值等统计指标上高度一致,分布特征相似。

根据 CV 值分析,花青素在校正集(86.60%)与验证集(70.64%)中的变异度最高,其次为还原糖(69.66%和 67.33%)、总酚(66.52%和 48.96%)、总黄

酮(62.33%和 47.60%)、粗蛋白(33.48%和 28.01%)和总淀粉(8.43%和 6.85%)。总体而言,验证集的 CV 值略低于校正集,但差异均在可接受范围内。

PCA 得分图进一步验证了这一结果(图 1-b)。2 个子集在主成分空间内呈现均匀混合分布,PC1 与 PC2 的累计方差贡献率为 94%,且未观察到明显的样本分离趋势,表明样本划分不仅保持了总体方差结构,而且确保了样本的分布一致性。因此,KS 算法能够有效地优化样本划分,为后续多元建模提供了可靠基础。

2.4 基于 CARS 与 RF 的光谱变量优化

为了提升模型的稳定性与预测性能,本研究分别采用竞争性自适应权重(CARS)与随机蛙跳(RF)算法对紫肉甘薯样品的光谱变量进行筛选,旨在减少高共线性和信息冗余,并降低模型过拟合的风险。CARS 与 RF 算法分别从 751 个光谱变量中筛选出不同规模的特征集,反映了 2 种算法在特征选择策略和理论依据上的差异(图 3)。CARS 筛选的特征波段数在 17 个(总酚)至 44 个(总淀粉)之间,而 RF 筛选的波段数则从 45 个(粗蛋白)至 112 个(花青素)不等。与 RF 相比, CARS 更注重特征精简和模型泛化能力的平衡,能够有效去除冗余信息;而 RF 算法则倾向于保留更多的变量,以确保信息的全面覆盖。

表 1 紫肉甘薯样品校正集和验证集化学成分的参考值
Table 1 Reference values for chemical components in purple-fleshed sweet potato samples within calibration and validation sets

参数 Parameter	校正集 Calibration set						验证集 Validation set					
	样本数 <i>n</i>	最小值 Mix.	最大值 Max.	平均值 Mean	标准差 SD	变异系数 CV (%)	样本数 <i>n</i>	最小值 Mix.	最大值 Max.	平均值 Mean	标准差 SD	变异系数 CV (%)
总淀粉 TS (g 100 g ⁻¹)	120	34.78	59.01	50.75	4.28	8.43	30	45.95	57.73	51.38	3.52	6.85
粗蛋白 CP (g 100 g ⁻¹)	120	2.13	9.50	4.51	1.51	33.48	30	2.13	7.01	4.57	1.28	28.01
还原糖 RS (g 100 g ⁻¹)	120	0.20	6.99	1.78	1.24	69.66	30	0.59	5.39	1.50	1.01	67.33
总黄酮 TFC (mg g ⁻¹)	120	0.72	42.25	15.29	9.53	62.33	30	5.40	28.87	16.22	7.72	47.60
总酚 TPC (mg g ⁻¹)	120	0.37	25.13	9.35	6.22	66.52	30	2.28	18.48	10.55	5.06	48.96
花青素 TAC (mg g ⁻¹)	120	0.00	2.63	0.97	0.84	86.60	30	0.00	2.76	1.09	0.77	70.64

TS: total starch; CP: crude protein; RS: reducing sugar; TFC: total flavonoid content; TPC: total phenolic content; TAC: total anthocyanin content; *n*: sample number; Min.: minimum; Max.: maximum; SD: standard deviation; CV: coefficient of variation.

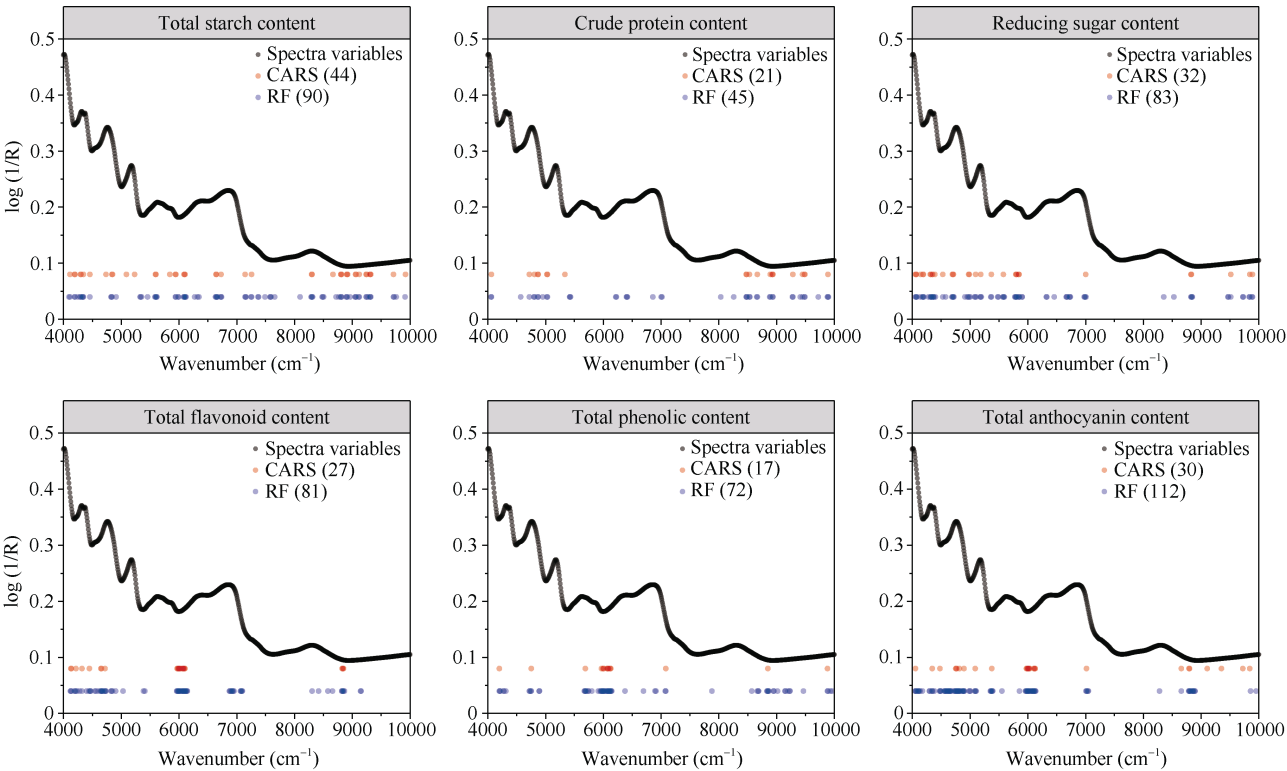


图 3 通过竞争性自适应权重(CARS)和随机蛙跳(RF)优化样品光谱变量选择
Fig. 3 Optimal variable selection for calibration sets using competitive adaptive reweighted sampling (CARS) and random frog (RF) algorithms
括号内数字为对应算法筛选出的特征变量数。Numbers in parentheses indicate spectral variables selected by each algorithm.

总体而言, 尽管近红外光谱数据存在较强的共线性, 通过这 2 种优化策略, 能够从全光谱中提取出与 C-H、C-C、N-H、O-H 等振动模式的倍频与合频吸收相关的特征波段子集, 从而显著提升 PLSR 模型的稳定性和预测精度。

2.5 紫肉甘薯品质性状的近红外光谱建模

为了确保建模结果的公正性, 本研究基于全波

段光谱建立了 6 个 PLSR 对照模型(CK), 并通过 CARS 和 RF 算法构建了 12 个优化模型, 用于预测紫肉甘薯的 6 项主要品质性状(总淀粉、粗蛋白、还原糖、总黄酮、总酚和花青素)。通过绘制蜘蛛图(图 4), 来对比不同 PLSR 模型在各项评价指标上的表现。

为了消除量纲差异的影响, 本研究将 6 项品质

性状分为第 1 组(总淀粉、粗蛋白和还原糖)和第 2 组(总黄酮、总酚和花青素)。结果表明, CK 模型对各项性状的基础预测能力较为稳定(RMSEC 和 RMSECV 值: 第 1 组 $0.275\sim 1.140\text{ g }100\text{ g}^{-1}$ 和 $0.365\sim 1.540\text{ g }100\text{ g}^{-1}$, 第 2 组 $0.136\sim 1.170\text{ mg g}^{-1}$ 和 $0.176\sim 1.670\text{ mg g}^{-1}$; 整体 R^2_c 和 R^2_{cv} 值: $0.929\sim 0.985$ 和 $0.871\sim 0.970$)。相较之下, CARS 和 RF 优化的 12 个 PLSR 模型性能显著提高, 其 RMSEC 和 RMSECV 值(第 1 组 $0.261\sim 1.050\text{ g }100\text{ g}^{-1}$ 和 $0.307\sim 1.280\text{ g }100\text{ g}^{-1}$, 第 2 组 $0.107\sim 1.170\text{ mg g}^{-1}$ 和 $0.127\sim 1.330\text{ mg g}^{-1}$)普遍低于 CK 模型, 整体 R^2_c 和 R^2_{cv} 值则普遍高于 CK 模型($0.936\sim 0.989$ 和 $0.912\sim 0.987$), 表明优化模型具有更强的稳定性和拟合能力。此外, 采用独立验证集评估模型的预测性能与稳健性。结果显示, 优化模型的 RMSEP 值(第 1 组 $0.264\sim 1.240\text{ g }100\text{ g}^{-1}$, 第 2 组 $0.115\sim 1.050\text{ mg g}^{-1}$)均低于 RMSECV 值, 表

明模型具有良好的稳健性, 并能够有效扩展至独立样本集。RPD 值($6.55\sim 19.80$)和 RER 值($21.9\sim 63.4$)进一步验证了模型的实用性和可靠性。

如图 4 所示, CARS-PLSR 模型在预测总淀粉、还原糖和花青素方面表现最佳, 而 RF-PLSR 模型则在粗蛋白、总黄酮和总酚的预测中表现出更高的精度。表 2 总结了表现最佳的 6 个 PLSR 模型的性能指标, 显示出这些模型在校正集($R^2_c = 0.936\sim 0.992$)、交叉验证集($R^2_{cv} = 0.918\sim 0.987$)以及验证集($R^2_v = 0.903\sim 0.987$)上的优异表现。图 5 进一步展示了模型预测值与化学测量值之间良好的线性拟合关系($r_c = 0.968\sim 0.996$ 和 $r_v = 0.951\sim 0.995$)。总体而言, 本研究构建的 6 个基于 CARS 与 RF 优化的 PLSR 模型能够高效、精准地预测紫肉甘薯的主要品质性状, 展现出良好的稳定性与实用性, 适用于紫肉甘薯样品的高通量筛选与快速品质评价。

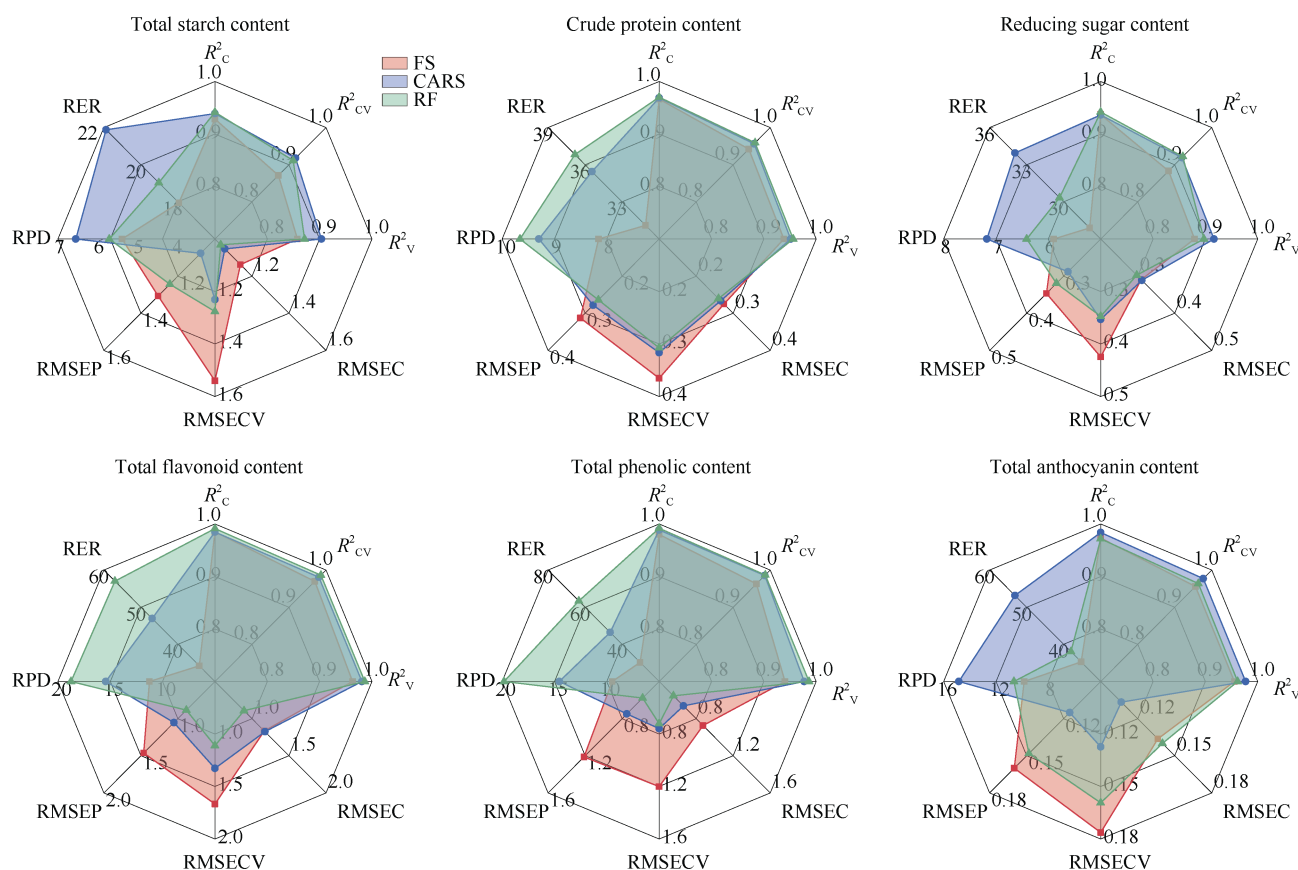


图 4 偏最小二乘回归模型对紫肉甘薯样品品质性状预测的性能

Fig. 4 Performance of partial least squares regression models for predicting quality components in purple-fleshed sweet potato samples

FS: 全光谱; CARS: 竞争性自适应权重; RF: 随机蛙跳; $R^2_c/R^2_{cv}/R^2_v$: 校正集/交叉验证/验证集决定系数; RMSEC/RMSECV/RMSEP: 校正/交叉验证/预测标准偏差; RPD: 预测偏差比; RER: 范围误差比。

FS: full spectrum; CARS: competitive adaptive reweighted sampling; RF: random frog; $R^2_c/R^2_{cv}/R^2_v$: coefficient determination of calibration/cross-validation/validation; RMSEC/RMSECV/RMSEP: root mean square error of calibration/cross-validation/prediction; RPD: ratio of prediction to deviation; RER: range error ratio.

表 2 预测紫肉甘薯样品中化学成分的最佳近红外光谱模型的特点
Table 2 Characteristics of the optimal near-infrared spectroscopy models for predicting chemical components in purple-fleshed sweet potato samples

参数 Parameter	最佳算法 Optimal algorithm	潜变量 LVs	校正集 Calibration		交叉验证 Cross-validation		验证集 External validation			
			校正集 决定系数	校正标 准偏差	交叉验证 决定系数	交叉验证 标准偏差	验证集 决定系数	预测标 准偏差	预测 偏差比	范围 误差比
			R^2_c	RMSEC	R^2_{cv}	RMSECV	R^2_v	RMSEP	RPD	RER
总淀粉 TS (g 100 g ⁻¹)	CARS	7	0.939	1.050	0.918	1.230	0.903	1.080	6.55	21.9
粗蛋白 CP (g 100 g ⁻¹)	RF	6	0.970	0.261	0.959	0.307	0.956	0.264	9.66	36.8
还原糖 RS (g 100 g ⁻¹)	CARS	7	0.936	0.311	0.919	0.353	0.917	0.288	7.18	34.0
总黄酮 TFC (mg g ⁻¹)	RF	9	0.991	0.894	0.987	1.110	0.986	0.885	18.70	57.0
总酚 TPC (mg g ⁻¹)	RF	8	0.992	0.555	0.987	0.721	0.987	0.579	19.80	63.4
花青素 TAC (mg g ⁻¹)	CARS	9	0.984	0.107	0.977	0.127	0.977	0.115	14.80	53.2

缩写同表 1。
Abbreviations are the same as those given in Table 1 and Fig. 4. LVs: latent variables.

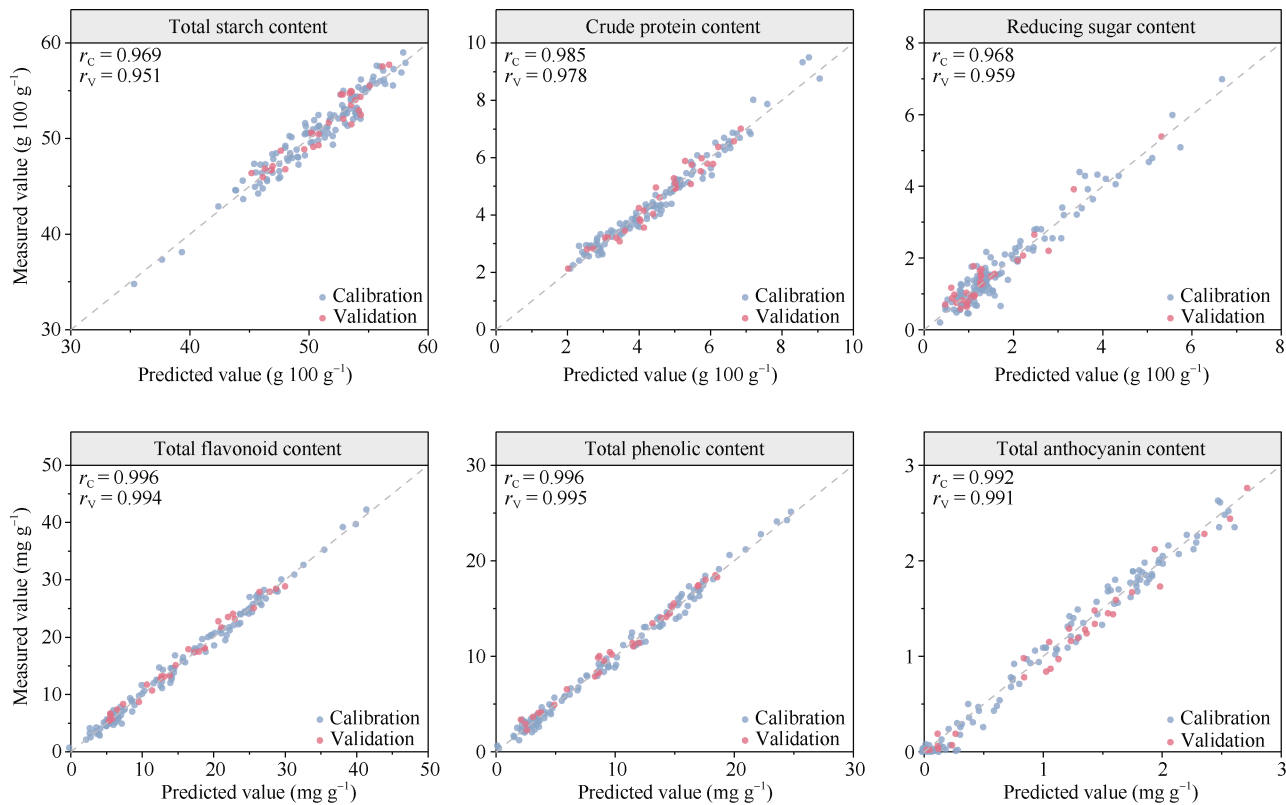


图 5 紫肉甘薯样品测量值与化学值散点图
Fig. 5 Scatter plots comparing predicted versus measured values of chemical components in purple-fleshed sweet potato samples
 r_c : 校正集相关性系数; r_v : 验证集相关性系数。
 r_c : correlation coefficient of calibration; r_v : correlation coefficient of validation.

3 讨论

随着现代作物育种和遗传学研究的深入, 针对大规模种质资源高效表征的需求日益增加^[38]。然而,

目前紫肉甘薯种质资源的评价体系仍存在诸多不足, 传统表征方法普遍面临效率低、通量有限等问题, 难以满足大规模育种需求^[39]。因此, 发展适用于紫肉甘薯品质性状的高通量表型分析体系, 已成为

当前亟待解决的关键课题。NIRS 凭借其非破坏性、快速响应和高通量等优势,已在相关领域得到广泛应用^[40-41],为品质性状的高效测定提供了极具前景的解决方案。NIRS 技术的应用,尤其在亲本克隆与杂交后代地筛选中,极大地提高了遗传资源的筛选效率^[16,42],并为突破紫肉甘薯品质性状渐进性改良的瓶颈提供了新的思路。

在构建可靠的 NIRS 预测模型时,充足的样品量($n > 100$)和具有广泛遗传多样性的代表性样品是基础^[12,43]。本研究基于国家甘薯种质资源圃(广州)的种质资源鉴定、优新品种筛选及配套栽培试验的系统结果,选取了 150 份具有代表性的紫肉甘薯样品。这些样品在品质性状上表现出显著的遗传多样性($H' = 4.83$)和较高的变异性($CV = 54.50\%$),这与甘薯的自交不亲和及其漫长的驯化历史密切相关^[44]。丰富的遗传变异为利用加性与非加性基因效应、实现等位基因分离与重组并获得异交优势提供了坚实的遗传基础^[45-46]。所选样品覆盖了紫肉甘薯的广泛性状变异范围,为构建高稳健性、精准性的 NIRS 预测模型提供了有力的材料支持和数据依据^[43,47]。与已有文献相比,本研究所选样品的规模和遗传多样性显著优于 Lu 等^[16] (93 个样品)、张北举等^[14] (112 个样品)以及王若楠等^[15] (总淀粉 $H' = 1.46 \sim 2.22$) 的研究。综上所述,本研究样品集在样品量和遗传多样性上的优势,确保了模型的普适性、预测精准性及稳健性^[48]。但需要指出的是,本研究采集的样品来源于单一试点,未将样品(或基因型)与种植环境(或生态区)互作对研究指标的影响纳入考量,开发的预测模型在普适性方面存在一定局限性。

样本子集的优化和光谱变量的筛选对模型的预测精度与稳健性至关重要^[34]。本研究通过 KS 算法优化样本划分,避免过拟合或欠拟合,进一步提升了模型的性能^[8]。为了有效克服 NIRS 数据中常见的高维冗余和多重共线性问题,并减少 PLSR 模型的过拟合风险,本研究进一步使用了 CARS 与 RF 算法对光谱变量进行筛选。优化后的模型表现出显著降低的 RMSEC、RMSECV 和 RMSEP 值,以及较高的 R^2_C 、 R^2_{CV} 和 R^2_V ,这与前人的研究结果一致^[8]。根据文献报道,RPD 值大于 3 和 RER 值大于 15 通常被视为评价模型性能优异的标准^[37,49]。在本研究中,所有品质性状的 RPD 值和 RER 值均显著高于这一标准,表明这些模型具备较强的预测能力和实际应用

价值。与同类研究相比,本研究的模型在多个品质性状上展现出明显的竞争优势。例如,淀粉的 R^2_V (0.903) 高于 Lu 等^[16] 的模型结果 ($R^2_V = 0.734$), 接近 Diaz 等^[17] 的结果 ($R^2_V = 0.910$)。蛋白质的 R^2_V (0.956) 与 Magwaza 等^[19] 的结果相近 ($R^2_V = 0.960$)。还原糖的 R^2_V (0.917) 和 RMSEP (0.288 g 100 g⁻¹) 略优于 He 等^[21] 的研究结果 ($R^2_V = 0.906$, RMSEP = 0.264 g 100 g⁻¹), 与 Tang 等^[25] 的糖分预测模型性能相似 ($R^2_V = 0.900 \sim 0.970$, RPD = 5.83~10.30)。花青素的 R^2_V (0.977) 和 RMSEP (0.115 mg g⁻¹) 明显优于卜晓朴等^[22] ($R^2_V = 0.888$, RMSEP = 0.226 mg g⁻¹) 和 Tian 等^[23] 的结果 ($R^2_V = 0.847$, RMSEP = 0.303 mg g⁻¹)。此外,本研究模型在总黄酮 ($R^2_V = 0.986$) 和总酚 ($R^2_V = 0.987$) 的预测结果上也优于 Tang 等^[8] 的高通量模型 ($R^2_V = 0.950$ 和 $R^2_V = 0.960$)。综上所述,模型的整体预测性能优异,得益于更加充足的样品量、丰富的化学变异性以及双重优化策略(样本子集和光谱变量优化)的应用。

本研究首次建立了紫肉甘薯中同步检测总淀粉、粗蛋白、还原糖、总黄酮、总酚和花青素 6 项主要品质性状的高通量 NIRS 预测模型,展现了极佳的模型适配度。所开发的 NIRS 检测方法为紫肉甘薯品质性状的高通量分析提供了强有力的技术支持,能够帮助育种者筛选优质种质资源,并为食品企业精准判定原料用途提供可靠依据。然而,当前研究主要基于探索性试验,样品的来源尚存在一定局限性,如国内外品种、多年份、多地点及不同栽培管理条件下的样品仍需补充。因此,未来的研究应纳入更广泛的遗传背景和环境因素,以构建更加普适的预测模型。此外,结合先进的化学计量学方法与人工智能技术,未来有望进一步提升模型的稳健性与预测能力。

4 结论

本研究通过对 150 份代表性样品进行近红外光谱分析,结合双重优化策略和机器学习算法,成功构建了 6 个高性能的 PLSR 预测模型。这些模型具有较高的 R^2_C (0.936~0.992)、 R^2_{CV} (0.918~0.987)、 R^2_V (0.903~0.987)、RPD (6.55~19.80) 和 RER (21.9~63.4),以及较低的 RMSEC、RMSECV、RMSEP 值。总之,本研究成功构建了紫肉甘薯主要品质性状的高通量近红外预测模型,能够实现多性状同步精准检测,

为种质资源的评价与遗传改良提供了重要的技术支撑。

References

- [1] 史光辉, 胡志和, 马科铭, 等. 紫薯花青素提取条件优化及淀粉等产物的制备. 食品科学, 2014, 35(22): 39–45.
Shi G H, Hu Z H, Ma K M, et al. Optimization of extraction conditions of anthocyanins from purple sweet potato and simultaneous preparation of starch and other by-products. Food Sci, 2014, 35(22): 39–45 (in Chinese with English abstract).
- [2] Song W H, Li C Y, Kou M, et al. Different regions and environments have critical roles on yield, main quality and industrialization of an industrial purple-fleshed sweetpotato (*Ipomoea batatas* (L.) Lam. “Xuzishu 8”). Heliyon, 2024, 10: e25328.
- [3] 朱大洲, 武宁, 张勇, 等. 营养导向型作物新品种选育与审定现状、问题与展望. 作物学报, 2023, 49: 1–11.
Zhu D Z, Wu N, Zhang Y, et al. Current situation, issues, and prospects of breeding and approval of new varieties of nutrition-oriented crops. Acta Agron Sin, 2023, 49: 1–11 (in Chinese with English abstract).
- [4] 谢一芝, 郭小丁, 贾赵东, 等. 紫心甘薯育种现状及展望. 植物遗传资源学报, 2012, 13: 709–713.
Xie Y Z, Guo X D, Jia Z D, et al. Progresses and prospects on purple-fleshed sweet potato breeding. J Plant Genet Resour, 2012, 13: 709–713 (in Chinese with English abstract).
- [5] 王欣, 李强, 曹清河, 等. 中国甘薯产业和种业发展现状与未来展望. 中国农业科学, 2021, 54: 483–492.
Wang X, Li Q, Cao Q H, et al. Current status and future prospective of sweet potato production and seed industry in China. Sci Agric Sin, 2021, 54: 483–492 (in Chinese with English abstract).
- [6] 何鸿举, 王婧茹, 刘红, 等. 近红外光谱技术在甘薯品质检测方面的应用研究进展. 食品科学, 2023, 44(21): 341–350.
He H J, Wang J R, Liu H, et al. Recent advances in the application of near-infrared spectroscopy in quality detection of sweet potato. Food Sci, 2023, 44(21): 341–350 (in Chinese with English abstract).
- [7] 魏良明, 严衍禄, 戴景瑞. 近红外反射光谱测定玉米完整籽粒蛋白质和淀粉含量的研究. 中国农业科学, 2004, 37: 630–633.
Wei L M, Yan Y L, Dai J R. Determining protein and starch contents of whole maize kernel by near infrared reflectance spectroscopy (NIRS). Sci Agric Sin, 2004, 37: 630–633 (in Chinese with English abstract).
- [8] Tang C C, Jiang B Z, Ejaz I, et al. High-throughput phenotyping of nutritional quality components in sweet potato roots by near-infrared spectroscopy and chemometrics methods. Food Chem X, 2023, 20: 100916.
- [9] Liu M, Wang R X, Shi D L, et al. Non-destructive prediction of tea polyphenols during Pu-erh tea fermentation using NIR coupled with chemometrics methods. J Food Compos Anal, 2024, 131: 106247.
- [10] Porep J U, Kammerer D R, Carle R. On-line application of near infrared (NIR) spectroscopy in food production. Trends Food Sci Technol, 2015, 46: 211–230.
- [11] 陈瀑, 戴嘉伟, 李敬岩, 等. 近红外光谱分析中的化学计量学方法进展. 化学试剂, 2023, 45(6): 105–112.
Chen P, Dai J W, Li J Y, et al. Progress of chemometric methods in near-infrared spectroscopy. Chem Reag, 2023, 45(6): 105–112 (in Chinese with English abstract).
- [12] 毕京翠, 张文伟, 肖应辉, 等. 应用近红外光谱技术分析稻米蛋白质含量. 作物学报, 2006, 32: 709–715.
Bi J C, Zhang W W, Xiao Y H, et al. Analysis for protein content in rice by near infrared reflectance spectroscopy (NIRS) technique. Acta Agron Sin, 2006, 32: 709–715 (in Chinese with English abstract).
- [13] 李栓明, 郭银巧, 王克如, 等. 小麦籽粒蛋白质光谱特征变量筛选方法研究. 中国农业科学, 2015, 48: 2317–2326.
Li S M, Guo Y Q, Wang K R, et al. Research on variable selection of wheat kernel protein content with near-infrared spectroscopy. Sci Agric Sin, 2015, 48: 2317–2326 (in Chinese with English abstract).
- [14] 张北举, 陈松树, 李魁印, 等. 基于近红外光谱的高粱籽粒直链淀粉、支链淀粉含量检测模型的构建与应用. 中国农业科学, 2022, 55: 26–35.
Zhang B J, Chen S S, Li K Y, et al. Construction and application of detection model for amylose and amylopectin content in sorghum grains based on near infrared spectroscopy. Sci Agric Sin, 2022, 55: 26–35 (in Chinese with English abstract).
- [15] 王若楠, 张颖星, 于筱菡, 等. 基于近红外快速检测技术的谷子淀粉多样性分析及模型构建. 作物学报, 2025, 51: 1757–1768.
Wang R N, Zhang Y X, Yu X H, et al. Near-infrared spectroscopic evaluation of starch diversity and model construction in foxtail millet. Acta Agron Sin, 2025, 51: 1757–1768 (in Chinese with English abstract).
- [16] Lu G Q, Huang H H, Zhang D P. Prediction of sweetpotato starch physiochemical quality and pasting properties using near-infrared reflectance spectroscopy. Food Chem, 2006, 94: 632–639.
- [17] Diaz J T, Veal M W, Chinn M S. Development of NIRS models to predict composition of enzymatically processed sweetpotato. Ind Crops Prod, 2014, 59: 119–124.
- [18] Gouveia C S S, Lebot V, de Carvalho M P. NIRS estimation of drought stress on chemical quality constituents of taro (*Colocasia esculenta* L.) and sweet potato (*Ipomoea batatas* L.) flours. Appl Sci, 2020, 10: 8724.
- [19] Magwaza L S, Messo Naidoo S I, Laurie S M, et al. Development of NIRS models for rapid quantification of protein content in sweetpotato [*Ipomoea batatas* (L.) Lam.]. LWT Food Sci Technol, 2016, 72: 63–70.
- [20] 唐忠厚, 李洪民, 李强, 等. 基于近红外光谱技术预测甘薯块根淀粉与糖类物质含量. 江苏农业学报, 2013, 29: 1260–1265.
Tang Z H, Li H M, Li Q, et al. Prediction of starch and sugar contents in sweet potato root by near-infrared spectroscopy (NIRS). Jiangsu J Agric Sci, 2013, 29: 1260–1265 (in Chinese with English abstract).
- [21] He H J, Wang Y Y, Zhang M, et al. Rapid determination of reducing sugar content in sweet potatoes using NIR spectra. J Food Compos Anal, 2022, 111: 104641.
- [22] 卜晓朴, 彭彦昆, 王文秀, 等. 生鲜紫薯花青素等多品质参数的可见-近红外快速无损检测. 食品科学, 2018, 39(16): 227–232.
Bu X P, Peng Y K, Wang W X, et al. Rapid nondestructive detection of multiple quality parameters of fresh purple sweet potato

- based on visible near infrared spectroscopy. Food Sci, 2018, 39(16): 227–232 (in Chinese with English abstract).
- [23] Tian X Y, Aheto J H, Bai J W, et al. Quantitative analysis and visualization of moisture and anthocyanins content in purple sweet potato by Vis–NIR hyperspectral imaging. J Food Process Preserv, 2021, 45: e15128.
- [24] Lebot V, Ndiaye A, Malapa R. Phenotypic characterization of sweet potato [*Ipomoea batatas* (L.) Lam.] genotypes in relation to prediction of chemical quality constituents by NIRS equations. Plant Breed, 2011, 130: 457–463.
- [25] Tang C C, Mo X Y, Ma Z M, et al. Machine learning-enhanced near-infrared spectroscopy for high-throughput phenotyping of sweet potato sugars across raw and cooked states. J Agric Food Res, 2025, 21: 101934.
- [26] Jeng T L, Lai C C, Liao T C, et al. Effects of drying on caffeoylquinic acid derivative content and antioxidant capacity of sweet potato leaves. J Food Drug Anal, 2015, 23: 701–708.
- [27] Chaukhande P, Luthra S K, Patel R N, et al. Development and validation of near-infrared reflectance spectroscopy prediction modeling for the rapid estimation of biochemical traits in potato. Foods, 2024, 13: 1655.
- [28] 唐忠厚, 李洪民, 马代夫. 甘薯蛋白质含量近红外反射光谱分析模型应用研究. 中国食品学报, 2008, 8(4): 169–173.
Tang Z H, Li H M, Ma D F. Studies on the application of analysis model for protein content in sweet potato by near infrared reflectance spectroscopy (NIRS). J Chin Inst Food Sci Technol, 2008, 8(4): 169–173 (in Chinese with English abstract).
- [29] 赵凯, 许鹏举, 谷广烨. 3,5-二硝基水杨酸比色法测定还原糖含量的研究. 食品科学, 2008, 29: 534–536.
Zhao K, Xu P J, Gu G Y. Study on determination of reducing sugar content using 3,5-dinitrosalicylic acid method. Food Sci, 2008, 29: 534–536 (in Chinese with English abstract).
- [30] Xu Y Y, Xiao Y D, Lagnika C, et al. A comparative evaluation of nutritional properties, antioxidant capacity and physical characteristics of cabbage (*Brassica oleracea* var. capitata var. L.) subjected to different drying methods. Food Chem, 2020, 309: 124935.
- [31] 陆国权, 赵文静. 甘薯多酚最佳提取工艺及多酚含量的基因型差异研究. 中国粮油学报, 2009, 24(3): 49–52.
Lu G Q, Zhao W J. Sweet potato polyphenol: optimal extraction condition and genotypic variation in polyphenol content. J Chin Cereals Oils Assoc, 2009, 24(3): 49–52 (in Chinese with English abstract).
- [32] Hosseinian F S, Li W D, Beta T. Measurement of anthocyanins and other phytochemicals in purple wheat. Food Chem, 2008, 109: 916–924.
- [33] Li M, Wang J, Du F, et al. High-throughput analysis of chemical components and theoretical ethanol yield of dedicated bioenergy sorghum using dual-optimized partial least squares calibration models. Biotechnol Biofuels, 2017, 10: 206.
- [34] Li M, He S Y, Wang J, et al. An NIRS-based assay of chemical composition and biomass digestibility for rapid selection of Jerusalem artichoke clones. Biotechnol Biofuels, 2018, 11: 334.
- [35] Li H D, Xu Q S, Liang Y Z. Random frog: an efficient reversible jump Markov Chain Monte Carlo-like approach for variable selection with applications to gene selection and disease classification. Anal Chim Acta, 2012, 740: 20–26.
- [36] Li H D, Liang Y Z, Xu Q S, et al. Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration. Anal Chim Acta, 2009, 648: 77–84.
- [37] Yang Z, Li K, Zhang M M, et al. Rapid determination of chemical composition and classification of bamboo fractions using visible-near infrared spectroscopy coupled with multivariate data analysis. Biotechnol Biofuels, 2016, 9: 35.
- [38] Zhong Y C, Zhang Z Q, Chen J, et al. Physicochemical properties, content, composition and partial least squares models of *A. trifoliata* seeds oil. Food Chem X, 2021, 12: 100131.
- [39] 赵凌霄, 王珏, 姚祝芳, 等. 我国甘薯种质资源研究现状及展望. 植物遗传资源学报, 2025, 26: 1068–1079.
Zhao L X, Wang Y, Yao Z F, et al. Research proceeding and prospects of sweet potato germplasm resources in China. J Plant Genet Resour, 2025, 26: 1068–1079 (in Chinese with English abstract).
- [40] 梁晓艳, 吉海彦. 近红外光谱技术在农作物品质分析方面的应用. 中国农学通报, 2006, 22(1): 366–371.
Liang X Y, Ji H Y. Applications of near infrared spectroscopy technology in analyzing the quality of crops. Chin Agric Sci Bull, 2006, 22(1): 366–371 (in Chinese with English abstract).
- [41] 王多加, 周向阳, 金同铭, 等. 近红外光谱检测技术在农业和食品分析上的应用. 光谱学与光谱分析, 2004, 24: 447–450.
Wang D J, Zhou X Y, Jin T M, et al. Application of near-infrared spectroscopy to agriculture and food analysis. Spectrosc Spectr Anal, 2004, 24: 447–450 (in Chinese with English abstract).
- [42] Laurie S M, Naidoo S I M, Magwaza L, et al. Assessment of the genetic diversity of sweet potato germplasm collections for protein content. S Afr N J Bot, 2020, 132: 132–139.
- [43] Sanchez P D C, Hashim N, Shamsudin R, et al. Applications of imaging and spectroscopy techniques for non-destructive quality evaluation of potatoes and sweet potatoes: a review. Trends Food Sci Technol, 2020, 96: 208–221.
- [44] Baccichet I, Chiozzotto R, Spinardi A, et al. Evaluation of a large apricot germplasm collection for fruit skin and flesh acidity and organic acids composition. Sci Hortic, 2022, 294: 110780.
- [45] Pfeiffer W H, McClafferty B. HarvestPlus: breeding crops for better nutrition. Crop Sci, 2007, 47: S-88–S-105.
- [46] Wang L, Wang L, Tan M L, et al. Rapid identification and preliminary evaluation of quality characters of oilseed sunflower by near infrared spectroscopy. Oil Crop Sci, 2022, 7: 142–148.
- [47] Xu F, Yu J M, Tesso T, et al. Qualitative and quantitative analysis of lignocellulosic biomass using infrared techniques: a mini-review. Appl Energy, 2013, 104: 801–809.
- [48] Ejaz I, He S Y, Li W, et al. Sorghum grains grading for food, feed, and fuel using NIR spectroscopy. Front Plant Sci, 2021, 12: 720022.
- [49] Li X R, Ma F M, Liang C P, et al. Precise high-throughput online near-infrared spectroscopy assay to determine key cell wall features associated with sugarcane bagasse digestibility. Biotechnol Biofuels, 2021, 14: 123.

附表 1 150 份甘薯样品清单

Table S1 List of 150 sweet potato samples

编号 No.	品种名称 Variety name	样品来源 Sample source
001	应菜种 Yingcaizhong	种质资源鉴定 Germplasm resource identification
002	宁紫薯 1 号 Ningzishu 1	种质资源鉴定 Germplasm resource identification
003	赣渝 3 号 Ganyu 3	品种筛选试验 Variety screening trial
004	鄂薯 3 号 Eshu 3	品种筛选试验 Variety screening trial
005	济薯 26 Jishu 26	品种筛选试验 Variety screening trial
006	万紫 56 Wanzi 56	品种筛选试验 Variety screening trial
007	胜利百号 Shenglibaihao	种质资源鉴定 Germplasm resource identification
008	福宁紫 3 号 Funingzi 3	品种筛选试验 Variety screening trial
009	白星 Baixing	种质资源鉴定 Germplasm resource identification
010	徐薯 22 Xushu 22	品种筛选试验 Variety screening trial
011	潮薯 1 号 Chaoshu 1	品种筛选试验 Variety screening trial
012	阳江烘烤薯 Yangjianghongkaoshu	种质资源鉴定 Germplasm resource identification
013	紫罗兰(越南紫) Ziluoan (Yuenan Zi)	种质资源鉴定 Germplasm resource identification
014	番薯香 Fanshuxiang	种质资源鉴定 Germplasm resource identification
015	浙薯 255 Zhesu 255	品种筛选试验 Variety screening trial
016	潮州(徐) Chaozhou (Xu)	种质资源鉴定 Germplasm resource identification
017	济南红 Jinan hong	种质资源鉴定 Germplasm resource identification
018	饶平薯 Raopingshu	品种筛选试验 Variety screening trial
019	紫 15-105 Zi 15-105	品种筛选试验 Variety screening trial
020	大鸡爪 Dajizhua	种质资源鉴定 Germplasm resource identification
021	HOK-HOK-TOT	种质资源鉴定 Germplasm resource identification
022	TIS70357	种质资源鉴定 Germplasm resource identification
023	广薯 111 Guangshu 111	品种筛选试验 Variety screening trial
024	新普紫 Xinpuzi	品种筛选试验 Variety screening trial
025	集 4 Ji 4	品种筛选试验 Variety screening trial
026	内原 Neiyuan	品种筛选试验 Variety screening trial
027	广紫薯 2 号 Guangzishu 2	种质资源鉴定 Germplasm resource identification
028	Y-6	种质资源鉴定 Germplasm resource identification
029	广薯 91 Guangshu 91	品种筛选试验 Variety screening trial
030	广紫薯 10 号 Guangzishu 10	种质资源鉴定 Germplasm resource identification
031	北京 553 Beijing 553	种质资源鉴定 Germplasm resource identification
032	勐腊紫红皮 Menglazi hongpi	种质资源鉴定 Germplasm resource identification
033	农林 38 号 Nonglin 38	种质资源鉴定 Germplasm resource identification
034	青藤 Qingteng	种质资源鉴定 Germplasm resource identification
035	平南红薯 Pingnan hongshu	种质资源鉴定 Germplasm resource identification
036	揭塘冬(接长冬) Jietangdong (Jiechangdong)	种质资源鉴定 Germplasm resource identification
037	桂经薯 8 号 Guijingshu 8	栽培试验 Cultivation experiment
038	红头 8 号 Hongtou 8	品种筛选试验 Variety screening trial
039	越南紫罗兰 Yuenanziluoan	品种筛选试验 Variety screening trial
040	广紫薯 1 号 Guangzishu 1	品种筛选试验 Variety screening trial
041	南薯 88 Nanshu 88	品种筛选试验 Variety screening trial
042	渝美 8 号 Yumei 8	品种筛选试验 Variety screening trial
043	褐顶凤凰夫 Hedingfenghuangfu	种质资源鉴定 Germplasm resource identification
044	双皮黄 Shuangpihuang	种质资源鉴定 Germplasm resource identification

(续附表 1)

编号 No.	品种名称 Variety name	样品来源 Sample source
045	金瓜黄 Jinguahuang	种质资源鉴定 Germplasm resource identification
046	沙根苕 Shagenshao	种质资源鉴定 Germplasm resource identification
047	台湾秋 Taiwanqiu	种质资源鉴定 Germplasm resource identification
048	广薯 98-7 Guangshu 98-7	种质资源鉴定 Germplasm resource identification
049	广薯 98 Guangshu 98	品种筛选试验 Variety screening trial
050	抗旱种 Kanghanzhong	种质资源鉴定 Germplasm resource identification
051	96-15-B-2	种质资源鉴定 Germplasm resource identification
052	渝薯 17 号 Yushu 17	品种筛选试验 Variety screening trial
053	广薯 95-1 Guangshu 95-1	种质资源鉴定 Germplasm resource identification
054	原基 Yuanji	种质资源鉴定 Germplasm resource identification
055	岩薯 5 号 Yanshu 5	品种筛选试验 Variety screening trial
056	农大红 Nongdahong	种质资源鉴定 Germplasm resource identification
057	红红 1 号 Honghong 1	品种筛选试验 Variety screening trial
058	三月白 Sanyuebai	种质资源鉴定 Germplasm resource identification
059	冀紫薯 2 号 Jizishu 2	品种筛选试验 Variety screening trial
060	冰淇淋(甘薯) Bingqilin (Ganshu)	种质资源鉴定 Germplasm resource identification
061	福建 60 日 Fujian 60 Ri	种质资源鉴定 Germplasm resource identification
062	木冬 Mudong	种质资源鉴定 Germplasm resource identification
063	渝美 43 Yumei 43	品种筛选试验 Variety screening trial
064	台农 27 Tainong 27	品种筛选试验 Variety screening trial
065	维多丽 Weiduoli	种质资源鉴定 Germplasm resource identification
066	北京 553 Beijing 553	品种筛选试验 Variety screening trial
067	勐腊紫红皮 Menglazihongpi	品种筛选试验 Variety screening trial
068	福宁紫 3 号 Funingzi 3	种质资源鉴定 Germplasm resource identification
069	Ayamurasaki	种质资源鉴定 Germplasm resource identification
070	生核 16-3-18 Shenghe 16-3-18	品种筛选试验 Variety screening trial
071	万紫 56 Wanzi 56	种质资源鉴定 Germplasm resource identification
072	徐紫 8 号 Xuzy 8	种质资源鉴定 Germplasm resource identification
073	阳江烘烤薯 Yangjianghongkaoshu	栽培试验 Cultivation experiment
074	日本紫罗兰 Ribenziluolan	栽培试验 Cultivation experiment
075	黑鬼 Heigui	品种筛选试验 Variety screening trial
076	饶平薯 Raopingshu	种质资源鉴定 Germplasm resource identification
077	广紫薯 10 号 Guangzishu 10	品种筛选试验 Variety screening trial
078	集 4 Ji 4	品种筛选试验 Variety screening trial
079	绵紫薯 9 号 Mianzishu 9	栽培试验 Cultivation experiment
080	内原 Neiyuan	种质资源鉴定 Germplasm resource identification
081	广薯 2k-29 Guangshu 2k-29	种质资源鉴定 Germplasm resource identification
082	紫罗兰(越南紫) Ziluolan (Yuenanzi)	栽培试验 Cultivation experiment
083	广紫薯 2 号 Guangzishu 2	品种筛选试验 Variety screening trial
084	紫 14-76 Zi 14-76	品种筛选试验 Variety screening trial
085	紫 12-007 Zi 12-007	品种筛选试验 Variety screening trial
086	紫 14-105 Zi 14-105	品种筛选试验 Variety screening trial
087	宁紫薯 1 号 Ningzishu 1	品种筛选试验 Variety screening trial
088	越南紫罗兰 Yuenanziluolan	栽培试验 Cultivation experiment
089	冀紫薯 2 号 Jizishu 2	种质资源鉴定 Germplasm resource identification

(续附表 1)

编号 No.	品种名称 Variety name	样品来源 Sample source
090	广紫薯 9 号 Guangzishu 9	栽培试验 Cultivation experiment
091	广紫薯 1 号 Guangzishu 1	种质资源鉴定 Germplasm resource identification
092	紫 15-105 Zi 15-105	栽培试验 Cultivation experiment
093	广薯 91 Guangshu 91	种质资源鉴定 Germplasm resource identification
094	宁紫薯 1 号 Ningzishu 1	栽培试验 Cultivation experiment
095	生核 16-3-18 Shenghe 16-3-18	栽培试验 Cultivation experiment
096	万紫 56 Wanzi 56	栽培试验 Cultivation experiment
097	福宁紫 3 号 Funingzi 3	栽培试验 Cultivation experiment
098	紫 12-007 Zi 12-007	栽培试验 Cultivation experiment
099	济紫薯 1 号 Jizishu 1	种质资源鉴定 Germplasm resource identification
100	阳江烘烤薯 Yangjianghongkaoshu	品种筛选试验 Variety screening trial
101	紫罗兰(越南紫) Ziluolan (Yuenanzi)	品种筛选试验 Variety screening trial
102	饶平薯 Raopingshu	栽培试验 Cultivation experiment
103	紫 15-105 Zi 15-105	栽培试验 Cultivation experiment
104	广紫薯 9 号 Guangzishu 9	种质资源鉴定 Germplasm resource identification
105	集 4 Ji 4	栽培试验 Cultivation experiment
106	内原 Neiyuan	栽培试验 Cultivation experiment
107	日本紫罗兰 Ribenziluolan	种质资源鉴定 Germplasm resource identification
108	广紫薯 2 号 Guangzishu 2	栽培试验 Cultivation experiment
109	广薯 91 Guangshu 91	栽培试验 Cultivation experiment
110	广紫薯 10 号 Guangzishu 10	栽培试验 Cultivation experiment
111	北京 553 Beijing 553	栽培试验 Cultivation experiment
112	勐腊紫红皮 Menglazi hongpi	栽培试验 Cultivation experiment
113	徐紫 8 号 Xu zi 8	品种筛选试验 Variety screening trial
114	绵紫薯 9 号 Mianzishu 9	种质资源鉴定 Germplasm resource identification
115	桂经薯 8 号 Guijingshu 8	品种筛选试验 Variety screening trial
116	越南紫罗兰 Yuenanziluolan	种质资源鉴定 Germplasm resource identification
117	黑鬼 Heigui	种质资源鉴定 Germplasm resource identification
118	紫 14-105 Zi 14-105	栽培试验 Cultivation experiment
119	紫 14-76 Zi 14-76	栽培试验 Cultivation experiment
120	广薯 2k-29 Guangshu 2k-29	品种筛选试验 Variety screening trial
121	Ayamurasaki	品种筛选试验 Variety screening trial
122	冀紫薯 2 号 Jizishu 2	栽培试验 Cultivation experiment
123	广紫薯 2 号 Guangzishu 2	栽培试验 Cultivation experiment
124	生核 16-3-18 Shenghe 16-3-18	品种筛选试验 Variety screening trial
125	福宁紫 3 号 Funingzi 3	栽培试验 Cultivation experiment
126	广紫薯 9 号 Guangzishu 9	品种筛选试验 Variety screening trial
127	日本紫罗兰 Ribenziluolan	品种筛选试验 Variety screening trial
128	阳江烘烤薯 Yangjianghongkaoshu	栽培试验 Cultivation experiment
129	桂经薯 8 号 Guijingshu 8	种质资源鉴定 Germplasm resource identification
130	广紫薯 1 号 Guangzishu 1	栽培试验 Cultivation experiment
131	广薯 2k-29 Guangshu 2k-29	栽培试验 Cultivation experiment
132	徐紫 8 号 Xu zi 8	栽培试验 Cultivation experiment
133	广薯 91 Guangshu 91	栽培试验 Cultivation experiment
134	紫罗兰(越南紫) Ziluolan (Yuenanzi)	栽培试验 Cultivation experiment

(续附表 1)

编号 No.	品种名称 Variety name	样品来源 Sample source
135	紫 15-105 Zi 15-105	品种筛选试验 Variety screening trial
136	宁紫薯 1 号 Ningzishu 1	栽培试验 Cultivation experiment
137	勐腊紫红皮 Menglazihongpi	栽培试验 Cultivation experiment
138	广紫薯 10 号 Guangzishu 10	栽培试验 Cultivation experiment
139	越南紫罗兰 Yuenanziluolan	栽培试验 Cultivation experiment
140	万紫 56 Wanzi 56	栽培试验 Cultivation experiment
141	冀紫薯 2 号 Jizishu 2	栽培试验 Cultivation experiment
142	黑鬼 Heigui	栽培试验 Cultivation experiment
143	内原 Neiyuan	栽培试验 Cultivation experiment
144	济紫薯 1 号 Jizishu 1	品种筛选试验 Variety screening trial
145	集 4 Ji 4	栽培试验 Cultivation experiment
146	饶平薯 Raopingshu	栽培试验 Cultivation experiment
147	北京 553 Beijing 553	栽培试验 Cultivation experiment
148	紫 12-007 Zi 12-007	品种筛选试验 Variety screening trial
149	Ayamurasaki	栽培试验 Cultivation experiment
150	绵紫薯 9 号 Mianzishu 9	品种筛选试验 Variety screening trial