

近红外光谱和人工神经网络的金苦荞氨基酸快速测定

朱丽伟, 严金欣, 黄娟, 石桃雄, 蔡芳, 李洪有, 陈庆富*, 陈其蛟*

贵州师范大学荞麦产业技术研究中心, 贵州 贵阳 550001

摘要 荞麦籽粒中富含谷类作物比较缺乏的赖氨酸,使其不同于其他作物,具有较高经济价值。传统氨基酸测定费时且昂贵,为满足金苦荞育种工作的需要,选用近红外光谱技术结合人工神经网络的算法建立快速检测金苦荞叶片中氨基酸含量的近红外模型。使用氨基酸含量差异较大的样品 255 个,扫描光谱后测定其化学值。研究发现样品苏氨酸(Thr)含量范围是 5.307~14.374 mg·g⁻¹;缬氨酸(Val)含量范围是 6.137~16.204 mg·g⁻¹;甲硫氨酸(Met)含量范围是 0.308~3.049 mg·g⁻¹;异亮氨酸(Ile)含量范围是 5.259~14.134 mg·g⁻¹;亮氨酸(Leu)含量范围是 9.730~26.061 mg·g⁻¹;苯丙氨酸(Phe)含量范围是 5.936~17.223 mg·g⁻¹;赖氨酸(Lys)含量范围是 6.640~17.280 mg·g⁻¹;谷氨酸(Glu)含量范围是 10.984~27.740 mg·g⁻¹;天冬氨酸(Asp)含量范围是 6.437~17.280 mg·g⁻¹;丝氨酸(Ser)含量范围是 3.467~8.312 mg·g⁻¹;精氨酸(Arg)含量范围是 4.937~14.772 mg·g⁻¹;丙氨酸(Ala)含量范围是 3.329~6.885 mg·g⁻¹;组氨酸(His)含量范围是 1.946~4.798 mg·g⁻¹;甘氨酸(Gly)含量范围是 4.196~9.264 mg·g⁻¹;脯氨酸(Pro)含量范围是 1.024~5.672 mg·g⁻¹;酪氨酸(Tyr)含量范围是 0.176~1.173 mg·g⁻¹;半胱氨酸(Cys)含量范围是 0.422~1.926 mg·g⁻¹。每次随机选取 50 个样品建设模型,以 4:1 的比例随机分为训练集和测试集。数据进行归一化处理后,使用神经网络结构 1102-9-1 进行模型建设。利用多次学习的方式建立了较优模型,其中 Arg 和 Asp 近红外模型的仿真测试结果最好,预测值与真实值的相关系数(R²)均大于 0.97,平均相对误差(RSD)也小于 10%;另外 Leu, Val, Tyr, Ile, Ser, Ala, Thr, His, Phe, Gly 和 Lys 模型的 R² 均大于 0.90,模型仿真测试数据的 RSD 小于 10%,模型均可用;Met 与 Cys 的模型进行仿真测试时,其预测值与真实值的 R² 均大于 0.78,但 RSD 大于 10%,模型不可用。结果表明,金苦荞叶片的氨基酸含量高,有极高应用价值,近红外光谱技术结合人工神经网络的分析方法可应用于金苦荞氨基酸含量的预测,为高品质荞麦育种工作提供了便利。

关键词 荞麦;近红外;人工神经网络;氨基酸;模型

中图分类号: S123 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2022)01-0049-07

引言

荞麦属于蓼科荞麦属的一年生草本双子叶植物,因生育期短、抗逆性强,在农业生产中经常用来补种。金荞麦是一种传统中药材,研究发现金荞麦根茎活性成分通过抑制肿瘤细胞增殖和迁移,诱导肿瘤细胞凋亡和自噬等作用而抑瘤,同时还有清热解毒、消炎抗菌等功效^[1-2]。金苦荞是金荞麦与苦荞杂交而培育的新型荞麦,特征介于双亲之间,再生力

强^[3],春季主要进行营养生长,生产大量苦荞叶。研究发现荞麦叶发酵茶可提高小鼠免疫功能,还具有防治急、慢性炎症的功能^[4-7]。荞麦籽粒中还含有丰富的蛋白质,其蛋白质中必需氨基酸含量充足,富含谷类作物中比较缺乏的赖氨酸,使其具有较高的营养价值。本实验室前期的研究发现,荞麦叶中各营养成分比籽粒中含量更高^[3],因此金苦荞叶片有极大研究开发价值。

目前常用的氨基酸测定方法预处理复杂、耗时长、价格昂贵。陶琳丽等研究发现,20 种氨基酸在 1 000~2 502 nm

收稿日期: 2020-12-01, **修订日期:** 2021-03-14

基金项目: 国家自然科学基金项目(31760430),国家重点研发计划项目(2019YFD1001300, 2019YFD1001304),贵州省科技计划项目(黔科合平台人才[2017]5726-20,黔科合基础[2016]1106号),贵州师范大学博士启动基金项目(11904/0516027),贵州省荞麦种质资源保育及创新重点实验室建设专项基金项目(黔教合 KY 字[2017]002)资助

作者简介: 朱丽伟,女,1985年生,贵州师范大学荞麦产业技术研究中心副教授 e-mail: liweib0401001@163.com

* 通讯作者 e-mail: cqf1966@163.com; qijiaochen@126.com

区域有非常明显的近红外光谱吸收且差异显著^[8]，这为利用近红外技术建立检测模型提供了可能性。Chang 等^[9]利用傅里叶变换-近红外分析技术建立了可快速测定菊花谷氨酸、天门冬酰胺和天冬氨酸的模型。通常采用偏最小二乘法 (partial least squares, PLS)、主成分回归 (principal component regression, PCR)、人工神经网络 (artificial neural network, ANN) 等化学计量学算法建立识别模型，我们前期应用线性相关系统的 PCR 法和 PLS 法尝试建立模型，所建立的模型维度太高，限制其应用。Viyona 等^[10]利用近红外光谱技术检测百香果中脂肪的方法，预处理光谱后，建立的前反馈 (back propagation, BP) 人工神经网络模型平均方差高达 0.959。因此本工作选用 BP 人工神经网络算法建立金苦荞氨基酸的近红外光谱预测模型，从内部交叉验证平均相对误差 (RSD) 和相关系数 (R^2) 等评价指标来看，多数氨基酸预测模型取得了令人满意的效果。

1 实验部分

1.1 材料

样品荞麦自交系叶片粉碎样采自贵州省贵阳市乌当区贵州师范大学荞麦中心实验基地，为得到化学值差异较大的样品，种植了 85 份遗传性状已稳定的金苦荞自交系，常规栽培管理，选择生长较为一致的植株，在其开花前采收上三叶、中三叶和下三叶，共 255 份样品，将其分别放于纸袋中，先于鼓风机 105 °C 杀青，然后 80 °C 烘 24 h，然后使用高速粉碎机粉碎并过 100 目筛，备用。

1.2 方法

1.2.1 光谱采集与预处理

采用德国布鲁克光谱仪器公司生产的 MPA 傅里叶变换近红外光谱仪，OPUS 光谱采集软件。采用漫反射的方法，扫描样品的近红外光谱，单次扫描 64 次，分辨率 4 cm^{-1} ，扫描范围 4 000~12 000 cm^{-1} ，每个样品均扫描多次，取平均光谱进行计算。原始谱数据中不仅包含与样品有关的信息，还包含各种干扰和无效信息，影响所建模型的可靠性。为降低仪器噪声和系统误差以提升光谱中的有效信息率，在建模前对原始光谱进行了数据预处理。本实验利用光谱分析软件 OPUS 执行自动优化程序，确定各组分建模的谱区范围和最佳预处理方法。

1.2.2 氨基酸的测定

参考程勇杰等的方法测定氨基酸的含量^[11]，略有改进。具体方法如下：称取 0.100 0 g 叶片粉末放于干净的水解管，加入 6 mol·L⁻¹ 盐酸 15 mL，使用氮吹的方法清除管内氧气后封口。然后将水解管放于烘箱中，调整温度慢慢升至 110 °C 后开始计时，水解 24 h 关机。取出冷却后的水解管，将水解液轻轻转移至 50 mL 容量瓶，去离子水冲洗水解管，冲洗液也倒入容量瓶，定容后混匀，然后吸取滤液 1 mL，使用旋转蒸发仪慢慢蒸干，再用 1 mL pH 2.2 柠檬酸钠缓冲液溶解，滤膜过滤后上机，所用仪器为日立 L-8900 氨基酸自动分析仪。

1.2.3 数据处理

采用 Matlab_R2014b 分析软件，在 iMac Pro(2017) 计算机上进行模型的建设。为了验证模型的性能，使用了交叉验证 (cross-validation) 的方法，根据模型的 R^2 和 RSD，选择最佳模型。当 RSD 小于 10% 时表明模型效果良好，预测精度较高，建立的模型可用。

2 结果与讨论

2.1 金苦荞叶片氨基酸含量

实验发现样品中氨基酸的含量差异较大 (表 1)，其中差异最明显的是必需氨基酸中的甲硫氨酸，所测样品的含量最低值为 0.308 $\text{mg} \cdot \text{g}^{-1}$ ，最高值达 3.049 $\text{mg} \cdot \text{g}^{-1}$ ，后者是前者的 9.9 倍；其次是非必需氨基酸中的酪氨酸，所测样品中含量最低为 0.176 $\text{mg} \cdot \text{g}^{-1}$ ，最高达 1.173 $\text{mg} \cdot \text{g}^{-1}$ ，后者是前者的 6.7 倍。说明所用建模样品间化学值差异明显的可满足建模需要。赖氨酸被称为谷物第一限制性氨基酸，高立成等 (2019) 研究发现，普通苦荞西农 9904 和黔苦 3 号籽粒的赖氨酸含量分别为 3.4 和 4.5 $\text{mg} \cdot \text{g}^{-1}$ ，萌发期其籽粒赖氨酸含量分别升至 3.8 和 5.2 $\text{mg} \cdot \text{g}^{-1}$ ，Bhinder 等研究不同品种苦荞粉中氨基酸的含量，发现苦荞粉中赖氨酸含量的最高值为 6.55 $\text{mg} \cdot \text{g}^{-1}$ ，其他必需氨基酸如缬氨酸、异亮氨酸、苏氨酸和亮氨酸的含量最高值分别是 3.5, 2.55, 8.79 和 12.4 $\text{mg} \cdot \text{g}^{-1}$ ，均远低于本实验中金苦荞叶片赖氨酸的含量^[12-13]，证明金苦荞叶片具有极高的营养价值和开发潜力。

表 1 金苦荞叶片氨基酸含量化学测定结果

Table 1 Contents of amino acid in golden tartary buckwheat leaves using chemical method

氨基酸	最小值/ ($\text{mg} \cdot \text{g}^{-1}$)	最大值/ ($\text{mg} \cdot \text{g}^{-1}$)	平均值/ ($\text{mg} \cdot \text{g}^{-1}$)
苏氨酸 (Thr)	5.307	14.374	7.483
缬氨酸 (Val)	6.137	16.204	8.238
甲硫氨酸 (Met)	0.308	3.049	0.909
异亮氨酸 (Ile)	5.259	14.134	7.083
亮氨酸 (Leu)	9.730	26.061	13.119
苯丙氨酸 (Phe)	5.936	17.223	8.212
赖氨酸 (Lys)	6.640	17.280	9.135
谷氨酸 (Glu)	10.984	27.740	20.185
天冬氨酸 (Asp)	6.437	17.280	9.135
丝氨酸 (Ser)	3.467	8.312	5.911
精氨酸 (Arg)	4.937	14.772	9.654
丙氨酸 (Ala)	3.329	6.885	5.208
组氨酸 (His)	1.946	4.798	3.202
甘氨酸 (Gly)	4.196	9.264	6.828
脯氨酸 (Pro)	1.024	5.672	2.607
酪氨酸 (Tyr)	0.176	1.173	0.647
半胱氨酸 (Cys)	0.422	1.926	1.296

2.2 网络结构

隐层节点数的选择与研究项目的要求、输入-输出节点的多少存在着直接的关系，其数目太少会使网络不能收敛，造成网络不能正确预测未知样本，容错性差，但太多又易造

成网络学习时间过长及过拟合现象的出现。一般隐层节点数是远小于训练样本数的。通过调整输入层、隐层的节点数可以优化网络结构。经过多次预测模型的建立，得到了相关系数较高的结果，其网络结构为(1102-9-1)。

2.3 数据归一化

为了提高建模效率，需要预处理光谱数据，常用的方法有一阶导数、二阶导数、矢量归一等。按照 4:1 的比例将化学值和光谱值随机生成训练集和测试集后，在处理的过程中发现光谱值以及化学值的特征值未在 $[0, 1]$ 的范围内，遂对其进行数据归一化处理。接着创建 9 个隐含层神经元的神经网络，并设置迭代次数为 1 000、MSE 均方根误差范围为 $[0, 1 \times 10^{-3}]$ 即 $[0, 0.001]$ 以及学习率为 0.01。开始训练网络。训练完成后进行仿真测试和数据的反归一化，进而进行性能评价，测试相对误差以及决定系数。

2.4 模型的分析及评价

采用神经网络建模的方法对 17 种氨基酸进行了数学建模后，可直观的看出运用近红外光谱模型对未知样品的预测效果，预测化学值与真实值的相关性及其精确度的提升空间。现对建立的 17 种氨基酸数学模型进行分析。

2.4.1 必需氨基酸的建模效果

从图 1 和表 2 可知，7 种必需氨基酸的近红外模型预测值与真实值的 R^2 均在 0.90 以上，其中苯丙氨酸、亮氨酸、

苏氨酸、缬氨酸、异亮氨酸和赖氨酸近红外模型测试过程，其预测值与真实值的 RSD 均低于 10%，模型可用。

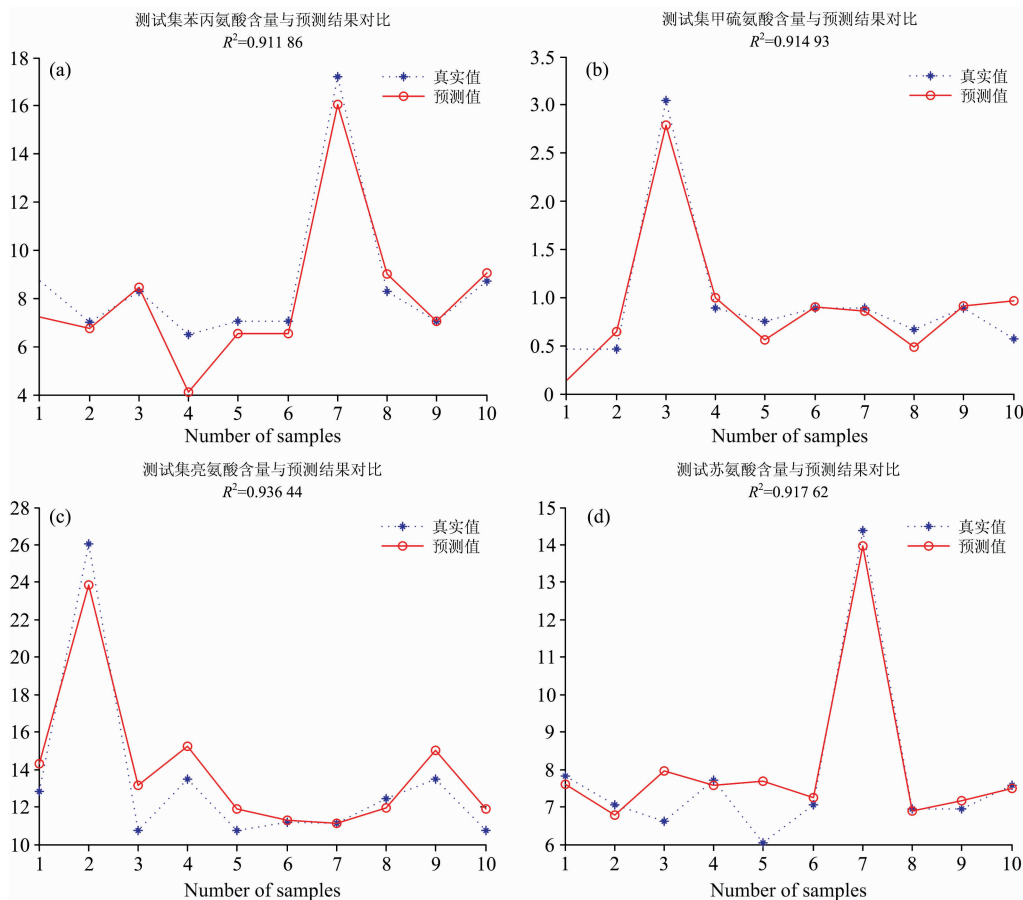
甲硫氨酸近红外模型的预测效果如图 1(b)和表 2 所示，从测试结果可知，供试样品的真实值和预测值的 R^2 虽然较高，达到 0.914 9，然而参与测试的 10 组数据 RSD 达到 25.98%，误差太大，模型不可用。

2.4.2 非必需氨基酸的建模效果

从表 2 和图 2 可知，天冬氨酸和精氨酸的建模效果最好，其模型真实值与预测值的 R^2 均大于 0.97，天冬氨酸模型仿真测试的 10 组数据，其 RSD 值为 6.67% [图 2(c)]；参与精氨酸模型测试的 10 组数据，其 RSD 为 4.22% [图 2(e)]，两种模型均较优秀。

从表 2 和图 2 还可知，丝氨酸、丙氨酸、组氨酸、甘氨酸和酪氨酸模型的预测值与真实值 R^2 均较高，达到 0.90 以上，且仿真测试过程，5 种氨基酸模型的 RSD 均低于 10%，模型较好。脯氨酸模型的仿真测试中，应试样品真实值与预测值的 R^2 为 0.899 1，稍低于其他氨基酸模型，10 组供试数据的 RSD 为 7.86%，模型也可用。

半胱氨酸近红外模型的预测效果如表 2 和图 2(b)所示，可知应试样品的预测值与真实值差异较大，二者的 R^2 只有 0.788 0，且 10 组数据的 RSD 为 33.23%，该模型预测结果不理想，有待进一步研究，可设计更加精细的实验，使用代表



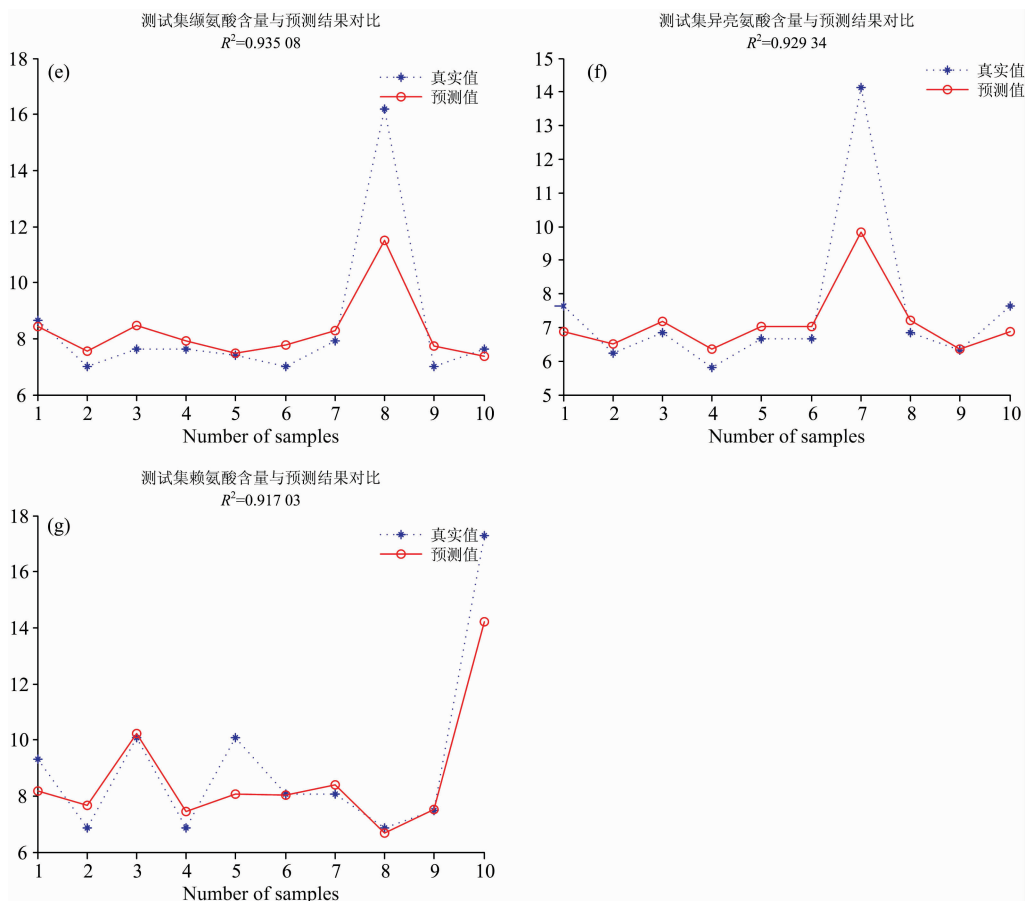


图 1 必需氨基酸的近红外模型仿真测试效果

(a): 苯丙氨酸; (b): 甲硫氨酸; (c): 亮氨酸; (d): 苏氨酸; (e): 缬氨酸; (f): 异亮氨酸; (g): 赖氨酸

Fig. 1 Simulation results of near-infrared model of essential amino acids

(a); Phenylalanine; (b); Methionine; (c); Leucine; (d); Threonine; (e); Valine; (f); Isoleucine; (g); Lysine

表 2 氨基酸近红外模型仿真测试效果

Table 2 Simulation results of near-infrared model of amino acids

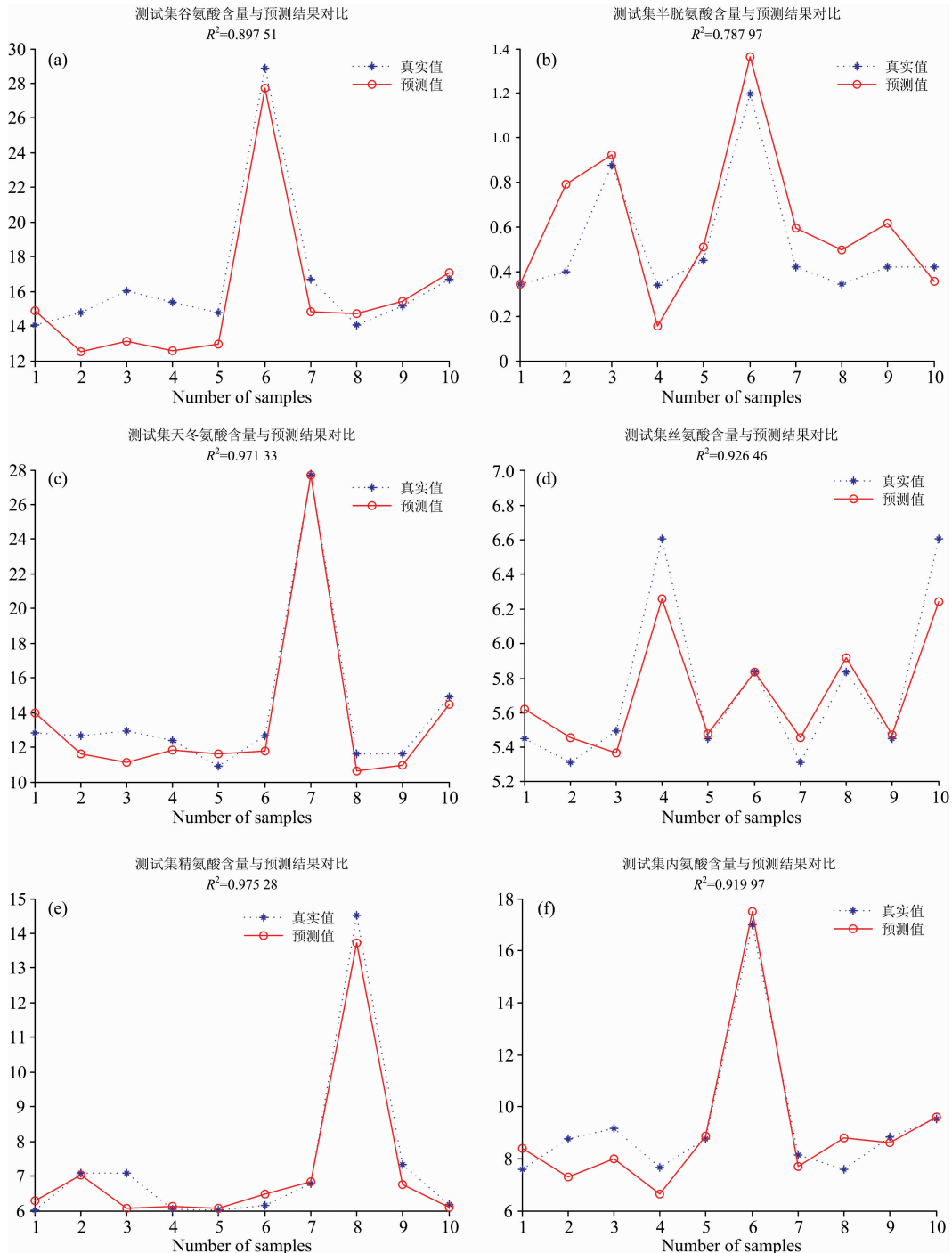
氨基酸模型	参试数据的相对误差										平均相对误差/%	相关系数
Phe	0.170 9	0.032 0	0.020 7	0.371 3	0.072 6	0.072 6	0.067 9	0.087 8	0.002 3	0.036 6	9.35	0.911 9
Met	0.694 9	0.388 7	0.085 4	0.119 6	0.253 6	0.010 2	0.043 0	0.269 0	0.024 5	0.709 6	25.98	0.914 9
Leu	0.117 5	0.085 6	0.227 7	0.127 9	0.107 1	0.013 0	0.002 6	0.038 4	0.114 7	0.107 1	9.41	0.936 4
Thr	0.026 3	0.037 8	0.201 7	0.016 2	0.271 7	0.028 5	0.027 9	0.006 5	0.030 4	0.011 8	6.59	0.917 6
Val	0.025 7	0.080 5	0.110 8	0.036 8	0.008 6	0.113 9	0.045 3	0.290 2	0.105 3	0.035 4	8.53	0.935 1
Ile	0.099 4	0.043 4	0.049 4	0.089 9	0.055 8	0.055 8	0.304 7	0.056 7	0.006 9	0.100 7	8.63	0.929 3
Lys	0.122 6	0.120 1	0.015 7	0.086 7	0.197 3	0.001 4	0.040 4	0.022 7	0.005 3	0.176 9	7.89	0.917 0
Ser	0.032 0	0.027 0	0.022 2	0.052 2	0.004 9	0.000 8	0.027 0	0.014 0	0.003 6	0.055 0	2.39	0.926 5
Arg	0.044 0	0.006 0	0.142 7	0.013 3	0.006 5	0.053 1	0.008 1	0.055 6	0.078 5	0.014 5	4.22	0.975 3
Ala	0.103 4	0.169 0	0.129 1	0.134 2	0.012 2	0.031 1	0.055 5	0.158 1	0.024 3	0.007 2	8.24	0.920 0
His	0.166 1	0.085 3	0.016 1	0.084 4	0.006 2	0.024 0	0.282 8	0.010 4	0.130 1	0.000 8	8.06	0.912 1
Gly	0.009 8	0.088 4	0.168 1	0.005 6	0.009 7	0.013 3	0.268 6	0.010 0	0.232 3	0.010 0	8.16	0.910 7
Glu	0.055 7	0.152 7	0.180 0	0.179 8	0.120 0	0.041 3	0.111 5	0.046 6	0.017 1	0.021 0	9.26	0.897 5
Cys	0.004 4	0.979 3	0.049 7	0.543 5	0.125 7	0.138 2	0.417 3	0.442 0	0.469 0	0.153 4	33.23	0.788 0
Asp	0.090 5	0.082 2	0.140 5	0.045 4	0.063 5	0.070 7	0.000 3	0.087 0	0.055 8	0.030 8	6.67	0.971 3
Pro	0.002 4	0.107 7	0.010 7	0.230 7	0.027 4	0.065 3	0.226 3	0.013 9	0.018 7	0.182 7	7.86	0.899 1
Tyr	0.008 4	0.022 0	0.025 5	0.299 5	0.028 3	0.064 8	0.071 9	0.215 1	0.082 1	0.008 4	8.26	0.931 2

性更强的样品，获得更加精准的化学值和光谱值数值，另外，采用更加合适的算法与光谱预处理方法，以达到更好的建模效果。

3 结 论

通过对苦荞叶氨基酸含量的测定，可知叶中的氨基酸，特别是必需氨基酸的含量远高于荞麦籽粒，使其具有极高的营养价值。常规氨基酸测定方法限制了高氨基酸苦荞的育种

与开发工作。本文建立的 17 种氨基酸的非线性预测模型中，精氨酸和天冬氨酸建模效果较好，预测值与真实值 R^2 均高于 0.97，其 RSD 均低于 7%；另外丙氨酸、苯丙氨酸、甘氨酸、谷氨酸、精氨酸、赖氨酸、酪氨酸、亮氨酸、丝氨酸、苏氨酸、缬氨酸、脯氨酸、异亮氨酸和组氨酸的模型预测值与真实值的 R^2 均大于 0.85，RSD 均低于 10%，可用于氨基酸的预测工作；而甲硫氨酸和半胱氨酸的近红外模型，RSD 均大于 10%，模型不可用，这两种氨基酸的近红外光谱非线性预测模型还需要进一步研究。



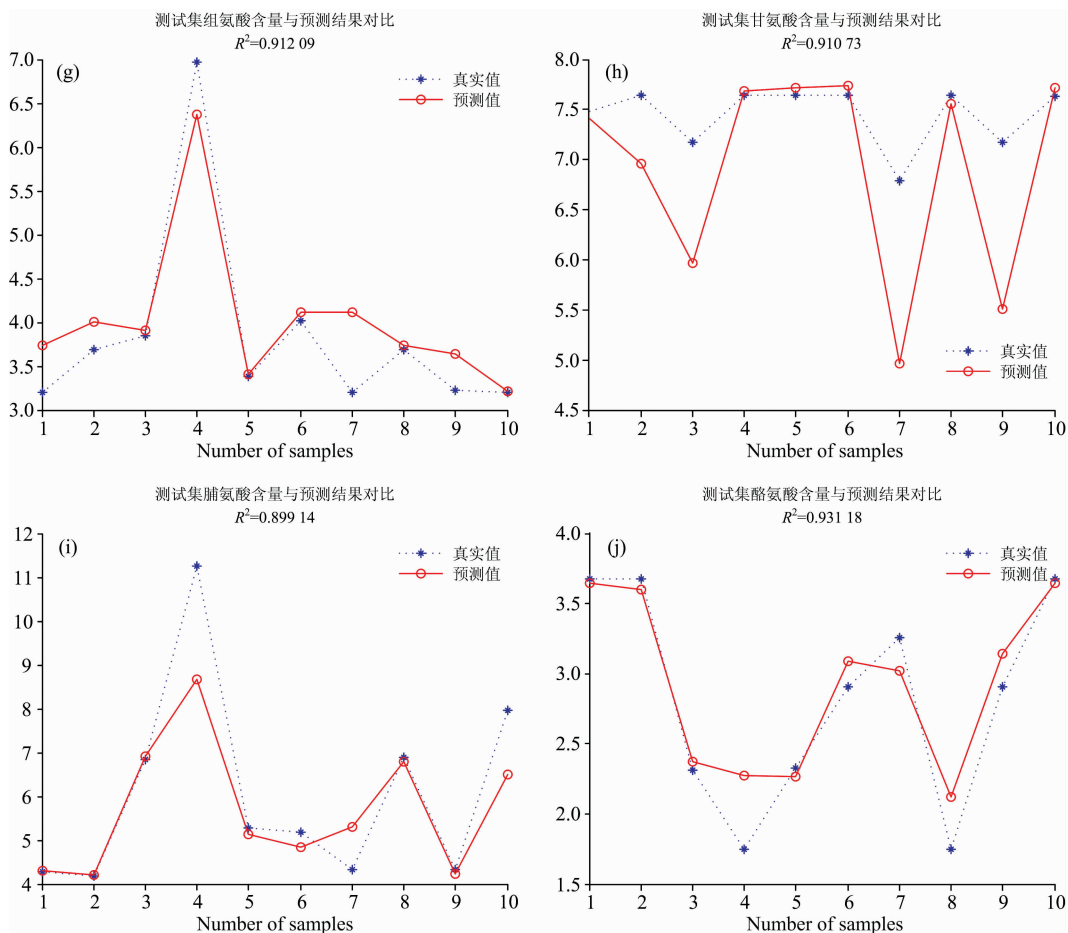


图 2 非必需氨基酸的近红外模型仿真测试效果

(a): 谷氨酸; (b): 半胱氨酸; (c): 天冬氨酸; (d): 丝氨酸; (e): 精氨酸;
(f): 丙氨酸; (g): 组氨酸; (h): 甘氨酸; (i): 脯氨酸; (j): 酪氨酸

Fig. 2 Near infrared simulation of non-essential amino acids

(a): Glutamic acid; (b): Cysteine; (c): Aspartic acid; (d): Serine;
(e): Arginine; (f): Alanine; (g): Histidine; (h): Glycine; (i): Proline; (j): Tyrosine

References

- [1] LI Hong-li, WEN Dan-dan, ZHOU Mei-liang, et al(李红丽, 文丹丹, 周美亮, 等). Chinese Journal of Clinical Pharmacology and Therapeutics(中国临床药理学与治疗学), 2019, 24(7): 833.
- [2] YANG Xi-wen, ZHANG Yan, LI Long-yun(杨玺文, 张 燕, 李隆云). Modern Chinese Medicine(中国现代中药), 2019, 21(6): 837.
- [3] Chen Qingfu, Huang Xiaoyan, Li Hongyou, et al. Sustainability, 2018, 10(2): 1.
- [4] HUANG Xiao-yan, HUANG Sha, CHEN Qing-fu(黄小燕, 黄 莎, 陈庆富). Journal of Anhui Agricultural University(安徽农业大学学报), 2015, 42(6): 854.
- [5] HUANG Xiao-yan, HUANG Sha, CHEN Qing-fu(黄小燕, 黄 莎, 陈庆富). Modernization of Traditional Chinese Medicine and Materia Medica-World Science and Technology(世界科学技术-中医药现代化), 2015, (5): 981.
- [6] Sytar O, Bruckova K, Kovar M, et al. Journal of Central European Agriculture, 2017, 18(4): 864.
- [7] Rebufa C, Pany I, Bombarda I. Food Chemistry, 2018, 261: 311.
- [8] TAO Lin-li, HUANG Wei, YANG Xiu-juan, et al(陶琳丽, 黄 伟, 杨秀娟, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2016, 36(9): 2766.
- [9] Chang Xiangwei, Wei Dandan, Su Shulan, et al. Microchemical Journal, 2020, 153: 104500.
- [10] Viyona M, Andasuryani, Putri R E, et al. Utilization of Artificial Neural Network (ANN) to Predict Fat Passion Fruit Seed Content (Passiflora Ligularis) Based on NIR-S Value. IOP Conference Series: Earth and Environmental Science, 2019, 327: 012017.
- [11] CHEN Yong-jie, CHEN Xiao-wei, ZHANG Sha-sha, et al(程勇杰, 陈小伟, 张沙沙, 等). Science and Technology of Food Industry(食

品工业科技), 2018, 39(6): 1.

- [12] GAO Li-cheng, XIA Mei-juan, BAI Wen-ming, et al(高立城, 夏美娟, 白文明, 等). Acta Nutrimenta Sinica(营养学报), 2019, 41(6): 103.
- [13] Bhinder S, Kaur A, Singh B, et al. Food Research International, 2019, 130(2): 108946.

Rapid Determination of Amino Acids in Golden Tartary Buckwheat Based on Near Infrared Spectroscopy and Artificial Neural Network

ZHU Li-wei, YAN Jin-xin, HUANG Juan, SHI Tao-xiong, CAI Fang, LI Hong-you, CHEN Qing-fu*, CHEN Qi-jiao*
Research Center of Buckwheat Industry Technology, Guizhou Normal University, Guiyang 550001, China

Abstract Buckwheat is rich in lysine, which is a lack in cereal crops, making it different from other cereal crops and has high economic value. Traditional determination of amino acids was time-consuming and expensive. In order to meet the needs of breeding of golden Tartary buckwheat, this study selected near-infrared spectroscopy combined with an artificial neural network algorithm to establish a near-infrared model for rapid determination of amino acid content in buckwheat leaves. A total of 255 samples with different amino acid contents were studied, and their chemical values were determined after scanning spectra. It was found that the content of threonine (Thr) in the samples ranged from 5.307 to 14.374 mg · g⁻¹. Valine (Val) content ranged from 6.137 to 16.204 mg · g⁻¹. The content of methionine (Met) ranged from 0.308 to 3.049 mg · g⁻¹. The content of isoleucine (Ile) ranged from 5.259 to 14.134 mg · g⁻¹. Leucine (Leu) content ranged from 9.730 to 26.061 mg · g⁻¹. The content of phenylalanine (Phe) ranged from 5.936 to 17.223 mg · g⁻¹. Lysine (Lys) content ranged from 6.640 to 17.280 mg · g⁻¹. The content of glutamic (Glu) ranged from 10.984 to 27.740 mg · g⁻¹. Aspartic (Asp) content ranged from 6.437 to 17.280 mg · g⁻¹. Serine (Ser) content ranged from 3.467 to 8.312 mg · g⁻¹. Arginine (Arg) content ranged from 4.937 to 14.772 mg · g⁻¹. The content of Alanine (Ala) ranged from 3.329 to 6.885 mg · g⁻¹. Histidine (His) content ranged from 1.946 to 4.798 mg · g⁻¹. The content of glycine (Gly) ranged from 4.196 to 9.264 mg · g⁻¹. Proline (Pro) content ranges from 1.024 to 5.672 mg · g⁻¹. The content of tyrosine (Tyr) ranged from 0.176 to 1.173 mg · g⁻¹. The content of cysteine (Cys) ranged from 0.422 to 1.926 mg · g⁻¹. During each modeling, 50 samples were randomly selected and randomly divided into the training set and test set at a ratio of 4 : 1. After data normalization, the neural network structure 1102-9-1 was used for model construction. The simulation results of Arg and Asp near-infrared models were the best, the correlation coefficient (R^2) between the predicted value and the real value was greater than 0.97, and the average relative error (RSD) was less than 10%. Simulation test process found, Val, Tyr, Ile, Ser, Ala, Thr, His, Phe, Gly and Lys of model prediction and the real value of R^2 are greater than 0.90, the RSD is less than 10%, models are available; When the models of Met and Cys were tested in simulation, the R^2 between the predicted value and the true value were both greater than 0.78, but the RSD was greater than 10%, so the model was not available. The results showed that golden Tartary buckwheat leaves had a high content of essential amino acids and had high application value. The analysis method of near infrared spectroscopy combined with an artificial neural network could be used to predict the amino acid content of buckwheat, which provided convenience for the breeding of high-quality buckwheat.

Keywords Buckwheat; Near infrared spectroscopy; Artificial neural network; Amino acid; Model

(Received Dec. 1, 2020; accepted Mar. 14, 2021)

* Corresponding authors