

花生籽仁含油量单粒近红外光谱预测模型的建立

任建军^{1,2} 乔亲哲¹ 李小蓓¹ 张开元¹ 邓立苗³

王轩慧³ 唐艳艳^{1,2} 王晶珊^{1,2} 乔利仙^{1,2}

(¹青岛农业大学农学院/中美花生品质改良国际合作联合实验室, 266109, 山东青岛;

²东营青农大盐碱地高效农业技术产业研究院, 257091, 山东东营; ³青岛农业大学理学与信息科学学院, 266109, 山东青岛)

摘要 花生含油量属于复杂性状, 容易受到环境条件的影响而发生变化, 在育种过程中对花生籽仁进行含油量检测至关重要。选取宇花 18 号 (高油) 与 AT215 (低油) 杂交构建的重组近交系群体中 150 个家系为材料, 利用 Antaris II 型傅立叶近红外仪对花生单粒籽仁进行光谱数据采集, 并对扫描后的籽仁进行含油量测定, 采用最小二乘法构建花生单粒籽仁含油量近红外检测模型。该模型内部验证均方差为 1.04, 相关系数 (R^2) = 0.9589。选取群体外 24 份花生样品进行外部验证, $R^2=0.9730$ 。选用宇花 14 号 (高油) 与花育 20 号 (低油) 杂交 F_2 群体继续进行验证, 对该模型与现有单粒模型测定值进行比较, 2 个模型含油量预测差值变化范围为 -0.52%~1.92%, 变异范围较小, 满足含油量预测要求, 可用于花生含油量筛选及品质检测。

关键词 花生; 籽仁; 含油量; 单粒检测; 近红外模型

花生 (*Arachis hypogaea* L.) 是重要的经济作物和油料作物之一, 榨油是我国花生的首要用途, 油用花生约占花生总产量的 52%, 在我国大宗油料作物中, 花生单位面积产油量居首位, 是提升我国油料自给率的首选作物^[1]。花生籽仁含油量作为衡量其品质的核心指标之一, 影响花生产业发展的多个环节, 普通花生含油量约为 50%, 而高油花生含油量则在 55% 以上^[2]。因此, 高油育种依然是目前花生育种的主要目标之一。花生含油量属于复杂性状^[3], 易受到环境条件的影响而发生变化, 在育种早代如 F_1 ~ F_2 代进行表型筛选, 有助于提高选择效率。同时, 在花生加工领域, 快速准确检测花生籽仁含油量, 可保证加工原料的品质, 有助于合理规划市场用途以及精准判定产品价值。因此, 建立花生含油量单粒检测模型, 可满足育种早代选择以及提升市场产值的双重需求。

含油量的检测方法大多基于化学测定法, 如索氏抽提法、直滴式抽提法、皂化法、粗脂肪快速测定法、粗脂肪萃取仪法、酸水解法和比重法等^[4]。利用化学抽提法进行单粒检测存在样本量少, 检测误差较大的弊端, 而且样品被破坏后无法进行后续繁殖和选择, 并不适用于早期材料的筛选检测, 因此建立一种简便快捷、高效且无损的测定方法至

重要。近红外光谱技术 (NIR) 作为分析化学领域迅速发展的分析技术, 可实现样品多成分快速分析, 样品无需破坏, 检测成本低。目前 NIR 已广泛应用于油料作物含油量的快速检测^[5-6], 甘莉等^[7]采用残余法对油菜种子的含油量进行测定, 利用近红外光谱分析仪进行光谱数据采集并建立了数学模型。纪红昌等^[8]选取 229 份花生样品, 通过近红外与索氏抽提法获得的含油量数据建立了花生籽仁含油量近红外检测混合模型。赵星等^[9]利用近红外光谱扫描仪建立了花生蔗糖含量近红外预测模型。近红外分析模型的建立大大加快了花生表型鉴定速度, 提高了选择效率, 但混合模型扫描检测需要大量样本, 并不能满足育种过程中 F_1 ~ F_2 等早代育种材料的检测需求, 因此建立一种通过检测单粒籽仁即可准确获取含油量表型的近红外模型非常重要。韩宏伟等^[10]建立了单粒花生饱和与不饱和脂肪酸近红外定量模型, 李建国等^[11]建立了单粒花生主要脂肪酸油酸、亚油酸和棕榈酸含量近红外预测模型, 胡美玲等^[12]建立了单粒花生蔗糖含量近红外预测模型。针对不同类型近红外仪建立的模型并不能相互兼容使用, 为满足花生品质快速检测的需求, 有必要建立适用于不同材料及不同仪器的专用检测模型。

作者简介: 任建军, 研究方向为作物遗传育种, E-mail: 1725584182@qq.com

乔利仙为通信作者, 研究方向为作物遗传育种, E-mail: lxqiao73@163.com

基金项目: 青岛市科技局中央引导地方科技发展专项 (23-1-3-8-zyyd-nsh); 山东省自然科学基金 (ZR2024GC021); 山东省重点研发计划 (2021LZGC026-07); 山东省盐碱地产业技术体系 (SDAIT-29-03)

收稿日期: 2025-01-16; 修回日期: 2025-05-12; 网络出版日期: 2025-05-22

本研究以高油和低油花生亲本杂交构建的重组自交系群体 (RIL) 中的 150 份家系作为材料, 针对 Antaris II 型近红外仪构建花生籽仁含油量单粒预测模型^[13], 以期花生含油量检测提供一种快速无损的方法, 并为花生加工及贸易等产业环节提供理论依据。

1 材料与方法

1.1 试验材料

选取高油花生品种宇花 18 号 (含油量 56.52%) 和低油花生品种 AT215 (含油量 50.64%) 杂交选育的 RIL 群体 150 个家系 (编号 1~150) 为建模材料, 24 份群体外材料 (编号 A1~A24) 为模型验证材料。宇花 14 号 (含油量 56.37%) 与花育 20 号 (含油量 50.99%) 杂交 F₂ 群体作为验证群体及含油量筛选群体。所有试验材料于 2023 年 5 月春播于山东青岛平度南村镇, 9 月正常收获, 经自然晾干后于 4 °C 冷库低温保存备用。

1.2 花生单粒籽仁含油量近红外光谱采集

通过 Antaris II 型近红外光谱仪 (赛默飞世尔科技, 美国) 对 150 份花生样品进行光谱数据的采集。使用前先打开光谱扫描仪预热 30 min, 预热结束后在 Thermo RESULT 工作软件中校准仪器, 检查各项指标是否正常并调节 workflow, 设置光谱数据保存路径, 选取单粒花生籽仁 (粒型饱满, 无破损) 放入样品杯中进行光谱数据采集。每个样品重复扫描 3 次, 取平均值作为建模数据。

1.3 花生籽仁含油量

采用油料种子含油量测定法 (残余法) (NY/T 1285-2007)^[14] 测定籽仁含油量, 测定仪器为索氏提取器, 根据国标法将花生籽仁放于烘箱中 100±2 °C 烘干 2 h, 待冷却后利用组织破碎仪磨碎花生样品, 随后根据标准称取适量样品放在滤纸筒中并于 105±2 °C 再烘干 3 h, 烘干结束放于干燥器中冷却至室温后, 放于萃取剂中浸泡过夜, 最后对样品抽提 8 h, 每个样品重复测定 3 次。

含油量计算公式: 含油量 (%) = $(m_2 - m_1) / m_2 \times 100$, 式中, m_2 代表索氏提取前样品的质量, m_1 为经过萃取后的样品质量, 称量记录数据后计算花生籽仁含油量。每个样品重复测定 3 次, 取平均值为花生籽仁含油量的化学值。

1.4 花生籽仁置信光谱区间的选择与优化

利用 TQ Analyst 进行模型构建, 建模时首先将

150 份花生材料化学测定值与光谱数据导入, 以软件目录中 Regions 命令对 150 份样品光谱区域进行选择, 选用恒定光源 (Constant) 对光谱数据预处理, 采用定量分析方法中偏最小二乘法 (PLS) 进行模型构建。模型初步建成后, 对比模型的决定系数 (R^2) 和内部交叉验证均方差 (RMSECV), 对模型中测得的异常值进行剔除, 经过反复验证对模型的测量准确性进行优化。

1.5 单粒近红外模型的外部验证

选取未参与建模的 24 份花生材料 (编号 A1~A24) 进行外部验证。含油量测定参照上述残余法标准, 每个样品重复测定 3 次, 对获得的化学测定值与预测值进行相关性分析。

1.6 单粒和多粒模型含油量差异分析

实验室前期已经建立花生含油量混合模型, 利用建好的花生籽仁单粒近红外模型与多粒模型进行数据采集对比。随机选取 20 份 (编号 B1~B20) 饱满无损的花生样品, 利用单粒及多粒模型分别进行光谱采集, 并对扫描后的结果进行比对分析, 进一步验证该单粒模型应用的预测精确度。

1.7 高低油杂交 F₂ 群体验证

选用宇花 14 号 (高油) 与花育 20 号 (低油) 杂交的 F₂ 群体, 利用原有单粒模型继续对该模型进行验证。分别使用 2 个模型对该 F₂ 群体进行预测, 然后对预测值进行比较, 根据比较结果验证该模型的准确性。

1.8 数据处理

利用 TQ Analysis 软件对光谱数据进行处理, 利用偏最小二乘法 (PLS) 进行模型构建, 用 Microsoft Excel 2021 和 IBM SPSS 进行数据统计分析。

2 结果与分析

2.1 花生籽仁含油量的统计分析

150 个样品表型变异丰富, 含油量变化范围为 47.81%~65.57% (表 1), 含油量表型呈较好的离散分布趋势 (图 1), 含油量变化从低到高均有不同分布, 表明化学测定值可用于构建花生单粒籽仁近红外预测模型。

2.2 花生籽仁近红外原始光谱分析

运用近红外光谱扫描仪对 150 个样品花生籽仁光谱数据展开分析, 结果如图 2。光谱值在 4000~10 000 /cm 内近红外漫反射光谱的吸收特征及光谱趋势上表现一致, 在 4500~6000 /cm 区间内表现为

表 1 150 份花生样品含油量的化学测定值
Table 1 Chemical determination of oil content in 150 peanut samples

编号 Number	含油量 Oil content	编号 Number	含油量 Oil content	编号 Number	含油量 Oil content	编号 Number	含油量 Oil content	编号 Number	含油量 Oil content	编号 Number	含油量 Oil content
1	63.03	23	56.07	45	61.81	67	56.86	89	57.62	111	48.75
2	57.84	24	59.15	46	55.54	68	56.96	90	53.15	112	53.25
3	64.13	25	51.23	47	52.98	69	55.20	91	51.23	113	47.88
4	52.37	26	52.15	48	54.06	70	51.77	92	55.63	114	49.46
5	53.91	27	48.51	49	58.40	71	53.53	93	61.33	115	49.57
6	52.43	28	56.52	50	55.24	72	54.41	94	54.33	116	49.77
7	56.30	29	54.36	51	57.33	73	58.97	95	51.28	117	49.91
8	56.59	30	55.30	52	51.39	74	54.53	96	52.81	118	49.95
9	53.36	31	57.61	53	49.43	75	60.49	97	54.19	119	50.03
10	58.71	32	60.25	54	52.77	76	59.25	98	50.64	120	50.06
11	63.23	33	57.15	55	56.16	77	62.78	99	54.34	121	52.38
12	65.57	34	59.93	56	58.85	78	58.77	100	51.42	122	53.15
13	58.19	35	57.34	57	56.78	79	64.97	101	52.83	123	52.79
14	56.98	36	57.93	58	50.09	80	60.15	102	60.77	124	53.32
15	57.18	37	57.32	59	55.07	81	63.76	103	60.22	125	53.42
16	58.25	38	52.59	60	49.88	82	50.05	104	56.82	126	54.45
17	56.48	39	52.00	61	50.77	83	48.24	105	55.06	127	54.60
18	55.87	40	49.03	62	51.83	84	53.19	106	47.81	128	54.63
19	50.80	41	51.03	63	50.64	85	50.81	107	48.36	129	54.76
20	48.93	42	49.28	64	56.23	86	54.26	108	57.63	130	54.77
21	54.11	43	58.54	65	55.16	87	53.70	109	48.70	131	54.81
22	53.19	44	54.91	66	55.32	88	50.07	110	48.73	132	58.89
										133	58.91
										134	58.93
										135	59.39
										136	59.74
										137	56.79
										138	59.94
										139	59.97
										140	60.06
										141	62.67
										142	61.89
										143	58.68
										144	60.32
										145	55.76
										146	54.38
										147	54.55
										148	56.77
										149	56.80
										150	61.09

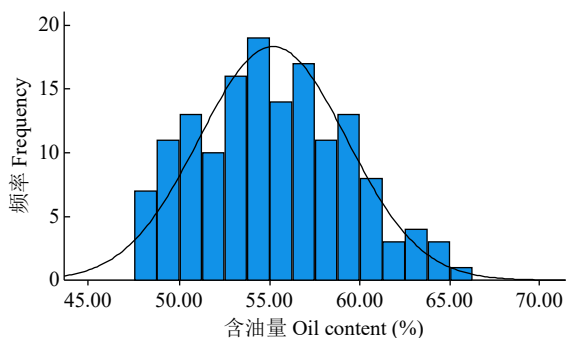


图 1 150 份花生样品含油量的方差分析
Fig.1 Analysis of variance for oil content in 150 peanut samples

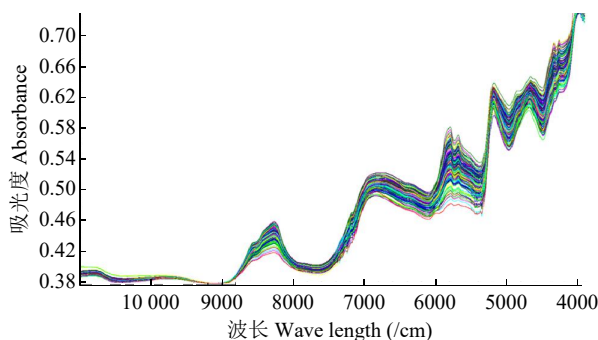


图 2 花生样品的近红外光谱区间
Fig.2 The near infrared spectrum of peanut samples

比较复杂的波状峰图，波动范围较大且每个样品都

展现多个吸收峰、但在 5800~10 000 /cm 高级倍频区，观察到花生吸收峰减弱，动态变化范围基本一致。不同特征光谱曲线的近红外漫反射光谱，表明不同花生含油量的不同，这也为花生籽仁含油量构建提供了丰富的光谱信息。

2.3 近红外原始光谱曲线的优化

利用一阶导数法 (first derivative, FD) 对 150 份样品光谱数据进行处理，对存在较大基线偏移的异常光谱曲线进行矫正，而经过处理后的光谱曲线精确程度有了明显的提升。如图 3 所示，在光谱分析中常出现的基线漂移现象也明显减弱，整体光谱动态变化范围处于正常水平，经修正后光谱变化峰轮廓变得更清晰，可辨识度更高，可以用作近红外模型的构建。

2.4 花生单粒含油量模型的构建与校正

利用光谱建模软件 TQ Analyst 对近红外模型扫描得到的光谱数据与采用索氏提取法测得的含油量化学值进行拟合，对偏离较大的异常值使用软件内部的交叉验证功能对数据进行剔除或矫正，挑选数值在合理范围内且质量较高的数据作为定标模型。由图 4 可知，对于异常值进行优化后模型内

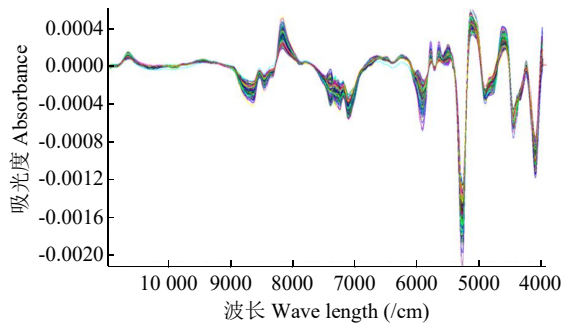


图3 原始光谱曲线优化
Fig.3 Optimization of original spectral curve

部交叉验证均方差为 1.04，相关系数 $R^2=0.9589$ ，模型准确性高^[15]，可用于花生籽仁含油量定量分析。

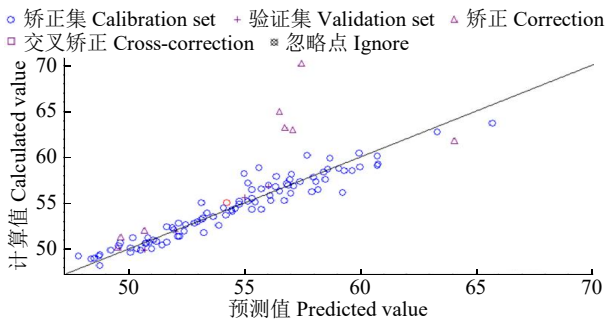


图4 花生单粒籽仁含油量近红外模型构建
Fig.4 Establishment of near-infrared model on oil content of single-kernel in peanut

2.5 单粒近红外模型的外部验证

为检验所构建模型在实际应用中的可靠性与有效性，选取 24 份建模样本之外的花生样品（编号 A1~A24）作为验证集，将利用索氏提取法测得的化学测定值与近红外模型预测值进行比对，两者之间绝对误差变化范围 ≤ 1.6 ，相对误差的变化范围为 $-0.05\% \sim 3.17\%$ 。对 24 份样品化学测定值与预测值进行相关性分析，其决定系数 $R^2=0.9730$ （图 5），表明所建立的单粒近红外模型准确性较高，可以满

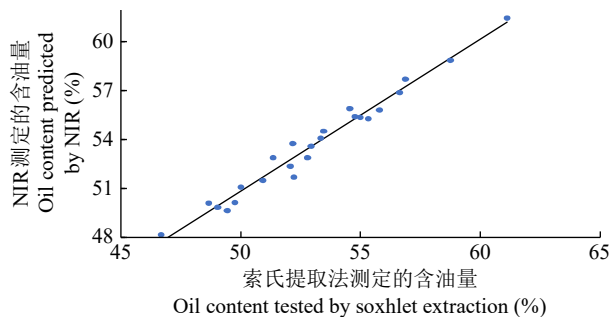


图5 验证材料化学值与预测值相关性分析
Fig.5 Correlation analysis between chemical value and predicted value of verification material

足花生含油量快速检测的实际需求，可用于花生育种过程中的早期材料筛选以及花生加工产业中对于原料品质的快速检测。

2.6 单粒模型与多粒模型比较

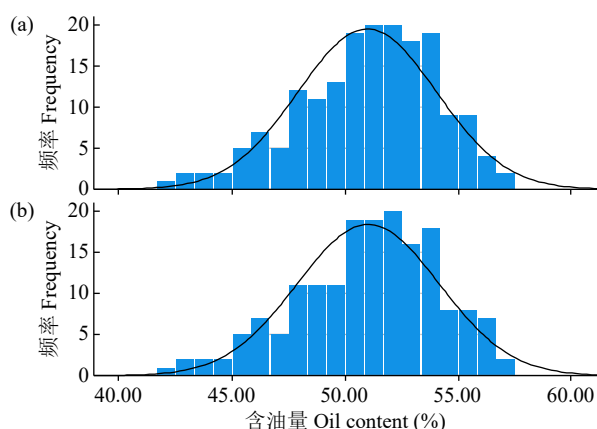
分别利用单粒及多粒模型对花生籽仁进行近红外光谱扫描，利用光谱分析软件进行分析，结果（表 2）表明 2 种模型所测含油量差值较小，含油量动态变化范围为 $0.05\% \sim 1.15\%$ ，其中含油量差异最大的为 B14，差异最小的为 B13，可满足单粒快速无损检测的需求，有望在育种选择及籽仁品质检测发面发挥作用。

表 2 单粒模型与多粒模型预测值比较
Table 2 Comparison of predicted values between single-kernel and multi-kernel model %

编号 Number	含油量（多粒） Oil content (multi-kernel)	含油量（单粒） Oil content (single-kernel)	含油量差值 Oil content difference
B1	50.08	49.05	1.03
B2	49.52	49.29	0.23
B3	50.47	49.36	1.11
B4	50.81	50.40	0.41
B5	48.51	48.66	-0.15
B6	48.85	49.78	-0.93
B7	49.08	49.78	-0.70
B8	50.24	49.82	0.42
B9	48.90	49.99	-1.09
B10	50.35	50.03	0.32
B11	50.25	50.04	0.21
B12	49.36	50.16	-0.80
B13	50.17	50.22	-0.05
B14	49.09	50.24	-1.15
B15	51.15	50.31	0.84
B16	50.14	50.32	-0.18
B17	49.82	50.38	-0.56
B18	51.12	50.41	0.71
B19	51.66	50.55	1.11
B20	51.60	50.63	0.97

2.7 F₂ 群体含油量统计分析

利用原有单粒模型与本研究所建模型分别对字花 14 号与花育 20 号杂交 F₂ 群体进行含油量预测，图 6 显示 2 个模型所预测含油量均符合正态分布规律。通过对 2 个模型含油量表型统计分析，发现含油量差值变化范围在 $-0.52\% \sim 1.92\%$ ，其中超过 64% 的样品含油量差值小于 0.5%，超过 90% 的样品含油量差值小于 1.0%。新模型含油量差异较小，可用于单粒花生含油量快速检测。



(a) 原有模型, (b) 新建模型。

(a) original model, (b) new model.

图 6 IBM SPSS 方差分析结果

Fig.6 IBM SPSS analysis of variance results

3 讨论

花生是我国重要的特色油料作物之一,花生籽仁富含脂肪和蛋白质,以及优质脂肪酸油酸和亚油酸等成分,具有极高的营养价值与经济价值^[16-17]。花生籽仁品质性状的快速无损检测一直都是科研及产业应用的有效手段,目前已针对脂肪、蛋白质、蔗糖含量以及脂肪酸、氨基酸和微量营养成分等构建了多个近红外模型^[18-21]。雷永等^[22]利用美国 Unity 公司的近红外扫描仪建立蔗糖含量相关的小样品杯近红外模型,杨军军等^[23]建立了花生油酸及亚油酸含量 2 个品质性状相关的近红外模型,曲艺伟等^[24]利用 NIR-FIEXN-500 近红外光谱仪构建了花生脂肪酸含量预测模型。但所建模型并不适用于非同一型号的检测仪器,因此各科研单位和加工企业都在尝试建立适合自己应用的检测模型。尤其单粒检测模型的建立对于样品选择及检测具有更高的要求。本研究拟利用重组近交系群体材料,利用 Thermo 公司(美国)生产的 Antaris II 型近红外光谱仪建立花生籽仁含油量单粒检测模型。

本研究所用亲本宇花 18 号为高油亲本,含油量变异范围在 55.50%~57.32%,AT215 为低油亲本,含油量变异范围在 49.65%~51.65%。利用两亲本杂交构建的 RIL 群体,其含油量变异范围广,适合用作建模材料。用于验证模型的 F₂ 群体,其高油亲本为宇花 14 号,含油量变异范围在 55.93%~57.48%,低油亲本为花育 20,含油量变异范围 49.30%~51.09%,F₂ 代群体含油量变异丰富,适合作为验证群体。利用单粒模型检测数据与已建立的多粒混合模型检测数据进行比较,同一样品 2 种模型预测值

存在较小差异,说明单粒模型可满足单粒快速无损检测的需求,有望在育种选择及品种品质检测方面发挥作用。

建模过程中发现还需注意的问题,首先,选择 3 粒籽仁进行光谱扫描,然后取平均值,再分别对 3 粒籽仁进行化学法测定含油量,然后取平均值作为建模数据,可以最大程度地减少单粒测定的误差值。其次,各样品应尽量选择大小均匀一致的籽仁,保证籽仁放置方向一致,使籽仁正好覆盖投射孔,避免漏光造成光谱数据不准确。由于花生不同家系及不同品系之间籽仁大小及形状均存在一定程度的差异,如果有足够多的家系和品系,则可以按大小、形状、或种皮颜色等进行分类建模,这样可以最大程度地提高模型的准确性。另外,花生含油量会随储存方式和储存时间不同而发生变化,故以当年收获自然晾干后当年秋天 11-12 月样品作为最佳取样时间,再考虑到环境条件的影响,光谱扫描及含油量测定均应在一定时间和条件下集中完成测定,以最大程度保证模型的准确度。

本实验室前期已经建立花生含油量近红外混合检测模型和花生籽仁花青素含量近红外模型^[25]等。在建模过程中发现近红外光谱仪虽然分析速度快,可多组分同时分析,操作简便。但是同时也存在分辨率及光谱扫描范围有限等不足之处。与近红外相比,高光谱可扫描光谱范围更广,数据信息丰富度高,可同时获取被检测样品的空间信息和光谱信息,光谱采集过程中可同时多个样品进行大批量扫描,目前,已有学者^[26]建立高光谱与花生品质相关扫描模型,后期将尝试建立高光谱单粒花生检测籽仁含油量的模型,并进一步提高检测的精准度和效率。

4 结论

选取宇花 18 号(高油)与 AT215(低油)杂交构建的重组自交系群体中 150 个家系为材料,建立花生籽仁含油量单粒近红外预测模型,相关系数(R^2)为 0.9589。经验证该单粒预测模型检测结果准确性较好,可用于花生早期世代单粒籽仁含油量检测,为高低油花生品系筛选提供快速无损的检测方法。

参考文献

- [1] 廖伯寿. 我国花生生产发展现状与潜力分析. 中国油料作物学报, 2020, 42(2): 161-166.

- [2] 任小平, 姜慧芳, 王圣玉, 等. 野生花生高油基因资源的发掘与鉴定. 中国油料作物学报, 2010, 32(1): 30-34.
- [3] 胡美玲. 花生籽仁含油量的遗传分析与 QTL 定位. 北京: 中国农业科学院, 2022.
- [4] 李帅, 贺群岭, 雷红霞, 等. 不同检测方法对花生粗脂肪含量的检测分析. 粮食科技与经济, 2020, 45(9): 85-88, 96.
- [5] 于伶俐, 谭亚军, 赵甲慧, 等. 近红外光谱技术在食品领域的研究进展. 食品安全导刊, 2022(29): 177-180.
- [6] 吕建伟, 饶庆琳, 姜敏, 等. 花生籽仁油酸、亚油酸含量近红外模型构建及育种应用. 中国油料作物学报, 2023, 45(2): 399-406.
- [7] 甘莉, 孙秀丽, 金良, 等. NIRS 定量分析油菜种子含油量、蛋白质含量数学模型的创建. 中国农业科学, 2003, 36(12): 1609-1613.
- [8] 纪红昌, 邱晓臣, 柳文浩, 等. 花生籽仁含油量近红外模型的构建及其应用. 中国油料作物学报, 2022, 44(5): 1089-1097.
- [9] 赵星, 张嘉楠, 张一鸣, 等. 花生籽仁蔗糖含量近红外光谱快速测定方法研究. 中国油料作物学报, 2025, 47(1): 226-233.
- [10] 韩宏伟, 王传堂, 符明联, 等. 11 个单粒花生脂肪酸近红外定量分析模型构建. 中国油料作物学报, 2023, 45(2): 407-412.
- [11] 李建国, 薛晓梦, 张照华, 等. 单粒花生主要脂肪酸含量近红外预测模型的建立及其应用. 作物学报, 2019, 45(12): 1891-1898.
- [12] 胡美玲, 鄧晨阳, 薛晓梦, 等. 单粒花生蔗糖含量近红外预测模型的建立. 作物学报, 2023, 49(9): 2498-2504.
- [13] Zhang S T, Xue L, Wang D, et al. Rapid determination of oil content of single peanut seed by near-infrared hyperspectral imaging. *Oil Crop Science*, 2024, 9(4): 220-224.
- [14] 中华人民共和国农业农村部. 油料种籽含油量的测定残余法: NY/T 1285—2007. 北京: 中国标准出版社, 2007.
- [15] Dubrovkin J. Linear transformations of multivariate calibration models in near infrared spectroscopy: a comparative study. *Journal of Near Infrared Spectroscopy*, 2017, 25(4): 223-230.
- [16] Lin M, Long M X, Li G L, et al. Analysis of peanut using near-infrared spectroscopy and gas chromatography-mass spectrometry; correlation of chemical components and volatile compounds. *International Journal of Food Properties*, 2016, 19(3): 508-520.
- [17] 赵思梦, 于宏威, 高冠勇, 等. 花生蛋白组分及其亚基含量近红外分析检测方法. 光谱学与光谱分析, 2021, 41(3): 912-917.
- [18] 张晓科, 王珊珊, 曲伟华. 近红外定量分析自然风干花生油酸和亚油酸及棕榈酸含量的探索. 农业开发与装备, 2017(7): 76-77.
- [19] 杨传得, 唐月异, 王秀贞, 等. 傅立叶近红外漫反射光谱技术在花生脂肪酸分析中的应用. 花生学报, 2015, 44(1): 11-17.
- [20] 张建成, 王传堂, 王秀贞, 等. 花生自然风干种子油酸、亚油酸和棕榈酸含量的近红外分析模型构建. 中国农学通报, 2011, 27(3): 90-93.
- [21] 张欣. 花生过氧化值、酸值和芥酸含量近红外分析模型构建及 FAE1 基因克隆. 郑州: 河南农业大学, 2018.
- [22] 雷永, 王志慧, 淮东欣, 等. 花生籽仁蔗糖含量近红外模型构建及在高糖品种培育中的应用. 作物学报, 2021, 47(2): 332-341.
- [23] 杨军军, 庞洪利, 韩宏伟, 等. 花生油油酸亚油酸含量近红外模型构建. 农业与技术, 2021, 41(15): 4-7.
- [24] 曲艺伟, 张鹤, 韩笑, 等. 花生脂肪酸近红外模型的建立. 分子植物育种, 2019, 17(2): 568-578.
- [25] 李鑫, 刘潇洋, 于复欣, 等. 花生籽仁花青素含量近红外模型的构建及应用. 中国粮油学报, 2024, 39(9): 191-197.
- [26] Yu H W, Liu H Z, Wang N, et al. Rapid and visual measurement of fat content in peanuts by using the hyperspectral imaging technique with chemometrics. *Analytical Methods*, 2016, 8(41): 7482-7492.

Establishment of a Near-Infrared Spectroscopy Model for Predicting Oil Content of Single Peanut Kernels

Ren Jianjun^{1,2}, Qiao Qinzhe¹, Li Xiaobei¹, Zhang Kaiyuan¹, Deng Limiao³,
Wang Xuanhui³, Tang Yanyan^{1,2}, Wang Jingshan^{1,2}, Qiao Lixian^{1,2}

¹College of Agriculture, Qingdao Agricultural University/International Cooperation Laboratory for Peanut Quality Improvement between China and US, Qingdao 266109, Shandong, China;

²Academy of Dongying Efficient Agricultural Technology and Industry on Saline and Alkaline Land in Collaboration with Qingdao Agricultural University, Dongying 257091, Shandong, China; ³College of Sciences and Information Sciences, Qingdao Agricultural University, Qingdao 266109, Shandong, China)

Abstract Peanut oil content is a complex trait, which is easily affected by environmental conditions. It is critically important to detect oil content of peanut kernels in peanut breeding. A total of 150 lines of recombinant inbred lines (RIL) derived from the cross between Yuhua 18 (high oil) and AT215 (low oil) were selected as materials, the spectral data of single-kernel was collected by Antaris II near infrared spectrometer, and the oil content was determined. A near-infrared (NIR) detection model on oil content for single-kernel oil content was established by partial least squares (PLS). The Root-Mean-Squares Error of Cross-Validation was 1.04 and the correlation coefficient (R^2) was 0.9589. Twenty-four peanut samples with obvious difference in oil content were selected for external validation, and the R^2 was 0.9730. Another F₂ population from the cross between Yuhua 14 (high oil) and Huayu 20 (low oil) was used as the test population, whose oil content was determined by this model and the present single-kernel model. The variation range of predicted values from these two models was from -0.52% to 1.92%, showing a narrow variation range. This model is relatively small, and is reliable for oil content confirmation, which can be used for the oil content selection and quality detection of peanut.

Key words Peanut; Kernel; Oil content; Detection on single kernel; NIR model