DOI: 10.13718/j. cnki. xdzk. 2015. 03. 006

甘蓝型油菜硫苷组分近红外检测模型的建立。

王耐红, 张 敏, 王 瑞, 李施蒙, 刘 川

西南大学 农学与生物科技学院,重庆 400716

摘要:为探索近红外光谱技术在快速检测硫苷组分质量摩尔浓度方面的可行性,该研究选取了甘蓝型油菜籽材料 266份,采用高效液相色谱法测定了硫苷组分的质量摩尔浓度,并利用近红外分析仪和 WinISI III 软件对数据进行 判断与处理,建立了甘蓝型油菜籽硫苷组分的近红外检测模型.其中,2-羟基-3-丁烯基脱硫硫苷,3-丁烯基脱硫 硫苷定标决定系数(RSQ)与交叉验证决定系数(1-VR)均超过了 0.90(越接近 1 越理想),说明模型预测性能较好, 可以用于油菜品种的快速筛选.但是,4-羟基-3-吲哚甲基脱硫硫苷,4-戊烯基脱硫硫苷,苯乙基脱硫硫苷模型的 RSQ 分别为 0.470,0.219 和 0.147,1-VR 分别为 0.304,0.128,0.175,建模效果欠佳,还需要进一步研究. 关键词:甘蓝型油菜;硫苷组分;近红外

中图分类号: S634.3 文献标志码: A 文章编号: 1673-9868(2015)03-0035-07

甘蓝型油菜的种植面积和产量在我国均居首位,是我国重要的油料作物.油菜籽加工后的菜籽饼粕约含40%的蛋白质,是重要的饲料蛋白资源^[1-3].然而,油饼中的硫苷降解会产生有毒物质,影响饼粕的饲用效果和价值,所以油菜品质育种的一个重要任务,就是把硫苷质量摩尔浓度减到最低水平^[4].现在国际上普遍采用高效液相色谱仪分析法测定硫苷质量摩尔浓度,此方法适用于硫苷的精确测定,并不适合在育种中进行大批量样品的快速筛选.因此,探索出一种高效、快速、准确的硫苷组分质量摩尔浓度的检测方法,为油菜品质育种提供参考很有实际意义.

近年来,近红外光谱技术(Near infrared reflectance spectroscopy technique,简称 NIRS)以其快速、无损伤、分析费用低等优点,在作物育种品质性状的大批量筛选中得到了广泛应用,现在已经成为作物育种中重要的选择手段^[5-12].国内外应用近红外光谱技术建立的油菜籽中硫苷质量摩尔浓度的定标方程,*R*² 高达 0.99^[13-16],且 Font^[15]以及 Velasco 等^[17]分别建立了油菜籽主要硫苷类型油菜籽饼质量摩尔浓度 (µmol/g 饼)和总硫苷百分比(mol/mol)NIRS 分析定标方程,但是国内还没有关于硫苷组分近红外模型方面的研究.

本研究测定了 266 份供试材料的硫苷组分质量摩尔浓度,应用近红外反射光谱技术建立甘蓝型油菜 硫苷组分质量摩尔浓度的定标方程,并对方程进行优化,旨在为油菜品质育种和改良中品种的大量筛选 提供参考.

1 材料与方法

1.1 供试材料

266 份供试材料均为稳定的甘蓝型油菜品种(高硫苷、中硫苷和低硫苷材料),包括 2010 年和 2011

① 收稿日期: 2013-09-15

基金项目:重庆市科技攻关重点项目(CSTC2012ggB80008);国家高新技术研究发展计划("863"计划,2011AA10A104). 作者简介: 王耐红(1987-),女,河北唐山人,硕士,主要从事油菜品质分析研究. 通信作者:王 瑞,副研究员. 年 2 年的材料,由西南大学重庆市油菜工程技术研究中心提供.样品硫苷质量摩尔浓度为 29.11~ 136.83 μmol/g饼,平均值为 71.84 μmol/g饼,表示样品的硫苷质量摩尔浓度分布范围较广,具有较强 的代表性.样品分为校正集和检验集,前者用于建立模型(内部交叉验证);后者用于检验所建立的定标 方程的预测性能(外部验证).

1.2 测试项目与方法

1.2.1 硫苷组分质量摩尔浓度测定

采用 ISO9167-1: 1992(E)高效液相色谱国际标准方法^[18-19]测试材料的硫苷组分质量摩尔浓度.

硫苷提取:预烘干的甘蓝型油菜籽粒打粉,迅速称量 200 mg 样品并转入 10 mL 离心管中,在 75 ℃水 浴锅中加热 1 min 灭活,加入 2 mL 的 70% 沸甲醇水溶液和 200 µL 内标,混匀后转入超声波清洗器 10 min,然后离心并保留上清液,再加入 2 mL 70% 沸甲醇水溶液重复以上步骤提取 1 次,最终将 2 次离心上清液混匀;取 2 mL 溶液上醋酸型 DEAE-Sephadex A-25 阴离子交换柱,加硫酸酯酶 500 µL,并在 35 ℃ 条件下酶促反应 16 h;用 1 mL 纯水冲洗微柱,待液体排干后再用 1 mL 纯水冲洗 1 次,将水溶液完全收集 后混匀,用 0.45 µm 的微孔滤膜过滤,得到待测液.

色谱进样分析条件:使用 Waters UPLC ACQUITY H-Class 超高效液相色谱系统(配紫外检测器, Empower 工作站),Waters BEH C18 柱(2.1 mm×50 mm,填料 1.7 μ m)色谱柱进行分析.流动相为 20% 乙腈水溶液:超纯水=15:85,流速为 0.8 mL/min,进样体积 10 μ L,柱温 30 ℃,检测波长为 229 nm,硫 苷各组分 30 min 内全部出峰、分离比较充分.

液相色谱数据处理:根据保留时间对硫苷分量色谱峰进行定性,对比内标峰面积和含量计算硫苷 组分质量摩尔浓度,采用 Microsoft Office Excel 软件对数据进行保留及处理,硫苷质量摩尔浓度结果 以 μmol/g 饼表示.

1.2.2 近红外扫描光谱的采集

使用 FOSS NIR Systems 5000 型近红外光谱仪收集样品的吸收光谱,光谱扫描范围为 1 100~2 498 nm,扫描次数为 64 次,分辨率 8 cm⁻¹,所有样品重复扫描 3 次,以减少仪器波动和装样对光谱扫描 的干扰.用 WinISI III软件对光谱进行判断,对 3 次扫描的光谱进行 AVERAGE 处理,平均值作为该样品 的吸收光谱.

1.3 近红外检测模型的建立

经由与近红外检测仪配套的化学计量学软件 WinISI Ⅲ分析, 剔除掉马氏距离 GH(Global H)值大于 3.0 的样品,并定义邻居距离 NH(Neighbor Hood)值在 0.8 范围内的样品为相似样品^[20],不参与最后定 标,由此将样品集分为校正集(建模集)和检验集.采用全光谱建立模型(Develop equations with full spectrum),回归技术采用改进最小二乘法(Modified partial least squares method,简称 MPLS)^[21],建模的光 谱区段设置为 1 308~2 393,8 nm.采用不同散射处理方式和导数处理方式对光谱数据进行预处理,得 到定标方程,验证集对模型进行验证.标准偏差 S_{SEC} (Standard Error of the calibration)、交叉验证误差 S_{SECV} (Standard Error of cross validation)所反映的是模型对自身校正样品的预测能力,由模型建立过程中 软件自动得出,值越接近 0 越好;定标相关系数 R_{RSQ} (R squared)与交叉验证相关系数 V_{FVR} (1 minus the variance ratio)为反映模型拟合情况的参数,也由软件在模型建立中得出,值越接近 1 越好.预测标准差 S_{SEP} (square error of prediction)是使用未参与校正的样品集(验证集)对模型评估得来的预测标准差,值越 接近 0 越好.

2 结果与分析

2.1 液相色谱检测结果

如图 1 所示,甘蓝型油菜硫苷组分在 30 min 内全部出峰、分离较为完全,满足试验分析要求.高效液相色谱数据处理结果见表 1,其中 2-羟基-3-丁烯基脱硫硫苷(Desulfoprogoitrin, DSPRO)平均

质量摩尔浓度最高,质量摩尔浓度变化幅度最大,平均质量摩尔浓度为 28.10 μ mol/g 饼,变幅为 0.60~83.44 μ mol/g 饼;其次是 3 -丁烯基脱硫硫苷(Desulfogluconapin, DSGNA),平均质量摩尔浓 度为 24.34 μ mol/g 饼,变幅为 0.66~73.51 μ mol/g 饼;而 4 -羟基-3 -吲哚甲基脱硫硫苷(Desulfo-4-hydroxyglucobrassicin, DS4-OHGBS)、4 -戊烯基脱硫硫苷(Desulfoglucobrassicanapin, DSGBN)和苯乙基脱硫硫苷(Desulfogluconasturtiin, DSGST)质量摩尔浓度较低,质量摩尔浓度变幅也很小.



保留时间参考: 4.19 min: 2-羟-3-丁烯基脱硫硫苷; 5.83 min: 丙烯基内标脱硫硫苷; 9.82 min: 3-丁烯基脱硫硫苷; 11.19 min: 4-羟-3-吲哚甲基脱硫硫苷; 12.75 min: 4-戊烯基脱硫硫苷; 14.49 min: 苯乙基脱硫硫苷.

图 1	硫苷组分的高效液相色谱图

组分	均值/(µmol•g ⁻¹)	变幅/(μ mol•g ⁻¹)	标准差	变异系数/%
DSPRO	28.10	0.60~83.44	23.48	83.56
DSGNA	24.34	0.66~73.51	18.49	75.97
DS4-OHGBS	5.17	0.51~24.13	3.64	70.41
DSGBN	3.41	0.25~17.33	3.35	98.24
DSGST	3.35	0.69~11.23	1.60	47.76
总 量	97.81	10.05~194.41	35.32	36.11

注:DSPRO:2-羟基-3-丁烯基脱硫硫苷;DSGNA:3-丁烯基脱硫硫苷;DS4-OHGBS:4-羟基-3-吲哚甲基脱硫硫苷;DSGBN:4-戊烯基脱硫硫苷;DSGST:苯乙基脱硫硫苷.

2.2 定标方程的建立

图 2 为甘蓝型油菜籽样品的近红外漫反射光谱.选择与近红外检测仪配套的化学计量学软件WinISI III 分析,将实验值输入对应的光谱位置,剔除超常样品并选择出代表性样品,采用改进最小二乘法建立甘蓝型油菜硫苷组分质量摩尔浓度的近红外检测模型.

本研究得到了 89 份具有代表性的样品组成校正集,分别建立不同散射处理方式(Detrend only:去散

射处理; None: 无散射处理; SNV[Standard Normal Variant]only: 校正正常化处理; SNV+Detrend: 标准正常化+散射处理.)和导数处理方式(1阶、2阶、3阶)下2-羟基-3-丁烯基脱硫硫苷、3-丁烯基脱硫 硫苷、4-羟基-3-吲哚甲基脱硫硫苷、4-戊烯基脱硫硫苷、苯乙基脱硫硫苷质量摩尔浓度的近红外检测模型,选择 V_{1-VR}最大、S_{SECV}最小的模型为该硫苷组分的最佳模型(表 2).



图 2 样品的近红外漫反射光谱

表 2 内部交叉检验结果

组分	校正数	平均值	校正标	定标相	内部交叉	内部交叉验	业游处理大学
	/个	/ %	准偏差	关系数	验证误差	证相关系数	兀谙处理万式
DSPRO	86	26.745 1	3.909 2	0.973	5.978 2	0.937 2	SNV+Detrend, 2,4,4,1
DSGNA	87	23.084 5	4.087 6	0.956	6.232	0.900 9	None, 1,4,4,1
DS4-OHGBS	86	5.615 7	2.701 1	0.676	3.2877	0.304 3	Detrend only, 2,4,4,1
DSGBN	80	3.421 4	1.945 3	0.295	2.120 3	0.128 1	SNV+Detrend, 1,4,4,1
DSGST	87	3.322 3	1.252 7	0.264	1.249 9	0.175 1	SNV+Detrend, 2,4,4,1

由表 2 可知, 2 -羟基-3 -丁烯基脱硫硫苷和 3 -丁烯基脱硫硫苷模型的内部交叉检验效果都较好, V_{1-VR} 分别为 0.937 2,0.900 9, S_{SEC} 分别为 3.909 2,4.087 6, S_{SECV} 分别为 5.978 2,6.232 0, 表明该方法可以准确地预测这两种硫苷组分的质量摩尔浓度;而 4 -羟基-3 -吲哚甲基脱硫硫苷、4 -戊烯基脱硫硫苷、苯乙基脱硫硫苷预测模型不太理想, V_{1-VR} 分别为 0.304 3,0.128 1 和 0.175 1, S_{SEC} 分别为 2.701 1,1.945 3 和 1.252 7, S_{SECV} 分别为 3.287 7,2.120 3 和 1.249 9.

2.3 定标方程的验证

利用 WinISI Ⅲ软件中的结果监控软件,将没有参与定标的 177 份样品作为验证集,检验所建立的模型的效果.校正效果的好坏常用外部检验相关系数 R_{RSQ}、预测标准偏差 S_{SEP}、相对误差和检验偏差 B_{Bias}等表示.其中 S_{SEP}是使用未参与校正的样品集(验证集)对模型评估得来的预测标准差,是模型评价的最直接标准.选择 R_{RSQ}趋近 1、S_{SEP}和 B_{Bias}最小的模型(表 3 和图 3),表示模型在实际应用中的可行性强^[22].

			夕	下 部 检	验		
组分	验证集	平均	值/%	外部检验	预测	相对误差	扒扒伯士
	/个	Lab	NIR	相关系数	标准差	/ 1/0	位短偏差
DSPRO	152	21.899	22.628	0.962	4.703	<3.32	-0.729
DSGNA	151	18.291	19.475	0.937	4.650	<6.47	-1.185
DS4-OHGBS	154	5.066	5.153	0.470	2.551	<1.72	-0.087
DSGBN	138	2.868	3.249	0.219	1.931	<13.28	-0.382
DSGST	154	3.223	3.448	0.147	1.275	<6.98	-0.225

表 3 外部检验结果

注:相对误差=(示值-标准值)/真实值(即绝对误差所占真实值的百分比).

从表 3 中可看出, 2 -羟基-3 -丁烯基脱硫硫苷和 3 -丁烯基脱硫硫苷质量摩尔浓度模型的 R_{RSQ}较高, 分别为 0.962 和 0.937, S_{SEP}分别为 4.703, 4.650, 相对误差分别小于 3.32%, 6.47%, 检验偏差均很 小,分别为-0.729, -1.185, 说明模型的结果比较令人满意. 而 4 -羟基-3 -吲哚甲基脱硫硫苷、4 -戊 烯基脱硫硫苷、苯乙基脱硫硫苷质量摩尔浓度模型的 R_{RSQ}较小,分别为 0.470, 0.219, 0.147, S_{SEP}分别 为 2.551, 1.931, 1.275, 相对误差分别小于 1.72%, 13.28%, 6.98%, 检验偏差分别为-0.087, -0.382和-0.225, 结果比较不理想.



图 3 甘蓝型油菜硫苷各组分质量摩尔浓度化学测试值和 NIRS 预测值的关系

观察所得样品散点图(图 3)可以看出,样品硫苷组分的模型预测值与化学测试值有明显的相关关系. 其中 2-羟基-3-丁烯基脱硫硫苷和 3-丁烯基脱硫硫苷质量摩尔浓度模型预测值与化学测试值的契合程度 较高,检验样品多集中在中心线附近,表明这 2个模型的分析结果具有较高的测试精确度,而 4-羟基-3-吲哚甲基脱硫硫苷、4-戊烯基脱硫硫苷、苯乙基脱硫硫苷质量摩尔浓度模型预测值和化学测试值的吻合程 度不高,检验样品分布也比较散,说明这3个模型在实际应用过程中预测性能较差.

3 讨 论

建立近红外检测模型主要采用多元线性回归(MLR)、逐步回归(SMR)、主成分回归(PCR)、改进最小 二乘法(MPLS)^[21]等方法.选用 MPLS 的主要原因:偏最小二乘法(PLS)是一种新型的多元统计数据分析 方法,它利用全部光谱信息,需要的样品数量较少,却可以同时对多个组分建模,适合于处理变量较多而 样本较少的分析,信息提取效率高^[23];改进最小二乘法(MPLS)还能对重点数据进行加权,MPLS 所得模 型具有较高的预测稳定性.

在本研究中,2-羟基-3-丁烯基脱硫硫苷和3-丁烯基脱硫硫苷质量摩尔浓度模型预测性能较好,可 以用于育种,而4-羟基-3-吲哚甲基脱硫硫苷、4-戊烯基脱硫硫苷、苯乙基脱硫硫苷质量摩尔浓度模型预 测性能较差,无法使用.这与Font^[15]以及 Velasco等^[17]分别建立的油菜籽主要硫苷类型油菜籽饼质量摩尔 浓度(μmol/g 饼)或者总硫苷百分比(mol/mol)NIRS 分析定标方程结果一致.在很多研究中,由于质量摩 尔浓度较低、变幅很小,棕榈酸、硬脂酸、花生烯酸等模型的精准度都会较差^[17,24].值得注意的是,Velasco等^[25]采用 21 种芸薹属植物种子混合校正,很好地解决了油菜籽某些类型硫苷因质量摩尔浓度低、变异 范围小而无法建模的问题.

近红外光谱分析技术是一种间接的相对分析,需要通过建立检测模型来测定未知样品成分,而且分析 结果受代表性样品的选择、近红外光谱准确性等因素的影响^[26].本研究借助 WinISI III 软件,利用样品光谱 之间的相似程度来选取样品,其效果优于浓度选取方法^[20].近红外检测模型有很强的专一性和地域性,本 研究所建立的硫苷组分模型仅适用于重庆地区.

参考文献:

[1] 唐春艳,齐德生.双低菜籽粕在饲料中的应用 [J]. 中国饲料, 2005(21): 11-14.

- [2] 高玉鹏, 薛雅蓉, 张士贤. 双低油菜籽粕营养成分及饲喂育成鸡研究 [J]. 西南农业大学学报, 1996, 18(3): 281-187.
- [3] 李培武,杨 湄,张 文,等.我国油菜产品质量安全现状及对策 [J].中国油料作物学报,2004,26(1):84-88.
- [4] 盖钧镒.作物育种学各论 [M].北京:中国农业出版社,2004.
- [5] 祝诗平,梁 晶,屠大伟,等. 基于近红外光谱与 DPLS 的潲水油快速鉴别方法 [J]. 西南大学学报:自然科学版, 2012, 34(5):1-6.
- [6] 高建芹,张洁夫,浦惠明,等.近红外光谱法在测定油菜籽含油量及脂肪酸组成中的应用 [J]. 江苏农业学报,2007, 23(3):189-195.
- [7] 李长治.分子光谱新技术 [M].北京:科学出版社,1986:97-98.
- [8] 李 伟,肖爱平,冷 鹃.近红外光谱技术及其在农作物中的应用[J].中国农学通报,2009,52(3):56-59.
- [9] 闫 龙,蒋春志,于向鸿,等.大豆粗蛋白、粗脂肪含量近红外检测模型建立及可靠性分析 [J].大豆科学,2008, 27(5):833-837.
- [10] 张哗晖,赵龙莲,李晓薇.用傅里叶变换近红外光谱法测定完整油菜籽三种品质性状的初步研究 [J].激光生物学报, 1998,7(2):138-141.
- [11] 杨燕宇,陈社员.油菜品质近红外反射光谱分析建模及应用研究进展 [J].作物研究,2007,21(2):152-156.
- [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 [J]. 作物学报,2006,32(5):709-715.
- [13] 陈文杰, 谭小力, 王竹云, 等. 用傅立叶变换近红外光谱仪测定油菜种子品质指标的研究 [J]. 陕西农业科学, 2002(8): 6-9.
- [14] 杨翠玲,陈文杰,赵兴忠,等.近红外光谱法同时分析油菜9种品质参数的研究[J].西北农林科技大学学报:自然科学版,2006,34(3):61-67.
- [15] FONT R, DEL. RÍO M, DOMINGUE Z J, et al. Using Nirs for Determining Glucosinolate Content in Brassica juncea Seed [C]. Canberra: Proceedings of the 10th International Rapeseed Congress, 1999: 51-53.
- [16] VELASCO L, BECKER H C. Estimating the Fatty Acid Composition of the Oil in Intact-Seed Rapeseed (Brassica napus L.)

by Near Infrared Reflectance Spectroscopy [J]. Euphytica, 1998, 101(2): 221-230.

- [17] VELASCO L, BECKER H C. Analysis of Total Glucosinolate Content and Individual Glucosinolates in Brassica spp. by Near-Infrared Reflectance Spectroscopy [J]. Plant Breeding, 1998, 117(2): 97-102.
- [18] 王 瑞,徐新福,李加纳,等.甘蓝型油菜硫苷组分的胚、细胞质和母体遗传效应分析 [J].作物学报,2007,33(12): 2001-2006.
- [19] 董 莉,任雪松,李成琼,等.甘蓝硫甙组分和含量分析 [J].西南大学学报:自然科学版,2012,34(12):34-38.
- [20] 吉海彦, 严衍禄. 用主成分回归和偏最小二乘法定量测定谷物成分 [J]. 北京农业大学学报, 1994, 20(1): 59-63.
- [21] 李 勇,魏益民,王 锋.影响近红外光谱分析结果准确性的因素 [J].核农学报, 2005, 19(3): 236-240.
- [22] Olsson G. Some Relations Between Number of Seeds Per Pod, Seed Size and Oil Content and the Effects of Selection for These Characters in Brassica and Sinapis [J]. Hereditas, 1960, 46(1-2): 29-70.
- [23] 严衍禄. 近红外光谱分析基础与应用 [M]. 北京:中国轻工业出版社, 2005: 182-183.
- [24] 尚国霞. 甘蓝型油菜高油酸性状的遗传研究及近红外检测模型的建立 [D]. 重庆: 西南大学, 2010.
- [25] VELASCO L, MOLLERS C, BECKER H C. Analysis of Individual Glucosinolates in Brassica spp. by Near-Infrared Reflectance Spectroscopy [C]. Canberra: Proceedings of 10th International Rapeseed Conyless, 1999: 26-29.
- [26] WILLIAMS P, NORRIS K. Near Infrared Technology in the Agriculture and Food Industries [M]. Washington, D. C. : American Association of Cereal Chemists, Inc, 1987: 35-37.

The Construction of a Near Infrared Inspection Model for Glucosinolate Constituent Contents in *Brassica napus* L.

WANG Nai-hong, ZHANG Min, WANG Rui, LI Shi-meng, LIU Chuan

School of Agronomy and Biotechnology, Southwest University, Chongqing 400716, China

Abstract: To explore the feasibility of rapidly detecting the contents of glucosinolate constituents with near infrared reflectance spectroscopy, we selected 266 representative samples, and analysed their glucosinolates constituent contents with high performance liquid chromatography (HPLC) and processed the data, using an NIR spectrophotometer and the WinISIII software. In the end, a near infrared inspection model was constructed for glucosinolates constituent contents in *Brassica napus* L. The results showed that the cross-validation correlation coefficient (1-VR) of the NIR spectroscopy model for desulfoprogoitrin (DSPRO) and desulfogluconapin (DSGNA) was 0.937 2 and 0.900 9 and the external validation correlation coefficient (RSQ) was 0.962 and 0.937, respectively, which suggested that the model was satisfactory and could be used for the rapid screening of rapeseed varieties. However, the RSQ of desulfo-4-hydrox-yglucobrassicin (DS4-OHGBS), desulfoglucobrassicanapin (DSGBN) and desulfogluconasturtiin (DSGST) was 0.147, 0.219 and 0.419 and their 1-VR was 0.175, 0.128 and 0.304, respectively, which result was unsatisfactory and needed further research.

Key words: Brassica napus L.; glucosinolate constituent; near infrared (NIR)