

Rapid Identification of Insecticide- and Herbicide-Tolerant Genetically Modified Maize Using Mid-Infrared Spectroscopy

Authors:

Xiaodan Liu, Yonghui Yu, Xiulin Bai, Xiaolong Li, Jun Zhang, Dun Wang

Date Submitted: 2023-02-17

Keywords: genetically modified maize, identification, mid-infrared spectroscopy, chemometric methods

Abstract:

Genetically modified (GM) technology is of great significance for increasing crop production, protecting biodiversity, and reducing environmental pollution. However, with the frequent occurrence of safety events regarding GM foods, more and more disputes have arisen over the potential safety of transgenic technology. It is particularly necessary to find a fast and accurate method for transgenic product identification. In this research, mid-infrared spectroscopy, coupled with chemometric methods, was applied to discriminate GM maize from its non-GM parent. A total of 120 GM maize and 120 non-GM maize samples were prepared, and the spectral information in the range of 4000–400 cm⁻¹ was collected. After acquiring the spectra, wavelet transform (WT) was used to preprocess the data, and k-means was carried out to split all samples into calibration and prediction sets in the ratio of 2:1. Principal component analysis (PCA) was then conducted to qualitatively distinguish the two types of samples, and an apparent cluster was observed. Since the full spectrum covered a large amount of data and redundant information, we adopted the successive projections algorithm (SPA) to select optimal wavelengths for further analysis. Chemometrics, including partial least squares-discriminant analysis (PLS-DA), the k-nearest neighbor algorithm (KNN), and the extreme learning machine (ELM), were performed to establish classification models based on full spectra and optimal wavelengths. The overall results indicated that ELM models based on full spectra and optimal spectra showed better accuracy and reliability, with a 100% recognition rate in the calibration set and a 98.75% recognition rate in the prediction set. It has been confirmed that mid-infrared spectroscopy, combined with chemometric methods, can be a novel approach to identify transgenic maize.

Record Type: Published Article

Submitted To: LAPSE (Living Archive for Process Systems Engineering)

Citation (overall record, always the latest version):

LAPSE:2023.0125

Citation (this specific file, latest version):

LAPSE:2023.0125-1

Citation (this specific file, this version):

LAPSE:2023.0125-1v1

DOI of Published Version: <https://doi.org/10.3390/pr11010090>

License: Creative Commons Attribution 4.0 International (CC BY 4.0)

Article

Rapid Identification of Insecticide- and Herbicide-Tolerant Genetically Modified Maize Using Mid-Infrared Spectroscopy

Xiaodan Liu ^{1,2}, Yonghui Yu ¹, Xiulin Bai ³ , Xiaolong Li ³, Jun Zhang ⁴ and Dun Wang ^{5,*}¹ College of Food Science and Engineering, Wuhan Polytechnic University, Wuhan 430023, China² Key Laboratory for Deep Processing of Major Grain and Oil (Wuhan Polytechnic University), Ministry of Education, Wuhan 430023, China³ College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310058, China⁴ College of Mechanical and Electrical Engineering, Jiaying Nanhu University, Jiaying 314001, China⁵ Xiangyang Academy of Agricultural Sciences, Xiangyang 441057, China

* Correspondence: wangdunzju@163.com; Tel.: +86-710-3085380

Abstract: Genetically modified (GM) technology is of great significance for increasing crop production, protecting biodiversity, and reducing environmental pollution. However, with the frequent occurrence of safety events regarding GM foods, more and more disputes have arisen over the potential safety of transgenic technology. It is particularly necessary to find a fast and accurate method for transgenic product identification. In this research, mid-infrared spectroscopy, coupled with chemometric methods, was applied to discriminate GM maize from its non-GM parent. A total of 120 GM maize and 120 non-GM maize samples were prepared, and the spectral information in the range of 400–4000 cm⁻¹ was collected. After acquiring the spectra, wavelet transform (WT) was used to preprocess the data, and k-means was carried out to split all samples into calibration and prediction sets in the ratio of 2:1. Principal component analysis (PCA) was then conducted to qualitatively distinguish the two types of samples, and an apparent cluster was observed. Since the full spectrum covered a large amount of data and redundant information, we adopted the successive projections algorithm (SPA) to select optimal wavelengths for further analysis. Chemometrics, including partial least squares-discriminant analysis (PLS-DA), the k-nearest neighbor algorithm (KNN), and the extreme learning machine (ELM), were performed to establish classification models based on full spectra and optimal wavelengths. The overall results indicated that ELM models based on full spectra and optimal spectra showed better accuracy and reliability, with a 100% recognition rate in the calibration set and a 98.75% recognition rate in the prediction set. It has been confirmed that mid-infrared spectroscopy, combined with chemometric methods, can be a novel approach to identify transgenic maize.

Keywords: genetically modified maize; identification; mid-infrared spectroscopy; chemometric methods

Citation: Liu, X.; Yu, Y.; Bai, X.; Li, X.; Zhang, J.; Wang, D. Rapid Identification of Insecticide- and Herbicide-Tolerant Genetically Modified Maize Using Mid-Infrared Spectroscopy. *Processes* **2023**, *11*, 90. <https://doi.org/10.3390/pr11010090>

Academic Editor: Ángeles Alonso-Moraga

Received: 4 December 2022

Revised: 26 December 2022

Accepted: 26 December 2022

Published: 29 December 2022



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Maize (*Zea mays L.*) is one of the most important grain and cash crops in the world, and it also serves as an important livestock feed and industrial raw material, playing an important role in agricultural and industrial production. To meet some specific requirements and improve the yield and quality of maize, the genetically modified (GM) technology has been applied to maize due to the advantage of being able to cultivate an elite crop variety with a high yield and high resistance [1,2]. High quality maize varieties exhibiting herbicide, disease, and insect resistance were developed over the past decades [3,4]. However, there are still some disputes regarding the use of GM technology due to the frequent occurrence of safety events regarding genetically modified food. Therefore, the screening and supervision of genetically modified products based on GM maize is very important to ensure food safety and guarantee the rights and interests of consumers. Consequently, it

is urgently required to develop a fast and reliable approach to discriminate GM products based on GM maize.

Traditional genetically modified organism detection methods are conducted based on exogenous DNA and protein, including polymerase chain reaction (PCR)/restriction enzyme assay, enzyme-linked immunosorbent assays, lateral flow strip, and microarray [5–7]. Generally, these methods are sensitive and accurate in most cases, but they are also time-consuming, environmentally unfriendly, and require highly skilled operators, making them unsuitable for fast online detection. Recently, with the advantages of fast, environmentally friendly, easy to operate, spectral techniques (requiring no complex sample pretreatments), including near-infrared spectroscopy, mid-infrared spectroscopy, hyperspectral imaging, and terahertz spectroscopy have been widely used in agriculture and food areas [5,8–13], including for the detection of GM products, and good results have been achieved. Mid-infrared spectroscopy provides the information about the fundamental vibrational frequency of the functional groups, which could be used as supplementary information for other spectral technologies. Moreover, the vibrational absorption frequencies of chemical bonds in organic molecules highly related to genes occur mainly in the mid-infrared band [14], which makes mid-infrared spectroscopy an advantage in analyzing the physical properties and chemical composition of GM samples [15]. To the best of our knowledge, no attempt has been made to identify GM maize by mid-infrared spectroscopy.

The purpose of this study is to explore the feasibility of mid-infrared spectroscopy for identifying GM maize, providing a potential method for the rapid and accurate identification of GM products. In this work, the mid-infrared spectra of GM maize and non-GM maize were directly compared. Optimal wavelengths were selected based on the successive projections algorithm (SPA). Chemometric methods of principal component analysis (PCA), partial least squares-discriminant analysis (PLS-DA), the k-nearest neighbor algorithm (KNN), and extreme learning machine (ELM) were conducted to establish the qualitative and quantitative identification models for GM maize. The performance of the chemometric models was further compared to select the most effective method for the identification of GM maize.

2. Materials and Methods

2.1. Analytical Procedure for the Identification of Maize

The specific analytical procedure is given in Figure 1. A series of five steps were conducted, including sample preparation, spectra acquisition, PCA analysis, variable selection, and modeling based on chemometric methods.

2.2. Maize Samples

The genetically modified (GM) maize kernels and their non-GM parents were acquired from the Institute of Insect Sciences, Zhejiang University, China. The GM maize kernels containing insecticide- and herbicide-tolerant traits were obtained by transferring the cry1Ab/cry2Aj-G10evo gene, and the detailed flow chart of gene insertion can be found in the reference [16]. Moreover, no other differences existed between the GM and non-GM maize kernels, according to the results of molecular biology experiments. In our experiment, 600 GM maize kernels and 600 non-GM maize kernels were prepared and dried in an oven at 60 °C for 6 h. Every 5 maize kernels were milled into flour by a grinder, and 0.02 g maize flour was selected and evenly mixed with 0.98 g of potassium bromide (KBr; SCR, Shanghai, China) powders [16,17], and the sample was dried in an oven at 80 °C for 6 h and cooled in a desiccator. A total of 0.2 g of the mixed sample was then crushed into a square tablet 10 mm in side length and 2 mm in thickness by a pressing machine. Thus, a total of 240 tablet samples were obtained for spectral acquisition.

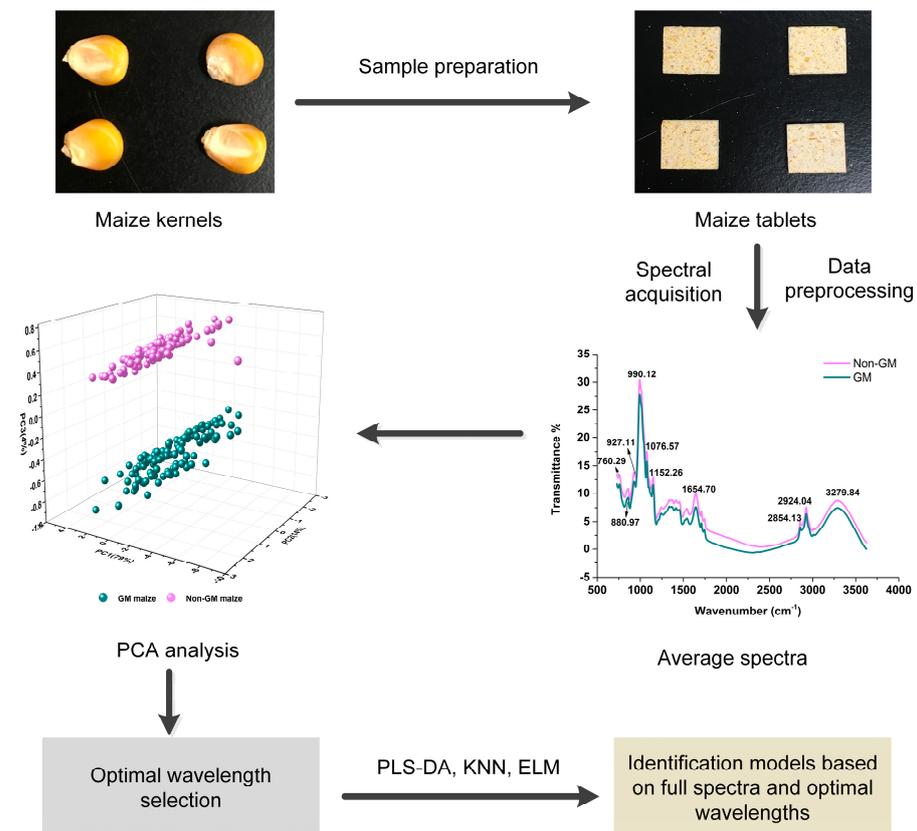


Figure 1. Analytical procedure for the identification of maize.

2.3. Spectral Acquisition

A Jasco FT/IR-4100 spectrometer (Japan) with a resolution of 4 cm^{-1} was used for the collection of the transmittance MIR spectral data of the samples. Specifically, the sample tablets were placed in the detection platform, and 32 scans were conducted in the spectral range of $400\text{--}4000\text{ cm}^{-1}$ for each sample. The average of the 32 spectra was used as the transmittance spectrum of the sample for further analysis.

2.4. Data Analysis

2.4.1. Data Preprocessing

Considering that the interference of the external environment during the experiment will negatively cause noise in the raw spectrum, the wavelet transform (WT) approach, as an efficient and widely used denoising method [18], was used to reduce the noise and improve the accuracy of data analysis. Daubechies 4 wavelets, with decomposition scale of 6, were optimized.

2.4.2. Chemometric Methods

Chemometric methods including variable selection, as well as qualitative and quantitative modeling, are widely used in spectral data analysis due to the large amount of original data with the interference of redundant information. In our propositions, SPA, PCA, PLS-DA, KNN, and ELM were applied to extract important spectral information and establish accurate qualitative and quantitative analysis models of GM maize.

SPA is a forward variable selection algorithm based on a projection analysis of the vector, which has been widely applied for variable selection in spectra [19]. The detailed introduction of SPA can be found in reference [20]. In this study, the combination of variables with the least redundant information and the least collinearity were selected for further modeling to improve model accuracy.

PCA is a data compression and dimensionality reduction algorithm which has been commonly used for the qualitative discrimination of spectral data [21]. Principal components (PCs) were important variables in PCA analysis obtained by orthogonal transformation of original variables [22]. Generally, the first few PCs with a large contribution rate can explain the main information of the original data. In this study, the score plot of the first three PCs was mapped to qualitatively distinguish GM maize and non-GM maize.

PLS-DA is a multivariate statistical analysis algorithm based on PCA, as well as a supervised pattern recognition method [23], which has been widely used in the classification analysis of spectral data due to its provision of intuitive and clear analysis results and its high prediction accuracy. Latent variables can reflect most information of origin data, and they are important variables in the PLS-DA model [24]. In this study, the number of latent variables is optimized by the cross validation method when the minimal predicted residual sum of squares (PRESS) is obtained. Besides, the PLS-DA model was established based on leave-one-out cross validation, and the spectral data of samples were used as input variables, while the category of the sample was used as a response variable. The threshold of 0.5 [15] was utilized to distinguish the prediction results and calculate the final prediction accuracy of the model.

KNN is one of the simplest classification methods in data mining, and it has been commonly applied in the cluster analysis of spectral data. The classification of the target sample was realized by calculating the distance between all samples and determining the main category of K samples with the smallest distance from the target sample; thus, the target sample will finally be classified as the main category of the K samples [25,26]. Generally, a good discriminant result is highly related to the K value, which is mainly selected by the continuous training method.

ELM is a machine learning method based on a feedforward neural network, and it is suitable for both supervised and unsupervised learning problems [27]. The structure of the ELM model consists of three layers: the input layer, the hidden layer, and the output layer. The number of hidden layer nodes is a critical parameter, and it has been optimized by continuous training in a predefined range that does not exceed the number of samples of the calibration set.

2.5. Software Tools

Spectra data acquisition was conducted in Spectra manager CFR software. Unscrambler X10.1 (CAMO, Process, AS, OSLO, Norway) was used for PCA and PLS-DA analysis. MATLAB R2009a (v7.8, The Math Works, Inc., Natick, MA, USA) was also used for data analysis including WT, SPA, KNN, and ELM. Additionally, graph designing was realized by Origin Pro 8.0 SR0 (Origin Lab Corporation, Northampton, MA, USA).

3. Results

3.1. Spectral Analysis

The average spectra of GM maize and non-GM maize are shown in Figure 2. As the front and rear parts of the spectra in the range of 400–4000 cm^{-1} showed high noise levels caused by the interference of the external environment, only the spectra of 728.48–3625.52 cm^{-1} were used for analyses, and 6010 variables were included. Besides, WT was conducted to reduce the noise. It is obvious that the trends of the spectra of the two varieties were similar, but the average reflectance value of the non-GM samples was always higher than that of the GM samples, which might be related to the differences in the physical and chemical components due to the genotypic difference of samples. The typical wavelengths were marked. According to previous research, the wavenumber around at 760.29 cm^{-1} and 880.97 cm^{-1} (in the range of 700–900 cm^{-1}) represents the C-H deformation vibration [28]. The peak at 927.11 cm^{-1} is associated with the C-H bending of carbohydrates, and the peak observed near 991 cm^{-1} (990.12 cm^{-1}) is characteristic of the C-O stretching vibration in the C-OH group and C-C stretching vibration in the carbohydrate [29]. The peak at 1076.57 cm^{-1} may be responsible for the C-C, C-O, C-O-C stretching vibration of carbohydrates [30],

and the band near 1152 cm^{-1} (1152.26 cm^{-1}) is related to C-O-C vibration in cellulose and hemicelluloses [31]. The band at 1654.70 cm^{-1} is due to the amide I peak of proteins [32], and the peak ranging between $2853\text{--}2855\text{ cm}^{-1}$ (2854.13 cm^{-1}) may be assigned to the symmetric CH(-CH₂-) stretching vibration [33]. In addition, the peak of 2924.04 cm^{-1} is related to the asymmetric CH(-CH₂-) stretching vibration, while the wavenumber of 3279.84 cm^{-1} is related to the O-H stretching vibration of the free hydroxyl group [33]. As a whole, the wavelengths reflected the differences of the two kinds of samples in carbohydrate, cellulose and hemicelluloses, and proteins, which further confirmed the difference in the metabolites of the introduced foreign genes, and the results have been proven by previous researches [34–37]. In spite of this, it was hard to discriminate GM samples from their non-GM parents using only the mid-infrared spectral transmittance. Therefore, chemometrics methods, in combination with mid-infrared spectra, were applied for identification analysis of GM maize samples.

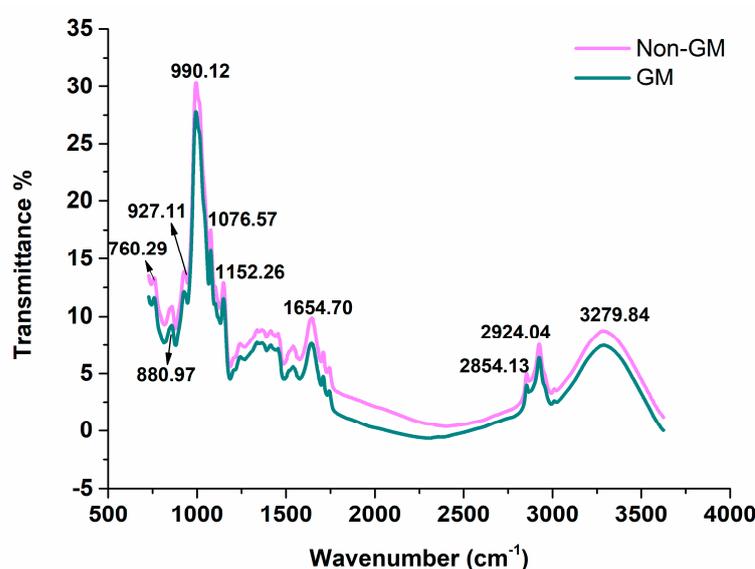


Figure 2. Average spectra transmittance of GM and non-GM maize ($728.48\text{--}3625.52\text{ cm}^{-1}$).

3.2. PCA Analysis

PCA was conducted to visualize the distribution of GM and non-GM maize based on mid-infrared spectral information. Since the first three principal components (PCs) explained 97% (79%, 14%, and 4% for PC1, PC2, and PC3, respectively) of the variables of the full spectra, they were selected as representative of the original data for further analysis. The score plots of the first three PCs for the maize samples are shown in Figure 3. It could be found that the two classes were well-separated by PC3, while there were overlaps in PC1 and PC2, which indicated that the spectral fingerprints of PC3 carried most of discriminant information for GM maize. Due to the fact that the value of the PCA loading is highly related to the degree of correlation between the PCs and the raw variables [15], the loading of PC3 is provided in Figure 4 to explain the variation. It is obvious that main peaks of PC3 loading were consistent with the peaks marked in Figure 2, which are attributed to the vibrational absorption of chemical bonds in the organic molecules. Besides, the unincluded peaks at 1480.58 cm^{-1} are also related to the amide II band of protein [38]. As a whole, the results confirmed the feasibility of mid-infrared spectroscopy for GM maize identification. Therefore, additional discriminant models for GM maize were built, and it is necessary to select relevant variables for modeling to improve the accuracy of the discriminant models.

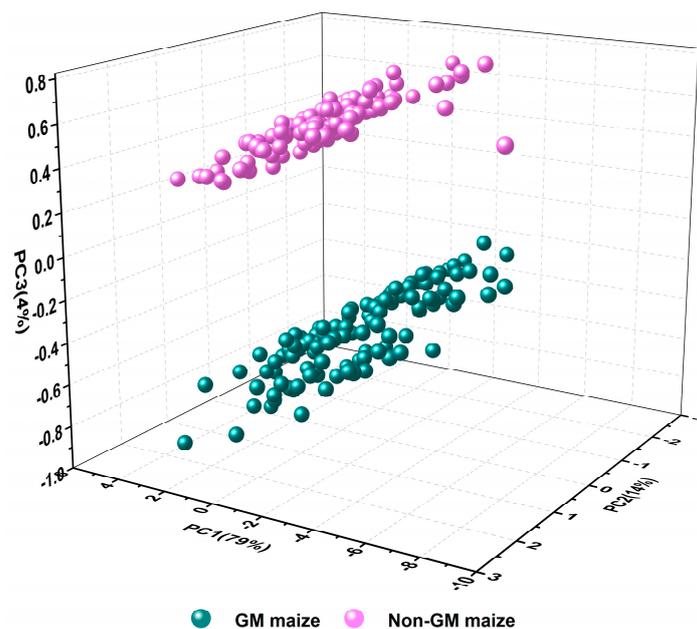


Figure 3. Three-dimensional principal component analysis (PCA) scores scatter plot.

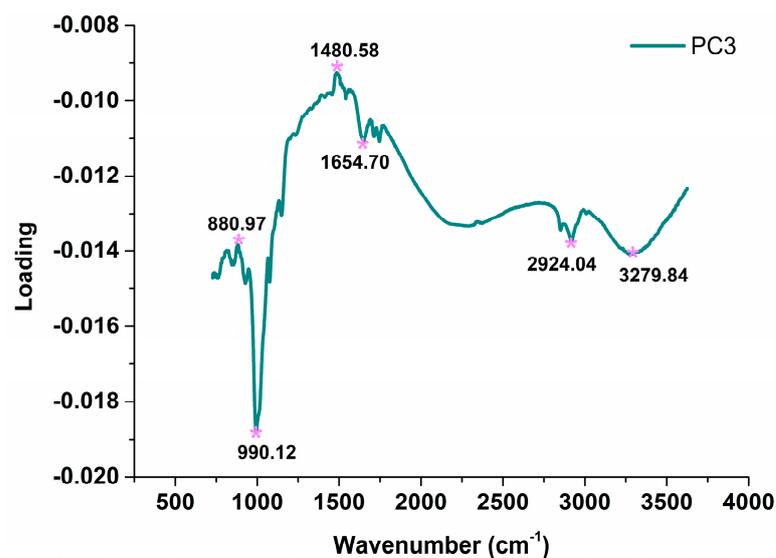


Figure 4. The loading plot of PC3.

3.3. Variable Selection

The selected full mid-infrared spectra contain over 6000 variables, not all of which are highly related to the identification of GM maize; moreover, redundant variables will affect the working speed and performance of the discriminant model. Thus, SPA was utilized to select the optimal wavelengths for modeling. The results of characteristic wavelength selection are given in Figure 5, and 14 optimal wavelengths were selected. The band at 2480.29 cm^{-1} was excluded due to the interference of H_2O [39]. The remaining 13 optimal wavelengths are mainly associated with the organic functional groups, including C-H, C-O, and C=O, which are presented in Table 1. Moreover, the average spectra transmittance of the optimal wavelengths of GM and non-GM maize were provided, and a statistical significance analysis was conducted. The results are listed in Table 1. The p -values of most of the optimal wavelengths of GM and non-GM maize were lower than 0.05, except for the band at 992.68 cm^{-1} and 3618.77 cm^{-1} , indicating that there was a significant difference in

most of the optimal wavelengths between GM and non-GM maize, which further confirmed the difference in metabolites caused by the genotypes.

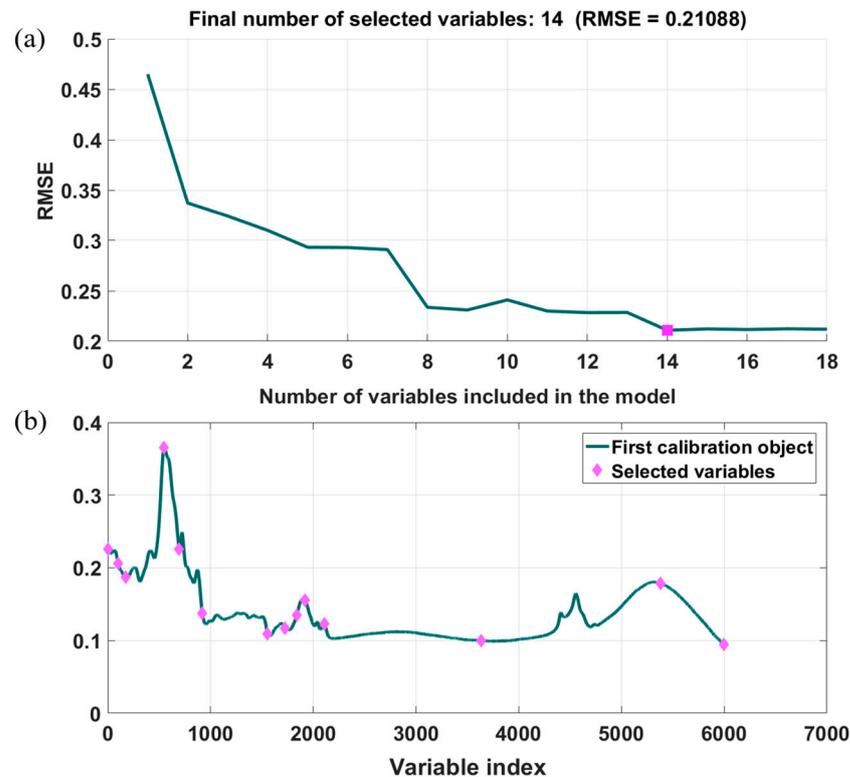


Figure 5. Optimal wavelength selection by the successive projections algorithm (SPA): (a) final number of selected variables; (b) selected variables.

Table 1. Optimal variables and average spectra transmittance.

Wavelength (cm ⁻¹)	Assignment	Average Spectra Transmittance % (GM Maize)	Average Spectra Transmittance % (non-GM Maize)	p-Value
730.41	deformation vibrations of the C-H group [28]	11.79 ± 2.89	13.06 ± 4.56	1.56 × 10 ⁻²
776.69	deformation vibrations of the C-H group [28]	9.85 ± 2.64	11.19 ± 4.32	6.51 × 10 ⁻³
811.40	deformation vibrations of the C-H group [28]	7.82 ± 2.41	9.24 ± 4.04	2.02 × 10 ⁻³
992.68	stretching of C-O in the C-OH group [40]	28.77 ± 6.08	29.35 ± 8.59	5.76 × 10 ⁻¹
1063.07	C-O stretching vibrations [28]	13.31 ± 2.87	14.41 ± 4.43	3.52 × 10 ⁻²
1171.54	CO-O-C asymmetric stretching vibrations [31]	5.71 ± 1.78	6.87 ± 3.15	1.30 × 10 ⁻³
1477.69	C=C stretching vibrations [28]	4.85 ± 1.43	6.09 ± 2.62	2.46 × 10 ⁻⁵
1559.17	C=C stretching vibrations [28]	5.05 ± 1.50	6.43 ± 2.70	5.82 × 10 ⁻⁶
1615.09	aromatic C=C stretching vibrations [28]	6.19 ± 1.67	7.74 ± 2.85	2.07 × 10 ⁻⁶
1653.18	stretching vibrations of C=O group [31]	7.66 ± 1.83	9.25 ± 3.03	6.34 × 10 ⁻⁶
1745.74	vibrations of the carbonyl group (-C=O) [41]	3.63 ± 1.80	5.32 ± 3.10	1.73 × 10 ⁻⁶
3320.34	O-H stretching vibrations [30]	7.19 ± 2.17	8.65 ± 3.82	7.25 × 10 ⁻⁴
3618.77	O-H stretching vibrations [30]	0.93 ± 0.45	0.94 ± 0.66	9.24 × 10 ⁻¹

3.4. Classification Models Based on Optimal Wavelengths and Full Spectra

PLS-DA, KNN, and ELM identification models, based on full spectra and optimal wavelengths, were further developed to classify GM and non-GM maize, and the correct recognition rate was used as the evaluation index of models. Before modeling, K-means was applied to divide a total of 240 maize samples into a calibration set (80 for both GM and non-GM maize) and a prediction set (40 for both GM and non-GM maize) in

a ratio of 2:1. The discrimination results of all identification models are summarized in Table 2. It can be clearly seen that models based on the full spectra obtained good results, with a recognition accuracy of over 80% in both the calibration and prediction sets. Compared with the models based on full spectra, the PLS-DA model, based on optimal spectra, showed better performance, with a recognition accuracy of over 96% in both the calibration and the prediction sets. However, the optimal wavelengths-based KNN model performed the worst. Compared with the PLS-DA and KNN models, the ELM model showed relatively better performance. Particularly, the ELM models based on full spectra and optimal spectra both yielded a 100% recognition rate in the calibration set and 98.75% recognition rate in the prediction set. Generally, these results indicated that GM maize could be accurately distinguished using mid-infrared spectroscopy combined with appropriate chemometric methods.

Table 2. Identification results of classification models based on optimal wavelengths and full spectra.

Input Variables	Discrimination Model	Calibration Identification Rate	Prediction Identification Rate
Full spectra	PLS-DA	90%	80%
	KNN	95.00%	93.75%
	ELM	100.00%	98.75%
SPA	PLS-DA	98.13%	96.25%
	KNN	79.38%	77.50%
	ELM	100.00%	98.75%

4. Conclusions

This research focused on the discrimination of GM maize kernels by using mid-infrared spectroscopy combined with chemometric methods. A three-dimensional PCA scores scatter plot based on the first three PCs showed the feasibility of GM maize identification, and 13 optimal wavelengths were selected based on SPA. Discrimination models based on optimal wavelengths and full spectra were established. It was found that the performances of full spectra-based models were better than those of the optimal wavelengths-based models. Moreover, the outstanding result was achieved using the ELM model based on both the full spectra and optimal wavelengths, with a 100% recognition rate in the calibration set and a 98.75% recognition rate in the prediction set. It has been demonstrated that mid-infrared spectroscopy, combined with the ELM chemometrics method, could be a potential tool for the identification of GM maize, and it could also provide references for the on-line detection of other GM products.

Author Contributions: D.W. and X.L. (Xiaodan Liu) conceived and designed the experiment; X.L. (Xiaodan Liu) and J.Z. performed the experiments; X.B. and X.L. (Xiaolong Li) contributed to the data analysis; X.L. (Xiaodan Liu) and Y.Y. analyzed the data and wrote the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Scientific Research Project of Wuhan Polytechnic University (532100308).

Data Availability Statement: The data will be provided upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Caradus, J.R. Intended and unintended consequences of genetically modified crops—myth, fact and/or manageable outcomes? *N. Z. J. Agric. Res.* **2022**, 1–101. [[CrossRef](#)]
- van Esse, H.P.; Reuber, T.L.; van der Does, D. Genetic modification to improve disease resistance in crops. *New Phytol.* **2020**, *225*, 70–86. [[CrossRef](#)] [[PubMed](#)]
- Beckie, H.J.; Ashworth, M.B.; Flower, K.C. Herbicide Resistance Management: Recent Developments and Trends. *Plants* **2019**, *8*, 161. [[CrossRef](#)]

4. Kaundun, S.S. Syngenta's contribution to herbicide resistance research and management. *Pest Manag. Sci.* **2021**, *77*, 1564–1571. [[CrossRef](#)] [[PubMed](#)]
5. Sohn, S.I.; Pandian, S.; Oh, Y.J.; Zaukuu, J.Z.; Kang, H.J.; Ryu, T.H.; Cho, W.S.; Cho, Y.S.; Shin, E.K.; Cho, B.K. An Overview of Near Infrared Spectroscopy and Its Applications in the Detection of Genetically Modified Organisms. *Int. J. Mol. Sci.* **2021**, *22*, 9940. [[CrossRef](#)] [[PubMed](#)]
6. Qian, C.; Wang, R.; Wu, H.; Ping, J.; Wu, J. Recent advances in emerging DNA-based methods for genetically modified organisms (GMOs) rapid detection. *TrAC Trends Anal. Chem.* **2018**, *109*, 19–31. [[CrossRef](#)]
7. Demeke, T.; Dobnik, D. Critical assessment of digital PCR for the detection and quantification of genetically modified organisms. *Anal. Bioanal. Chem.* **2018**, *410*, 4039–4050. [[CrossRef](#)]
8. Rocha, P.D.; Medeiros, E.P.; Silva, C.S.; da Silva Simoes, S. Chemometric strategies for near infrared hyperspectral imaging analysis: Classification of cotton seed genotypes. *Anal. Methods* **2021**, *13*, 5065–5074. [[CrossRef](#)]
9. Wei, X.; Zheng, W.; Zhu, S.; Zhou, S.; Wu, W.; Xie, Z. Application of terahertz spectrum and interval partial least squares method in the identification of genetically modified soybeans. *Spectrochim. Acta A Mol. Biomol.* **2020**, *238*, 118453. [[CrossRef](#)]
10. Pereira, C.G.; Luiz, L.V.; Bell, M.J.; Anjos, V. Near And Mid Infrared Spectroscopy To Assess Milk Products Quality: A Review Of Recent Applications. *J. Dairy Res. Tech.* **2020**, *3*, 1–10.
11. Aykas, D.P.; Rodrigues Borba, K.; Rodriguez-Saona, L.E. Non-Destructive Quality Assessment of Tomato Paste by Using Portable Mid-Infrared Spectroscopy and Multivariate Analysis. *Foods* **2020**, *9*, 1300. [[CrossRef](#)] [[PubMed](#)]
12. Zuk, M.; Dyminska, L.; Kulma, A.; Boba, A.; Prescha, A.; Szopa, J.; Maczka, M.; Zajac, A.; Szoltysek, K.; Hanuza, J. IR and Raman studies of oil and seedcake extracts from natural and genetically modified flax seeds. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2011**, *78*, 1080–1089. [[CrossRef](#)] [[PubMed](#)]
13. Tiplady, K.M.; Lopdell, T.J.; Littlejohn, M.D.; Garrick, D.J. The evolving role of Fourier-transform mid-infrared spectroscopy in genetic improvement of dairy cattle. *J. Anim. Sci. Biotechnol.* **2020**, *11*, 39. [[CrossRef](#)] [[PubMed](#)]
14. Bec, K.B.; Grabska, J.; Huck, C.W. Biomolecular and bioanalytical applications of infrared spectroscopy—A review. *Anal. Chim. Acta* **2020**, *1133*, 150–177. [[CrossRef](#)]
15. Feng, X.; Zhao, Y.; Zhang, C.; Cheng, P.; He, Y. Discrimination of Transgenic Maize Kernel Using NIR Hyperspectral Imaging and Multivariate Data Analysis. *Sensors* **2017**, *17*, 1894. [[CrossRef](#)]
16. He, Y.; Zhao, Y.; Zhang, C.; Sun, C.; Li, X.; Zhao, Y.; Zhang, C.; Sun, C.; Li, X. Determination of β -Carotene and Lutein in Green Tea Using Fourier Transform Infrared Spectroscopy. *Trans. ASABE* **2019**, *62*, 75–81. [[CrossRef](#)]
17. Zhang, C.; Wang, C.; Liu, F.; He, Y. Mid-Infrared Spectroscopy for Coffee Variety Identification: Comparison of Pattern Recognition Methods. *J. Mol. Spectrosc.* **2016**, *2016*, 7927286. [[CrossRef](#)]
18. Zhang, C.; Zhou, L.; Zhao, Y.; Zhu, S.; Liu, F.; He, Y. Noise reduction in the spectral domain of hyperspectral images using denoising autoencoder methods. *Chemom. Intell. Lab. Syst.* **2020**, *203*, 104063. [[CrossRef](#)]
19. Wang, L.; Wang, R. Determination of soil pH from Vis-NIR spectroscopy by extreme learning machine and variable selection: A case study in lime concretion black soil. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2022**, *283*, 121707. [[CrossRef](#)]
20. Xia, K.; Xia, S.; Shen, Q.; Yang, B.; Song, Q.; Xu, Y.; Zhang, S.; Zhou, X.; Zhou, Y. Moisture spectral characteristics and hyperspectral inversion of fly ash-filled reconstructed soil. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2021**, *253*, 119590. [[CrossRef](#)]
21. Zhang, L.; Huang, X.; Fan, X.; He, W.; Yang, C.; Wang, C. Rapid fingerprinting technology of heavy oil spill by mid-infrared spectroscopy. *Environ. Technol.* **2021**, *42*, 270–278. [[CrossRef](#)] [[PubMed](#)]
22. Liu, X.; Feng, X.; Liu, F.; Peng, J.; He, Y. Rapid Identification of Genetically Modified Maize Using Laser-Induced Breakdown Spectroscopy. *Food Bioprocess Technol.* **2018**, *12*, 347–357. [[CrossRef](#)]
23. Kaushal, S.; Nayi, P.; Rahadian, D.; Chen, H.-H. Applications of Electronic Nose Coupled with Statistical and Intelligent Pattern Recognition Techniques for Monitoring Tea Quality: A Review. *Agriculture* **2022**, *12*, 1359. [[CrossRef](#)]
24. Khodabakhshian, R.; Bayati, M.R.; Emadi, B. An evaluation of IR spectroscopy for authentication of adulterated turmeric powder using pattern recognition. *Food Chem.* **2021**, *364*, 130406. [[CrossRef](#)]
25. Jimenez-Carvelo, A.M.; Gonzalez-Casado, A.; Bagur-Gonzalez, M.G.; Cuadros-Rodriguez, L. Alternative data mining/machine learning methods for the analytical evaluation of food quality and authenticity—A review. *Food Res. Int.* **2019**, *122*, 25–39. [[CrossRef](#)]
26. Al-Hashedi, K.G.; Magalingam, P. Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. *Comput. Sci. Rev.* **2021**, *40*, 100402. [[CrossRef](#)]
27. Cocco Mariani, V.; Hennings Och, S.; dos Santos Coelho, L.; Domingues, E. Pressure prediction of a spark ignition single cylinder engine using optimized extreme learning machine models. *Appl. Energy* **2019**, *249*, 204–221. [[CrossRef](#)]
28. Abbas, O.; Compère, G.; Larondelle, Y.; Pompeu, D.; Rogez, H.; Baeten, V. Phenolic compound explorer: A mid-infrared spectroscopy database. *Vib. Spectrosc.* **2017**, *92*, 111–118. [[CrossRef](#)]
29. Tewari, J.C.; Irudayaraj, J.M. Floral classification of honey using mid-infrared spectroscopy and surface acoustic wave based z-Nose sensor. *J. Agric. Food. Chem.* **2005**, *53*, 6955–6966. [[CrossRef](#)]
30. Hazarika, S.; Hebb, R.L.; Rizvi, S.S.H. Prediction of Ripening Stage of Cameo Apple Using Fourier-Transform Infrared Spectroscopy. *Int. J. Fruit Sci.* **2017**, *18*, 188–198. [[CrossRef](#)]

31. Via, B.; Fasina, O.; Pan, H. Assessment of pine biomass density through mid-infrared spectroscopy and multivariate modeling. *BioResources* **2013**, *6*, 807–822. [[CrossRef](#)]
32. Matouke, M.M. FTIR study of the binary effect of titanium dioxide nanoparticles (nTiO₂) and copper (Cu²⁺) on the biochemical constituents of liver tissues of catfish (*Clarias gariepinus*). *Toxicol. Rep.* **2019**, *6*, 1061–1070. [[CrossRef](#)] [[PubMed](#)]
33. San-Blas, E.; Cubillan, N.; Guerra, M.; Portillo, E.; Esteves, I. Characterization of Xenorhabdus and Photorhabdus bacteria by Fourier transform mid-infrared spectroscopy with attenuated total reflection (FT-IR/ATR). *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2012**, *93*, 58–62. [[CrossRef](#)] [[PubMed](#)]
34. Rayan, A.M.; Abbott, L.C. Compositional analysis of genetically modified corn events (NK603, MON88017xMON810 and MON89034xMON88017) compared to conventional corn. *Food Chem.* **2015**, *176*, 99–105. [[CrossRef](#)]
35. Agapito-Tenfen, S.; Guerra, M.; Wikmark, O.-G.; Nodari, R. Comparative Proteomic Analysis of Genetically Modified Maize Grown under Different Agroecosystems Conditions in Brazil. *Proteome Sci.* **2013**, *11*, 46. [[CrossRef](#)]
36. Tan, Y.; Yi, X.; Wang, L.; Peng, C.; Sun, Y.; Wang, D.; Zhang, J.; Guo, A.; Wang, X. Comparative Proteomics of Leaves from Phytase-Transgenic Maize and Its Non-transgenic Isogenic Variety. *Front. Plant Sci.* **2016**, *7*, 1211. [[CrossRef](#)]
37. Fornale, S.; Capellades, M.; Encina, A.; Wang, K.; Irar, S.; Lapierre, C.; Ruel, K.; Joseleau, J.P.; Berenguer, J.; Puigdomenech, P.; et al. Altered lignin biosynthesis improves cellulosic bioethanol production in transgenic maize plants down-regulated for cinnamyl alcohol dehydrogenase. *Mol. Plant* **2012**, *5*, 817–830. [[CrossRef](#)]
38. Tonolini, M.; Sorensen, K.M.; Skou, P.B.; Ray, C.; Engelsen, S.B. Prediction of alpha-Lactalbumin and beta-Lactoglobulin Composition of Aqueous Whey Solutions Using Fourier Transform Mid-Infrared Spectroscopy and Near-Infrared Spectroscopy. *Appl. Spectrosc.* **2021**, *75*, 718–727. [[CrossRef](#)]
39. Xia, Y.; Ugarte, C.M.; Guan, K.; Pentrak, M.; Wander, M.M. Developing Near- and Mid-Infrared Spectroscopy Analysis Methods for Rapid Assessment of Soil Quality in Illinois. *Soil Sci. Soc. Am. J.* **2018**, *82*, 1415–1427. [[CrossRef](#)]
40. Cengiz, M.F.; Durak, M.Z. Rapid detection of sucrose adulteration in honey using Fourier transform infrared spectroscopy. *Spectrosc. Lett.* **2019**, *52*, 267–273. [[CrossRef](#)]
41. Daoud, S.; Bou-Maroun, E.; Waschatko, G.; Horemans, B.; Mestdagh, R.; Billecke, N.; Cayot, P. Detection of Lipid Oxidation in Infant Formulas: Application of Infrared Spectroscopy to Complex Food Systems. *Foods* **2020**, *9*, 1432. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.