Rapid and Noninvasive Typing and Assessment of Nutrient Content of Maize Kernels Using a Handheld Raman Spectrometer

Mark Krimmer,†,§ Charles Farber,†,§ and Dmitry Kurouski*†‡

†Department of Biochemistry and Biophysics and ‡The Institute for Quantum Science and Engineering, Texas A&M University, College Station, Texas 77843, United States

ABSTRACT: To thrive as a global civilization, food production must meet the demands of our ever-growing population. There are more than a billion people on the planet suffering from malnutrition through poor quality or lack of food. Nutrient content of food can be determined by a variety of methods, which have issues such as slow analysis or sample destruction. Near-infrared (NIR) spectroscopy is a long-standing alternative to these methods. In this work, we demonstrated that Raman spectroscopy (RS), another spectroscopic method, can also be used to assess the nutrient content of maize (Zea mays), one of the most widely cultivated grains in the world. Using a handheld Raman spectrometer, we predicted the content of carbohydrates, fibers, carotenoids, and proteins in six different varieties of maize. This analysis requires only a single maize kernel and is fast (1s), portable, noninvasive, and nondestructive. Moreover, we showed that RS in combination with chemometric methods can be used for highly accurate (approximately 90%) spectroscopic typing of maize, which is important for plant breeders and farmers. Finally, we demonstrate that Raman-based approach is as accurate as NIR analysis. These findings suggest that portable Raman systems can be used on combines and grain elevators for autonomous control of grain quality.

INTRODUCTION

Food quality has a direct impact on the economic growth of nations and global food security. Therefore, food quality is strictly regulated in many developed countries in the world.1 There are many parameters, such as nutrient content, texture, flavor, and visual appearance, which are used to evaluate the quality of food. However, two of them, the presence of pathogens2,3 and nutrient content, have major impacts on human well-being.

The nutrient content of food is determined by the amount of carbohydrates, proteins, fibers, and oils, as well as by vitamins and minerals present in it. Several colorimetric and chromatographic methods can be used for quantitative assessment of these nutrients. These methods are destructive, time-consuming, and labor-intensive. Also, these assays require sophisticated laboratory equipment, which is often not available in developing countries.† Development of near-infrared (NIR) analyzers enabled completely noninvasive and nondestructive assessment of nutrient content in agricultural crops.5−9 Some nutritional qualities assayable by NIR include moisture, starch, protein, and many others.10 This method is also capable of determining the nutrient content of single grains of crops such as maize or wheat.11−13

A potential alternative or complementary method to NIR is Raman spectroscopy (RS). The Raman effect is based on the inelastic scattering of photons by sample molecules that are being excited to higher vibrational or rotational states.14 Thus, RS provides information about molecular vibrations and consequently the structure and composition of the analyzed sample. Our group recently demonstrated that RS could be used for confirmatory, noninvasive, and nondestructive detection and identification of fungal diseases in wheat, maize, and sorghum seeds.15,16 We have also shown that RS could detect insects inside intact cowpeas with high statistical accuracy.17 Also, using RS, we were able to diagnose nutrient deficiency and asymptomatic Huanglongbing disease on orange and grapefruit trees.18

In this study, we investigate whether RS could be used for noninvasive, nondestructive, confirmatory assessment of the nutrient content of maize, one of the most widely cultivated cereals in the world.19 The commercial impact of maize exceeds 50 billion U.S. dollars.20 Maize has a wide variety of applications, including as livestock feed, raw material in industry, biofuel, and a human food. We show that Raman-based approach is fast and self-sufficient for the rapid assessment of nutrients in food. We also show that in addition to nutrient content assessment, RS can be used for the identification or typing of maize.

RESULTS AND DISCUSSION

We collected more than 600 Raman spectra from six different varieties of maize (Zea mays). Maize varieties that have been chosen for our experiment had a distinctly different phenotypic appearance, Figure 1. Specifically, we have chosen three light color (small yellow (SY), large yellow (LY), and small white (SW)) and three dark color (purple (PP), blue (BL), and red (RD)) maize varieties.

Raman-Based Assessment of Nutrient Content. We have found that all six varieties of maize had similar spectral profiles with different intensities of vibrational bands, which can be assigned to carbohydrates, carotenoids, fibers (lignin),

Received: June 6, 2019
Accepted: September 17, 2019
compared to the light color maize varieties. Therefore, intense spectra (under the same experimental conditions) scattering, dark color maize varieties should produce less pale color kernels. Since RS is based on inelastic light and consequently scatter less light compared to the yellow or kernels. One can imagine that dark maize kernels absorb more the spectra acquired from yellow or pale color (SY, LY, SW) a much lower intensity of these vibrational bands compared to collected from colored maize varieties (BL, PP, RD) exhibited of BL, SW, SY, PP, RD, and LY maize kernels. The 1458 cm large (XL), blue (BL), purple (PP), and red (RD) maize kernels.

Figure 1. Photographs of small white (SW), small yellow (SY), extra-large (XL), blue (BL), purple (PP), and red (RD) maize kernels.

Figure 2. Baseline-corrected (A) and normalized (B) Raman spectra of BL, SW, SY, PP, RD, and LY maize kernels. The 1458 cm$^{-1}$ peak, which was used for spectral normalization, is indicated by an asterisk (*).

Table 1. Vibrational Bands and Their Assignments for Maize Kernels

<table>
<thead>
<tr>
<th>band</th>
<th>vibrational mode</th>
<th>assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1660</td>
<td>C=O stretching (amide I)</td>
<td>proteins$^{14}$</td>
</tr>
<tr>
<td>1632</td>
<td>C=C−C (ring)</td>
<td>lignin$^{19}$</td>
</tr>
<tr>
<td>1600</td>
<td>$\nu(\text{C} \equiv \text{C})$ + $\nu(\text{CH})$</td>
<td>lignin$^{19}$</td>
</tr>
<tr>
<td>1527</td>
<td>$\delta(\text{CH})$ + $\nu(\text{CH}_2)$ + $\delta(\text{C} \equiv \text{O} - \text{H})$ CH, CH$_2$, and COH deformations.</td>
<td>carotenoids$^{25}$</td>
</tr>
<tr>
<td>1460</td>
<td>$\delta(\text{C} \equiv \text{C})$</td>
<td>carbohydrates$^{16}$</td>
</tr>
<tr>
<td>1381</td>
<td>$\delta(\text{C} \equiv \text{O} - \text{H})$ - coupling of the CCH and COH deformation modes</td>
<td>carbohydrates$^{16}$</td>
</tr>
<tr>
<td>1340</td>
<td>$\nu(\text{C} \equiv \text{O})$, $\delta(\text{C} \equiv \text{O} - \text{H})$</td>
<td>carbohydrates$^{16}$</td>
</tr>
<tr>
<td>1261</td>
<td>$\delta(\text{C} \equiv \text{C})$ + $\delta(\text{O} \equiv \text{C} - \text{H})$ + $\delta(\text{C} \equiv \text{O} - \text{H})$ carbohydrates$^{16}$</td>
<td>carbohydrates$^{16,27}$</td>
</tr>
<tr>
<td>1153</td>
<td>C−C stretching; $\nu(\text{COC})$, $\nu(\text{CC})$ in glycosidic linkage, asymmetric ring breathing</td>
<td>carotenoids$^{28}$, carbohydrates$^{26}$</td>
</tr>
<tr>
<td>1125</td>
<td>$\nu(\text{C} \equiv \text{O})$ + $\nu(\text{C} \equiv \text{C})$ + $\delta(\text{C} \equiv \text{O} - \text{H})$ carbohydrates$^{26}$</td>
<td>carbohydrates$^{26}$</td>
</tr>
<tr>
<td>1085</td>
<td>$\nu(\text{C} \equiv \text{O})$ + $\nu(\text{C} \equiv \text{C})$ + $\delta(\text{C} \equiv \text{O} - \text{H})$ carbohydrates$^{26}$</td>
<td>carbohydrates$^{26}$</td>
</tr>
<tr>
<td>1052</td>
<td>$\nu(\text{C} \equiv \text{O})$ + $\nu(\text{C} \equiv \text{C})$ + $\delta(\text{C} \equiv \text{O} - \text{H})$ carbohydrates$^{26}$</td>
<td>carbohydrates$^{26}$</td>
</tr>
<tr>
<td>1009</td>
<td>phenylalanine ring stretching mode</td>
<td>proteins$^{14}$</td>
</tr>
<tr>
<td>939</td>
<td>$\delta(\text{C} \equiv \text{O} - \text{C})$ + $\delta(\text{C} \equiv \text{O} - \text{H})$ + $\nu(\text{C} \equiv \text{O})$ C1,4 glycosidic linkages carbohydrates$^{26}$</td>
<td>carbohydrates$^{26}$</td>
</tr>
<tr>
<td>865</td>
<td>$\delta(\text{C} \equiv \text{C})$ + $\delta(\text{C} \equiv \text{O})$ glycosidic bond; anomic region carbohydrates$^{26}$</td>
<td>carbohydrates$^{26}$</td>
</tr>
<tr>
<td>768</td>
<td>$\delta(\text{C} \equiv \text{O})$ carbohydrates$^{26}$</td>
<td>carbohydrates$^{26}$</td>
</tr>
<tr>
<td>717</td>
<td>$\delta(\text{C} \equiv \text{C})$ related to glycosidic ring skeletal deformations carbohydrates$^{26}$</td>
<td>carbohydrates$^{26}$</td>
</tr>
<tr>
<td>576</td>
<td>$\delta(\text{C} \equiv \text{O})$ + $\nu(\text{C} \equiv \text{O})$ carbohydrates$^{26}$</td>
<td>carbohydrates$^{26}$</td>
</tr>
<tr>
<td>614</td>
<td>$\delta(\text{C} \equiv \text{C})$ carbohydrates$^{26}$</td>
<td>carbohydrates$^{26}$</td>
</tr>
<tr>
<td>521</td>
<td>S−S gauche-gauche-trans</td>
<td>protein$^{30}$</td>
</tr>
<tr>
<td>479</td>
<td>CCO and CCC deformations; related to glycosidic ring skeletal deformations $\delta(\text{C} \equiv \text{C})$ + $\nu(\text{C} \equiv \text{O})$ Scissoring of C−C−C and out-of-plane bending of C−O carbohydrates$^{26}$</td>
<td>carbohydrates$^{26}$</td>
</tr>
<tr>
<td>442</td>
<td>skeletal modes of pyranose ring</td>
<td>carbohydrates$^{26}$</td>
</tr>
</tbody>
</table>

and proteins, Figure 2A and Table 1. Interestingly, the spectra collected from colored maize varieties (BL, PP, RD) exhibited a much lower intensity of these vibrational bands compared to the spectra acquired from yellow or pale color (SY, LY, SW) kernels. One can imagine that dark maize kernels absorb more and consequently scatter less light compared to the yellow or pale color kernels. Since RS is based on inelastic light scattering, dark color maize varieties should produce less intense spectra (under the same experimental conditions) compared to the light color maize varieties. Therefore, observed variations in spectral intensities likely originate from different light absorption and scattering properties of these maize kernels.

This observation indicates that raw spectra collected from colored maize kernels cannot be used for direct assessment of their nutrient content. For instance, the intensity of the 479 cm$^{-1}$ band, which can be used to estimate the carbohydrate content in the maize, appears lower in the spectra collected from PP maize compared to that in LY or SY maize. However, this low intensity of this band originates from the poor scattering properties of PP maize kernels relative to those of LY or SY maize. This problem can be solved by spectral normalization. However, normalizing spectra without an internal standard is a challenging task. Normalization on one particular band that can be assigned to the specific class of molecules, such as carbohydrates, may not be appropriate. Such normalization would bias spectral interpretation in regard to the nutrient content of that class of molecules. Because the 1458 cm$^{-1}$ band, assigned to CH$_2$ vibrations, cannot be associated with any specific class of biomolecule, we have chosen to normalize our spectra to 1458 cm$^{-1}$.

This unbiased spectral normalization can be used to access the nutrient content of maize varieties. The analysis of variance (ANOVA) of the 479 cm$^{-1}$ band, which can be assigned to carbohydrates such as monomeric sugars and starch, revealed two statistically significant groups of maize (Figure 3). We found that the SY and RD carbohydrate contents were significantly different from those of BL and PP. At the same
time, SW and LY were not significantly different from the first and second groups of maize, respectively.

Carotenoids exhibit bands around 1530 cm\(^{-1}\), which can be assigned to in-plane \(-C=\pi C-\) vibrations. ANOVA revealed three statistically significant groups of maize based on the intensities of these carotenoid bands. SW and RD had the lowest carotenoid content and could be assigned to the first group. The carotenoid content of this group is statistically different from that of BL. Finally, SY, PP, and LY belong to the third group with the highest carotenoid content. From Figure 3B, it is apparent that the darkness of a kernel is not correlated with the intensity of the carotenoid band. Our results support a previous hypothesis that the color of maize kernels is determined by anthocyanins.\(^{21,22}\)

Fibers are polyphenolic molecular assemblies that have two distinct vibrational bands in Raman spectra.\(^{23,24}\) The band at 1600 cm\(^{-1}\) can be assigned to \(C=C\) ring stretching and symmetric \(C-H\) vibration.\(^{13}\) The 1632 cm\(^{-1}\) band originates from \(C=\pi C\) aromatic ring vibration.\(^{24}\) ANOVA revealed at least two groups of maize based on the fiber content. Specifically, SY and RD can be assigned to the first group with the lowest fiber content. PP showed the highest fiber content that is statistically significant from all other groups. Based on the intensity of 1600 cm\(^{-1}\) band, the fiber content in BL, SW, and LY is higher compared to that in SY and RD, but lower than in PP. However, it is not statistically different from SY, RD, and PP. At the same time, the analysis of the intensity of 1632 cm\(^{-1}\) band showed that the fiber content of LY is significantly different from both first (SY and RD) and second (PP) groups, whereas BL and SW were not significantly different from PP.

Proteins exhibit a carbonyl vibration of the peptide bond at 1640–1670 cm\(^{-1}\), known as the amide I band, as well as a band at 1000 cm\(^{-1}\), which can be assigned to phenylalanine.\(^{14}\) Both of these bands were observed in the Raman spectra collected from maize kernels. At the same time, the band around 1650 cm\(^{-1}\) could be attributed to \(C=\pi C\) bond of unsaturated fatty acids, whereas 1000 cm\(^{-1}\) could originate from carotenoid vibrations.\(^{25,31}\) Our results reveal a very similar pattern for both 1000 and 1660 cm\(^{-1}\) bands indicating that these bands represent the same chemical structure. Since the ANOVA pattern of both 1000 and 1660 cm\(^{-1}\) bands is distinctly different from the ANOVA pattern that was observed from 1527 cm\(^{-1}\) band, which can be unambigously assigned to carotenoids, we can conclude that both 1000 and 1660 cm\(^{-1}\) bands in the Raman spectra of maize can be used to reveal the protein content. We found that RD maize has the lowest protein content, whereas BL, SW, SY, and PP have statistically higher protein content.

These results clearly demonstrate that RS can be used for fast (1s spectral acquisition), noninvasive, and nondestructive assessment of carbohydrates, proteins, fibers, and carotenoids.

Figure 3. Means (circles) and confidence intervals for the intensities of the maize kernel spectra, normalized to 1458 cm\(^{-1}\), at the indicated Raman shift. Colors indicate significantly different groups. Multiple colors indicate a member of a group that has overlap between two separate groups. Each of (A)–(F) corresponds to a different selected Raman shift discussed in the text.
in the intact grain. When coupled to ANOVA, RS provides a highly accurate content of these four classes of nutrients. Also, handheld spectrometer enables the performance of this analysis directly in the field with only one kernel required for quantitative elucidation of the grain nutrient content.

The question to ask is how this information can be related to nutrient content analysis made by already established techniques. To answer this question, we performed the analysis of the same grain samples using NIR and determined protein and carbohydrate contents using the Dumas combustion method and megazyme total starch content assay (subsequently megazyme assay), respectively (Tables 2 and 3).

### Table 2. Results of NIR Analysis of Nutrient Content of Six Different Maize Varieties

<table>
<thead>
<tr>
<th></th>
<th>starch (%)</th>
<th>protein (%)</th>
<th>moisture (%)</th>
<th>oil (%)</th>
<th>ash (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL</td>
<td>63.1</td>
<td>10.0</td>
<td>13.4</td>
<td>4.6</td>
<td>0.8</td>
</tr>
<tr>
<td>SW</td>
<td>62.7</td>
<td>10.7</td>
<td>13.2</td>
<td>3.4</td>
<td>0.8</td>
</tr>
<tr>
<td>SY</td>
<td>63.1</td>
<td>11.9</td>
<td>13.0</td>
<td>4.0</td>
<td>0.9</td>
</tr>
<tr>
<td>PP</td>
<td>62.2</td>
<td>11.4</td>
<td>13.5</td>
<td>3.5</td>
<td>0.8</td>
</tr>
<tr>
<td>RD</td>
<td>60.2</td>
<td>12.0</td>
<td>14.1</td>
<td>3.7</td>
<td>0.9</td>
</tr>
<tr>
<td>LY</td>
<td>62.4</td>
<td>10.6</td>
<td>13.1</td>
<td>3.8</td>
<td>0.9</td>
</tr>
</tbody>
</table>

According to NIR analysis, the amount of starch was found to vary in the maize varieties from 60.2 to 63.1%. Considering the fact that the accuracy of such analysis is ±5%, NIR analysis suggests that there is no significant difference in starch content among the maize varieties. According to the results of megazyme assay, starch content varies from 54.6 to 59.3% (±3%). These results can also be used to reference Raman readings to the starch content. Thus, 2.55–2.9 scale in the Raman can be assigned to 57% of starch by the dry weight.

### Raman-Based Typing of Maize Varieties

We conducted partial least-squares discriminant analysis (PLS-DA) to demonstrate that RS can be used for typing of these maize varieties. The model was built using all 613 spectra with the mean offset removed at each wavenumber and cross-validated with the same data partitioned by the venetian blinds method. This model contained 19 latent variables (LVs) and was used to generate a misclassification table (Table 4).

### Table 4. Misclassification Table of Cross-Validation for the PLS-DA Model

<table>
<thead>
<tr>
<th></th>
<th>members</th>
<th>correct (%)</th>
<th>BL</th>
<th>SW</th>
<th>SY</th>
<th>PP</th>
<th>RD</th>
<th>LY</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL</td>
<td>113</td>
<td>98</td>
<td>111</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>SW</td>
<td>100</td>
<td>98</td>
<td>98</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SY</td>
<td>98</td>
<td>94</td>
<td>94</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP</td>
<td>100</td>
<td>89</td>
<td>89</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>89</td>
<td>1</td>
</tr>
<tr>
<td>RD</td>
<td>100</td>
<td>99</td>
<td>99</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>99</td>
<td>0</td>
</tr>
<tr>
<td>LY</td>
<td>102</td>
<td>94</td>
<td>94</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>96</td>
</tr>
</tbody>
</table>

The first three LVs explain 94.6, 1.7, and 0.7% of the variation, respectively. The misclassification table reports the true positive rate (accuracy) of the model; green cells in the table indicate members of a class that were correctly assigned to their class during cross-validation. Overall, we can see that the model performs with an 89% classification accuracy at minimum.

## CONCLUSIONS

Our findings indicate that RS in combination with advanced statistical analysis can be used to predict the nutrient content of carbohydrates, carotenoids, fibers, and protein in intact maize kernels. We have also shown that RS is capable of highly accurate typing of maize grain. Thus, this study shows that RS is a highly efficient multifunctional method for the analysis of grain.

## MATERIALS AND METHODS

### Maize

Six different varieties of maize (*Zea mays*) were purchased from Amish Country Popcorn (Berne, Indiana) and used as received.

### Raman Spectroscopy

Raman spectra were collected with a handheld Resolve Agilent spectrometer equipped with an 831 nm laser and a spectral resolution of 15 cm⁻¹. The following experimental parameters were used for all collected spectra: 1s acquisition time, 495 mW power, surface scanning mode, and baseline spectral subtraction by device software. Spectra shown in the manuscript are baseline-corrected by the instrument software without smoothing. One spectrum was collected per one maize kernel.

### Spectral Processing

Spectral processing (described below) and averaging were conducted using PLS_Toolbox 8.6.2 (Eigenvector Research, Inc., Manson, WA).

### Statistical Analysis

Raman spectra were imported into Matlab (Mathworks) and assigned a class based on their visual phenotype (color). For ANOVA, spectra were normalized to the 1458 cm⁻¹ band. For PLS-DA, the mean offset was removed at each wavenumber before analysis.

### NIR Analysis

PerkinElmer DA 7250 NIR analyzer was used to determine the amount of protein, starch, oils, moisture, and ash in six varieties of maize. For each maize variety, ~30 g of material was submitted to the analyzer.
Chemical Content Analysis. Phenolic content was measured according to the Folin–Ciocalteu method.\textsuperscript{37} Total protein concentration was determined by combustion using LECO instrument and a nitrogen-to-protein conversion factor. Protein concentration was determined by combustion using the microassay kit, based on the RT-SO4 procedure from the AOAC 76.13 and AOAC 999.11 methods (K-TSTA; Megazyme, Bray, Ireland).

\section*{AUTHOR INFORMATION}

\subsection*{Corresponding Author}
*E-mail: dkurouski@tamu.edu, Tel: 979-458-3778.

\subsection*{ORCID}
Dmitry Kurouski: 0000-0002-6040-4213

\subsection*{Author Contributions}
\textsuperscript{§}M.K. and C.F. contributed equally to this work.

\subsection*{Notes}
The authors declare no competing financial interest.

\section*{ACKNOWLEDGMENTS}
The authors are grateful to Joseph M. Awika and Audrey Girard from the Department of Soil and Crop Sciences of Texas A\&M University for their help with NIR analysis of maize varieties. This study was supported by funds from Texas A\&M AgriLife Research, Texas A\&M University Governor’s University Research Initiative (GURI) grant program of 12-1961/M1700437.

\section*{REFERENCES}
(5) Osborne, B. G.; Fearn, T.; Hindle, P. H.; Hindle, P. T. \textit{Practical NIR Spectroscopy with Applications in Food and Beverage Analysis}; Longman Scientific & Technical, 1993


