Food Chemistry 134 (2012) 1165–1172

Contents lists available at SciVerse ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem

Analytical Methods

Feasibility of conventional and Roundup Ready® soybeans discrimination by different near infrared reflectance technologies

Lidia Esteve Agelet [a,*], Aoife A. Gowen [b], Charles R. Hurburgh Jr. [a], Colm P. O'Donell [b]

[a] Department of Agriculture and Biosystems Engineering, Iowa State University, Ames, IA 50014, United States
[b] Agriculture and Food Science Centre, University College Dublin, Belfield, Dublin 4, Ireland

ARTICLE INFO

Article history:
Received 18 November 2011
Accepted 22 February 2012
Available online 9 March 2012

Keywords:
Chemical imaging
Discrimination
Single seed
Genetically modified organisms

ABSTRACT

Identification and proper labelling of genetically modified organisms is required and increasingly demanded by legislation and consumers worldwide. In this study, the feasibility of three near infrared reflectance technologies (a chemical imaging unit, a commercial diode array instrument, and a light tube non-commercial instrument) were compared for discriminating Roundup Ready® and not genetically modified soybean seeds. Over 200 seeds of each class (Roundup Ready® and conventional) were used. Principal Component Analysis with Artificial Neural Networks (PCA–ANN) and Locally Weighted Principal Component Regression (LW-PCR) were used for creating the discrimination models. Discrimination accuracies when new tested seeds belonged to samples included in the training sets achieved accuracies over 90% of correctly classified seeds for LW-PCR models. The light tube performed the best, while the imaging unit showed the worse accuracies overall. Models validated with new seeds from samples not included in the training set had accuracies of 72–79%.

1. Introduction

In 1994, the first genetically modified (GM) soybean (*Glycine max* L.) variety was introduced in the United States market. They were the first generation of Roundup Ready® soybeans that incorporates a gene from the bacterium *Agrobacterium tumefaciens* conferring resistance to the glyphosate-based Roundup herbicide. Although other GM soybean varieties are expected in the future – with genes conferring high oleic acid or resistance to other herbicides – the most widespread GM soybean varieties at present are Roundup Ready®, comprising close to 60% of worldwide soybean crops (Konduru, Kruse, & Kalaitzandonakes, 2008). Despite rapid acceptance of this variety, controversy remains. The uncertainty of GM effects on human health, environmental safety, and ecological quality (i.e. varietal preservation) are just some of the concerns raised with GM technologies.

Low-level mixtures of GM in conventional batches of soybeans, for instance by incomplete cleaning of machinery during harvesting operations, are common and virtually impossible to eliminate (Hanna, Quick, & Jarboe, 2010). With the increasing demand for organic or natural products, worldwide governments have created regulations for labelling and control of GM products. Acceptance thresholds for adventitious GM presence in soybeans range from 0.9% to 5%. Below the thresholds in respective markets, there is no need for specific labelling of soybean batches destined for both human and animal feeding purposes; in addition, no market has a rule for GM-fed animal products for human consumption (Friends of the earth, briefing resource, 2006). The European Union has the smallest tolerance for GM admixture. The European Novel Food Regulation EC 1829/2003 sets a threshold of 0.9% of GM contamination in food and animal feed without labelling if the adventitious presence comes from one of the fully accepted GM varieties (2 soybean varieties so far, Roundup Ready® being one of them). A 0.5% tolerance limit is applied for unauthorized GM varieties if they are proven safe by relevant scientific committees, even if commission approval has not been granted.

There is a need for GM detection methods that are accurate, fast, and inexpensive. Currently, there are two recognised classes of GM identification methods, both based on detecting molecules: DNA and expressed protein. Polymerase Chain Reaction (PCR) methods are the most sensitive, with lower limits of GM DNA detection of 0.1% (Miljuš-Djukić et al., 2010). Protein-based detection methods such as Enzyme-Linked Immunosorbent Assay (ELISA) are less accurate; more sample is needed, and prior knowledge of the protein produced by a specific GM event is required. Protein-based methods are faster and simpler to perform; other methods are seed germination, tetrazolium tests, insect resistance bioassays, biosensors, chromatography, use of microfabricated devices, and nanoscale analysis (Rizzi, Sorlini, & Dalfonchio, 2004). All require sample destruction, considerable time, and human resource

* Corresponding author. Address: 1545 Food Science Building, Iowa State University, Ames, IA 50014, United States. Tel.: +1 515 294 8319.
E-mail address: esteve.lidia@gmail.com (L. Esteve Agelet).

0308-8146/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.foodchem.2012.02.144

expenses. Because only a few seeds can be taken per analysis, the accuracy of the method becomes dependant on the sampling procedure. The distribution of GM seeds in a conventional batch of any grain can create up to a 20% sampling error (Hübner, Waiblinger, Pietsch, & Brodmann, 2001).

Roussel, Hardy, Hurburgh, and Rippke (2001) introduced the use of Near Infrared Spectroscopy for discrimination of conventional and Roundup Ready® bulk soybeans by Near Infrared Spectroscopy (NIRS) transmittance. Three classification methods (Partial Least Squares Discrimination Analysis (PLS-DA), Artificial Neural Networks (ANN), and Locally Weighted Regression (LWR)) were tested. Non-linear methods, ANN and LWR, achieved the highest classification accuracies of 88% and 93%, respectively, having spectra from two instruments involved and a single crop year. Near Infrared measurements do not measure compounds at trace levels, but could measure the physical or chemical expression of a genetic trait as long as the expression includes a change in biochemical structure. In this case, it was suspected that differences among Roundup Ready® and conventional soybeans arose from fibre structure, after observing the relevance of the regression coefficients from the PLS–DA models in the carbohydrate absorbance region (894–950 nm).

Near Infrared technologies provide fast and non-destructive analysis, which offers an attractive way to measure whole batches of seeds and reduce sampling limitations. A previous novel study (Roussel et al., 2001) was carried out on bulk samples. A whole sample of 250–500 g of seeds was scanned and was classified as either Roundup Ready® or conventional. The threshold or percentage of sample impurity was unknown. The small spectral differences suggested that low-level mixtures would not be identifiable by bulk sample methods. A more recent paper by Lee and Choung (2011) analyzed the feasibility of using the range of visible (VIS) and NIR radiation for classifying herbicide-resistant soybean seeds, reporting high accuracies (97% of the seeds correctly classified). However, only 20 samples were involved and the model was not validated with independent samples. Moreover, the major differences among GM seeds and conventional arised in the VIS region - indicating that colour differences among groups contributed in the classification. While the previous study successfully utilised Partial Least Squares Discriminant Analysis algorithms (PLS-DA), Roussel et al. (2001) suggested that discrimination of conventional and Roundup Ready® seeds based only on the NIR region was a complex problem when including a large number of samples. Non-linear algorithms able to model complex relationships were needed.

This paper tests the feasibility of NIRS reflectance to discriminate Roundup Ready® herbicide resistant single soybean seeds from conventional, utilising a large sample pool of over 200 samples for each class (each sample containing >500 seeds, from which 15 were randomly selected). Two algorithms utilised in the novel study carried by Roussel et al. (Artifical Neural Networks and Locally Weighted Principal Component Regression) were tested. We also compared the classification models accuracy when validated with more independent seeds (new seeds belonging to samples not included in the classification model but belonging to the same crop year range) and when validated with new seeds belonging to samples already represented in the classification set. The paper groups two studies in which three different technologies were employed and compared. In the first study, we utilised Near Infrared Reflectance Imaging (NIR-CI); two single point instruments with different measurement approaches were used in the second. The main objectives in this study can be summarised in three points: (1) to study the overall feasibility of NIRS reflectance technologies for discrimination of Roundup-Ready® and conventional soybean seeds, either belonging to samples represented in the classification model or nearly independent samples, (2) to compare possible differences in classification accuracies from existent technologies for single seed analysis, and (3) to compare the performance of two classification algorithms for this application.

2. Materials and methods

2.1. Study utilising a NIR chemical imaging unit

Chemical imaging technologies combine the advantages of conventional NIRS with digital mapping, which is useful for analyzing heterogeneous samples. The resulting data structure can be thought of as a cube or a stack of cards, where two spatial dimensions are combined with a third dimension corresponding to the chemical information or spectra (wavelengths). The mapping capability of imaging systems is brought by digital cameras with 2 dimensional arrays of detectors (pixels) such as CCDs, which are effective in capturing the lower light intensities of these technologies. However, the chemical mapping advantage was not relevant in our study since soybean seeds are homogeneous in shape and composition. The main advantage of chemical imaging in this study was the possibility of analyzing several seeds simultaneously.

2.1.1. The instrument

The reflectance imaging system used for this study was a line-scanning instrument (DV Optics Ltd., Padua, Italy) with an InGaS camera with low resolution (320 × 240 pixels, 12 bits resolution) and a Specim N17E spectrograph (Spectral Imaging Ltd., Oulu, Finland). The wavelength range covered was from 880 nm to 1720 nm, taking data points every 7 nm. The translation stage located under the camera where seeds were placed was set to a speed equal to 20 μm/s, obtaining images of 350 lines by 320 columns. The seeds were arranged in batches of 60 seeds on the instrument translation stage.

2.1.2. Samples and seeds

Two hundred sixteen Roundup Ready® (RR) and 202 conventional soybean samples from the Grain Quality Laboratory storage bank at Iowa State University (Ames, IA) covering crop years from 1984 to 2008 were measured in the first set of images. Fifteen individual seeds were randomly taken from each sample and mixed, obtaining two bags of 3240 conventional and 3030 RR seeds. A first set of 92 images were obtained from 92 draws of 30 seeds from each bag (30 RR and 30 conventional seeds per image), randomly arranged on the instrument translation stage utilising a Matlab v.7.4 (Mathworks, Natick, MA) random generator number function. Since light on the sampling surface could be hetereogeneous, the random arrangement of seeds was employed to avoid any interaction between light intensity and sample type which could bias the discrimination. The distance between seeds was kept constant and large enough to avoid light scattering interactions.

We later collected a second set of 31 images with a small and variable proportion of Roundup Ready® seeds and conventional soybean seeds from new samples, simulating a real situation of screening for Roundup Ready® contamination of conventional batches (Table 1) with accidental presence of conventional seeds from soybean samples not included in the first images. Again, a total of 60 seeds were randomly arranged per image. Twenty new RR samples, not used in the previous 92 images, were selected. Thirteen seeds were drawn from each new RR sample and mixed in a bag. The conventional seeds used in those 31 images were taken from a mixture of 45 previously used samples (20 new seeds drawn per sample) plus a set of 25 new conventional samples not used in the previous 92 images (22 seeds/sample). Summarising, all RR seeds and approximately ¼ of the conventional seeds in