Industrial Crops and Products 43 (2013) 301–305

Contents lists available at SciVerse ScienceDirect

Industrial Crops and Products

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

Grain quality determination by means of near infrared spectroscopy in *Jatropha curcas* L.

J.M. Montes[a,*], F. Technow[b], B. Bohlinger[c], K. Becker[a,c,d]

[a] JatroSelect GmbH, Wollgrasweg 49, 70599 Stuttgart, Germany
[b] Institute of Plant Breeding, Seed Science and Population Genetics, University of Hohenheim, Fruwirthstr. 21, 70599 Stuttgart, Germany
[c] Institute of Animal Production in the Tropics and Subtropics, University of Hohenheim, Fruwirthstr. 12, 70593 Stuttgart, Germany
[d] JatroSolutions GmbH, Wollgrasweg 49, 70599 Stuttgart, Germany

ARTICLE INFO

Article history:
Received 4 April 2012
Received in revised form 31 May 2012
Accepted 13 June 2012

Keywords:
Jatropha
Grain quality
Bulk samples
NIRS

ABSTRACT

Jatropha curcas L. (JCL) is an important energy plant which has received great attention in recent years for its utilization in biodiesel production and its high potential for greening and rehabilitation of wastelands. Breeding, farming and industrial processing of JCL need rapid and efficient methods to determine quality parameters associated with the value of the oil, press cake and biodiesel. Wet chemistry analyses are time consuming and expensive and near-infrared spectroscopy (NIRS) might be a suitable method for the analysis of large numbers of samples. Our main goal was to investigate the potential of NIRS to determine grain quality parameters in bulk JCL samples. We concluded that NIRS has a high potential for determining concentrations of oil, protein, ash, stearic, oleic, linoleic, linolenic and arachidic acids in the kernel fraction of JCL grains. The precision of NIRS-based phorbol ester determination could be improved by excluding samples with phorbol ester <0.1 mg g^{-1}. NIRS-based prediction of shell quality traits proved to be less accurate than for kernel traits.

1. Introduction

Jatropha curcas L. (JCL) is an important energy plant which has received great attention in recent years for its utilization in biodiesel production (Ceasar and Ignacimuthu, 2011). JCL is a small tree or large shrub, which belongs to the Euphorbiaceae family and has a life expectancy of up to 50 years (Divakara et al., 2010). The plant has a high potential for greening and rehabilitation of wastelands and the grains have a high oil concentration with excellent quality for conversion into biodiesel (Francis et al., 2005). Although various efforts have been made to develop JCL as an industrial crop (Fairless, 2007; Sanderson, 2009), the absence of improved cultivars and lack of agronomic knowledge represent the main bottleneck that limits the full exploitation of this plant's potential (King et al., 2009).

Grain quality is an important aspect to be considered in breeding, crop production and industrial processing of JCL because quality parameters affect the value of the oil, press cake and biodiesel as well as the production efficiency of the processes involved (Ramos et al., 2009). The ideal vegetable oil should have low saturation and low polyunsaturation, i.e., be high in monounsaturated fatty acids (Gunstone, 2004). Due to its favorable oil composition, JCL is of particular interest as an alternative, sustainable and renewable source for biodiesel production (Francis et al., 2005).

The JCL oil is usually extracted by screw presses with about 60–70 °C heat treatment (Ofori-Boateng et al., 2012). The press cake remaining after oil extraction is rich in proteins but it contains toxic compounds (Herrera et al., 2012) and a high percentage of cellulose and lignin originated from the grain shell. The toxicity is caused by phorbol esters (Makkar et al., 1998). Some Mexican accessions have been reported to be free from phorbol esters (Makkar and Becker, 1999) and they are an interesting germplasm source for breeding cultivars that can be used for biodiesel production and fodder.

Breeders, growers and traders need to estimate grain quality in large quantities (from a few kilos to many tons). While the breeder is interested in measuring the quality on plant or plot basis (0–50 kg), the grower and trader are interested in quality determination of much larger quantities (0.1–1000 tons). In these situations, representative bulk samples need to be taken to estimate the grain quality of the plant, plot or grain-lot and the results must be available as soon as possible to speed breeding progress and processing through the commercialization chain.

Near-infrared spectroscopy (NIRS) has been used in various crops as a method to assist breeding and industrial processing (Roberts et al., 2004). In JCL, the potential of NIRS to estimate concentrations of oil and protein and oil composition in single intact kernels was investigated by Vaknin et al. (2011). They concluded

* Corresponding author. Tel.: +49 0 711 451 017425; fax: +49 0 711 451 017421.
E-mail address: juan.montes@jatroselect.com (J.M. Montes).

0926-6690/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.indcrop.2012.06.054

that NIRS was a reliable and accurate method to estimate concentrations of oil, protein, oleic and linoleic acids. To our knowledge, the potential of NIRS to determine grain quality parameters in bulk grain samples has not been reported which represents a lack of knowledge and tools for breeders, growers and traders.

Our main goal was to investigate the potential of NIRS for determination of grain quality parameters in bulk JCL samples. The objective was to assess the predictive ability of NIRS for determination of dry matter, oil, protein, phorbol ester, ash and fatty acids (palmitic, palmitoleic, stearic, oleic, linoleic, linolenic and arachidic) concentrations in the kernel fraction, and dry matter, acid detergent fiber (ADF), acid detergent lignin (ADF) and ash concentrations in the shell fraction of JCL grains.

2. Materials and methods

2.1. Plant material

A total of 370 grain samples from the JatroSelect GmbH germplasm bank were used. The samples originated from a wide range of geographical regions from around the world. Each sample comprised 40 seeds. Seeds were de-shelled, resulting in kernel and shell fractions, which were analyzed separately. Kernels and shells were ground using a IKA-A11 basic grinder to 1.5 mm (kernel) and 1 mm (shell) particle size. The ground materials were subsequently used for NIRS and wet chemistry analysis. The NIRS and wet chemistry analysis of any particular sample were both done on the same day to avoid variations caused by moisture or storage.

2.2. Near-infrared reflectance spectroscopy (NIRS)

NIR measurements were made on 740 samples (370 kernel and 370 shell samples). Before being placed in the spectrometer, samples were prepared by compressing about 1.5 g of the homogenized ground material lightly into a small petri dish with the back of a spoon to produce a layer approximately 2 mm thick. NIR spectra were collected with a NIRFLEX N-500 spectrometer (Büchi Labortechnik AG, Switzerland) in the wavelength range 1000–2500 nm with 1 nm spectral resolution by following the standard procedures recommended by the manufacturer. Three replicate NIRS measurements were made on each sample.

2.3. Wet chemistry analysis

2.3.1. Kernel fraction

Dry matter concentration was determined by difference in weight before and after drying. About 0.5 g of kernel material was dried overnight at 100 °C and then put in a desiccator to cool to room temperature (AOAC, 1990).

Ash concentration was determined by difference in weight before and after incineration in a muffle furnace at 500 °C for 6 h (AOAC, 1990).

Oil was extracted from the seed samples (1.5 mm particle size) with petroleum ether in a Tecator Soxhlet type extractor for 6 h until the evaporation of residual ether revealed that it contained <0.2% lipid. The oil concentration of the samples was determined gravimetrically (AOAC, 1990).

The nitrogen content of the samples was determined using a C/N-Macro-Analyzer (C/N vario MAX, Elementar Analysensysteme, Hanau, Germany) according to the Dumas principle. The crude protein was calculated as nitrogen content × 6.25 (ISO 16634-1: 2008).

Phorbol ester concentration was determined in at least two duplicates of each sample according to the methanol procedure outlined by Makkar et al. (2007) and Devappa et al. (2010) which in turn was based on the method of Makkar et al. (1997). Briefly, 0.5 g of each test sample was extracted four times with 2% tetrahydrofuran in methanol. A suitable aliquot was loaded into a HPLC (Merck Hitachi L 7000) fitted with a reverse-phase C_{18} LiChrospher 100, 5 μm column 4 mm diameter × 250 mm long. The separation was done using three gradients: (A) 1.75 ml o-phosphoric acid (85%) in 1 L distilled water, (B) Acetonitrile and (C) Methanol Tetrahydrofuran (95/5) and was carried out at room temperature (22 °C) at a flow rate of 1.3 mL/min using gradient elution. Phorbol esters peaks (four) appeared between 25 and 31 min and were detected at 280 nm. The area of each peak was measured using a Merck-Hitachi L-7000 photodiode array detector. The results were expressed as equivalents of phorbol-12-myristate 13-acetate (PMA, Sigma, St. Louis, MO), which was used as a marker and appeared between 31 and 32 min. The areas of each of the four PE-peaks were converted to their PMA equivalent by comparing their peak areas with that of the standard solution of PMA (Makkar et al., 2007; Devappa et al., 2010).

A subset of 220 samples from the 370 samples was randomly chosen for the analysis of fatty acids. Separation and identification of fatty acid methyl esters (FAMEs) were carried out in at least two duplicates of each sample. The FAMEs were prepared by the boron trifluoride method (AOAC, 1990; Schlechtriem et al., 2004) and analyzed in a Shimadzu GC-14A gas chromatograph (Shimadzu, Tokyo, Japan) equipped with a fused silica capillary column, 50 m × 0.25 mm I.D., coated with Permabond FFAP-DF 0,10 using nitrogen as carrier gas (1.3 kg cm^{-2}). The oven thermal gradient was programmed to increase from an initial 160 °C to 198 °C at a rate of 2.5 °C min^{-1}, then from 198 °C to 218 °C at 2 °C min^{-1}, from 218 °C to 240 °C at 1.5 °C min^{-1} and finally from 240 °C to 250 °C at 1 °C min^{-1}. Temperatures were maintained for 5, 15, 10 and 2 min at 198 °C, 218 °C, 240 °C and the final temperature respectively. Individual FAMEs were identified by comparison with a known standard mixture (Sigma 47885-U) and quantified by means of a Shimadzu LabSolution software (Schlechtriem et al., 2004).

2.3.2. Shell fraction

Dry matter and ash concentrations were measured by following the same procedures as for the kernel fraction. ADF and ADL were determined as described by Van Soest et al. (1991).

2.4. Statistical analysis

The whole analysis was based on the full spectrum of 1501 wavelengths from 1000 to 2500 nm, averaged over the three spectral replications per sample. We used partial least squares regression (PLSR) to build the prediction models. The optimal number of partial least squares components was determined by a 10-fold cross validation (CV) using a subset of 50% of the samples. To prevent overfitting, we restricted ourselves to a maximum of 10 components. This CV was also used to determine the optimal mathematical pretreatment of the spectra. As pre-treatments we considered scaling and centering, of the original spectra or of its first or second derivative. Using only a subset of the samples for tuning these parameters (number of components, mathematical pre-treatment) was a further guard against overfitting, because then the tuning data set was at least partly independent of the full data set used later on for evaluating the prediction performance. The tuning process was done separately for each trait. After determining the optimal parameter combinations, we employed a 10-times repeated 5-fold CV to assess the prediction accuracy. In one CV repetition a traditional 5-fold CV was performed. In traditional 5-fold CV the data set is split randomly into 5 separate subsets. Then each subset is in turn predicted by a model built from the other 4 subsets. In each calculation the statistics of interest (Pearson correlation (r) between predicted and reference values and Root mean square error of prediction (RMSEP)) are recorded