Quantification of Natural Rubber in *Taraxacum kok-saghyz* by Near Infrared Reflectance Spectroscopy

Ming Ma, Sarah K. McNulty, Zinan Luo, Stephen E. Kopicky, Katrina Cornish

ABSTRACT

Taraxacum kok-saghyz (TK), is a herbaceous perennial that produces high molecular weight rubber within its root, which is a potential rubber resource for industry use. In order to accelerate the TK breeding process, rapid methods to quantify natural rubber in TK are needed.

Near Infrared Reflectance (NIR) Spectroscopy is a rapid, non-destructive method to quantify the natural rubber in TK. In this study, 290 TK root samples were used to construct an NIRS model. We created a rubber prediction model (r^2 =0.88, standard error of cross validation = 8.3 mg rubber/g dry root). This NIR rubber model covers a wide range of rubber contents in TK from 14-205 mg rubber/g dry root.

We have uses the NIR model to predict the rubber content of 303 unknown samples from 35 - 151 mg rubber/g dry root. This method has no recurring costs, and is much faster (approximately 10 samples/h including grinding time) than the reference method, accelerated solvent extraction (23 samples/2 days).

INTRODUCTION

Natural rubber (NR), high molecular weight *cis*-1,4-polyisoprene, is used in the manufacture of over 40,000 consumer and industrial products, and cannot be substituted for by synthetic rubbers . *Taraxacum kok-saghyz* (TK) produces almost identical *cis*-1,4-polyisoprene rubber polymer as *Hevea brasiliensis*,. TK has an annual life cycle and can be grown across a wide range of environments in the northern US. In order to improve the efficiency of TK breeding, rubber quantification methods are essential, but rubber quantification methods, like accelerated solvent extractions (ASE) [1] and latex quantification (LQ) [2], are time consuming, expensive and limit the breeding process.

Near Infrared Reflectance (NIR) spectroscopy is a spectroscopic method that uses the near-infrared region of the electromagnetic spectrum (from about 700 nm to 2500 nm), which is based on molecular overtone and combination vibrations[3]. When a sufficient number of samples have been collected to represent all the available variation and the samples have been analyzed by the reference method, a mathematical model can be constructed that describes the relationship between specific spectral features and the sample characteristic of interest. Thereafter, a researcher may quickly predict that same characteristic in a new target sample by applying the chemometric model to the spectrum of the new sample. NIR is a rapid, cost-effective, non-destructive, accurate and efficient analytical method.

METHODS

Rubber Prediction Model

To develop a NIR quantification prediction model for TK roots, 290 TK root samples were collected. Roots were dried and finely ground with an analytical mill (A11 basic, IKA). NIR data were collected with an Analytical Spectral Devices (ASD) Field Spec 3using IndicoPro. The spectral files were converted to a GRAMS SPC file format with the equation Absorbance=log(1/Reflectance). The rubber contents of the ground samples were determined by ASE. The NIR predictive model was developed by pairing the accelerated solvent extraction (ASE) data with the spectral data by GRAM IQ. About one fifth of the scans (61) were removed from 290 to generate a validation set. The NIR model was developed with the spectral region, 1078-1780nm.

Fig. 1 Near-Infrared Spectroscopy (NIR)

Model Validation

To evaluate the NIR model, the validation data set was predicted with the software IQ Predict. The reference rubber concentration was plotted against the predicted rubber concentration for the validation set and compared to the GRAMS IQ Actual vs Predicted plot, and the r^2 value of the test set, the standard error of prediction (SEP) and the ratio performance deviation (RPD) were calculated.The SEP should never be more than 1.5 times greater than the SECV. If RPD values are less than 1.75, it would indicate that the model is less reliable could not be used for quantitation[4].

Unknown Samples Prediction

303 large TK plants from six different locations were scanned using the ASD Field Spec 3, an their rubber contents predicted using the NIR model. The software will also produce an M-distance value which, if above 3, indicates that the sample characteristics are not comparable to the calibration samples and that the predicted value may not be accurate[4].

RESULTS

Rubber Prediction Model

With the use of mean centering and a 31 point Savitzky-Golay Second derivative, an NIR rubber prediction model was developed for TK roots. After removal of 21 outliers, the model had an r^2 value of 0.88 and a standard error of cross validation (SECV) of 8.3 mg/g dry weight rubber. The SECV plot indicated that most of the spectral variation was explained by the first few factors. The addition of each addded factor improved the prediction model's r^2 (1- 0.31, 2 – 0.51, 3 – 0.71, 4 – 0.82 5 - 0.85, 6 – 0.86, 7 – 0.87, 8 – 0.87, 9 – 0.88)., The dry rubber content range was of 14-205 mg rubber/g dry root in the 290 samples, a wide range of rubber contents.

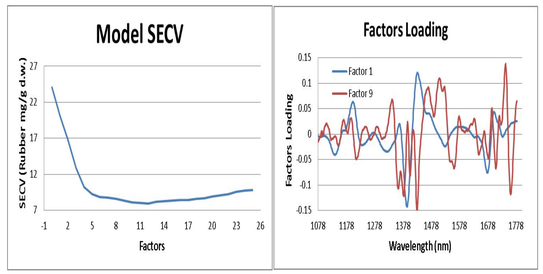

Figure 2 (left): Standard error of cross validation for each factor. Figure 3 (right): Plot shows the spectral characteristics for first and ninth factors.

Model Validation

Comparisons of the predicted rubber content to the known reference values led to the determination of a standard error of prediction of 6.6 mg rubber/g dry root. The proximity of the rubber model's SEP to the calibration SECV (8.3mg/g) indicates that the model tested accurately on the validation set. Additionally a validation r^2 value of 0.88 was achieved, indicating the model's ability to predict rubber contents with a high degree of accuracy. A ratio of performance to deviation of 2.9 was calculated for the model, meaning that the model was moderately successful and considered as a useful prediction model. Because high rubber samples were much rarer than low rubber samples, their precise rubber content may be less accurate than the rubber in the low rubber samples. All sample predicted to have high rubber are confirmed with ASE, and added back into the model to continually improve its accuracy.

Prediction the unknown samples

In the 303 predicted samples, the samples with high M-distance will be rescanned in order to get more accurate spectral data. Finally, 21 samples (6.9%) with high M-distance (over 3) were rejected. The highest rubber content was 151 mg rubber/g dry root and the lowest one was 35 mg rubber/g dry root.

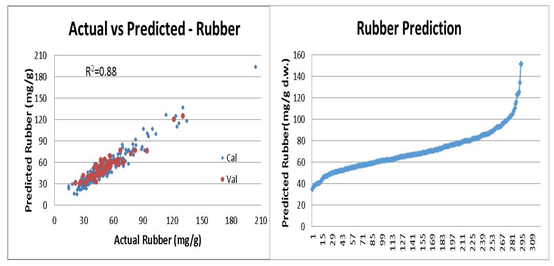

Fig. 4 (left): Predicted rubber against actual rubber concentration for the calibration and validation data sets. Fig. 5 (right): Rubber content of 303 large plants determined by NIR

CONCLUSIONS

The TK NIR rubber prediction model can be used to estimate the rubber content in unknown samples. It reduces sample processing time and the cost of accelerated solvent extractions, so improves the speed of phenotyping and TK breeding. However, it still requires dry ground roots. Therefore, the next step is to attempt to create a new NIR model with fresh roots, or perhaps rosettes, to further accelerate the breeding process.

When predicting unknown samples with the NIR model, high M-distance samples still occur, so more samples should be collected to update the NIR model, especially with samples containing with high rubber concentrations.

ACKNOWLEDGEMENTS

We would like to thank the PENRA Consortium, Ohio Third Frontier, OARDC, and the USDA National Institute of Food and Agriculture, Hatch project 230837, for financial support.

BIBLIOGRAPHY

1 Pearson, C.H.,. Cornish K, Rath D.J.. "Extraction of Natural Rubber and Resin from Guayule Using an Accelerated Solvent Extractor." Industrial Crops and Products 43 (2013): 506-10.

2 Cornish, K., Chapman, M. H., Nakayama, F. S., Vinyard, S. H. and Whitehand, L. C. Latex quantification in guayule shrub and homogenate. Industrial Crops and Products 10 (1999): 121-136.

3 Font, R. Rio-Celestion, M. Haro-bailon,."Near Infrared Reflectance Spectroscopy" Phytoremediation: Methods and Reviews. Chapter 17.(2007)

4 PANalytical. "Grams IQ/IQ Predict v 9.1 Chemometric Training Manual." ASD Document(2012): 1-75. Boulder, CO.