

Assessment of Genetic Variability of Cell Wall Degradability for the Selection of Alfalfa with Improved Saccharification Efficiency

Marc-Olivier Duceppe · Annick Bertrand ·
Sivakumar Pattathil · Jeffrey Miller · Yves Castonguay ·
Michael G. Hahn · Réal Michaud · Marie-Pier Dubé

© Her Majesty the Queen in Right of Canada 2012

Abstract Alfalfa (*Medicago sativa* L.) has a high potential for sustainable bioethanol production, particularly because of its low reliance on N fertilizer. We assessed near-infrared reflectance spectroscopy (NIRS) as a high-throughput technique to measure cell wall (CW) degradability in a large number of lignified alfalfa stem samples. We also used a powerful immunological approach, glycome profiling, and chemical analyses to increase our knowledge of the composition of CW polysaccharides of alfalfa stems with various levels of degradability. NIRS accurately predicted CW degradability in four different alfalfa cultivars, as assessed by glucose released following enzymatic saccharification ($R^2=0.94$). There was a large genetic diversity for enzyme-released glucose. The 10 genotypes with the highest (D+) and 10 genotypes with the lowest (D-) amounts of enzyme-released glucose of a biomass-type (Orca) and a winterhardy-type (54V54) cultivar were further characterized. Glycome profiling showed that there were very few differences in CW polysaccharide composition between the two groups, although the D+ genotypes were at least 35 % more degradable than the D- genotypes.

Electronic supplementary material The online version of this article (doi:10.1007/s12155-012-9204-4) contains supplementary material, which is available to authorized users.

M.-O. Duceppe · A. Bertrand (✉) · Y. Castonguay · R. Michaud ·
M.-P. Dubé
Soils and Crops Research and Development Centre,
Agriculture and Agri-Food Canada,
2560 Hochelaga Boulevard,
Québec, Québec, Canada G1V 2J3
e-mail: annick.bertrand@agr.gc.ca

S. Pattathil · J. Miller · M. G. Hahn
BioEnergy Science Center,
Complex Carbohydrate Research Center, University of Georgia,
315 Riverbend Rd.,
Athens, GA 30602, USA

Determination of CW composition by chemical analyses showed that a higher lignin content of the D- genotypes was closely related to their lower enzyme-released glucose ($R=-0.83$). In each cultivar tested, 20 D+ genotypes and 20 D- genotypes were intercrossed to generate D+ and D- populations. Assessment of CW enzyme-released glucose in the progenies showed that this trait is genetically inherited. The large genetic diversity for enzyme-released glucose and its potential for selection support the huge potential of alfalfa for the sustainable production of bio-ethanol.

Keywords Alfalfa · Cell wall degradability ·
Enzyme-released glucose · Near-infrared spectroscopy ·
Glycome profiling · Selection

Introduction

Alternative sources of fuel to replace petroleum are urgently needed because worldwide reserves are decreasing while global demand is exponentially increasing. Bioethanol production from lignocellulosic biomass represents a sustainable alternative to fossil fuels because of its low overall emission of greenhouse gases [1], its abundance and its high content of energy-rich polysaccharides [2]. On the other hand, cell wall (CW) resistance to enzymatic degradation is currently the principal barrier to ethanol production from lignocellulosic feedstock [2].

With its high biomass production and its capacity to symbiotically fix atmospheric nitrogen, perennial alfalfa (*Medicago sativa* L.) has considerable potential as a lignocellulosic feedstock for the sustainable production of bioethanol [3]. Additionally, alfalfa cultivation, harvesting and storage methods are well established. These established

processing procedures convey a strong advantage to using alfalfa as a bioenergy feedstock if we consider that the cost for biomass collection, storage, preprocessing and transport to the biorefinery gate can reach between 35 % and 65 % of the total production costs of cellulosic ethanol [4]. The value of alfalfa as a lignocellulosic feedstock depends largely on its chemical composition. More specifically, the fermentable CW sugar concentration, mainly glucose from cellulose, should be larger than that of compounds that restrict enzymatic access to fermentable carbohydrates, such as lignin and hemicellulose, during the saccharification process [3]. The relative proportion of CW fermentable polysaccharides and interfering compounds also varies with stem maturity. While the cellulose content increases, the lignin and xylose concentrations also increase, resulting in a reduction of CW degradability in mature alfalfa stems [5]. Delaying harvest from late flowering to green pod maturity stages has been shown to increase the overall availability of fermentable carbohydrates because the increased yield surpasses the loss of degradability [6]. This production system also has the advantage of minimizing crop management and harvest costs. Another advantage of alfalfa as bio-ethanol production feedstock is its relatively high degradability. For instance, we found that up to 40 % of highly lignified stems are available for ethanol production after enzymatic saccharification without any pre-treatments (unpublished results).

Alfalfa is an open-pollinated autotetraploid species ($2n=4x=32$) with extensive genetic variability that can be exploited for genetic improvement [7, 8]. Significant differences in CW composition among alfalfa cultivars and individual genotypes have been reported [5, 9], but this potential has not been fully exploited, as shown by the lack of systematic selection for this trait [3]. In a recent study, we showed that a large genetic variability of CW degradability in response to enzymatic hydrolysis exists within and among multiple alfalfa genetic backgrounds [10]. Glucose released from cell wall following enzymatic hydrolysis is a function of both glucose concentration and the efficiency with which that glucose is released. Although this previous study included only a limited number of genotypes, we found that enzyme-released glucose, a direct measure of CW degradability, varied from 70 to 120 mg g⁻¹ CW. In addition, the biomass-type cultivars of Flemish origin (Orca and Europe) were less degradable, on average, than the winterhardy-type cultivars (54V54, Amerigraze and Perfect) found in North America. This previous determination of CW degradability was based on the evaluation of a limited number of genotypes because the enzymatic screening test developed was time-consuming and labour intensive. In order to use the natural genetic diversity of alfalfa to develop new biofuel-type cultivars, high-throughput CW degradability screening methods are needed.

Near-infrared reflectance spectroscopy (NIRS) technology is based on the correlation of near-infrared reflectance spectra of scanned samples with actual analytical measurements of the constituents by other analytical methods [11, 12]. NIRS is a high-throughput method requiring minimal sample preparation. It has been successfully applied in alfalfa stem samples to predict the concentration of many structural carbohydrates [6, 13], soluble carbohydrates [14] and, with less accuracy, total lignin [13, 15]. Furthermore, NIRS has been shown to successfully predict glucose released by enzymatic saccharification following pretreatment in corn stover [16] as well as the ethanol yield in switchgrass [17]. Thus, NIRS allows the screening of a large number of genotypes, increasing the likelihood of identifying genotypes with extreme phenotypes, which are highly valuable for genetic improvement.

Characterization of CW composition of genotypes varying in degradability is essential for unravelling the underlying biochemical factors associated with the variability of stem degradability. Acid hydrolysis of CWs is a procedure commonly used to separate the complex CW matrix into simple, analysable fractions. For instance, polymeric carbohydrates are turned into monomers and total lignin is separated into acid-soluble and acid-insoluble fractions. Combined with analytical methods such as liquid or gas chromatography, spectroscopy and thermogravimetry, chemical composition and relative abundance of CW constituents can accurately be determined [18]. More recently, “glycome profiling”, a powerful tool to study plant CW structure and function was developed [19]. This technique consists of determining and comparing glycan epitope compositions amongst sequential CW extracts by ELISA using a comprehensive toolkit of 147 glycan-directed monoclonal antibodies [20].

The objectives of this study were: (a) to validate NIRS prediction of enzyme-released glucose for the high-throughput assessment of genetic variability in alfalfa, (b) to identify genotypes with high (D+) and low (D-) stem degradability, (c) to identify underlying factors associated with the differences in their degradability using glycome profiling and acid hydrolysis analysis of CWs of pooled D+ and D- genotypes and (d) to intercross genotypes with superior (D+) or inferior (D-) enzyme-released glucose and assess the selection potential for this trait.

Materials and Methods

Plant Material

Field Nursery

Alfalfa was seeded in a greenhouse on March 27, 2008 and transplanted to a field in Lévis QC, Canada (Kamouraska

clay loam; 46° 9' 15" N, 71° 12' 00" W; altitude ≈45 m) on June 4, 2008. Individual plants were transplanted at 90 cm intervals. Weeds were controlled by a combination of pre-emergence herbicides and hand weeding. A number of genotypes from each of three winterhardy-type cultivars (136 genotypes of the cultivar Perfect, 119 genotypes of the cultivar Amerigraze and 136 genotypes of the cultivar 54V54) and one biomass-type cultivar (228 genotypes of the cultivar Orca) were individually harvested on September 22, 2009 at the green pod maturity stage. Plants were harvested at this late stage of maturity to ensure that stems were lignified. At that stage, internode elongation and radial growth have ceased, and cellulose, hemicellulose and pectin concentrations are stabilized [5]. Only the bottom parts (25 cm) of the plants were harvested in order to further minimize plant maturity differences [21].

Samples were dried for 48 h at 55°C in a forced-air oven and leaves were removed by shaking. Dried stems were ground a first time using a Wiley mill (model Digital ED-5 midsized mill; Arthur H. Thomas Co., Philadelphia PA, USA) fitted with a 2-mm screen, and a second time with a Cyclone Lab Sample mill (UDY Corporation, Fort Collins, CO, USA) fitted with a 1-mm screen.

Plant Crosses

For two cultivars, Orca and 54V54, 20 genotypes with high (D+) and low (D-) degradability, as assessed by enzyme-released glucose, were identified based on NIRS predictions and a subsequent validation by enzymatic saccharification. These genotypes were transplanted in 15-cm pots and transferred to a greenhouse at the end of the 2009 growing season. Plants within each group (D+ and D-) were intercrossed to generate four populations: Orca D+, Orca D-, 54V54 D+ and 54V54 D-. Crosses were carried out by hand with toothpicks. Mature seeds were harvested from individual plants and identical amounts from each genotype within each D+ and D- group were pooled. All seed lots used and produced in this study were kept under optimal storage conditions (4°C, 10 % relative humidity).

Progeny Assessment

Approximately 110 plants of each progeny population, Orca D-, Orca D+, 54V54 D- and 54V54 D+, were sown in a greenhouse under a 25/15°C (day/night) temperature regime and a 16-h photoperiod provided by a combination of natural and artificial irradiance using mercury vapour lamps (H36GW-1000/DX; Osram Sylvania, Mississauga, ON, Canada). Plants were transplanted to the field in spring 2010. At the end of October 2010, plants were cut, transplanted into 15-cm pots and transferred to a greenhouse. At the green

pod stage, the lower 25 cm of the stems was assessed for enzyme-released glucose by NIRS [21].

Analysis of Alfalfa Stem Composition

Non-Structural Carbohydrate Quantification and CW Preparation

Removal of sugar and starch from the alfalfa biomass was done based on the method of Theander et al. [22]. Approximately 200 mg of ground stem material was incubated in 7 mL of deionised H₂O at 100°C for 90 min. Tubes were subsequently centrifuged 10 min at 1,500×g and 500 µL of the supernatant was collected for quantification of soluble carbohydrates (sucrose, glucose, pinitol and fructose) by high-performance liquid chromatography (HPLC) using a Model 515 pump, a Model 717^{plus} autosampler and a Model 2410 refractive index detector, controlled by Empower 2 software (Waters, Milford, MA, USA). Carbohydrates were separated on an Aminex HPX-87P column (300 mm×7.8 mm×9 µm) preceded by a Carbo-P pre-column (Bio-Rad, Mississauga, ON, Canada) and eluted isocratically at 80°C, at a flow rate of 0.5 mL min⁻¹, with deionised H₂O.

Tubes containing the remaining supernatant and pellet were subsequently vortexed and starch was hydrolyzed by adding 3 mL of digestion buffer (200 mM sodium acetate, pH 4.5) containing amyloglucosidase (15 U mL⁻¹; Sigma-Aldrich, Oakville, ON, Canada). Tubes were incubated 60 min at 55°C and then centrifuged 10 min at 1,500×g. The supernatant was collected for quantification of glucose by HPLC, as described above. Starch was quantified by subtracting the amount of soluble glucose from the total amount of glucose measured following digestion with amyloglucosidase. Non-structural carbohydrates (NSC) were estimated by the sum of soluble carbohydrates and starch. Pellets remaining after extraction of soluble carbohydrates and starch were washed three times with methanol at 60°C and air dried to obtain CW preparations. Approximately 115 mg of CW was obtained from 200 mg starting material.

Structural Carbohydrate and Lignin Quantification

Structural carbohydrates (glucose, xylose, galactose, arabinose, mannose, rhamnose, fucose and galacturonic acid) were quantified by HPLC following enzymatic saccharification or acidic hydrolysis.

Enzymatic saccharification was carried out according to the modified method of Selig et al. [23]. Briefly, CW preparations were incubated for 48 h at 50°C with shaking in 10 mL of buffer containing 50 mM sodium citrate, 0.02 % (w/v) sodium azide and 25 µL of each of the following enzyme preparations: Accellerase 1500 (2638 CMC U/g cellulase activity and 674 p-NPG U/g β-glucosidase activity, Genencor, Rochester, NY,

USA); Accellerase XY (22454 ABXU/g xylanase activity, Genencor), Accellerase XC (2500–3800 ABXU/g xylanase activity and 1207 CMC U/g cellulase activity, Genencor) and Pectinex 3XL (3000 FD U/g pectinase activity, Sigma-Aldrich, Oakville, ON, Canada).

Acid hydrolysis was done with H_2SO_4 according to the National Renewable Energy Laboratory (NREL) analytical protocols [18]. A 20-mL aliquot of the hydrolysate was neutralized with CaCO_3 prior to the quantification of sugars. Glucose, xylose, galactose, arabinose and mannose were quantified by HPLC using an Aminex HPX-87P column (Bio-Rad, Hercules, CA, USA) preceded by a Deashing pre-column (Bio-Rad) and eluted isocratically at 80°C , at a flow rate of 0.5 mL min^{-1} with H_2O . Galacturonic acid, rhamnose and fucose were separated using a HPX-87H column preceded by a Cation H pre-column (Bio-Rad) and eluted at 55°C with $0.01 \text{ N H}_2\text{SO}_4$ at a flow rate of 0.6 mL min^{-1} . Peak identity and quantity were determined by comparison to standards. Carbohydrate concentrations were expressed either on a 105°C dry matter (DM) basis or on a 105°C cell wall (CW) basis.

Total lignin was determined as the sum of Klason lignin and acid-soluble lignin. Klason lignin was estimated as the ash-free, acid-insoluble residue remaining after acidic hydrolysis, using a thermogravimetric analyzer (Model TGA 701, Leco, St-Joseph, MI, USA). Acid soluble lignin was quantified by spectroscopy (Lambda 35; Perkin Elmer, Waltham, MA, USA) with an absorptivity constant of $25 \text{ L g}^{-1} \text{ cm}^{-1}$ at a wavelength of 240 nm.

Near Infrared Reflectance Spectroscopy

Dried and ground alfalfa stem samples were scanned over 400 to 2498 nm, at 2 nm intervals, using a NIR-Systems 6500 monochromator (Foss, Silver Spring, MD, USA), as described in Nie et al. [14]. Seventy-five samples were selected based on the scores to form a calibration set ($n=60$) and a validation set ($n=15$). The 75 samples were chemically analysed for carbohydrate fractions (non structural and structural carbohydrates, lignin and enzyme-released glucose) as previously described.

Glycome Profiling

Sequential Extraction of CW Samples

The milled dried biomass samples were first suspended in 80 % (v/v) ethanol (100 mL g^{-1}) and stirred for 1 h followed by centrifugation at $3,000\times g$ for 15 min at room temperature. The supernatants were discarded and the pellets were sequentially washed with absolute ethanol and acetone. The resultant biomass residues were dried overnight under a hood to make the CW preparations [alcohol-insoluble residues (AIR)].

Sequential extraction of the AIR was carried out with reagents of increasing harshness on 0.3 g of four samples of alfalfa stems. Each sample was a pool of 20 genotypes of contrasting degradability representing: (a) Orca, highly degradable (D+); (b) Orca, poorly degradable (D-); (c) 54V54, highly degradable (D+); (d) 54V54, poorly degradable (D-). All extractions were done in 10 mg mL^{-1} suspensions, based on the starting AIR weight used. Each sample was first suspended in 50 mM ammonium oxalate (pH 5) with constant agitation for 24 h at room temperature. The suspension was then centrifuged at $3,400\times g$ for 15 min at room temperature to pellet the insoluble residue, and the clear supernatant was decanted and saved. The pellet was re-suspended with the same volume of deionized H_2O , centrifuged again and the supernatant was discarded. Subsequent extractions were carried out in a similar manner using sequentially the following reagents, 50 mM Na_2CO_3 containing 0.5 % (w/v) NaBH_4 (pH 10), 1 M KOH and 4 M KOH each containing 1.0 % (w/v) NaBH_4 . Following the 4 M KOH step, each sample was extracted using sodium chlorite (100 mM) [24] to break down the lignin polymers into small fragments. The insoluble residue after the chlorite extraction was once again subjected to a 4 M KOH [containing 1.0 % (w/v) NaBH_4] extraction as explained above. All 1 M KOH and 4 M KOH extracts were neutralized using glacial acetic acid. All extracts were dialyzed against deionized H_2O four times (with sample to water ratio of 1:60) at room temperature and freeze-dried individually.

Total Sugar Estimation and ELISA

All extracts were dissolved in deionised H_2O at a concentration of 0.2 mg mL^{-1} . The total sugar estimation was done using a phenol-sulphuric acid colorimetric assay [25, 26]. The ELISA plates (Costar 3598) were coated with $50 \mu\text{L well}^{-1}$ of CW extracts ($60 \mu\text{g sugar mL}^{-1}$) and allowed to evaporate to dryness overnight at 37°C . ELISAs were performed as described in Pattathil et al. [19]. Briefly, the glycome profiling analyses involves ELISA-based screening of CW materials following the sequential extractions, with a toolkit of 147 monoclonal antibodies (See Supplemental Table 1) directed against diverse plant CW glycan epitopes. The collection of 147 monoclonal antibodies used comprised 19 defined groups that recognize most major classes of plant CW glycans including non-fucosylated xyloglucans, fucosylated xyloglucans, pectic backbones, arabinogalactans and xylans [19]. ELISA results are presented as heat maps of binding intensities of antibodies to the extracts [20].

Monoclonal Antibodies

Monoclonal antibodies were obtained as hybridoma cell culture supernatants from laboratory stocks at the Complex Carbohydrate Research Center (CCRC, JIM and MAC

series antibodies are available from CarboSource Services, www.carbosource.net).

The assessment of CW components of the four pools of samples was also done by chemical analysis following acid hydrolysis as described above. Furthermore, the degradability of the four pools of samples was measured by enzymatic saccharification and was predicted by NIRS.

Statistical Analysis

Concentrations of structural carbohydrates of alfalfa genotypes were assessed individually and results for each genetic background and levels of degradability were compared statistically. The $P < 0.05$ probability level was used to determine significance. When model parameters had a significant F test, the least-significant difference method was used to compare means. All biochemical analyses were done in duplicate.

Results

NIRS Prediction of Carbohydrate Fractions

The validation statistics of the NIRS calibration (Table 1) showed that most non structural carbohydrates could be successfully predicted by NIRS. For instance, sucrose, glucose and fructose had R^2 values of 0.90, 0.93 and 0.96, respectively. Total non-structural carbohydrates were the most successfully predicted, with a R^2 value of 0.97, while starch alone was less predictable ($R^2 = 0.78$). The concentrations of the CW structural carbohydrates rhamnose, mannose, xylose and galactose, measured by HPLC after acid hydrolysis, were not well predicted, with R^2 varying from 0.09 for galactose to 0.57 for mannose. The R^2 for the prediction of soluble and insoluble lignins were of 0.78 and 0.68, respectively. CW degradability, measured as glucose released by enzymatic saccharification, was successfully predicted by NIRS with a R^2 of 0.94 (Table 1).

Assessment of CW Degradability

NIRS prediction of enzyme-released glucose of 619 highly lignified alfalfa stem samples revealed a large genetic variability for CW degradability among and within the four cultivars studied (Fig. 1). On average, degradability of winterhardy-type cultivars (54V54, Amerigraze and Perfect) tended to be higher than the biomass-type cultivar Orca ($P = 0.1$). On average, enzyme-released glucose amounts were 130.0, 136.5, 140.5 and 140.6 mg g^{-1} CW for Orca, 54V54, Amerigraze and Perfect, respectively. Within each cultivar, enzyme-released glucose amounts varied extensively: from 89.8 to 169.8 mg g^{-1} CW for Orca, from 98.0 to 168.8 mg g^{-1} CW for 54V54, from 106.3 to 177.1 mg g^{-1} CW for Amerigraze

Table 1 Statistics on carbohydrate and lignin concentrations (mg g^{-1} DM) in samples of the calibration sets used to develop near-infrared reflectance spectroscopy (NIRS) equations and statistics on the performance of each NIRS prediction

Type of components	Constituents	Statistics of sample concentrations			Calibration	
		<i>n</i>	Mean	SD	SEC	R^2
Non structural carbohydrates	Total	66	36.9	14.1	2.3	0.97
	Sucrose	57	19.6	7.1	2.2	0.90
	Glucose	56	1.8	1.3	0.4	0.93
	Pinitol	59	8.7	1.8	1.6	0.30
	Fructose	58	2.3	2.0	0.4	0.96
	Starch	63	2.5	1.7	0.8	0.78
Structural carbohydrates (acidic hydrolysis)	Total	52	778.5	30.4	19.4	0.59
	Glucose	57	437.6	18.9	11.3	0.64
	Xylose	60	142.7	15.4	11.6	0.43
	Galactose	55	22.8	5.8	5.6	0.09
	Mannose	59	75.1	22.9	15.1	0.57
	Arabinose	58	9.6	4.5	2.0	0.79
	Rhamnose	57	5.9	1.0	0.9	0.24
Lignin	Total	64	208.1	21.6	12.9	0.64
	Acid soluble lignin	59	44.9	2.8	1.3	0.78
	Klason lignin	63	163.6	22.2	12.5	0.68
CW degradability	Enzyme-released glucose	59	179.7	29.1	11.3	0.94

n number of samples, *SD* standard deviation, *SEC* standard error of calibration, R^2 coefficient of determination of calibration, *CW* cell wall

and from 104.1 to 176.8 mg g^{-1} CW for Perfect. Enzyme-released glucose amounts showed a strong negative correlation ($R = -0.83$) with total lignin concentrations in mature alfalfa stems (Fig. 2).

The 10 most (D+) and the 10 least (D-) degradable Orca and 54V54 genotypes, established by NIRS, were individually analysed to determine their CW composition as determined by acid hydrolysis (Table 2). CW degradability (enzyme-released glucose) was also confirmed by submitting the 40 genotypes to the enzymatic saccharification assay (Pearson rank correlation value of 0.84). The assay showed that for both cultivars the D+ genotypes had significantly higher amounts of enzyme-released glucose (>35 %) than the D- genotypes ($P < 0.001$). Comparing the results of enzymatic and acidic digestions showed that 36.6 % of total CW glucose was released during the enzymatic assay for the Orca D+ genotypes, as compared to only 28.9 % for the Orca D- genotypes. Similar results were obtained for the 54V54 cultivar, with 34.6 % of available CW glucose released for the D+ genotypes and only 26.4 % for the D- genotypes.

Acid hydrolysis also showed that xylose concentrations from the hemicellulose fraction of CW, as well as the galactose

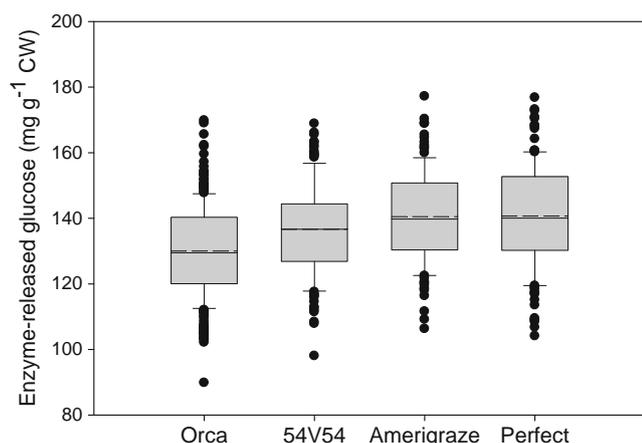


Fig. 1 NIRS prediction of glucose released from alfalfa stems by enzymatic saccharification for the biomass-type cultivar Orca (228 samples), and the three winterhardy-type cultivars 54V54 (136 samples), Amerigraze (119 samples) and Perfect (136 samples). The *full line in boxes* represents the 50th percentile (the median) and the *dotted line* represents the mean. A *box* represents the interquartile range (IQR) in which 50 % of the ranked data are found (the *bottom and top of the box* are the 25th and 75th percentile, respectively). *Dots* represent the outliers which values exceed ± 1.5 IQR

concentrations from the pectin fraction of CW, differed significantly between Orca and 54V54 cultivars: concentrations of both components were higher in 54V54 (Table 2; $P=0.004$). Within each cultivar, lignin and enzyme-released glucose differed significantly between the two classes of degradability, D+ and D- ($P<0.0001$). Average lignin concentrations were higher in D- (222 mg g⁻¹ CW) than in D+ genotypes (198 mg g⁻¹ CW). Conversely, amounts of enzyme-released glucose were higher in D+ (169 mg g⁻¹ CW) than in D- genotypes (124 mg g⁻¹ CW).

Cellulose concentrations, assessed as acid hydrolysis-released glucose, differed significantly between both cultivars

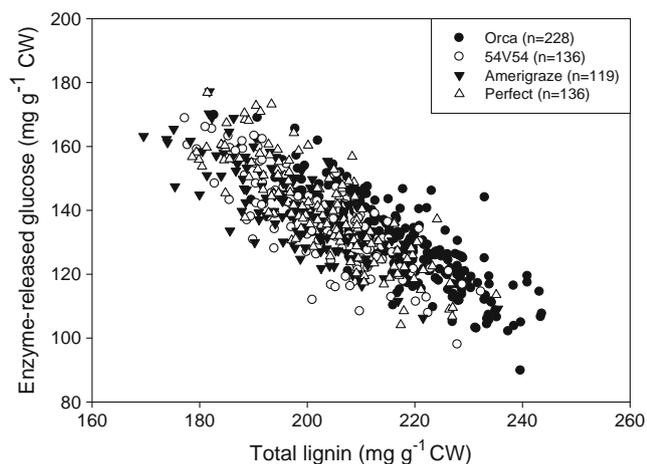


Fig. 2 Relationship between the amounts of enzyme-released glucose and total lignin in alfalfa stems of four cultivars, as predicted by NIRS ($y=-0.7077x+308.88$, $R^2=0.6914$)

($P=0.002$) and degradability groups ($P=0.02$), with higher concentrations in 54V54 than in Orca and in D+ than in D- (Table 2).

Glycome Profiling Studies

Ground stem samples of the 20 most (D+) and 20 least (D-) degradable Orca and 54V54 genotypes were pooled and submitted to glycome profiling. Overall, the glycome profiles of Orca D+, Orca D-, 54V54 D+ and 54V54 D- pools were similar, indicating that there were no major differences in CW carbohydrate extractability or CW epitope composition between the samples tested (Fig. 3). Small differences between the two genetic backgrounds (Orca vs 54V54) were nevertheless present. For instance, the binding patterns of the xylan-3 and the xylan-4 groups of antibodies showed slight differences in the carbonate and 1 M KOH extracts when comparing the two genetic backgrounds.

The winterhardy type 54V54 pools exhibited slightly more oxalate extractable pectin (as indicated by the higher binding of HG-backbone and RG-I backbone groups of antibodies) and arabinogalactan and pectic arabinogalactan epitopes (as indicated by the higher binding of RG-I/AG, AG-1, AG-2, AG-3 and AG-4 groups of antibodies) than do the Orca pools. These differences are marked in Fig. 3 as dotted white blocks. There are other subtle differences in the glycome profiles between the different plant pools. The significance of these differences, if any, in terms of altered recalcitrance, will require additional studies.

Selection for Stem Degradability

The 20 most (D+) and the 20 least (D-) degradable Orca and 54V54 genotypes identified by NIRS were pooled into four groups (Orca D-, Orca D+, 54V54 D- and 54V54 D+) and submitted to the enzyme-released glucose assay to confirm the predicted values and to ensure we had truly contrasting material to proceed with the crosses. These four samples were the same ones used for glycome profiling studies. Predicted and measured enzyme-released glucose amounts obtained from the pools were very similar (Fig. 4).

For each cultivar, there was a significant difference of degradability between the D+ and the D- pools ($P<0.001$, Figs. 4 and 5, Parents), with average enzyme-released glucose amounts of 156 mg g⁻¹ CW for the D+ pools and 116 mg g⁻¹ CW for the D- pools. This contrasting material was then suitable to start a first cycle of recurrent selection. No significant difference was observed between the two cultivars.

Assessment of enzyme-released glucose of 110 progenies issued from each of the four crosses (Orca D-, Orca D+, 54V54 D- and 54V54 D+) revealed significant differences in enzyme-released glucose amounts between the classes of degradability and the cultivars ($P<0.001$; Fig. 5, Progenies).

Table 2 Average CW composition (\pm SD) of 10 genotypes of Orca D+, Orca D-, 54V54 D+ and 54V54 D-^a

CW fraction	Constituents	Cultivar and degradability groups			
		Orca D+ mg g ⁻¹ CW	Orca D-	54V54 D+	54V54 D-
Cellulose	Glucose	470.2 \pm 46.0a ^{1*}	434.0 \pm 36.9a ²	488.1 \pm 19.7b ¹	472.4 \pm 13.2b ²
Hemicellulose	Xylose	141.6 \pm 25.2a	147.8 \pm 19.4a	170.9 \pm 9.2b	182.9 \pm 10.2b
	Mannose	40.9 \pm 2.3	38.5 \pm 2.1	39.1 \pm 2.4	39.3 \pm 1.8
Pectin	Arabinose	32.7 \pm 5.4	30.8 \pm 3.4	34.4 \pm 9.1	33.1 \pm 6.3
	Galactose	48.0 \pm 8.3a	43.6 \pm 9.5a	52.4 \pm 7.4b	52.0 \pm 6.9b
	Galacturonic acid	18.6 \pm 3.1a ¹	16.8 \pm 2.0a ²	20.6 \pm 2.5b ¹	18.7 \pm 3.2b ²
Lignin	Total	204.4 \pm 11.8 ¹	221.7 \pm 10.5 ²	190.8 \pm 18.7 ¹	222.7 \pm 8.4 ²
CW degradability (measured)	Enzyme-released glucose	170.9 \pm 19.6 ¹	124.5 \pm 14.5 ²	167.1 \pm 16.0 ¹	123.2 \pm 9.8 ²
CW degradability (predicted)	Enzyme-released glucose	160.9 \pm 5.8 ¹	111.9 \pm 4.1 ²	163.1 \pm 3.0 ¹	111.5 \pm 5.3 ²

Cellulose, hemicelluloses and pectin constituents were quantified by HPLC following acid hydrolysis. Total lignin was determined as the sum of Klason lignin and acid-soluble lignin. Predicted CW degradability values were obtained by NIRS while measured CW degradability was assessed by HPLC following enzymatic hydrolysis

^a Values followed by different letters indicate significant differences between genotypes and values followed by different numbers have significantly different levels of degradability (*t* test, $\alpha=0.05$)

The average enzyme-released glucose amounts of 54V54 D+ and D- progenies were of 196.5 mg g⁻¹ CW and 185.6 mg g⁻¹ CW, respectively. The difference between

D+ and D- Orca progenies was even larger, with average amounts of 193.3 mg g⁻¹ CW for the D+ and of 163.3 mg g⁻¹ CW for the D-.

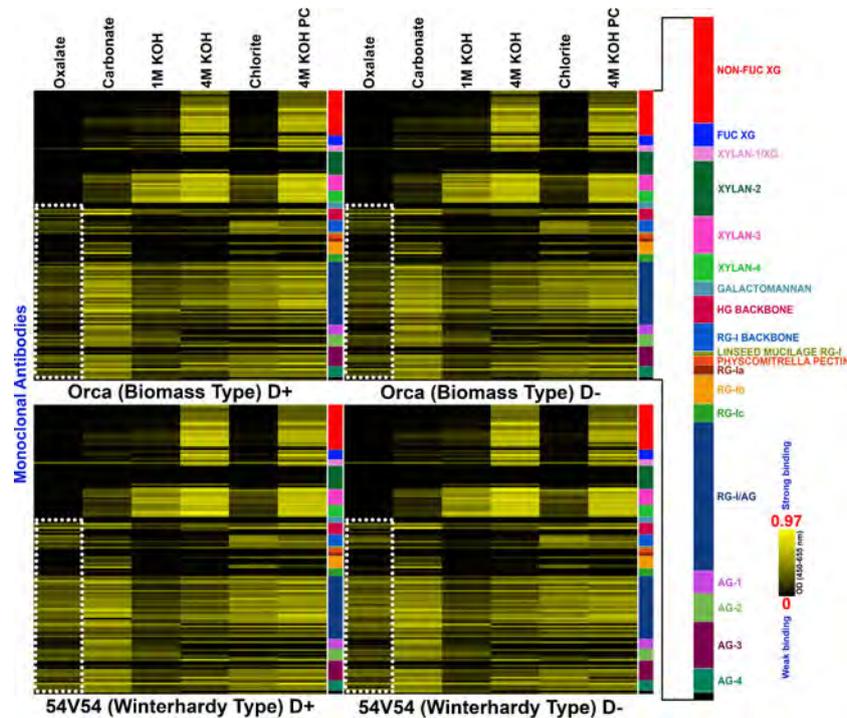


Fig. 3 Glycome profiling of sequential cell wall extracts of four alfalfa biomass samples pooled from 20 plants: (a) high degradability (D+) biomass-type cultivar Orca, (b) low degradability (D-) cultivar Orca, (c) high degradability (D+) winterhardy-type cultivar 54V54, (d) low degradability (D-) cultivar 54V54. The 18 groups of antibodies against CW glycan epitopes (See Supplemental Table 1) are identified on the right panel of the figure, while the cell wall extractants used are

indicated at the top of the figure in order of increasing harshness (left to right). Binding intensities of antibodies to the extracted CW glycans, as measured in ELISAs, are depicted as heatmaps with black to bright yellow colors representing no binding to strongest binding, respectively. The boxes outlined in white dots highlight the largest differences observed in the glycome profiles between the four alfalfa samples

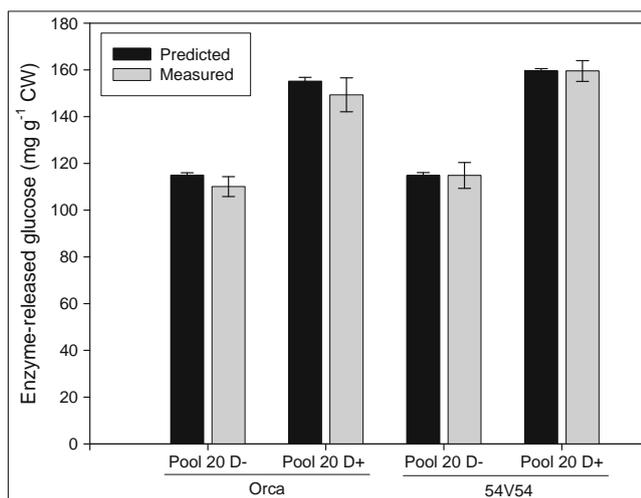


Fig. 4 Amounts of enzyme-released glucose as assessed by NIRS (black) and by HPLC (gray). Pools of the 20 most degradable and the 20 least degradable genotypes of the biomass-type cultivar Orca and the winterhardy-type 54V54 were used. Each bar represents the mean \pm SE ($n=3$)

Discussion

NIRS Predictions of CW Degradability

High-throughput screening methods for CW degradability are needed to identify genetic material prone to enzymatic degradation and, consequently, to develop a biofuel-type alfalfa germplasm. In the current report, we document that

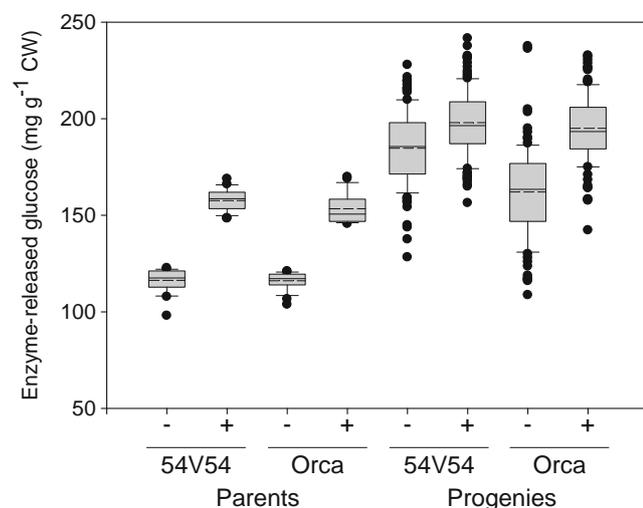


Fig. 5 NIRS prediction of glucose amounts released by enzymatic saccharification of alfalfa stems in parent plants (20 plants for each group of parents) and in progenies obtained after one cycle of recurrent selection (110 plants for each group of progenies). Parent plants were field-grown and the 20 most (D+) and 20 least (D-) degradable genotypes of both Orca and 54V54 cultivars were selected. Progenies were derived from crosses within the 20 genotypes of each 54V54 D-, 54V54 D+, Orca D- and Orca D+ and were greenhouse-grown. See Fig. 1 for box plot explanations

NIRS can successfully predict glucose released by enzymatic saccharification of lignified alfalfa stems, the core step in bio-ethanol production from cellulosic feedstock. We gathered evidence that this screening approach is sensitive, robust and reliable and that it can readily predict enzyme-released glucose of hundreds of alfalfa genotypes of various cultivars. NIRS has been shown to successfully predict enzyme-released glucose in maize stovers [16] as well as ethanol yield in switchgrass [17]. In alfalfa, in vitro neutral detergent fibre digestibility (dNDF) has been successfully predicted [27]. dNDF reflects the availability of cell wall carbohydrates to rumen enzymes, a trait closely associated to the convertibility of biomass into ethanol. Using NIRS screening, we found that an extensive genetic variability for CW degradability exists within and between Orca and winterhardy-type alfalfa cultivars. All cultivars considered, enzyme-released glucose amounts varied from 90 to 180 mg g⁻¹ CW. We also found that enzyme-released glucose was generally greater in winterhardy-type cultivars than in the cultivar Orca. Finally, we observed some morphological differences at harvest between the cultivars under study: Orca had larger stem diameter with hollow pith and longer internodes than the winter-hardy type cultivars.

Assessment of CW Composition and Degradability

The strong negative correlation between enzyme-released glucose and lignin content support the idea that cross-linking of lignin with CW polysaccharides interferes with enzymatic degradation of cellulose and limits the conversion of herbaceous plant material into ethanol, as reported previously [3, 28]. Similar results have been observed with transgenic alfalfa featuring low lignin content [29]. Suppression of early lignin biosynthesis genes led to a 50 % reduction in lignin content in comparison to the wild-type phenotype. In comparison, using our NIRS-based screening approach, with all cultivars considered, we found that the D+ genotypes had on average 20 % less lignin than the D- genotypes, with total lignin concentrations of 185 mg g⁻¹ CW and 231 mg g⁻¹ CW, respectively. Low-lignin alfalfa plants, such as the D+ lines identified here, could reduce the reliance on high-cost pre-treatments that are commonly used with lignocellulosic feedstock prior to enzymatic saccharification, the products of which can inhibit fermentation [30]. On the other hand, a lower lignin content could also reduce biomass yield [31]. For instance, Chen and Dixon [29] observed a 40 % yield reduction in low-lignin alfalfa transgenic lines compared to wild-type alfalfa. However, this yield loss was offset by a large increase in fermentable sugar content. Taken together, our results suggest that NIRS screening of broad-based populations is a promising tool for the development of alfalfa having low lignin content while maintaining high biomass yield.

Alfalfa is a good potential bioethanol feedstock [6], but much of the effort to develop new biomass-type crops has focussed on grasses because they produce high yields of cellulose-rich biomass [31]. With concentrations of glucose in the cell wall, determined by acid hydrolysis, exceeding 430 mg g^{-1} CW (Table 2), mature alfalfa stems can yield higher cellulose concentrations than grasses [31]. After making a conversion from mg g^{-1} DW to mg g^{-1} of CW and considering an average of 70 % CW, the glucose concentration of alfalfa stems that we measured is comparable to values around 420 mg g^{-1} CW reported by others for alfalfa [15, 32]. The slightly higher cellulose concentrations observed here are likely due to a more advanced stage of maturity (green pod) at harvest as compared to the full flower stage used in other reports, as well as to the fact that we only used the highly lignified lower part of the stems (25 cm) for our assessments. Our rationale for using more mature plants was based on the fact that, while the recovery of enzyme-released glucose decreases with maturity due to higher lignin concentrations, glucose yield may actually increase because of higher cellulose content [32]. Larger cellulose contents are advantageous for ethanol production because glucose is efficiently converted to ethanol by industrial yeast strains [33].

Glycome Profiling Studies

Glycome profiling was previously done on a number of different plant biomass samples. The profiles obtained for alfalfa pools in our study are distinct from other plant species such as poplar [34], nicotiana [35] and sorghum [36]. The glycome profiling approach showed that CW composition of alfalfa stems from 54V54 and Orca as well as of D+ and D- groups share many similarities. Acidic hydrolysis revealed that D+ and D- genotypes had significantly different concentrations of glucose, galacturonic acid and total lignin. Also, Orca and 54V54 had different amounts of glucose, xylose, galactose and galacturonic acid. Interestingly, those significant differences in CW composition did not result in major shifts in CW carbohydrate extractability, as shown by glycome profiling. However, glycome profiling revealed a few differences between the samples analysed, with more differences observed between the two genetic backgrounds (Orca versus 54V54) than between the two levels of degradability (D+ versus D-). The glycome profile differences observed between the two cultivars suggest marginally higher polysaccharide extractability for the winterhardy-type than for the biomass-type cultivar, which is in accordance with the results obtained from acid hydrolysis. On the other hand, although chemical analyses showed that lignin concentrations differed markedly between D+ and D-, CW polysaccharide extractability

was not significantly affected. Altogether, these results suggest that the >35 % difference in CW degradability (enzyme-released glucose) between D+ and D- genotypes, as predicted by NIRS and validated by biochemical assays, could be associated with subtle differences in CW extractable polysaccharides.

Selection for CW Degradability

One of the major roadblocks in the improvement of lignocellulosic biomass conversion to ethanol relates to the difficulty in exploiting a large heterogeneity of stem degradability to develop improved genetic material. Here, we showed that NIRS can be fully integrated in a breeding program for identifying germplasm with highly contrasting CW degradability as assessed by enzyme-released glucose. This method was applied to identify 20 high-degradable D+ genotypes that were >35 % more degradable (154 mg g^{-1} CW of enzyme released glucose) than the 20 low-degradable D- genotypes (112 mg g^{-1} CW). The assessment of enzyme-released glucose from progenies issued from these crosses confirmed that this trait is genetically inherited. Around 110 progenies from each cross were assessed for enzyme-released glucose by NIRS prediction, confirming significant differences between both cultivars and groups of genotypes. On average, enzyme-released glucose differed by 6 % between the D+ and the D- progenies for 54V54 while it differed by 18 % between these groups in Orca, showing a strong response to divergent selection for both cultivars. Interestingly, when assessing enzyme-released glucose of the 25 most degradable progenies at late flowering, we obtained values greater than 220 mg g^{-1} CW for both Orca and 54V54 without pre-treatment (data not shown). Yield of enzymatically released glucose is the result of both glucose concentration in the biomass and the efficiency with which that glucose is released from polysaccharides. Although glucose concentration is higher in D+ selections due to lower lignin content, our results clearly show that D+ lines had significantly higher glucose released yield than D- lines. Whether the two germplasms achieved improved glucose released in D+ lines by the same mechanism remains to be investigated.

Conclusions

NIRS can accurately predict CW degradability (enzyme-released glucose) in mature alfalfa stems of various cultivars. The use of this high-throughput screening method on a large number of genotypes revealed a large genetic variability for this trait in alfalfa. Through recurrent selection, we were able to exploit this heterogeneity of stem degradability to produce a new germplasm more prone to

enzymatic saccharification. Assessment of CW degradability (enzyme-released glucose) of progenies, after one cycle of divergent selection, clearly showed the possibility to select for this trait. Progenies with contrasting CW degradability characteristics are currently being used to develop molecular markers for this trait.

Acknowledgments The authors sincerely thank Mrs. Josée Bourassa for her technical assistance and Monique Arts for proof-reading the article. This work was supported by the Cellulosic Biofuel Network of the Agricultural Bioproducts Innovation Program of Agriculture and Agri-Food Canada. The glycome profiling was supported by the Bio-Energy Science Center administered by Oak Ridge National Laboratory and funded by a grant (DE-AC05-00OR22725) from the Office of Biological and Environmental Research, Office of Science, United States, Department of Energy. The generation of the CCRC series of cell wall glycan-directed monoclonal antibodies used in this work was supported by the United States National Science Foundation Plant Genome Program (DBI-0421683).

References

- Charles D (2009) Corn-based ethanol flunks key test. *Science* 324:587
- Gomez LD, Whitehead C, Barakate A, Halpin C, McQueen-Mason SJ (2010) Automated saccharification assay for determination of digestibility in plant materials. *Biotechnol Biofuels* 3:1–12
- Samac DA, Jung HG, Lamb JFS (2006) Development of alfalfa (*Medicago sativa* L.) as a feedstock for production of ethanol and other bioproducts. In: Minteer S (ed) *Alcoholic Fuels*. CRC Press, Boca Raton, pp 79–98
- Feedstock Logistics Interagency Working Group (2010) Biofuel feedstock logistics: recommendations for research and commercialization. Available at http://www.usbiomassboard.gov/pdfs/biomass_logistics_2011_web.pdf (accessed 18 April 2012)
- Jung H-JG, Engels FM (2002) Alfalfa stem tissues: cell wall deposition, composition, and degradability. *Crop Sci* 42:524–534
- Lamb JFS, Jung H-JG, Sheaffer CC, Samac DA (2007) Alfalfa leaf protein and stem cell wall polysaccharide yields under hay and biomass management systems. *Crop Sci* 47:1407–1415
- Jones JS, Bingham ET (2010) Inbreeding depression in alfalfa and cross-pollinated crops. In: Janick J (ed) *Plant Breeding Reviews*, vol 13. John Wiley & Sons Inc, Oxford, pp 209–233
- Julier B, Huyghe C, Ecalte C (2000) Within- and among-cultivar genetic variation in alfalfa: forage quality, morphology, and yield. *Crop Sci* 40:365–369
- Sheaffer CC, Martin NP, Lamb JFS, Cuomo GR, Jewett JG, Quering SR (2000) Leaf and stem properties of alfalfa entries. *Agron J* 92:733–739
- Bertrand A, Duceppe M-O, Castonguay Y, Michaud R, Dubé M-P (2010) Near-infrared reflectance spectroscopy prediction of cell wall degradability in alfalfa stems. *Proceedings 2nd Pan American Congress on Plants and Bioenergy*. Sao Pedro, Brasil p. 23
- Sanderson MA, Agblevor F, Collins M, Johnson DK (1996) Compositional analysis of biomass feedstocks by near infrared reflectance spectroscopy. *Biomass Bioenergy* 11:365–370
- Jung H-JG (1997) Analysis of forage fiber and cell walls in ruminant nutrition. *J Nutr* 127:S810–S813
- Jung H-JG, Lamb JFS (2003) Identification of lucerne stem cell wall traits related to *in vitro* neutral detergent fibre digestibility. *Anim Feed Sci Technol* 110:17–29
- Nie Z, Tremblay GF, Belanger G, Berthiaume R, Castonguay Y, Bertrand A et al (2009) Near-infrared reflectance spectroscopy prediction of neutral detergent-soluble carbohydrates in timothy and alfalfa. *J Dairy Sci* 92:1702–1711
- Jung H-JG, Lamb JFS (2004) Prediction of cell wall polysaccharide and lignin concentrations of alfalfa stems from detergent fiber analysis. *Biomass Bioenergy* 27:365–373
- Lewis MF, Lorenzana RE, Jung H-JG, Bernardo R (2010) Potential for simultaneous improvement of corn grain yield and stover quality for cellulosic ethanol. *Crop Sci* 50:516–523
- Schmer MR, Vogel KP, Mitchell RB, Dien BS, Jung HG, Casler MD (2012) Temporal and spatial variation in switchgrass biomass composition and theoretical ethanol yield. *Agron J* 104:54–64
- Sluiter A, Hames B, Ruiz R, Scarlata C, Sluiter J, Templeton D, et al. (2008) Determination of structural carbohydrates and lignin in biomass. In: NREL (ed) *Laboratory Analytical Procedure*, Midwest Research Institute, Golden, pp 1–13
- Pattathil S, Avci U, Baldwin D, Swennes AG, McGill JA, Popper Z et al (2010) A comprehensive toolkit of plant cell wall glycan-directed monoclonal antibodies. *Plant Physiol* 153:514–525
- Pattathil S, Avci U, Miller JS, Hahn MG (2012) Immunological approaches to cell wall and biomass characterization: Glycome profiling of plant cell wall/biomass. In: Himmel M (ed.) *Biomass Conversion: Methods and Protocols*, Humana Press, New York (in press).
- Buxton DR, Hornstein JS, Marten GC (1987) Genetic variation for forage quality of alfalfa stems. *Can J Plant Sci* 67:1057–1067
- Theander O, Aman P, Westerlund E, Andersson R, Petersson D (1995) Total dietary fiber determined as neutral sugar residues, and Klason lignin (The Uppsala Method): Collaborative study. *J AOAC Int* 78:1030–1044
- Selig M, Weiss N, Ji Y (2008) Enzymatic saccharification of lignocellulosic biomass. In: NREL (ed) *Laboratory Analytical Procedure*, Midwest Research Institute, Golden, pp 1–5
- Ahlgren PA, Goring DAI (1971) Removal of wood components during chlorite delignification of black spruce. *Can J Chem* 49:1272–1275
- Dubois M, Gilles KA, Hamilton JK, Rebers PA, Smith F (1956) Colorimetric method for determination of sugars and related substances. *Anal Chem* 28:350–356
- Masuko T, Minami A, Iwasaki N, Majima T, Nishimura S-I, Lee YC (2005) Carbohydrate analysis by a phenol-sulfuric acid method in microplate format. *Anal Biochem* 339:69–72
- Brogna N, Pacchioli MT, Immovilli A, Ruozzi F, Ward R, Formigoni A (2009) The use of near-infrared reflectance spectroscopy (NIRS) in the prediction of chemical composition and *in vitro* neutral detergent fiber (NDF) digestibility of Italian alfalfa hay. *Ital J Anim Sci* 8:271–273
- Mansfield SD, Mooney C, Saddler JN (1999) Substrate and enzyme characteristics that limit cellulose hydrolysis. *Biotechnol Prog* 15:804–816
- Chen F, Dixon RA (2007) Lignin modification improves fermentable sugar yields for biofuel production. *Nat Biotechnol* 25:759–761
- Chapple C, Ladisch M, Meilan R (2007) Loosening lignin's grip on biofuel production. *Nat Biotechnol* 25:746–748
- Casler MD, Buxton DR, Vogel KP (2002) Genetic modification of lignin concentration affects fitness of perennial herbaceous plants. *Theor Appl Genet* 104:127–131
- Dien BS, Jung H-JG, Vogel KP, Casler MD, Lamb JFS, Iten L et al (2006) Chemical composition and response to dilute-acid pretreatment and enzymatic saccharification of alfalfa, reed canarygrass, and switchgrass. *Biomass Bioenergy* 30:880–891

33. Yamada R, Taniguchi N, Tanaka T, Ogino C, Fukuda H, Kondo A (2011) Direct ethanol production from cellulosic materials using a diploid strain of *Saccharomyces cerevisiae* with optimized cellulase expression. *Biotechnol Biofuels* 4:1–8
34. DeMartini JD, Pattathil S, Avci U, Szekalski K, Mazumder K, Hahn MG et al (2011) Application of monoclonal antibodies to investigate plant cell wall deconstruction for biofuels production. *Energy Environ Sci* 4:4332–4339
35. Zhu X, Pattathil S, Mazumder K, Brehm A, Hahn MG, Dinesh-Kumar SP et al (2010) Virus-induced gene silencing offers a functional genomics platform for studying plant cell wall formation. *Mol Plant* 3:818–833
36. Lee SJ, Warnick TA, Pattathil S, Alvelo-Maurosa JG, Serapiglia MJ, McCormick H et al (2012) Biological conversion assay using *Clostridium phytofermentans* to estimate plant feedstock quality. *Biotechnol Biofuels* 5:5