Abstract: Marine macroalgal biomass is a promising sustainable feedstock for biorefineries. However, the development of macroalgal biomass for industrial cultivation and processing has been slow. In terrestrial plants, high-throughput phenotyping provides rapid imaging methods to select specimens with required properties, rapidly transforming traditional breeding techniques. To foster the development of macroalgal biomass for biorefinery applications, we developed a near-infrared spectrometry-based method for rapid phenotyping of the macroalga Ulva fasciata based on its glucose, rhamnose, xylose and glucuronic acid contents. Spectral slopes were calculated as indicative of major carbohydrate content change. In addition, different spectral indices were generated to distinguish between low and high contents of glucose, rhamnose, xylose and glucuronic acid in wet and dry biomass. Since glucose is a major monosaccharide in Ulva that is fermentable to bioethanol, as an example of future application, we developed a multivariate data analysis based on partial least squares regression to predict its content in dry and wet biomass samples solely from reflectance data. These methods could provide a useful, high-throughput tool to rapidly select thalli with high carbohydrate content for further propagation and to be used for feedstock development for marine biorefineries.

Keywords: biorefinery; macroalga Ulva; near-infrared spectroscopy; phenotyping; single thallus cultivation.

Introduction

Biorefinery is a collective term for the complex system that includes biomass production, transportation, conversion into products and product distribution. Concerns over net energy balance, potable water use, environmental hazards and processing technologies call into question the potential for terrestrial biomass, such as cereal crops and lignocellulose biomass, to provide sustainable feedstock for biorefineries (Gerbens-Leenes et al. 2009). In addition, cost-effective cultivation and further processing difficulties currently prevent the implementation of large-scale microalgal technologies (Hannon et al. 2010). Alternatively, an expanding body of evidence has demonstrated that marine macroalgae can provide a sustainable alternative source of biomass for food, animal feed, fuel and production of chemicals (Bruhn et al. 2011, Wargacki et al. 2012, Kraan 2013, van der Wal et al. 2013, Aitken et al. 2014, Golberg et al. 2014). Green macroalgae from the genus Ulva are of particular interest because of their high growth rates and fermentable carbohydrate content (Golberg et al. 2014, Korzen et al. 2015a,b, Vitkin et al. 2015).

In addition to technological improvements needed for large-scale cultivation, significant effort is required to develop and select macroalgal species and strains with specific properties tailored for food, chemicals or fuel applications (Robinson et al. 2012). Industrial strain development in agricultural crops was transformed by the introduction of high-throughput phenotyping with imaging. Imaging methods, such as imaging spectroscopy, thermal infrared imaging, fluorescence imaging, 3D imaging and tomographic imaging, have been used to determine the required properties of plants for robust, cost-effective downstream processing (Cabrera-Bosquet
et al. 2012, Li et al. 2014, Mutka and Bart 2015, Walter et al. 2015). However, to the best of our knowledge, none of these methods have been used for macroalgae strain selection.

Near-infrared spectroscopy (NIRS) is a widely used technology with broad applications in virtually all fields of science (Blanco and Villarroya 2002) and has been extensively used for high-throughput breeding of plants. In macroalgae, spectral signatures based on absorption features can be used as indicative of the macroalgal type and condition (Ramsey et al. 2002a,b, Malta et al. 2003). Furthermore, thallus absorption in the visible (VIS) wavelength region is the optical feature most closely aligned with pigments and concentration changes (Ramsey et al. 2002a,b). Jiao and Liu (1999) proposed a simple spectrophotometric method to estimate microgram quantities of algal polysaccharide. The method was based on the binding of a dye to algal polysaccharides that causes the absorption maximum at 664 nm to decrease linearly over the range of 0–30 μg. In a previous NIRS work, Robic et al. (2009) developed a partial least squares (PLS) model to characterise the chemical composition of ulvan, a major polymer of Ulva, with functional properties making it interesting for industrial production. That work focused on determining the chemical composition of the extracted ulvan (Robic et al. 2009). However, to the best of our knowledge, NIRS has not been used for complete Ulva biomass phenotyping based on its content of major monosaccharides, which are critically important for the rapid selection of macroalgal feedstock for biofuel and biorefinery applications (Golberg et al. 2014, Korzen et al. 2015a,b, Vitkin et al. 2015).

The current study explored direct Ulva biomass phenotyping based on spectral slopes and spectral index analyses of monosaccharides. Specifically, we investigated the potential of NIRS to distinguish between low and high contents of glucose, rhamnose, xylose and glucuronic acid from fresh and dried tissues of Ulva fasciata L. based on the differences in the spectral signatures of thalli. We decided to focus on these molecules as they are the major carbohydrates in the Ulva species. In addition, in our previous work, we showed that the glucose, rhamnose and xylose contents of U. fasciata predict the potential of this feedstock for bioethanol fermentation (Vitkin et al. 2015). In addition to examining the potential use of reflectance spectroscopy for deriving different sugar contents, we developed PLS to predict the contents of specific monosaccharides based on whole thalli spectrum features. The methodology and tools developed in this study will facilitate the high-throughput phenotyping of macroalgal biomass, enabling its use as a major feedstock for sustainable biorefinery.

Materials and methods

Cultivation experiments

The model species used in this study was Ulva fasciata L., a green marine macroalga of worldwide distribution found in the intertidal and shallow waters within the Israeli Mediterranean shores. For the current study, specimens were taken from stocks maintained at a seaweed collection at Israel Oceanographic and Limnological Research, Haifa, Israel (IOLR). Cultivation trials were conducted in an outdoor setting at IOLR. One selected thallus was cut into 16 pieces, measuring approximately 5 cm² [weighing 0.4–0.6 g fresh weight (FW) each]. This single mother plant was used to produce 16 clones with identical genetic backgrounds needed for the comparative nature of this study. Single pieces were placed in 15 × 15-cm plastic net baskets. A total of 16 baskets, with a single thallus per basket, were divided into two groups, eight baskets each, and tied to 40-l fibreglass tanks supplied with running seawater and aeration. The first group of eight baskets was supplied with 1 mM NH₄Cl and 0.1 mM NaH₂PO₄, every week. With each nutrient application, the water exchange was stopped for 24 h to allow for absorption. The second group of eight baskets served as controls and had no nutrient supplements. The total cultivation time for both groups during March 2015 was 4 weeks under seawater temperatures ranging from 23 to 25°C and irradiances of 100–200 μmol photons m⁻² s⁻¹. Each week, each thallus was weighed and returned to the basket. Growth rate was defined as the biomass accumulation for each of 16 thalli per week at the above-specified experimental conditions.

Drying and thermochemical hydrolysis

Following the growth period in the outdoor tanks, each thallus was dried separately at 60°C for 48 h until constant weight and ground into powder manually in a mortar with liquid nitrogen. For hydrolysis, 0.05 g of Ulva fasciata powder was mixed with 2 ml of sulphuric acid (5%) in 10-ml plastic tubes and autoclaved at 120°C for 45 min. Next, 500 mM phosphate buffer (Sigma, Israel) was added into the above mix and the resulting hydrolysate was neutralised with 3 M sodium hydroxide (NaOH) to pH 7.
Monosaccharide and sugar acid content determination with ion chromatography

Dionex ICS-5000 (Thermo Fischer Scientific, CA, USA) was used to quantify the monosaccharides glucose, rhamnose, xylose and glucuronic acid in the hydrolysates. Carbopac MA1 (Thermo Fischer Scientific, MA, USA) and its corresponding guard column were used for separation. An electrochemical detector with AgCl as reference electrode was used for detection. A trinary solvent system (Table 1) was used for elution. The column temperature was kept at 30°C and flow rate 0.25 ml min⁻¹. Calibration curves were made for rhamnose, glucose, xylose and glucuronic acid (0–100 μg ml⁻¹) to determine the concentration of corresponding substances in the hydrolysates.

Spectral reflectance measurements

Wet (n=16) and dry (n=16) samples, derived from the same initial thallus, were scanned by a FieldSpec Analytical Spectral Devices (ASD) full-range (FR) spectrometer (Analytical Spectral Devices, CO, USA) at three locations on each thallus. The FR spectrometer samples a spectral range of 350–2500 nm (http://www.asdi.com). The instrument uses three detectors spanning the visible, near-infrared (VNIR, comprising a Si photodiode array) and shortwave infrared (SWIR1 and SWIR2, comprising two separate InGaAs photodiodes). All samples were measured in the laboratory by attaching the high-intensity ASD contact probe (“potato”) device to the sample and extracting an average of 40 readings, using bare fibre and self-probed illumination. The “potato” was set on a stable tripod base and maintained in a constant position at a nadir-looking angle. For all measurements, we used a Spectralon standard white reference panel (Spectralon, Labsphere Inc., www.labsphere.com) in the same geometry as a white reference to enable conversion of the measurement data into reflectance values.

NIRS data analyses: slope and spectral index calculation

In this study, the spectral behaviour vs. chemical information on the wet and dry Ulva was visually inspected. The differences in spectral behaviour as a function of wavelength, which are indicative of changes in chemical properties, were identified. Importantly, the main criterion for spectral slope selection across the entire spectral range was similar for both wet and dry samples (Lugassi et al. 2015a,b). The slope was calculated using the following equation (Equation 1):

\[
\text{Slope} = \frac{y_2 - y_1}{x_2 - x_1}
\]

where the change in \( y \) signifies the change in spectral reflectance between two wavelengths and the change in \( x \) is the spectral range. We studied the dependence of spectral shape parameters on different sugar content distributions for all samples.

In addition, we calculated the normalised difference index (NDI) as follows:

\[
\text{NDI} = \frac{R_{\lambda_2} - R_{\lambda_1}}{R_{\lambda_1} + R_{\lambda_2}}
\]

where \( R_{\lambda_1} \) and \( R_{\lambda_2} \) are the values of spectral reflectance at two selected wavelengths. This index aims to identify the region of a rapid change in reflectance and to differentiate across the electromagnetic spectrum between Ulva thalli growth with and without external nutrient supply. To that end, the differences between two wavelengths were visually inspected and identified.

Feasibility of carbohydrate quantification in Ulva thalli using PLS

Partial least squares (PLS) regressions were run, with the goal being to define a relationship between different spectra and the concentration of each chemical constituent (e.g. rhamnose, xylose and glucuronic acid)

\[
Y = A_1X_1 + A_2X_2 + A_3X_3 + ... + A_nX_n
\]

where \( Y \) is the chemical constituent, \( A \) is an empirical coefficient and \( X_{i,n} \) are wavelengths.

Due to the limited number of samples (n = 16), leave-one-out cross-validation analyses (e.g. only one sample at a time is kept out of the calibration and used for prediction) were run (Mark 2000, Esbensen 2002). The difference between the predicted and measured carbohydrate content...
contents was expressed as a root mean square error of cross-validation (RMSECV). RMSECV is defined as the square root of the average of the squared differences between the predicted and measured values of the validation objects (Esbensen 2002):

\[
RMSECV = \left( \frac{\sum (X_m - X_p)^2}{n_v} \right)^{1/2}
\]

(4)

where \(X_m\) is the chemically measured value of a sample, \(X_p\) is the predicted value of the sample on the basis of spectral analysis and \(n_v\) is the number of samples in the calibration stage. The accuracy of each calibration model was evaluated based on the coefficient of determination (\(R^2\)) for predicted vs. measured carbohydrate content in cross-validation and by slope. In addition, values of the ratio of prediction to deviation (RPD) were calculated to verify the applicability of the calibration (Williams and Norris 2001). RPD is defined as the ratio of the standard deviation of the \(Y\) values (e.g. sugar content) to the RMSECV (Mouazen et al. 2005). An RPD value was defined using the following ranges: 0–2 as not recommended and between 2.0 and 2.5 translated to mean that quantitative prediction is possible. RPD values above 3.0 were considered to indicate that the predictive capability of the model is excellent. Note, however, that the classification of RPD values of prediction accuracy slightly varies among different studies (Viscarra Rossel 2007).

To identify significant \(X\) variables, we ran Martens’ uncertainty test (Unscrambler, Version 9.1, Camo, Norway, 2004; Esbensen 2002). Therefore, we first ran our models on the whole wavelength region, and thereafter only significant wavelengths were kept and each model was reassessed using Martens’ uncertainty test (Esbensen 2002). To simplify the spectral signals, thereby ensuring stable calibration, we applied first and second different orders of derivatives (Heise and Winzen 2002, Duckworth 2004), with the Savitzky and Golay method (Savitzky and Golay 1964). Finally, models that gave the lowest RMSECV, highest \(R^2\), slope and RPD value were considered to be optimal for future concentration estimates.

Results and discussion

External nutrient supply increases biomass growth rate, but decreases monosaccharide content in \textit{Ulva fasciata} thalli

External nutrient supply increased the thallus growth rate (Figure 1A). The growth rate of fertilised thalli was 0.730±0.095 at week 1, 0.676±0.278 at week 2 and 5.186±0.870 g FW week
\(^{-1}\) at week 3. The growth of thalli without added nutrients was 0.118±0.027 at week 1, 0.063±0.031 at week 2 and 0.240±0.068 g FW week
\(^{-1}\) during week 3.

We found that adding nutrients decreased (\(p<0.05\)) glucose and rhamnose concentrations in \textit{Ulva fasciata} tissues but had no effect on the concentrations of xylose and glucuronic acid (Figure 1B). The contents of rhamnose, glucose, xylose and glucuronic acid in the thalli without external nutrients were 11.93±0.91 mg g
\(^{-1}\) dry weight (DW), 18.85±1.78 mg g
\(^{-1}\) DW, 1.42±0.20 mg g
\(^{-1}\) DW and 7.05±0.40 mg g
\(^{-1}\) DW, respectively. The contents of rhamnose, glucose, xylose and glucuronic acid in the thalli with external nutrient supply were 8.93±0.73 mg g
\(^{-1}\) DW, 13.94±1.11 mg g
\(^{-1}\) DW, 1.08±0.15 mg g
\(^{-1}\) DW and 6.39±0.39 mg g
\(^{-1}\) DW, respectively. These results are consistent with previous findings that the C:N ratio in species of \textit{Ulva} reduces under nitrogen-rich conditions.
and increases during nitrogen starvation (Pinchetti et al. 1998).

Changes in spectral features between clones of Ulva thalli grown with and without external nutrient supply

Generally, vegetation reflectance spectra are quite informative regarding vegetation chlorophyll absorption bands in the VIS (400–700 nm) and NIR (700–1100 nm) and the effects of plant water absorption in the SWIR (1100–2500 nm). Figure 2 shows spectral reflectance variability for thalli grown with and without external nutrient supply (Figure 2; blue vs. red spectrum lines, respectively) for both wet (A) and dry (B) Ulva fasciata. Samples differ not only in absorbance depths across a spectrum (due to the presence of chemical chromophors) but also in spectral intensity, curve shape and spectral slopes across the measured spectral range.

The change in absorption depth around 1400 and 1900 nm between dry and wet U. fasciata is particularly deep for wet samples. Water has a broad absorption range centred around 1400 and 1940 nm that masks other absorption features associated with constituents such as nitrogen, lignin, sugar and cellulose.

For both wet and dry U. fasciata biomass, as shown in Figure 2C and D, several indices and slopes across similar spectral ranges were generated with an aim to distinguish between thalli with and without external nutrient supply. NDIs that separate the samples best lie between the following pairs of wavelengths: 550–686 nm, 425–850 nm and 700–900 nm. In addition, spectral slopes calculated between 470–530 nm and 686–770 nm are efficient for distinguishing between the two groups. Low carbohydrate content (no external nutrients) and high carbohydrate content groups (external nutrient supply) clearly stand out (p < 0.001). Importantly, these spectral ranges are not impacted directly by strong water absorbance.

Exploratory assessment of the potential of spectral-based chemometrics of monosaccharide content in Ulva fasciata

Table 2 presents the best results obtained for the cross-calibration set based on 16 samples. Several models predicted glucose content efficiently. We found that for

![Figure 2: (A–B) Spectral reflectance of wet and dry samples of Ulva fasciata. (C–D) Normalised difference spectral reflectance index (NDI) and spectral slope between two wavelengths; note that spectral index differentiates between samples with and without nutrient supply.](image-url)
prediction of different sugar components based on spectral measurements, the resulting quality is strongly dependent on the wavelength selection. This supervised model provides better results than when using the entire spectrum. During the selected cross-validation modelling, significant improvement in predictive accuracy, lower value of RMSECV, higher RPD and R² values was achieved when all pre-processed models were reassessed on a reduced number of spectral variables (according to Martens test; Esbensen 2002). Due to the limited number of samples (e.g. limited dynamic range of reference values), Y values (concentrations) for several models were log-transformed in order to get a normal distribution.

Figure 3A shows measured vs. predicted glucose content obtained for the best-fit model for dry biomass. The linear plot exhibits a slope of 0.87 and R²=0.88. A much clearer picture of spectroscopically determined glucose content changes compared with original reflectance is obtained from the plot of factor loadings [loading vectors (LV), or regression coefficients] vs. wavelength (Figure 3B). The first three LV components in the PLS model explained 95% of the X variance (spectra) and 90% of the Y variance (glucose content). This indicated that most of the spectral variation that was modelled by PLS is related to the glucose content.

Figure 3C shows the X-(PC1) and Y-(PC2) scores of a PLS regression plot of samples from the PLS model for glucose content for dry samples. The score plot indicated that a significant part of the spectral variations observed in the dry samples are indeed related to glucose as predicted by the PLS model. Specifically, there was an increasing trend in glucose content among samples from the bottom to the top along PC2, following a slight diagonal, presumably due to the difference in water content. Furthermore, there is a separation between the two groups of samples along PC1, samples with no external nutrient supply are located on the left of the score plot, whereas samples with external nutrient supplies are located on the right of the plot.

Figure 4 shows the best-fit model results for wet biomass. Similar to dry biomass, the best pre-processing technique was to consider reflectance as the second derivative, followed by a smoothing function and run on selected wavelengths. Here we used the Savitzky-Golay smoothing method to reduce high-frequency noise. This is an averaging algorithm that fits a polynomial to the data points. The relatively high predictive ability at the
cross-validation set level enables us to propose that our PLS models can be used to predict glucose content precisely. The highest regression coefficients were for the spectral range 1100–1300 nm (Figure 4B), which probably relates to the presence of sugar, starch and cellulose (Curran 1989).

Furthermore, as is evident from the analyses of the score plot of the wet biomass data set (Figure 4C), there is a separation between the two groups of samples along PC1, samples with no external nutrients are located on the left of the score plot, whereas samples with external nutrients are on the right. Similar to the dry model (Figure 3C), there is an increasing trend in glucose content among samples from the bottom to the top along PC2.

The major limitation of the current study is the relatively small number of samples used, which limits the PLS model. This limitation, however, does not change our main findings that spectral reflectance depends on glucose content. Importantly, this dependence is apparent for both dry and wet Ulva fasciata biomass. In addition, the same spectra and similar modelling approaches could be used in further studies for quantification of other important biorefinery constituents of Ulva biomass, such as starch, proteins, pigments and special chemicals. For example, Lugassi et al. (2015b) concluded that crude protein and neutral detergent fibre contents could be predicted using the spectral slope method for dry and wet vegetation.

Conclusions

This study shows that the change in spectral slope can be used in rapid phenotyping of Ulva fasciata as an indicator for high glucose content. Glucose is the most promising
source for fermentation and production of sustainable chemicals and biofuels. We have shown that it is possible to discriminate between high and low contents of glucose, rhamnose, xylose and glucuronic acid for both dry and wet samples of *Ulva fasciata* using similar spectral ranges across visible and near-infrared wavelengths. This result is important for the development of a future field device intended for rapid macroalgal profiling for breeding and strain development for marine biorefineries. As an example, we estimated the content of major monosaccharides of *U. fasciata* using PLS analyses based on spectral reflectance measurements. We found that the wavelength range of NIR-SWIR spectra for calibration and/or prediction affect the reliability of NIRS analysis. This study should be expanded to a larger calibration set of values.

**Acknowledgments:** The authors acknowledge the TAU Center for Innovation in Transportation and The Israel Ministry of National Infrastructures, Energy and Water Resources for support. The authors acknowledge and appreciate Prof. Matthew J. Dring’s scientific comments.

**References**


Bionotes

Shai Shefer
Porter School of Environmental Studies, Tel Aviv University, Israel

Shai Shefer is currently an MSc student at Tel Aviv University, Porter School of Environmental Studies, focusing on macroalgae profiling for breeding and strain development for marine biorefinery methods. He completed his MBA at the College of Management Academic Studies and BSc at the Hebrew University of Jerusalem.
Alvaro Israel is currently a senior scientist at the Israel Oceanographic and Limnological Research Institute at Haifa, Israel, and has been engaged for over 20 years in the study of seaweed biology and ecology. He has been involved in projects related to carbon fixation, photosynthesis, taxonomy and applied phycology.

Alexander Golberg is currently a faculty member at Tel Aviv University, Porter School of Environmental Studies. He completed his PhD in bioengineering at the Hebrew University of Jerusalem, focusing on low-cost bioenergy technologies for low-income countries. He continued working on a variety of projects that involved engineering and biology at UC Berkeley and Harvard Medical School. In 2012, he won the Green Talent 2012 Award for the design of distributed marine biorefineries for the production of feed and transportation fuels for rural areas in low-income counties. His main research interests are in bioenergy conversion efficiency and bioprocess development.

Alexandra Chudnovsky is currently a faculty member at Tel Aviv University, Department of Geography and Human Environment, School of Geosciences and a visiting Scientist at the Harvard T.H. Chan School of Public Health. Much of her research is concerned with developing modeling methods to predict and evaluate the contribution of local and regional pollution sources to urban air, water, vegetation and soil quality using passive remote-sensing technologies.