Methods in Ecology and Evolution

British Ecological Society

Methods in Ecology and Evolution 2016 doi: 10.1111/2041-210X.12596

RESEARCH ARTICLE

Spectroscopic determination of ecologically relevant plant secondary metabolites

John J. Couture[1]*, Aditya Singh[1], Kennedy F. Rubert-Nason[2], Shawn P. Serbin[1]†, Richard L. Lindroth[2] and Philip A. Townsend[1]

[1]*Department of Forest and Wildlife Ecology, University of Wisconsin-Madison, Madison, WI 53706, USA; and* [2]*Department of Entomology, University of Wisconsin-Madison, Madison, WI 53706, USA*

Summary

1. Spectroscopy has recently emerged as an effective method to accurately characterize leaf biochemistry in living tissue through the application of chemometric approaches to foliar optical data, but this approach has not been widely used for plant secondary metabolites. Here, we examine the ability of reflectance spectroscopy to quantify specific phenolic compounds in trembling aspen (*Populus tremuloides*) and paper birch (*Betula papyrifera*) that play influential roles in ecosystem functioning related to trophic-level interactions and nutrient cycling.

2. Spectral measurements on live aspen and birch leaves were collected, after which concentrations of condensed tannins (aspen and birch) and salicinoids (aspen only) were determined using standard analytical approaches in the laboratory. Predictive models were then constructed using jackknifed, partial least squares regression (PLSR). Model performance was evaluated using coefficient of determination (R^2), root-mean-square error (RMSE) and the per cent RMSE of the data range (%RMSE).

3. Condensed tannins of aspen and birch were well predicted from both combined (R^2 = 0·86, RMSE = 2·4, %RMSE = 7%)- and individual-species models (aspen: R^2 = 0·86, RMSE = 2·4, %RMSE = 6%; birch: R^2 = 0·81, RMSE = 1·9, %RMSE = 10%). Aspen total salicinoids were better predicted than individual salicinoids (total: R^2 = 0·76, RMSE = 2·4, %RMSE = 8%; salicortin: R^2 = 0·57, RMSE = 1·9, %RMSE = 11%; tremulacin: R^2 = 0·72, RMSE = 1·1, %RMSE = 11%), and spectra collected from dry leaves produced better models for both aspen tannins (R^2 = 0·92, RMSE = 1·7, %RMSE = 5%) and salicinoids (R^2 = 0·84, RMSE = 1·4, %RMSE = 5%) compared with spectra from fresh leaves. The decline in prediction performance from total to individual salicinoids and from dry to fresh measurements was marginal, however, given the increase in detailed salicinoid information acquired and the time saved by avoiding drying and grinding leaf samples.

4. Reflectance spectroscopy can successfully characterize specific secondary metabolites in living plant tissue and provide detailed information on individual compounds within a constituent group. The ability to simultaneously measure multiple plant traits is a powerful attribute of reflectance spectroscopy because of its potential for *in situ–in vivo* field deployment using portable spectrometers. The suite of traits currently estimable, however, needs to expand to include specific secondary metabolites that play influential roles in ecosystem functioning if we are to advance the integration of chemical, landscape and ecosystem ecology.

Key-words: condensed tannins, paper birch, phenolics, phytochemistry, plant defence, salicinoids, salicortin, spectroscopy, trembling aspen, tremulacin

Introduction

Field-based spectroscopy on live foliage has emerged as an effective method to estimate specific chemical constituents of plants due to improvements in the sensitivity and portability of spectrometers, as well as advances in computation power and chemometric modelling methods. To date, a number of foliar biochemical, physiological, structural and morphological properties have been successfully quantified using reflectance spectroscopy (Gillon, Houssard & Joffre 1999; Petisco *et al.* 2006; Asner & Martin 2008, 2011; Kleinebecker *et al.* 2009; Serbin *et al.* 2012, 2014). However, the application of field-based spectroscopy to estimate detailed concentrations of plant secondary metabolites has lagged, especially in living tissue.

Plant secondary metabolites play critical roles in ecosystem functioning and contribute greatly to phytochemical diversity. Rapid, non-destructive determination of these compounds

*Correspondence author. E-mail: jjcouture@wisc.edu
†Present address: Biological, Environmental & Climate Sciences Department, Brookhaven National Laboratory, Upton, NY 11973

© 2016 The Authors. Methods in Ecology and Evolution © 2016 British Ecological Society

in vivo and *in situ* using spectroscopy reduces the need to collect large amounts of material in the field, decreases processing time, lessens costly chemical analyses and eliminates sampling that could itself alter experimental conditions (Couture, Serbin & Townsend 2013). These benefits can thus facilitate the design of larger, more complex experiments with greater sample sizes. In addition, this approach can help provide rapid assessments of plant function over large geographic regions if scaled to remote sensing collections from air- or spaceborne platforms.

Trembling aspen (*Populus tremuloides*) and paper birch (*Betula papyrifera*) are broadly distributed tree species in the Northern Hemisphere, and their foliar chemical composition plays an influential role in forest functioning. The secondary metabolites mediating ecological interactions in aspen and birch are predominately phenolic products from the shikimic acid pathway (Constabel & Lindroth 2010). Two groups of phenolic compounds that have been extensively studied for their roles in influencing ecological interactions are condensed tannins and salicinoids. Condensed tannins are polymeric compounds composed of flavan-3-ol subunits found almost ubiquitously in the plant kingdom and play prominent roles in a number of ecosystem processes. For example, tannins prevent stress through protection from photodamage (Close & McArthur 2002) and alter nutrient cycling by reducing litter decomposition rates, forming protein complexes, microbial priming, or directly inhibiting microbial functioning (Hättenschwiler & Vitousek 2000; Kraus, Dahlgren & Zasoski 2003; Kraus *et al*. 2004; Madritch, Jordan & Lindroth 2007; Schweitzer *et al*. 2008; Madritch & Lindroth 2015). Foliar concentrations of condensed tannins vary considerably among tree species and genotypes, and in both aspen and paper birch are influenced by environmental variation (Lindroth *et al*. 2001; Osier & Lindroth 2001; Donaldson & Lindroth 2007). Tannins also comprise a significant carbon fraction in plants, behind only structural compounds (Hernes & Hedges 2000), and comprise upwards of 35% of foliar dry mass in aspen and paper birch. In aspen, simple phenolic compounds, structurally similar to salicylates (i.e. salicinoids), confer resistance to herbivory and are toxic to most lepidopteran herbivore species (Boeckler, Gershenzon & Unsicker 2011; Lindroth & St Clair 2013). Salicinoids also constitute a significant proportion of foliar dry mass (1–35%) with two generally correlated compounds (salicortin and tremulacin) generally representing >90% of total foliar salicinoid concentrations, although other salicinoids are present in smaller concentrations (Lindroth, Hsia & Scriber 1987; R.L. Lindroth, unpublished data).

Reflectance spectroscopy has been used to estimate concentrations of the major groups of secondary compounds, including alkaloids, glucosinolates, terpenoids and phenylpropanoids and related phenolic compounds (Schulz *et al*. 1999; Ebbers *et al*. 2002; Font *et al*. 2005; Carvalho *et al*. 2013; Couture, Serbin & Townsend 2013; Rubert-Nason *et al*. 2013). Of the major groups of secondary metabolites, phenolic compounds are a group of secondary metabolites universally distributed in the plant kingdom and play influential roles in plant physiology and ecosystem functioning. To date, however, few studies have explored the capacity of spectroscopy to estimate concentrations of specific phenolic compounds, instead focusing on bulk concentrations, a gap we address in this research.

The estimation of biochemical concentrations from reflectance spectroscopy relies on variations in absorption as a consequence of vibrational excitation of molecular bonds, primarily C–H, N–H and O–H bonds at specific wavelengths in the visible (400–700 nm), near-infrared (NIR, 700–1100 nm) and short-wave infrared (SWIR, 1100–2400 nm). In practice, spectral measurements of organic material collected in a consistent manner using a uniform and stable illumination source provide the foundation for the estimation of the chemical composition of a sample (Shenk, Workman & Westerhaus 1992). Calibration is accomplished by pairing reflectance spectra (e.g. across the full range 400–2500 nm) with independent chemical measurements, and then modelling chemical concentration as a function of the spectra using multivariate (chemometric) methods such as partial least squares regression (PLSR; Wold, Sjostrom & Eriksson 2001). Models are validated using independent samples and then applied to unknown samples using their spectral reflectance.

Recent research has shown that condensed tannins in aspen and birch and salicinoids in aspen are predictable on dry foliar samples using laboratory-based reflectance spectroscopy (Rubert-Nason *et al*. 2013). Here, we test the capacity of reflectance spectroscopy to characterize specific metabolites in both fresh and dry plant tissue, focusing on phenolic compounds in foliage of trembling aspen and paper birch. Importantly, models made using fresh leaves avoid the extra steps of processing collected tissue and provide the foundation for using field-based spectroscopy measurements to make real-time estimates of secondary metabolite concentrations, non-destructively capturing shifts in plant chemistry that influence ecological processes. Specifically, we (i) utilize reflectance spectroscopy to quantify specific secondary metabolites in green and dried leaf material, (ii) examine the ability of reflectance spectroscopy to quantify individual compounds (i.e. salicortin and tremulacin) within a chemical class (i.e. salicinoids), and (iii) compare models built using dry- and fresh-leaf material to determine whether predictability is sacrificed when moving from dry to fresh leaves.

Methods

FOLIAR COLLECTIONS

Accurate and replicable chemometric models require capturing a dynamic range of chemical variation, and foliar condensed tannin and salicinoid concentrations exhibit considerable variation among genotypes in trembling aspen (Lindroth & St Clair 2013). As such, we collected leaf samples from genotypes having a wide range of biochemical variation. In July 2013, we collected leaves from 139 trees of 116 genotypes of aspen, originally planted (2010) in a common garden (WisAsp) located at the Arlington Agricultural Research Station, near Arlington, WI, USA. From each tree, 5–6 leaves were collected from the middle and top levels of the canopy by removing the leaf from the petiole using scissors, after which spectra were immediately (<2 min) collected using a full-range spectrometer and leaves were stored on ice in the field. The leaves were flash-frozen with liquid N within 2 h of sampling,

lyophilized and then ground with a Wiley minimill equipped with a #60 mesh screen (Swedesboro, NJ, USA).

Birch leaves used in these analyses were collected from saplings grown from seed collected from a single tree in Madison, WI. Seeds were planted in 60 × 30 cm plastic flats in MetroMix potting medium #300 in a glasshouse with a constant temperature of 25°C. Thirty days after germination, 76 individual seedlings were transferred into 3-L pots containing the same growing medium. Concentrations of condensed tannins in trees are generally responsive to environmental variation, including nitrogen fertilization (Kraus, Dahlgren & Zasoski 2003); to maximize variation in condensed tannin levels in birch, one-half of the plants received fertilizer amendments equalling 50 kg ha^{-2} y^{-1}, simulating a high level of anthropogenic nitrogen deposition (Galloway *et al.* 2004). After scaling this amount of fertilizer to the pot size used, we administered weekly additions of 22 and 0 mg of NH_4NO_3 to fertilized and unfertilized plants, respectively, for 4 weeks, 6 months after transfer to individual pots. After the fourth week of fertilization treatments concluded, 3 leaves were collected from multiple levels of the saplings and spectral measurements were immediately (<2 min) made. Leaves were stored and processed following the same procedures as described above for aspen.

SPECTRAL COLLECTIONS

To examine absorption features of specific phenolic compounds, generate predictions of these compounds using foliar optical properties and examine the influence of water on prediction performance, we collected reflectance of purified aspen and birch condensed tannin standards, aspen salicortin and tremulacin standards, and fresh and dry leaves using a high-spectral-resolution FieldSpec 3 Full-Range (350–2500 nm) spectroradiometer (Analytical Spectral Devices, Boulder, CO, USA). Reflectance measurements of standards were collected on 250–300 mg of purified condensed tannins from aspen or birch or the aspen salicinoids salicortin and tremulacin using a plant probe attached to a Dremel® drill press to assure constant measurement geometry and pressure (Serbin *et al.* 2014). Standard material was loosely packed and levelled inside a custom-machined aluminium, flat matte-black-painted, sample cup which was fit underneath the vertically mounted plant probe. Each sample was compressed uniformly to 2 N m^{-1} during spectral collection through the use of a precision torque wrench (Effetto Mariposa, Giustaforza™ Professional) custom-machined to act as the handle for the drill press assembly. Nine spectral measurements were collected for each sample in three sets of three measurements, with the material being mixed and repacked in the sampling cup and the sample cup rotated by 90° between each step. These steps were conducted to minimize potential bias arising from sample cup orientation or from sensor characteristics. All nine spectra were averaged to determine the mean dry material reflectance and converted to first-derivative reflectance to highlight spectral absorption features.

Fresh-leaf measurements were taken from the leaf adaxial surface using a leaf-clip assembly attached to the ASD plant probe with an internal halogen light source. Five reflectance measurements were collected from three different areas of each green leaf and averaged to determine mean leaf reflectance. Dry-leaf measurements were taken from lyophilized, ground leaves following the same protocol used for measurements of chemical standards, except that each sample consisted of approximately 500 mg of tissue and reflectance measurements were not first-derivative-transformed. Dry-leaf spectra collections and analyses were restricted to aspen due to insufficient mass of leaf material in birch.

CHEMICAL ANALYSES

Condensed tannins of aspen and birch were extracted into acetone:water (7:3) with ascorbic acid (1·8 g L^{-1}). Tannins were quantified colorimetrically following formation of a coloured adduct by reaction with ferric ammonium sulphate under acidic conditions (Porter, Hrstich & Chan 1986) using a SpectraMax 340pc spectrophotometer (Molecular Devices, Sunnyvale, CA, USA). Condensed tannins, purified from aspen and birch following Hagerman & Butler (1980), served as standards.

Salicinoids were determined using ultrahigh-performance liquid chromatography and negative electrospray ionization single-quadrupole mass spectrometry (UHPLC-MS) on a Waters integrated Acquity I-Class UPLC® system (Milford, MA, USA). Pulverized plant tissue (20–30 mg) was extracted into methanol (containing 10 mg mL^{-1} ß-resorcylic acid as a control standard) with sonication for 15 min at <10°C, followed by centrifugation at 3450 *g* to separate the particulates. Extracts were diluted 1:1 with methanol, spiked with an internal standard (d_6-salicylic acid) to a concentration of 100 mg L^{-1}, filtered on a 0·45-μm polytetrafluoroethylene membrane, transferred to a 0·5-mL polypropylene autosampler vial, capped and maintained at −20°C or 4°C prior to analysis. Quantification of salicinoids was adapted from Abreu *et al.* (2011). Prepared extracts (2 μL) were separated on a Waters Acquity® CSH C-18 (2·1 × 100 mm, 1·7 μm) column over 26 min using a gradient of water and acetonitrile (with 0·1% formic acid) at 40°C and a flow rate of 0·5 mL min^{-1}. The mass spectrometer was operated in negative ionization mode, with selective ion recording of the salicinoid-formate adducts. Operating and data acquisition conditions for the mass spectrometer were as follows: cone potential, 35 V; capillary potential, 2500 V; extractor potential, 3 V; RF lens potential, 0·1 V; source temperature, 120°C; desolvation temperature, 250°C; desolvation gas flow, 500 L h^{-1}; cone gas flow, 10 L h^{-1}; infusion rate, 5 μL min^{-1}; dwell time, 0·025 s. Analytes were quantified as negative ion formate adducts. Calibrations were based on internal standardization, using 6-point (15–1500 mg L^{-1}) quadratic models with salicortin and tremulacin, purified following Lindroth, Hsia & Scriber (1987), serving as standards. Quality assurance was demonstrated in three ways: monitoring the peak area of the internal standard, monitoring the peak area of the process control standard and monitoring the salicinoid concentrations in reference samples (consisting of extracted *P. tremuloides* foliage) that were analysed at intervals of 15–20 samples.

MODEL CALIBRATION AND VALIDATION

We generated models to predict condensed tannin and salicinoid concentrations (% dry mass) from fresh and dry, untransformed leaf spectral reflectance characteristics using partial least squares regression (PLSR; Wold *et al.* 1984; Wold, Sjostrom & Eriksson 2001). When predictor variables are highly correlated, as is the case with spectroscopic data, classical regression techniques can produce unreliable coefficients as collinear predictor variables lead to bias in beta coefficients and error estimates (Grossman *et al.* 1996). In contrast to standard regression techniques, PLSR reduces a large number of collinear predictor variables into relatively few, uncorrelated latent variables, and has become the preferred method for chemometric analyses (Bolster, Martin & Aber 1996; Asner & Martin 2008, 2011; Atzberger *et al.* 2010; Serbin *et al.* 2014). The number of latent variables used was based on reduction of the predicted residual sum of squares (PRESS) statistic (Chen *et al.* 2004) using leave-one-out cross-validation. The final set of latent variables was then combined into a linear model

© 2016 The Authors. Methods in Ecology and Evolution © 2016 British Ecological Society, *Methods in Ecology and Evolution*

predicting chemical concentrations. Examination of prediction residuals was used to identify three outliers that were subsequently removed from the analyses.

Model performance was evaluated by conducting 500 randomized permutations of the data set using 70% of the data for internal calibration and the remaining 30% for external validation. These randomized analyses generated a distribution of fit statistics allowing for the assessment of the overall stability of the models as well as uncertainty in model predictions. For each permutation, we tracked the model goodness-of-fit (R^2) and root-mean-square error (RMSE) to assess model performance when applied to the validation data set. Here, we report both the RMSE and RMSE as a percentage of the range of data (% RMSE) for each dependent variable, with the latter metric being useful to assess likely error within the data range. %RMSE as a function of the data range is used because interpretation of RMSE related to the mean is highly sensitive to the actual value of the mean. For example, for two constituents with very different means but similar ranges, the RMSE as a percentage of the mean will overinflate the accuracy on one measure and underinflate the accuracy on the lower value, even though the accuracy is essentially the same. RMSE as a per cent of the range conveys the average resolving power of the PLS model within the estimation space. %RMSE is functionally similar to the classical residual prediction deviation (RPD) statistic for assessing model error within the context of the reference data, but avoids the distribution assumptions and subjective model classifications associated with RPD. We further determined the strength of the contribution of PLSR loadings by individual wavelengths using the variable important to the projection (VIP) statistic (Wold *et al.* 1984; Wold, Sjostrom & Eriksson 2001). The VIP statistic indicates the importance of individual wavelengths in explaining the variation in the response and predictor variables, with larger weightings confer greater value of contribution of individual wavelengths to the predictive model (Wold, Sjostrom & Eriksson 2001; Chong & Jun 2005).

Wavelengths between 1100 and 2400 were used in model building for both fresh and dry leaves. We focused on the SWIR region because (i) it contains wavelengths known to be characteristic of plant phenolics (Shenk, Workman & Westerhaus 1992; Flinn *et al.* 1996; Kokaly & Skidmore 2015), (ii) our previous work showed strong relationships at these wavelengths for predictive models of phenolic compounds using dried foliage of the same tree species (Rubert-Nason *et al.* 2013), and (iii) it avoids leveraging correlations between phenolics and pigments in the visible wavelengths.

Pearson's correlation coefficients were used to identify relationships among foliar nitrogen, lignin and water content and the secondary metabolites predicted in the current study. Nitrogen and lignin were predicted using spectroscopy using coefficients from Serbin (2012) and water content (NDWI) was determined following Gao (1996) and calculated as the relative difference between reflectance at wavelengths 857 and 1241 nm. The modelling approach and data analyses were performed using the *pls* package (Mevik & Wehrens 2007) in R (www.r-project.org) and correlations were conducted in JMP Pro v11 (SAS Institute Inc., 2013, Cary, NC, USA).

Results

VARIATION IN THE REFLECTANCE OF FOLIAGE AND OF PURIFIED STANDARDS

Major reflectance peaks in spectra obtained from purified aspen and birch condensed tannin standards had significant overlap with wavelengths known to interact with phenolic compounds (Figs 1a, b, S1, Supporting information). Major reflectance peaks in the spectra of purified salicinoid standards (salicortin and tremulacin) also overlapped with wavelengths known to interact with phenolics (Figs 1c, d, S1, Supporting information). Substantial variation existed among spectra collected from different aspen genotypes (Fig. S2a, Supporting information) and from birch leaves from different soil nutrient properties (Fig. S2b, Supporting information).

PREDICTION OF CONDENSED TANNINS

Condensed tannin concentrations varied over 12- and 9-fold for aspen and birch, respectively, and were accurately predicted using PLSR models from fresh leaves for both combined species and individual species (Figs 2 and 3). Fresh-leaf PLSR models including both species combined produced a mean R^2 of 0·86 (range: 0·63–0·96), RMSE of 2·46 (range: 1·64–3·45) and %RMSE of 7% for external validation. Fresh-leaf models of only aspen leaves produced mean R^2 of 0·86 (range 0·72–0·95), RMSE of 2·43 (range 1·58–3·27) and % RMSE of 6% for external validation. The prediction performance metrics of aspen fresh-leaf condensed tannin models were slightly lower than the models built from dried leaf material (Fig. 3). Aspen dry-leaf models had a mean R^2 of 0·92 (range 0·81–0·96), RMSE of 1·76 (range 1·08–2·33) and % RMSE of 5% for external validation. Birch fresh-leaf models produced a mean R^2 of 0·81 (range 0·43–0·95), RMSE of 1·92 (range 1·22–2·89) and %RMSE of 10% for external validation. All calibration models showed zero-centred Gaussian bias profiles with higher predicted variances in externally validated models (Fig. S3, Supporting information). Standardized PLSR coefficients and VIP measures for combined species, aspen (fresh and dry leaf) and birch (fresh only) models exhibited similar profiles and matched with wavelength regions of known phenolic absorption features (Fig. 4a–f).

PREDICTION OF SALICINOIDS

Salicinoid levels varied over 14-fold among aspen genotypes and both total and individual salicinoids were accurately predicted using PLSR models with fresh (Figs 5 and 6) and dry leaves (Figs 6 and S4, Supporting information). Models predicting total salicinoids, salicortin and tremulacin from fresh aspen leaves had mean R^2 of 0·76 (range 0·45–0·92), 0·57 (range 0·20–0·79) and 0·72 (range 0·49–0·92), mean RMSE of 2·44 (range 1·71–3·20), 1·96 (range 1·35–2·51) and 1·08 (range 0·81–1·44), and %RMSE of 8, 11 and 11%, respectively, for external validation (Fig. 6). Similar to models for aspen tannins, fresh-leaf models predicting salicinoids did not perform as well as models built from dried leaf material (Fig. 6). Aspen dry-leaf models for total salicinoids, salicortin and tremulacin had mean R^2 of 0·84 (range 0·65–0·95), 0·78 (range 0·45–0·91) and 0·78 (0·49–0·92), mean RMSE of 1·97 (range 1·14–2·80), 1·43 (range 0·93–1·95) and 0·73 (range 0·54–1·29), and % RMSE of 5, 8 and 7%, respectively, for external validation. Bias for all models followed zero-centred Gaussian profiles and, similar to tannin models, became more variable when

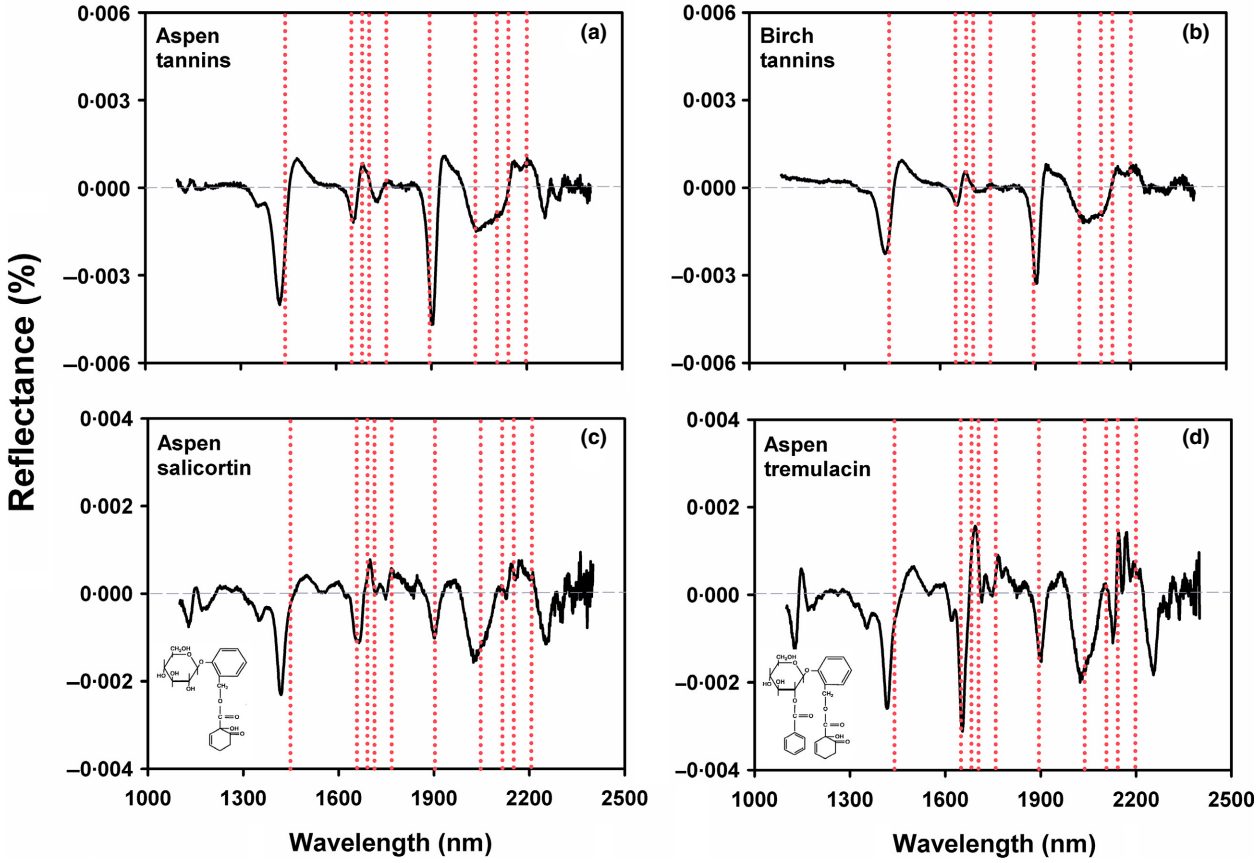

Fig. 1. First-derivative reflectance spectra of purified aspen (a) and birch (b) condensed tannins and the purified aspen salicinoids salicortin (c) and tremulacin (d). Red vertical dotted lines represent wavelengths reported in the literature having absorption features associated with phenolic compounds.

moving from internally calibrated to externally validated models (Fig. S3, Supporting information). In addition, examination of residuals revealed that models tended to over predict at low and under predict at high salicinoids values (Fig. 5). Similar to fresh- and dry-leaf tannin models, aspen standardized coefficients and VIP values were most pronounced near wavelengths with absorption features associated with phenolic compounds (Fig. 4g, h).

CORRELATIONS AMONG PLANT TRAITS

Tannins were strongly negatively correlated with foliar lignin in birch and exhibited statistically significant, yet weak, relationships with nitrogen for both aspen and birch, and with lignin in aspen (Table 1). Tannins were not related to foliar water content for either tree species (Table 1). Aspen salicinoids were weakly positively related to foliar lignin and water, negatively related to aspen tannins, but not correlated with nitrogen levels (Table 1).

Discussion

Ecologically relevant secondary metabolites in living plant tissue can be accurately characterized using reflectance spectroscopy. Furthermore, we present a robust methodology by which plant traits can be predicted using reflectance spectroscopy and that utilizes multiple permutations of the data, providing explicit estimates of model uncertainty (i.e. error bars in Figs 2 and 5). By combining high-fidelity reflectance measurements, standard chemical analyses and robust statistical modelling, we demonstrate the potential to expand the prediction capabilities of spectroscopic data for secondary metabolites.

Important spectral features for purified standards of condensed tannins and salicinoids overlapped with regions of wavelengths with known absorption features associated with phenolic compounds, including areas close to or in the range of 1400, 1650–1700, 1900, 2100–2200, 2226–2228 and 2400 nm (Curran *et al.* 1992; Shenk, Workman & Westerhaus 1992; Bian *et al.* 2013; Rubert-Nason *et al.* 2013; Kokaly & Skidmore 2015). Moreover, coefficient weightings and VIP statistics from models predicting tannins and salicinoids were most pronounced in these literature-reported regions, corroborating the conclusion that the bending and stretching of C–H and O–H bonds are important for spectroscopic prediction of plant phenolics (Shenk, Workman & Westerhaus 1992; Kokaly & Skidmore 2015).

Covariation among plant traits may aid or detract from the ability of spectroscopy to predict specific traits (Curran *et al.* 1992; Soukupova, Rock & Albrechtova 2002). For example,

© 2016 The Authors. Methods in Ecology and Evolution © 2016 British Ecological Society,

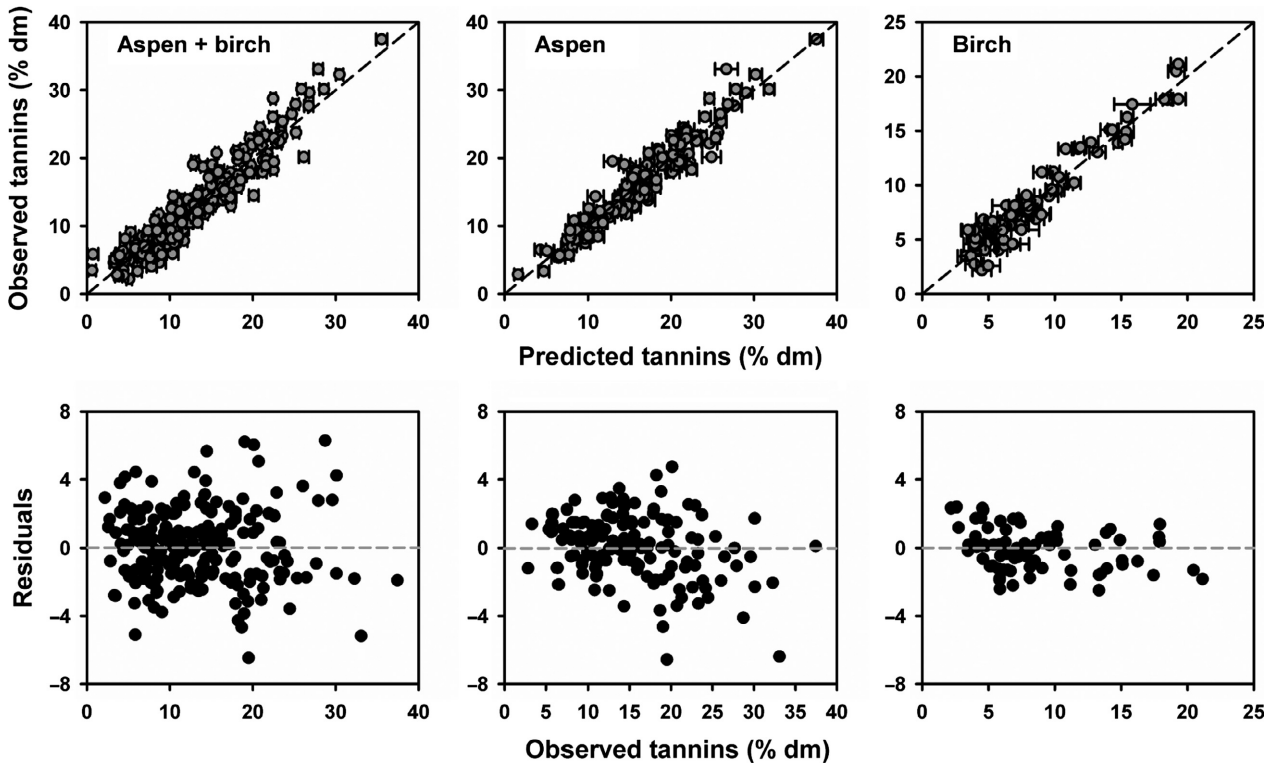

Fig. 2. Observed vs. predicted condensed tannin concentrations in fresh leaves (top panel, % dm = % dry mass). Error bars for predicted values represent the standard deviations generated from the 500 simulated models. 1:1 line in dashed black. Residuals plotted against observed values (bottom panel).

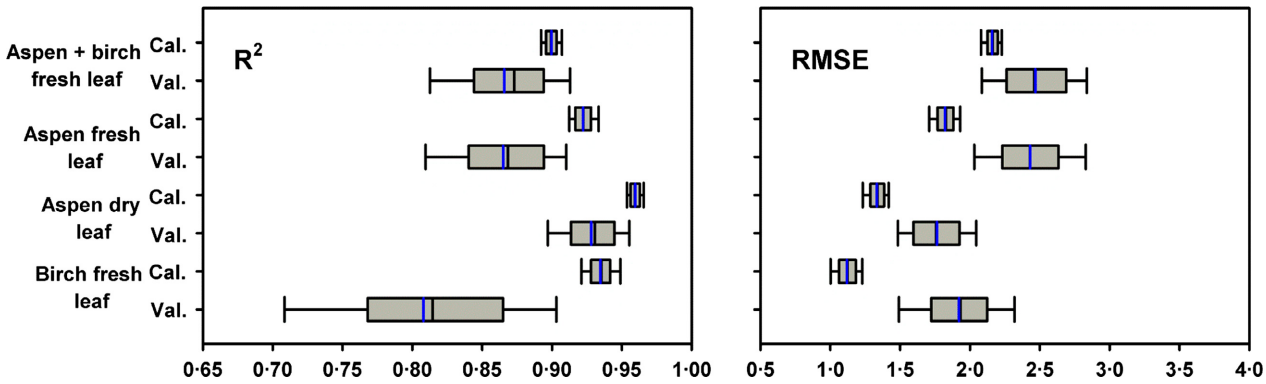

Fig. 3. Distribution (95% confidence intervals) of internal calibration (Cal.) and external validation (Val.) statistics from 500 simulations for models predicting condensed tannins in aspen fresh and dry leaves and birch fresh leaves. Each model permutation included 70% of the data for internal calibration and the remaining 30% for external validation. Black vertical line represents median value; blue vertical line represents mean values.

lignin retrievals using reflectance data are positively related to foliar phenolic levels due to similar absorption features (Soukupova, Rock & Albrechtova 2002). We found negative relationships between tannins and lignin levels for both aspen and birch, and a non-significant relationship between salicinoids and lignin, suggesting that models predicting tannins and salicinoids in the current study acted independently of the potential positive association of increased phenolic absorption features when lignin and tannins co-occur. General trends of coefficients and VIP statistics were relatively coordinated for both fresh- and dry-leaf models in aspen, and neither tannin nor salicinoid levels were correlated with foliar water content, suggesting that spectral features important for predicting tannins and salicinoids in aspen are not strongly obscured by water absorption features.

Condensed tannins in both aspen and birch were well predicted using either a combined-species or individual-species model, indicating that spectroscopic prediction of tannins via a generalized multispecies model is achievable. Progress of a generalizable spectroscopic model accurately predicting tannins has lagged, largely due to the lack of standards across many plant taxa that can be used in chemical analyses to capture the qualitative variation in specific condensed tannin profiles. A commonly used approach for chemical analysis of tannins is to

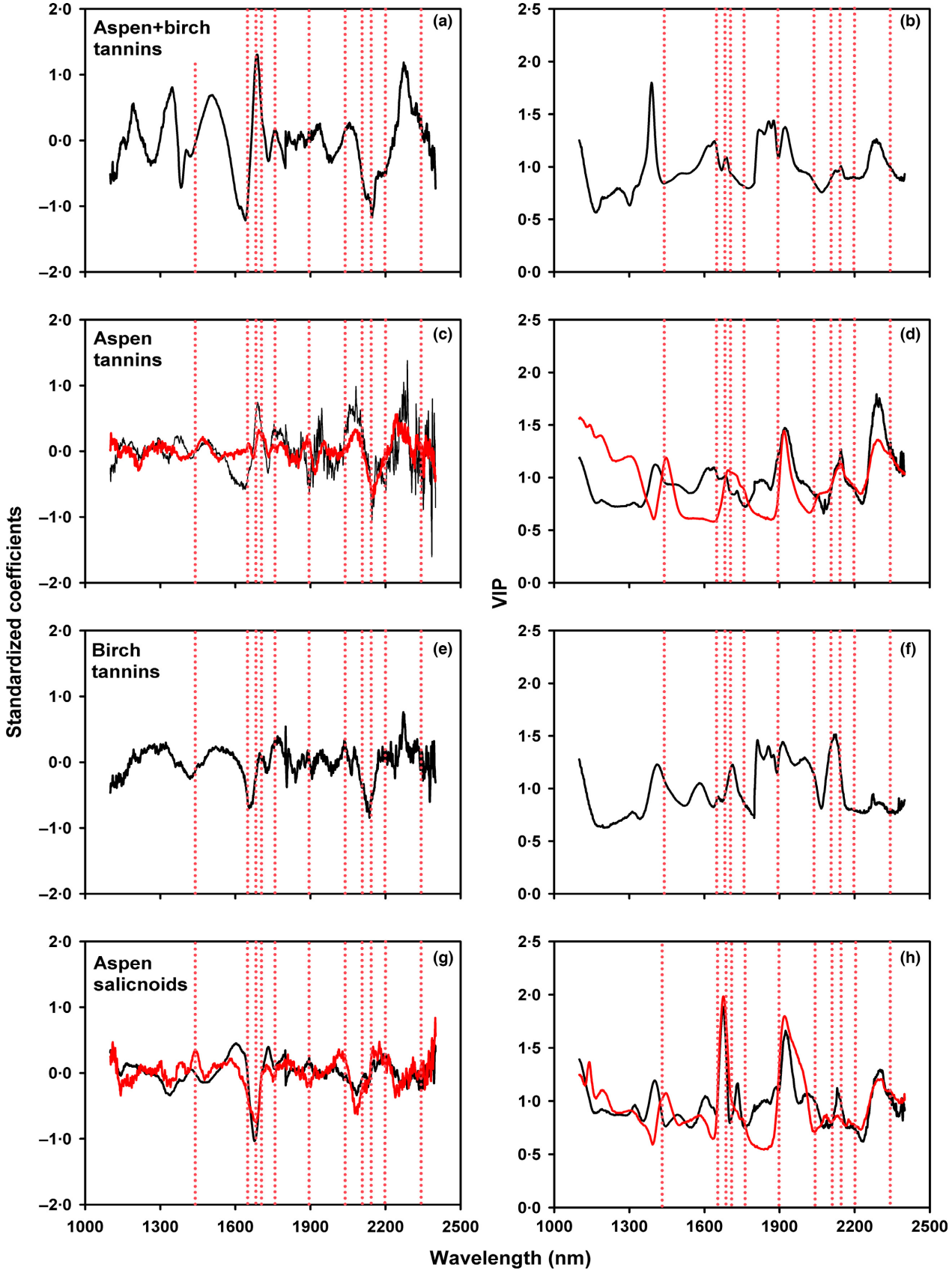

Fig. 4. Standardized coefficients (left column) and variable importance for projection values (VIP; right column) for a combined-species model and individual-species models predicting aspen and birch tannins and aspen salicinoids. For aspen tannins and salicinoids, black lines represent fresh-leaf model data, and red lines represent dry-leaf model data. Red vertical dotted lines represent wavelengths reported in the literature having absorption features associated with phenolic compounds.

© 2016 The Authors. Methods in Ecology and Evolution © 2016 British Ecological Society, *Methods in Ecology and Evolution*

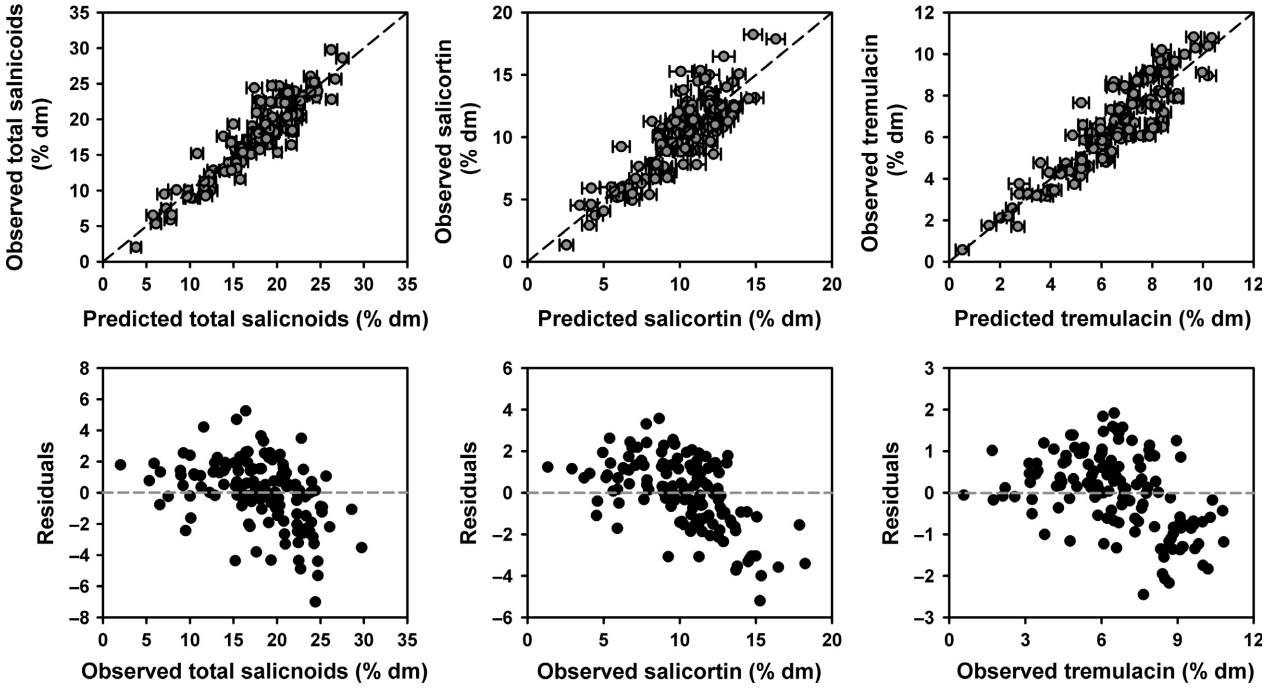

Fig. 5. Observed vs. predicted salicinoid concentrations in aspen fresh leaves (top panel, % dm = % dry mass). Error bars for predicted values represent the standard deviations obtained from the 500 simulated models. 1:1 line in dashed black. Residuals plotted against observed values (bottom panel).

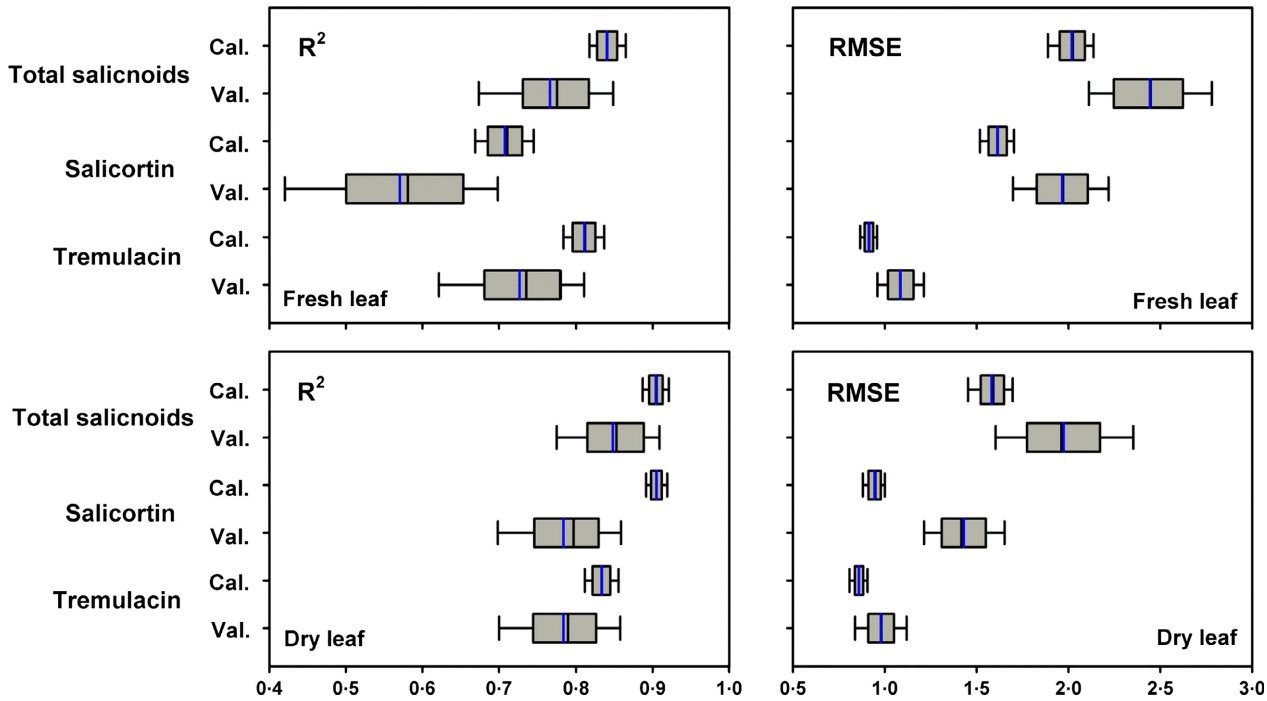

Fig. 6. Distribution (95% confidence intervals) of internal calibration (Cal.) and external validation (Val.) statistics from 500 simulations for models predicting salicinoids in aspen fresh (top panels) and dry (bottom panels) leaves. Each model permutation included 70% of the data for internal calibration and the remaining 30% for external validation. Black vertical line represents median value; blue vertical line represents mean value.

employ commercially available standards, resulting only in measurements relative to that standard for all plant species analysed, regardless of whether or not the specific tannins present in the standard are present in the species being studied. This common approach thus misses the qualitative variation in tannins that comprise the full tannin profile within a species, in which individual tannins differ in polymer length and functional groups (Giner-Chavez *et al.* 1997; Kraus, Dahlgren & Zasoski 2003). The use of a standard not purified from closely related taxa in chemical assays likely weights the quantification

Table 1. Correlation matrix describing relationships among secondary metabolites and nitrogen, lignin and water content of fresh aspen and birch leaves

Aspen	Nitrogen	Lignin	NDWI	Tannins
Lignin	**−0·29**			
NDWI	**−0·34**	**0·60**		
Tannins	**−0·19**	**−0·35**	−0·06	
Salicinoids	−0·03	0·10	0·11	**−0·67**
Birch	**Nitrogen**	**Lignin**	**NDWI**	
Lignin	**−0·34**			
NDWI	−0·10	0·07		
Tannins	**−0·18**	**−0·72**	0·07	

Nitrogen and lignin were predicted using spectroscopy using coefficients from Serbin (2012) and water content (NDWI) was determined following Gao (1996) and calculated as the relative difference between reflectance at wavelengths 857 and 1241 nm. Bolded values are statistically significant at $P < 0{\cdot}05$.

of tannins in favour of those plant taxa whose tannin profiles are more similar to the standard, and promotes the potentially erroneous assumption that all tannins in all species have similar absorption features. Methods that use commercial standards for estimates across unrelated species therefore bias the concentration estimates to those tannins present in the standard.

An alternative approach for tannin analysis is the purification of tannin standards directly from the species being measured, as was done in the current study, or from closely related species. This approach provides more information about the specific tannin profile within a species, but is more labour-intensive and can become logistically difficult in speciose communities. Purifying standards from individual or closely related species for use in the chemical analyses of reference material used in spectroscopic modelling, however, provides the most accurate chemical measurement of the tannin concentration in the leaf and thus the most accurate spectroscopic retrieval. In addition, models built using accurate chemical analyses can provide more insight into the optical properties of tannins within specific species and can subsequently be combined to produce multispecies models. Finally, the ability to make ecological inferences about the function of tannins in ecosystems is stronger if the tannin profile used as the basis for quantification, either in the laboratory or from spectral retrievals, matches that of the taxa being studied. Because tannin profiles vary across plant species, tannin estimates based on the use of standards not related to the taxa of interest may not be reliable indicators of the relative roles of the differences in tannin concentrations in ecosystem functioning.

Comparison of the standardized coefficients from the separate aspen and birch models predicting tannins in this study reveals a statistically significant positive correlation (Fig. S5a, Supporting information). The lack of a perfect correlation and differing magnitudes of bandwise model coefficients (Fig. S5b, Supporting information), however, illustrate the sensitivity of our models to qualitative variation in tannin profiles between plant species used in this study (i.e. the two species consist of different 'varieties' of condensed tannins) and substituting coefficients among the models results in poor prediction results (e.g. aspen coefficients for birch predictions and vice versa; Fig. S5c, d, Supporting information). Nevertheless, a generalized cross-taxa model performed well (Fig. 3) and common reflectance features (e.g. at 1650–1700, 2150 and 2400 nm) of importance were found in the two species-specific models (Fig. 4c–f), suggesting that these spectral regions that can be leveraged to yield a robust cross-taxa model. Because of the important roles of tannins in ecosystem functioning and their ubiquity among plant taxa, developing a generalized model for accurate tannin estimates using reflectance spectroscopy is needed to facilitate the integration of chemical, landscape and ecosystem ecology.

Models predicting aspen tannins and salicinoids from dried, ground leaves were more accurate and less variable than models built using spectra of fresh, green leaves. Fresh-leaf spectra exhibit confounding factors in water absorption regions (approximately 1350–1450 and 1850–1975 nm) that may mask optical features utilized in models built with dry-leaf spectra (Curran *et al.* 1992; Gao & Goetz 1994; Ramoelo *et al.* 2011), including phenolics (Bian *et al.* 2013; Kokaly & Skidmore 2015). We found no relationship, however, between foliar water content and salicinoid levels. Numerous phenolic absorption features are present outside or persist in the presence of water bands in the SWIR region, making it possible to predict phenolics in fresh leaves (Bian *et al.* 2013; Kokaly & Skidmore 2015). The decrease in model performance from dry- to fresh-leaf spectra was marginal, however, given the additional necessary steps of drying and grinding the leaves for analysis. Our findings reinforce the idea that prediction of specific phenolic compounds in fresh leaves can be a relatively fast and accurate, as well as non-destructive, surrogate for standard chemical analyses.

Total salicinoid concentrations were better predicted than individual salicinoid concentrations. This outcome demonstrates the additive effects of co-aligned spectral features from multiple, similarly structured compounds. Specifically, the combination of individual compounds increases the probability that a detection threshold of concentration, and its subsequent effect on leaf spectra, will be surpassed (Rubert-Nason *et al.* 2013). We found that the prediction of tremulacin in fresh leaves across 500 permutations of the data was more stable than that of salicortin. This result may be a product of the additional benzoate moiety on tremulacin (Lindroth, Hsia & Scriber 1987; Philippe & Bohlmann 2007), potentially making the signature of tremulacin more pronounced than salicortin. This interpretation is supported by the detection of more prominent spectral features in purified tremulacin, compared with those of salicortin, in spectral regions of known phenolic absorption (Fig 1c and d; see features at 1420, 1650–1700, 1900 and 2140 nm). Although prediction performances decreased for individual compared with total salicinoids, the prediction of individual constituents represents a substantial gain in valuable biochemical information relevant to plant function, within an

© 2016 The Authors. Methods in Ecology and Evolution © 2016 British Ecological Society, *Methods in Ecology and Evolution*

acceptable range of uncertainty (7–8% for individual salicinoids vs. 5% for total salicinoids), compared to estimating the bulk concentration of salicinoids.

The capability to repeatedly measure plant secondary metabolites *in vivo* using spectroscopy is an important advancement in our ability to understand numerous aspects of plant ecology, including responses to stress, chemodiversity, genetic variation and ecosystem functioning. Recurrent spectral measurements can provide information about the responses of plant secondary metabolite responses to damage (Couture, Serbin & Townsend 2013) and environmental change (Couture, Serbin & Townsend 2015). In addition, chemodiversity in plants is dependent on secondary metabolites (Weng, Philippe & Noel 2012) and specific constituents within a broad category of secondary metabolites can perform vastly different functions. For example, within the phenolic groups in aspen, salicinoids are generally regarded to provide defence against herbivory (Boeckler, Gershenzon & Unsicker 2011; Lindroth & St Clair 2013), while condensed tannins have little influence on insect herbivores, but function as important regulators of nutrient cycling (Madritch & Lindroth 2015). When scaled to landscape levels using airborne platforms and hyperspectral remote sensing, our ability to detect intra- and interspecific chemical diversity in plants, and how this variation influences ecosystem functioning, would be enhanced if specific secondary metabolites were estimated, opposed to estimation of broad, general classes.

Methodologically, our results show that we can potentially use spectroscopy to more broadly characterize the spatial and temporal variation in important secondary metabolites. Moreover, leaf-level estimates provide the basis for scaling plant traits to canopy and landscape levels through interpolation or using remote sensing platforms (Singh *et al*. 2015). The ability to simultaneously measure multiple plant traits is a powerful attribute of reflectance spectroscopy. The suite of traits currently estimable, however, needs to expand by including specific secondary metabolites that play influential roles in ecosystem functioning if we are to advance the integration of chemical, landscape and ecosystem ecology.

Acknowledgements

We are grateful to Nick Grout for glasshouse assistance, Eric Kruger for paper birch seeds and Hillary Grabner for assistance with fieldwork. This research was supported by USDA NIFA AFRI Fellowship grant 2012-67012-19900 to JJC, NASA grants NNX10AJ94G to PAT and NXN09AK15G to PAT and RLL, and USDA NIFA McIntire-Stennis projects WIS01651 to RLL and WIS01531 and WIS01599 to PAT.

Data accessibility

Spectral data used in this study and the partial least squares regression code used for model building are archived in the Ecosystem Spectral Information System (EcoSIS; www.ecosis.org) and can be found at https://ecosis.org/#result/d5445eb9-f334-4ee7-90a9-1fe07e67a20c.

References

Abreu, I.N., Ahnlund, M., Moritz, T. & Albrectsen, B.R. (2011) UHPLC-ESI/TOFMS determination of salicylate-like phenolic glycosides in *Populus tremula* leaves. *Journal of Chemical Ecology*, **37**, 857–870.

Asner, G.P. & Martin, R.E. (2008) Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels. *Remote Sensing of Environment*, **112**, 3958–3970.

Asner, G.P. & Martin, R.E. (2011) Canopy phylogenetic, chemical and spectral assembly in a lowland Amazonian forest. *Remote Sensing of Environment*, **189**, 999–1012.

Atzberger, C., Guerif, M., Baret, F. & Werner, W. (2010) Comparative analysis of three chemometric techniques for the spectroradiometric assessment of canopy chlorophyll content in winter wheat. *Computers and Electronics in Agriculture*, **73**, 165–173.

Bian, M., Skidmore, A.K., Schlerf, M., Wang, T., Liu, Y., Zeng, R. & Fei, T. (2013) Predicting foliar biochemistry of tea (Camellia sinensis) using reflectance spectra measured at powder, leaf and canopy levels. *ISPRS Journal of Photogrammetry and Remote Sensing*, **78**, 148–156.

Boeckler, G.A., Gershenzon, J. & Unsicker, S.B. (2011) Phenolic glycosides of the Salicaceae and their role as anti-herbivore defenses. *Phytochemistry*, **72**, 1497–1509.

Bolster, K.L., Martin, M.E. & Aber, J.D. (1996) Determination of carbon fraction and nitrogen concentration in tree foliage by near infrared reflectance: a comparison of statistical methods. *Canadian Journal of Forest Research*, **26**, 590–600.

Carvalho, S., Macel, M., Schlerf, M., Moghaddam, F.E., Mulder, P.P.J., Skidmore, A.K. & Van Der Putten, W.H. (2013) Changes in plant defense chemistry (pyrrolizidine alkaloids) revealed through high-resolution spectroscopy. *ISPRS Journal of Photogrammetry and Remote Sensing*, **80**, 51–60.

Chen, S., Hong, X., Harris, C.J. & Sharkey, P.M. (2004) Spare modeling using orthogonal forest regression with PRESS statistic and regularization. *IEEE Transactions on Systems, Man and Cybernetics*, **34**, 898–911.

Chong, I.-G. & Jun, C.-H. (2005) Performance of some variable selection methods when multicollinearity is present. *Chemometrics and Intelligent Laboratory Systems*, **78**, 103–112.

Close, D.C. & McArthur, C. (2002) Rethinking the role of many plant phenolics - protection from photodamage not herbivores. *Oikos*, **99**, 166–172.

Constabel, P.C. & Lindroth, R.L. (2010) The impact of genomics on advances in herbivore defense and secondary metabolism in *Populus*. *The Genetics and Genomics of Populus. Plant Genetics and Genomics, vol 8* (eds S. Jansson, R. Bhalerao & A. Groover), pp. 279–305. Springer-Verlag, Inc., New York.

Couture, J.J., Serbin, S.P. & Townsend, P.A. (2013) Spectroscopic sensitivity of real-time, rapidly induced phytochemical change in response to damage. *New Phytologist*, **198**, 311–319.

Couture, J.J., Serbin, S.P. & Townsend, P.A. (2015) elevated temperature and periodic water stress alter growth and quality of common milkweed (*Asclepias syriaca*) and monarch (*Danaus plexippus*) larval performance. *Arthropod-Plant Interactions*, **9**, 149–161.

Curran, P.J., Dungan, J.L., Macler, B.A., Plummer, S.E. & Peterson, D.L. (1992) Reflectance spectroscopy of fresh whole leaves for the estimation of chemical concentration. *Remote Sensing of Environment*, **39**, 153–166.

Donaldson, J.R. & Lindroth, R.L. (2007) Genetics, environment, and their interaction determine efficacy of chemical defense in trembling aspen. *Ecology*, **88**, 729–739.

Ebbers, M.J.H., Wallis, I.R., Dury, S., Floyd, R. & Foley, W.J. (2002) Spectrometric prediction of secondary metabolites and nitrogen in fresh Eucalyptus foliage: towards remote sensing of the nutritional quality of foliage for leaf-eating marsupials. *Australian Journal of Botany*, **50**, 761–768.

Flinn, P.C., Edwards, N.J., Oldham, C.M. & McNeil, M. (1996) Near-infrared analysis of the fodder shrub tagasaste (*Chamaecytisus proliferus*) for nutritive value and anti-nutritive factors. *Near Infrared Spectroscopy: The Future Waves* (eds A.M.C. Davies & P. Williams), pp. 576–580. NIR, Chichester.

Font, R., Del Río-Celestino, M., Rosa, E., Aires, A. & Haro-Balión, De (2005) Glucosinolate assessment in *Brassica oleracea* leaves by near-infrared spectroscopy. *Journal of Agricultural Sciences*, **145**, 65–73.

Galloway, J.N., Dentener, F.J., Capone, D.G., Boyer, E.W., Howarth, R.W., Seitzinger, S.P. *et al*. (2004) Nitrogen cycles: past, present, and future. *Biogeochemistry*, **70**, 153–226.

Gao, B. (1996) NDWI-a normalized difference water index for remote sensing of vegetation liquid from space. *Remote Sensing of Environment*, **58**, 257–266.

Gao, B.-C. & Goetz, A.F.H. (1994) Extraction of dry leaf spectral features from reflectance spectra of green vegetation. *Remote Sensing of Environment*, **47**, 369–374.

Gillon, D., Houssard, C. & Joffre, R. (1999) Using near-infrared reflectance spectroscopy to predict carbon, nitrogen and phosphorus content in heterogeneous plant material. *Oecologia*, **118**, 173–182.

Giner-Chavez, B.I., Van Soest, P.J., Robertson, J.B., Lascano, C., Reed, J.D. & Pell, A.N. (1997) A method for isolating condensed tannins from crude extracts with trivalent ytterbium. *Journal of Science of Food and Agriculture*, **74**, 359–368.

Grossman, Y.L., Ustin, S.L., Jacquemoud, S., Sanderson, E.W., Schmuck, G. & Verdebout, J. (1996) Critique of stepwise multiple linear regression for the extraction of leaf biochemistry information from leaf reflectance data. *Remote Sensing of Environment*, **56**, 182–193.

Hagerman, A.E. & Butler, L.G. (1980) Condensed tannin purification and characterization of tannin-associated proteins. *Journal of Agricultural and Food Chemistry*, **28**, 947–952.

Hättenschwiler, S. & Vitousek, P.M. (2000) The role of polyphenols in terrestrial ecosystem nutrient cycling. *Trends in Ecology and Evolution*, **15**, 238–243.

Hernes, P.J. & Hedges, P.I. (2000) Determination of condensed tannin monomers in environmental samples by capillary gas chromatography of acid depolymerization extracts. *Analytical Chemistry*, **72**, 5115–5124.

Kleinebecker, T., Schmidt, S.R., Fritz, C., Smolders, A.J.P. & Hoelzel, N. (2009) Prediction of delta 13C and delta 15N in plant tissues with near-infrared reflectance spectroscopy. *New Phytologist*, **184**, 732–739.

Kokaly, R.F. & Skidmore, A.K. (2015) Plant phenolics and absorption features in vegetation reflectance spectra near 1.66 μm. *International Journal of Applied Earth Observation and Geoinformation*, **43**, 55–83.

Kraus, T.E.C., Dahlgren, R.A. & Zasoski, R.J. (2003) Tannins in nutrient dynamics of forest ecosystems - a review. *Plant and Soil*, **256**, 41–66.

Kraus, T.E.C., Zasoski, R.J., Dahlgren, R.A., Horwath, W.R. & Preston, C.M. (2004) Carbon and nitrogen dynamics in forest soil amended with purified tannins from different plant species. *Soil Biology and Biochemistry*, **36**, 309–321.

Lindroth, R.L., Hsia, M.T.S. & Scriber, J.M. (1987) Characterization of phenolic glycosides from quaking aspen. *Biochemical Systematics and Ecology*, **15**, 677–680.

Lindroth, R.L. & St Clair, S.B. (2013) Adaptations of quaking aspen (*Populus tremuloides* Michx.) for defense against herbivores. *Forest Ecology and Management*, **299**, 14–21.

Lindroth, R.L., Kopper, B.J., Parsons, W.F.J., Bockheim, J.G., Karnosky, D.F., Hendrey, G.R., Pregitzer, K.S., Isebrands, J.G. & Sober, J. (2001) Consequences of elevated carbon dioxide and ozone for foliar chemical composition and dynamics in trembling aspen (*Populus tremuloides*) and paper birch (*Betula papyrifera*). *Environmental Pollution*, **115**, 395–404.

Madritch, M.D., Jordan, L.M. & Lindroth, R.L. (2007) Interactive effects of condensed tannin and cellulose additions on soil respiration. *Canadian Journal of Forest Research*, **37**, 2063–2067.

Madritch, M.D. & Lindroth, R.L. (2015) Condensed tannins increase nitrogen recovery by tree species following insect defoliation. *New Phytologist*, **208**, 410–420.

Mevik, B.-H. & Wehrens, R. (2007) The *pls* package: principal component and partial least squares regression in R. *Journal of Statistical Software*, **18**, 1–24.

Osier, T.L. & Lindroth, R.L. (2001) Effects of genotype, nutrient availability, and defoliation on aspen phytochemistry and insect performance. *Journal of Chemical Ecology*, **27**, 1289–1313.

Petisco, C., Garcia-Criado, B., Mediavilla, S., Vazquez De Aldana, B.R., Zabalgogeazcoa, I. & Garcia-Ciudad, A. (2006) Near-infrared reflectance spectroscopy as a fast and non-destructive tool to predict foliar organic constituents of several woody species. *Analytical and Bioanalytical Chemistry*, **386**, 1823–1833.

Philippe, R.N. & Bohlmann, J. (2007) Poplar defense against insect herbivores. *Canadian Journal of Botany*, **85**, 1111–1126.

Porter, L.J., Hrstich, L.N. & Chan, B.G. (1986) The conversion of procyanidins and prodelphinidins to cyanidin and delphinidin. *Phytochemistry*, **25**, 223–230.

Ramoelo, A., Skidmore, A.K., Schlerf, M., Mathieu, R. & Heitköing, I.M.A. (2011) Water-removed spectra increase the retrieval accuracy when estimating savannah grass nitrogen and phosphorous concentrations. *ISPRS Journal of Photogrammetry and Remote Sensing*, **66**, 408–417.

Rubert-Nason, K.F., Holeski, L.M., Couture, J.J., Gusse, A., Undersander, D.J. & Lindroth, R.L. (2013) Rapid phytochemical analysis of birch (Betula) and poplar (Populus) foliage by near-infrared reflectance spectroscopy. *Analytical and Bioanalytical Chemistry*, **405**, 1333–1344.

Schulz, H., Engelhardt, U.H., Wegent, A., Drews, H.H. & Lapczynski, S. (1999) Application of near-infrared reflectance spectroscopy to the simultaneous prediction of alkaloids and phenolic substances in green tea leaves. *Journal of Agricultural and Food Chemistry*, **47**, 5064–5067.

Schweitzer, J.A., Madritch, M.D., Bailey, J.K., LeRoy, C.J., Fischer, D.G., Rehill, B.J. *et al.* (2008) From genes to ecosystems: the genetic basis of condensed tannins and their role in nutrient regulation in a *Populus* model system. *Ecosystems*, **11**, 1005–1020.

Serbin, S.P. (2012) *Spectroscopic Determination of Leaf Nutritional, Morphological, and Metabolic Traits*, PhD dissertation. UWMadison, Madison, WI, USA.

Serbin, S.P., Dillaway, D.N., Kruger, E.L. & Townsend, P.A. (2012) Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature. *Journal of Experimental Botany*, **63**, 489–502.

Serbin, S.P., Singh, A., Mcneil, B.E., Kingdon, C.C. & Townsend, P.A. (2014) Spectroscopic determination of leaf morphological and biochemical traits for northern temperate and boreal tree species. *Ecological Applications*, **24**, 1651–1669.

Shenk, J.S., Workman, J.J. Jr & Westerhaus, M.O. (1992) Application of NIR spectroscopy to agricultural products. *Handbook of Near-Infrared Analysis, Vol 13* (eds D.A. Burns & E.W. Ciuczak), pp. 383–431. Marcel-Dekker, New York.

Singh, A., Serbin, S.P., McNeil, B.E., Kingdon, C.C. & Townsend, P.A. (2015) Imaging spectroscopy algorithms for mapping canopy foliar chemical and morphological traits and their uncertainties. *Ecological Applications*, **25**, 2180–2197.

Soukupova, J., Rock, B.N. & Albrechtova, J. (2002) Spectral characteristics of lignin and soluble phenolics in the near infrared – a comparative study. *International Journal of Remote Sensing*, **23**, 3039–3055.

Weng, J.-K., Philippe, R.N. & Noel, J.P. (2012) The rise of chemodiversity in plants. *Science*, **366**, 1667–1670.

Wold, S., Sjostrom, M. & Eriksson, L. (2001) PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, **58**, 109–130.

Wold, S., Ruhe, A., Wold, H. & Dunn, W. (1984) The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM Journal on Scientific and Statistical Computing*, **5**, 735–743.

Received 25 November 2015; accepted 11 May 2016
Handling Editor: Matthew Davey

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Figure S1. Raw reflectance of purified aspen and birch phenolic compounds.

Figure S2. SWIR-region variation in fresh-leaf spectra of (a) aspen and (b) birch.

Figure S3. Distribution (95% confidence intervals) of internal calibration (Cal.) and external validation (Val.) model bias from 500 simulations of models predicting fresh-leaf model combining both aspen and birch, fresh- and dry-leaf aspen-only models, and fresh-leaf birch-only models (top panel) and salicinoids in aspen fresh (middle panel) and dry (bottom panel) leaves.

Figure S4. Observed vs. predicted condensed tannin concentrations (left panel) and salicinoid concentrations (right panel) in aspen dry leaves.

Figure S5. Top panel: a) relationship between standardized coefficients for separate models built from aspen and birch and b) absolute proportion change between aspen and birch model coefficients across wavelengths. Bottom panel: c) observed aspen tannins vs aspen tannins predicted using birch coefficients and d) observed birch tannins vs birch tannins predicted using aspen coefficients.

© 2016 The Authors. Methods in Ecology and Evolution © 2016 British Ecological Society, *Methods in Ecology and Evolution*