Association of spectroscopically determined leaf nutrition related traits and breeding selection in *Sassafras tzumu*

Yanjie Li, Wenjian Liu, Zifeng Tan, Jingmin Jiang and Jun Liu*

Research Institute of Subtropical Forestry, Chinese Academy of Forestry, Fuyang 311400, Zhejiang, China

*Correspondence: Jun Liu, ucjackley@gmail.com

Abstract：

Background: The nutrition related to traits is an influential role in tree growth, tree production and nutrient cycling. Therefore, the influence of genetic parameters on leaf nutrition traits ought to take account of optimal tree breeding selection. However, the measurement methods are seriously affected by the progress of breeding selection program. In this study, we tested the ability of spectroscopy to quantify the specific leaf nutrition traits including Anthocyanins (ANTH), flavonoids (FLAV) and Nitrogen balance index (NBI), and estimated the genetic variation of these leaf traits based on the spectroscopic predicted data. Live fresh leaves of *Sassafras tzumu* were selected for spectral collection, after which concentrations of ANTH, FLAV and NBI were analyzed by standard analytical methods. Partial least squares regression (PLSR), five spectra pre-processing methods, and four variable selection algorisms were conducted for the optimal prediction model selection. Each trait model was simulated 200 times for error estimation.

Results: The stander normal variation (SNV) to the ANTH model and 1st derivatives to the FLAV and NBI models, combined with significant Multivariate Correlation (sMC) algorithm variable selection are finally regarded as the best performance model. The ANTH model produced the highest accuracy of prediction with a mean R^2 of 0.72 and mean RMSE of 0.10 %, followed by FLAV and NBI model (mean R^2 =0.58, mean RMSE = 0.11 % and mean R^2 =0.44,

mean RMSE = 0.04 %). High heritability was found of ANTH FLAV and NBI with h^2 of 0.78, 0.58 and 0.61 respectively. It shows that it is benefitting and possible of breeding selection for the improvement of leaf nutrition traits.

Conclusions: Spectroscopy can successfully characterize the leaf nutrition traits in living tree leaves and the ability to simultaneous multiple plant traits provides a promising and high-throughput tool for the quick analysis of large size samples and serves for genetic breeding program.

Keyword: spectroscopy; Anthocyanins (ANTH); flavonoids (FLAV); Nitrogen balance index (NBI), breeding selection

Background

Nitrogen (N) is one of the most essential nutrients in plant growth, which is needed to improve grain yield and quality (Cánovas *et al.*, 2018). While excessive N fertilizer application brings severe environment problems, its importance for plant productivity still cannot be underestimated. The precise N application in the plant has been an important demand for a long time (Srinivasan, 2006). The nutrition state in plant should be precisely measured at all stage to set up a guide for precise fertilization to avoid N deposition (Payne *et al.*, 2017). N is the most common limiting factor for the individual, natural and artificial ecosystems growth of the plant. Plant requires N to maintain plant metabolism for growth mainly through external and internal sources, including Soil organic matter, fertilizers, atmospheric deposition and stored N by plant themselves (Millard, 1996). Plants, such as boreal species, store N seasonally through the process of internal cycling and it is a major source of N supplement for tree growth especially when the external availability of N is limited (Nambiar and Fife, 1991; Millard and Proe, 1993). Trees store N in the pattern of proteins mainly in their perennial wood and bark tissues in summer and winter. In addition, other parts of foliage trees, like roots and leaves, also store N which provides nutrition for young roots and needles development. Tree N

remobilisation often occurs during the growth season. The stored N mainly determines the amount of N remobilised and plays an important role for the tree seasonal growth. The dynamics and mobilization of N stored in trees have been widely studied (Cyr *et al.*, 1990; Malagoli *et al.*, 2005). The variation of plant species, genotype, soil and environment leads to the diversity of leaf nitrogen content (Sinclair and Horie, 1989; Famula *et al.*, 2019). It is reported that the chlorophyll content has a strong positive correlation with N content which is an estimative index for N status in leaf (Wood *et al.*, 1993). A portable chlorophyll meter, which collects the leaf light interaction that related to greenness and relative chlorophyll content, is developed for the estimation of N status (Moreau *et al.*, 2004; Hardin *et al.*, 2012). In addition to chlorophyll content, flavonoids (FLAV), one of the main polyphenol components, which have antioxidant activity, are confirmed as biological activity and potential health benefit and also applied to the estimation of N status (Tremblay *et al.*, 2012). Evidence shows that the rise of N fertilization will lead to flavonoid content decreasing and chlorophyll content increasing (Padilla *et al.*, 2014). Another N status index, N balance index (NBI),which is the ratio of chlorophyll to flavonoid, is verified that it has a better and more reliable correlation with leaf N concentration than chlorophyll content alone (Tremblay *et al.*, 2012).

Anthocyanins (ANTH) are a group of water soluble flavonoid pigments that occur in all plant tissues. Anthocyanins are mostly related to a wide range of plant colour but often appear as red (Croft and Chen, 2017). In addition, unfavourable conditions will transiently have an impact on anthocyanins accumulation in both juvenile and senescent observable plant leaves (Garriga *et al.*, 2014; Naing *et al.*, 2017; Trojak and Skowron, 2017). Thus, Anthocyanins are taken as an indicator of plant leaf senescence and stresses (Paul *et al.*, 2017; Liu *et al.*, 2019).

However, the dearth of research on plant growth, the variation of N storage and remobilization driving are mainly ascribed to the high consuming methods and the labour to measure the N concentration and index properties (NBI, chlorophyll content, ANTH, and FLAV), such as

Atomic absorption spectrometry (Borges and Holcombe, 2017; Hu *et al.*, 2019), chromatography (Rivero-Villar *et al.*, 2018) and so on. These analytical methods will limit the breeding selection of tree growth with a large number of samples.

Alternatively, Near-infrared spectroscopy (NIRS) is a rapid, high-throughput technique that has been used for chemical components analysis in many fields. NIRS is a promising and reliable method that can be used for the assessment of a large number of samples (Forina *et al.*, 2015; Ramirez *et al.*, 2015; Guillemain *et al.*, 2017; Malegori *et al.*, 2017; Li *et al.*, 2018a). It basically relies on the spectra variation of absorption because of the vibration, stretching and bending of molecular bonds. Special bonds, including C–H, N–H and O–H bonds (Bokobza, 2002), will interact with the specific wavelengths in the NIR spectroscopy. A stable and uniform illumination source of NIR spectra with collected samples will provide a better platform for the organic chemicals (Siesler *et al.*, 2008). To obtain a better predicting result, the multivariate methods such as partial least squares regression (PLSR) (Wold *et al.*, 2001) will be used by pairing the NIR spectra and the independent chemical measurements together to calibrate a high accurate prediction model. PLSR holds the advantages of producing reliable coefficients, reducing the bias and estimated error, and consuming fewer PLSR components, all of which make it one of the most popular methods for chemometric analyses (Bolster *et al.*, 1996; Asner *et al.*, 2011). The satisficed model will then be applied to unknown samples by their spectra data for independent chemical prediction. Our recent research shows that leaf chlorophyll content and colour parameters are predictable on fresh leaf samples with field near infrared spectrophotometry (Li *et al.*, 2019). The total FLAV and ANTH concentration also have been predicted by a general calibration model in *Ginkgo biloba* leaf and four Indonesian herbal plant species, including *Syzigium oleana*, *Piper betle*, *Jasminum* and *Graptophyllum pictum* with NIR reflectance spectroscopy. NIR also serves as an reliable and promising tool

for tree breeding selection programs during its robust prediction. And with robust calibration and skill, it is able to screen massive samples (Gebreselassie *et al.*, 2017; Li *et al.*, 2019).

The robustness and reliability of model accuracy are largely determined by the spectra quality and feature selection. The massive overtones, combinations vibrations information and noise of the raw NIR spectra (Yang *et al.*, 2018) will result in overlapping and difficulty to directly distinguish the target plant properties (Inagaki *et al.*, 2018). Spectra pre-processing methods, can efficiently reduce the overlapping and noise influence, such as stander normal variation (SNV), 1st and 2nd derivatives and so on (Jin *et al.*, 2017; Park *et al.*, 2018). To yield a robust and reliable model and avoid the influence of irrelevant variables and noise, it is essential to carry out variable selection methods to pick the most relevant variables responding to the target properties instead of the full length of spectra (Fernández *et al.*, 2019; Liang *et al.*, 2020).

Nowadays, the joint analyses of chemometric statistics and variable selection algorithms is on a rise to eliminate the irrelevant variables and improve the model accuracy (Caliari *et al.*, 2017; Mancini *et al.*, 2018). The most common methods of variable selection are Genetic algorithm (Ga) (Zhao and Cao, 2016), Regularized elimination procedure (Rep) algorithm (Mehmood *et al.*, 2012), Iterative predictor weighting (Ipw) (Forina *et al.*, 1999) and significant Multivariate Correlation (sMC) algorithm (Tran *et al.*, 2014). However, the comparison of variable selection algorithms along with PLSR for prediction of multiple leaf nutrition traits is less studied.

Sassafras tzumu is a deciduous tree species that has colourful leaves in autumn. Zhejiang province in China is vigorously promoting the cultivation of colourful species and makes the *S. tzumu* become a famous tree species which has been widely planted in Zhejiang province to develop the urban and mountain landscape (Jiang aiping *et al.*, 2016).

Our latest study addresses that the leaf colour and chlorophyll are under genetic control (Li *et al.*, 2019). However, the relationship between inheritance and variation of leaf nutrition traits (NBI, ANTH and FLAV) in *S. tzumu* is less discovered.

Therefore, the aims of this research are to 1) test the capacity of reflectance spectroscopy to characterize the NBI, ANTH and FLAV with PLSR model; 2) find out the most optimal pre-processing method for these three leaf traits. 3) identify the most important wavelength that related to NBI, ANTH and FLAV by four variable selection methods, including significant multivariate correlation (sMC), regularized variable elimination procedure (Rep), iterative predictor weighting (Ipw), and Genetic algorithm (Ga) variable selection; 4) estimate genetic parameters and correlations of NBI, ANTH and FLAV in *S. tzumu*.

Methods and materials

Materials

50 half-sib families of *S. tzumu* were selected for our study from 6 different regions. Trees were planted in a randomised complete block by a 2 m × 3 m spacing Changle Forest Farm Nursery (30°27' N, 119°48' E), Hangzhou, Zhejiang, China in 2016 with 5 replications and 6 individual trees.

NIR spectra collection

Samples spectra data was collected through 5-6 leaves from each tree in October 2018. The NIR spectra data was taken from the upside surface of the leaves for three times with a handheld fibre optic contact prob from a field-based spectrometer (LF-2500, Spectral evolution, USA). Each spectrum took averagely 32 scans with a range of 1100 to 2500 nm by a 6 nm resolution. All spectra were obtained from the leaves of 1500 trees and 500 trees leaves were placed in a marked paper bag and transferred to the refrigerator immediately for chemical measurement.

Leaf FLAV measurement

Each leaf was ground into powder and being mixed with methanol for 24hrs. 0.5 ml (1mg/mL) extract of each sample was taken to mixed with methanol (1.5 ml), 10% aluminium chloride

(0.1 ml), 1 M potassium acetate (0.1 ml) and distilled water (2.8 ml). The mixture was being placed under room temperature for 30 mins and then measured at 415 nm for the absorbance by UV–Visible spectrophotometer (UV-1280, Shimadzu, Japan). The flavonoid content of the sample was accessed by the value of absorbance density (Eom *et al.*, 2007).

Pigment extraction and NBI estimation

A weighed circular piece cutting from each leaf was place into a mortar by a pestle ground with 100% methanol until the colour changed into white. The extract was being centrifuged for 6 mins by 14,000 rpm at 4 ℃ and subsequently assayed by a UV–Visible spectrophotometer (UV-1280, Shimadzu, Japan). It conducted the equation and specific absorption in the wavelength which was reported by Wellburn (1994). The solution was mixed with 3 ml acidified methanol (1 % HCl) at 4 ℃ with moderate shaking for 12 hrs and then being centrifuged for 10 mins at 14,000 rpm. The extraction was then placed into the spectrophotometer, and it took the absorption at 530 and 657 nm wavelengths to determine the ANTH concentration (Strack and Wray, 1989). The NBI index was figured as the ratio of chlorophyll to flavonoid content.

Model calibration and validation

The original five different types of pre-processing spectra (SNV, 1^{st}, 2^{nd} derivatives, SNV+1^{st} derivatives, SNV+ 2^{nd} derivatives) combined with PLSR (Wold *et al.*, 2001) algorithm were compared in our study. The Savitzky-Golay smoothing (Press and Teukolsky, 1990) with a window size of 15 data points was applied in both 1^{st} and 2^{nd} derivatives spectra. PLSR models were generated with leave-one-out cross-validation for the prediction of ANTH, NBI, and FLAV content. Data were randomly split 200 times into calibration (80%) for model building and validation (20%) for model test respectively. Therefore, the PLSR model has been conducted 200 times for the evaluation of model performance. Each model combined with four variable selections (sMC, Ipw, Rep and Ga) was conducted to find out the most important

spectral variables. The coefficient of determination (R^2) and root-mean-square error (RMSE) in each model derived from both calibration (Cal) and validation (Val) were applied for the evaluation model performance.

Statistical analysis

The estimation of genetic parameters were measured by A multivariate restricted maximum likelihood (REML) linear mixed model, details can be found in (Li *et al.*, 2018a). The narrow sense heritability (h^2) of trait i and genetic correlations($r_{g_{ij}}$) and phenotypic correlation($r_{p_{ij}}$) between trait i and trait j were calculated as:

$$h_i^2 = \frac{2.5\sigma_{f_i}^2}{\sigma_{f_i}^2 + \sigma_{e_i}^2}$$

$$r_{g_{ij}} = \frac{\sigma_{fifj}}{\sqrt{\sigma_{f_i}^2 + \sigma_{f_j}^2}}$$

$$r_{p_{ij}} = \frac{\sigma_{fifj} + \sigma_{eiej}}{\sqrt{\left(\sigma_{f_i}^2 + \sigma_{e_i}^2\right)\left(\sigma_{f_j}^2 + \sigma_{e_j}^2\right)}}$$

where $\sigma_{f_i}^2$ is the estimated family variance for trait i, and $\sigma_{f_j}^2$ is the estimated family variance for trait j, $\sigma_{e_i}^2$ and $\sigma_{e_j}^2$ are the residual variances for trait i and j, and σ_{fifj} and σ_{eiej} are the family and residual covariances between traits i and trait j. The realized genetic gain (ΔG_R) was calculated by the difference between the mean breeding values of selected top ratio leaf traits and the total mean of the leaf traits.

R software (version 3.1.2) (R Core Team, 2017) was taken for all of the data analysis. The *pls* package (Mevik *et al.*, 2015) in R was carried out for PLSR model building, the *plsVarSel* (Mehmood *et al.*, 2012) for variables selection, the *prospectr* package (Stevens and Ramirez–Lopez, 2014) for NIR spectra manipulation, the *lme4* package (Bates *et al.*, 2015) for estimation of genetic parameters, and the *ggplot2* package (Wickham, 2016) for visualization plot.

Results

Model performance

Three leaf traits ANTH, FLAV and NBI constructed the NIR spectral PLSR model. The results are shown in Figure 1. ANTH model has the highest accuracy, followed by FLAV and NBI model. The average of R^2 and RMSE for these three models in calibration sets (Cal) is 0.54 (range: 0.43-0.63), 0.47 (range: 0.35-0.58) and 0.36 (range: 0.26-0.45), in validation sets (Val) is 0.54 (range: 0.28-0.75), 0.47 (range: 0.28-0.69) and 0.38 (range: 0.25-0.64) respectively. As for all spectral pre-processing models, SNV+2nd derivative prediction model is found to be the highest well-performing for predicting ANTH concentration than the other pre-processing methods, with a mean R^2_{Cal} and $RMSE_{Cal}$ of 0.59 (range: 0.55-0.63), 0.11% (range: 0.11-0.12%), a mean R^2_{Val} and $RMSE_{Val}$ of 0.57 (range: 0.38-0.72), 0.11% (range: 0.09-0.13%), followed by 2nd, SNV+1st, 1st, original with the mean of R^2 in the calibration sets (Cal) is 0.56 (range: 0.42-0.75), 0.56 (range: 0.51-0.60), 0.53 (range: 0.48-0.59), 0.52 (range: 0.47-0.56), and RMSE 0.11% (range: 0.11-0.12%),), 0.11% (range: 0.11-0.12%), 0.11% (range: 0.11-0.12%), 0.12% (range: 0.11-0.12), and in the validation sets (Val) is 0.57 (range: 0.42-0.75), 0.54 (range:0.30-0.70), 0.53 (range: 0.36-0.67), 0.52 (range: 0.32-0.72), and RMSE 0.11% (range: 0.09-0.13%), 0.12% (range: 0.10-0.14%), 0.12% (range: 0.10-0.14%), 0.11% (range: 0.09-0.14%) respectively. SNV shows the worst effect with the mean of R^2 and RMSE for calibration sets (Cal) and validation sets (Val) 0.49 (range: 0.44-0.54), 0.49 (range: 0.28-0.63), and 0.13% (range: 0.13-0.14%), 0.13% (range: 0.12-0.16%) respectively. However, 1st yields the best PLSR model in the prediction of FLAV and NBI than the other pre-processing model, with high mean R^2_{CV}, R^2_V of 0.51 (range: 0.46-0.58), 0.52 (range: 0.29-0.68), and low mean of $RMSE_{CV}$, $RMSE_V$ of 0.12% (range: 0.11-0.13%), 0.12 (range: 0.10-0.12%) in FLAV model and high mean $R^2_{CV,}$ R^2_V of 0.39 (range: 0.33-0.45), 0.41 (range: 0.26-0.60), and low mean of $RMSE_{CV}$, $RMSE_V$ of 0.05% (range: 0.05-0.05%), 0.05 (range: 0.04-0.06%) in NBI model

respectively. The effect of SNV shows a poor prediction in the FLAV and NBI as well. The mean of R^2_V is 0.40 (range: 0.26-0.64) and 0.47 (range: 0.29-0.64) respectively.

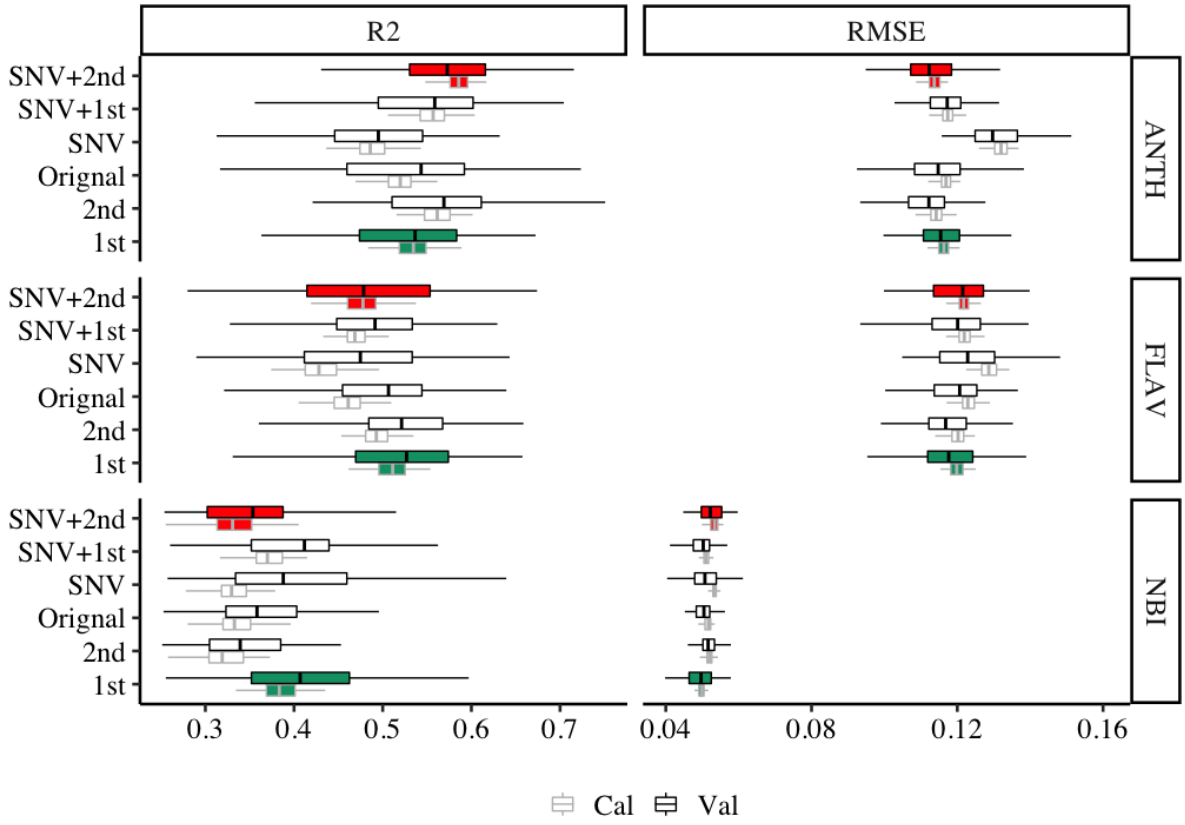
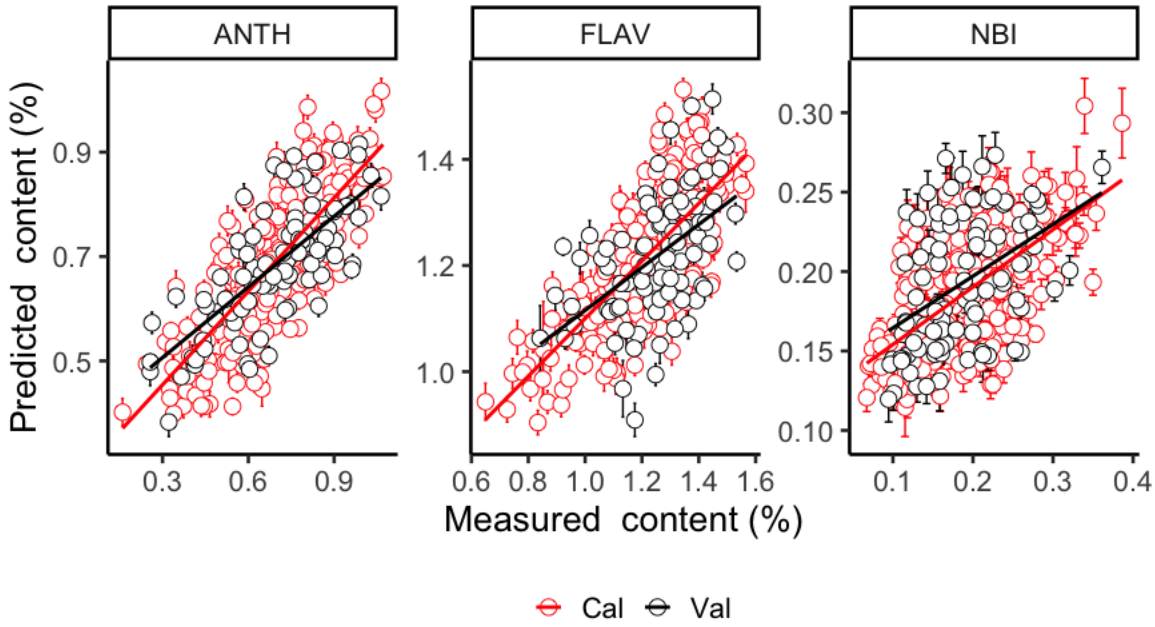


Figure 1 Distribution (95% confidence intervals) of calibration and validation statistics from 200 simulations of models predicting ANTH, FLAV and NBI with full length NIR spectra. Each model permutation included 80% of the data for internal calibration and the remaining 20% for validation. R^2: coefficient of determination of cross-validation; RMSE: root-mean-square error of cross-validation; The black vertical line in each box represents median value, the red colour box represents the SNV+ 2^{nd} model. the green colour box represents the 1^{st} model.

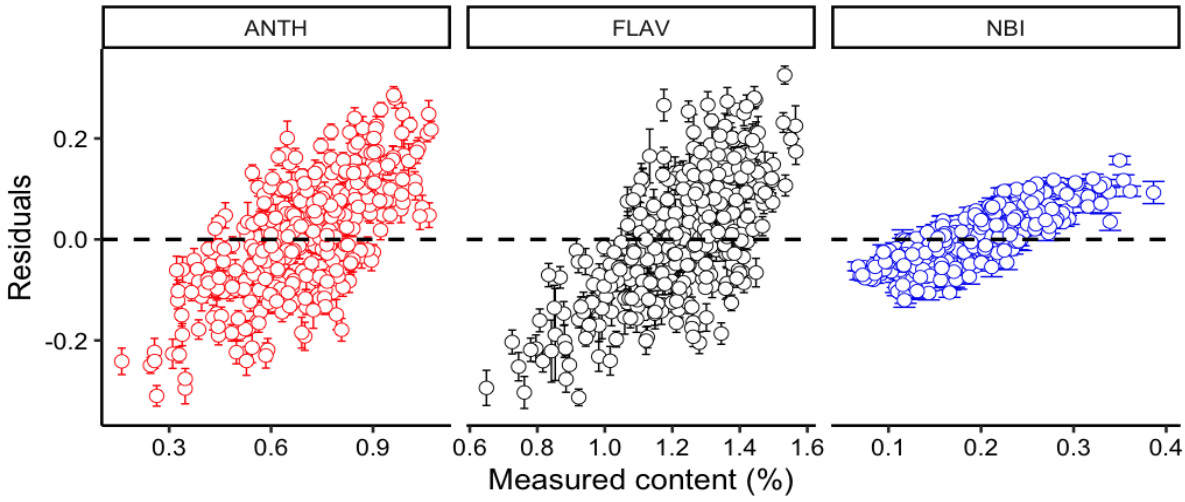
The relationship between the predicted and measured content of Cal and Val datasets by ANTH model with SNV+ 2^{nd} derivative spectra, FLAV and NBI model with 1^{st} derivative spectra was plotted in Figure 2. The error bar represents the prediction error of 200 times per sample. It shows that due to the high accuracy of the ANTH and FLAV models, the predicted values are more correlated with the measured values, while the relationship between predicted and measured values of NBI model is relatively poor. Although the prediction accuracy of each model is different, the prediction error of the Cal and Val data sets is still little.



236

237 Figure 2 Measured and predicted ANTH, FLAV and NBI contents with full length of NIR spectra. Error bars for predicted
238 values represent the standard deviations obtained from the 200 simulated models.

239 The residual of the best processing spectra model for each leaf trait shows that all of these three
240 models tend to be underpredicted when the measurement value is small. With the rise of the
241 measurement value, the prediction value has the tendency of overprediction. The residual value
242 of ANTH, FLAV and NBI model is between an acceptable range from -0.3 to 0.3.



243

244 Figure 3 Residuals plotted against measured ANTH, FLAV and NBI with full length of spectra. Error bars for predicted values
245 represent the standard deviations obtained from the 200 simulated models.

Variable selection and model optimization

Four types of variable selection methods were compared to test the performance of ANTH, FLAV, and NBI PLSR models (Figure 4). The prediction accuracy of ANTH, FLAV, and NBI PLSR models was enhanced much better than the full-length spectra models by these four different variable selection methods. ANTH model still holds the highest R^2 and RMSE value in both Cal and Val data, followed by the FLAV and NBI model. The highest prediction model for ANTH, FLAV and NBI was found through sMC-selected NIR spectra variables with the mean R^2_{Val} of 0.72 (ranged: 0.69 to 0.75), 0.58 (ranged from: 0.54 to 0.62), 0.44 (ranged from: 0.26 to 0.67), and of the mean $RMSE_{Val}$ of 0.10 % (range: 0.09-0.10 %), 0.11 % (range: 0.10-0.12 %), 0.04 % (range: 0.04-0.05 %) respectively. The sMC_PLSR models reached a more stable prediction with less than 16% of full length of spectra on each leaf trait (Figure 5), and having a similar residual range to the model with full spectral information (Figure 6).

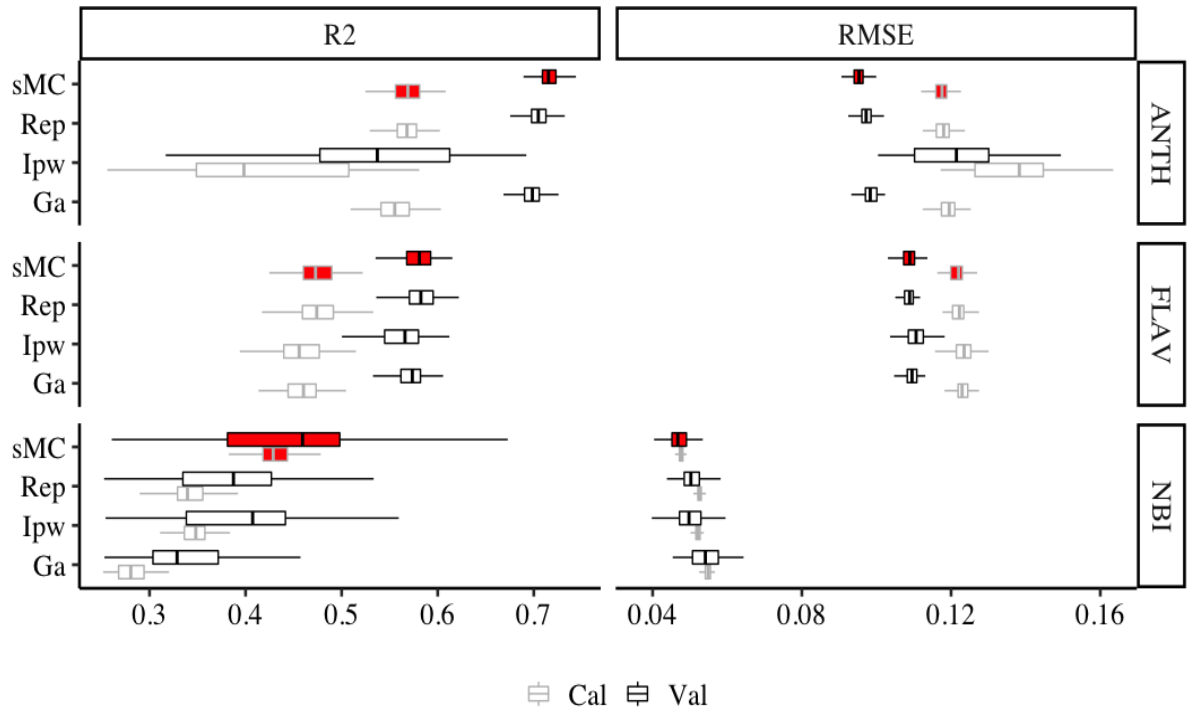


Figure 4 Distribution (95% confidence intervals) of calibration and validation statistics from 200 simulations for models predicting ANTH, FLAV and NBI contents using sMC, Rep, Ipw and Ga variable selection. Each model permutation included

80% of the data for calibration and the remaining 20% for validation. R^2: coefficient of determination of cross-validation; RMSE: root-mean-square error of cross-validation; The black vertical line in each box represents median value, the red colour box represents the sMC model.

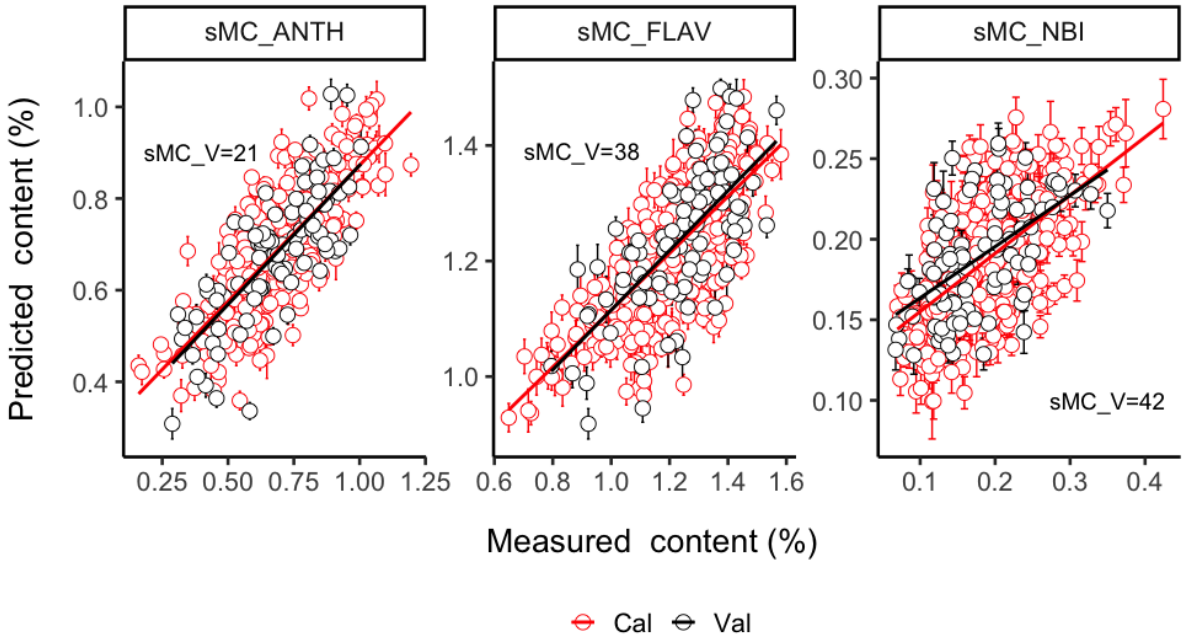


Figure 5 Measured and predicted ANTH, FLAV and NBI contents with sMC selected NIR spectra. Error bars for predicted values represent the standard deviations obtained from the 200 simulated models. sMC_V: the total selected number of variables.

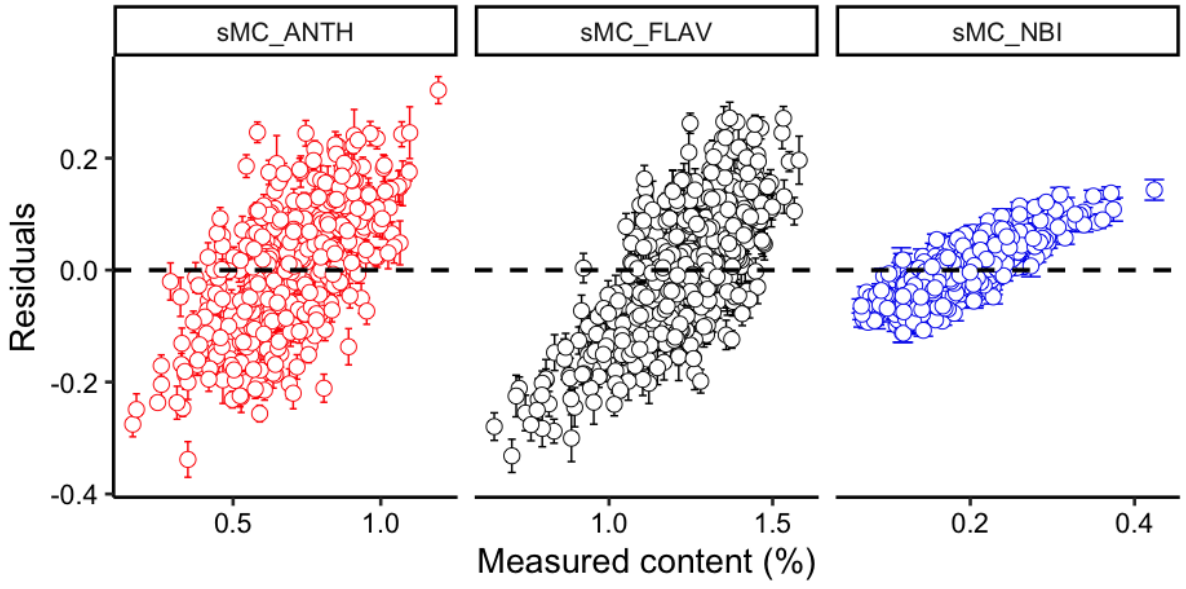


Figure 6 Residuals plotted against measured ANTH, FLAV and NBI with sMC selected spectra. Error bars for predicted values represent the standard deviations obtained from the 200 simulated models.

Figure 7 displays the important variable information area selected by sMC variable selection method in the ANTH, FLAV and NBI model which conducted 200 times on each model. Even the predicted model of three leaf traits was being run 200 times, sMC variable selection brought out stability for the selected important variable areas with a few relative spectral region in prediction models. The variables at 2060，2180，2270，2330 and 2440nm are considered as the vital roles in the construction of ANTH prediction model. As for FLAV, 1070，1235，1950 and 2220 nm are more important areas. Spectroscopic variables at 1100, 1220, 1465, 1950 and 2220 nm make a critical difference in the NBI predictive model.

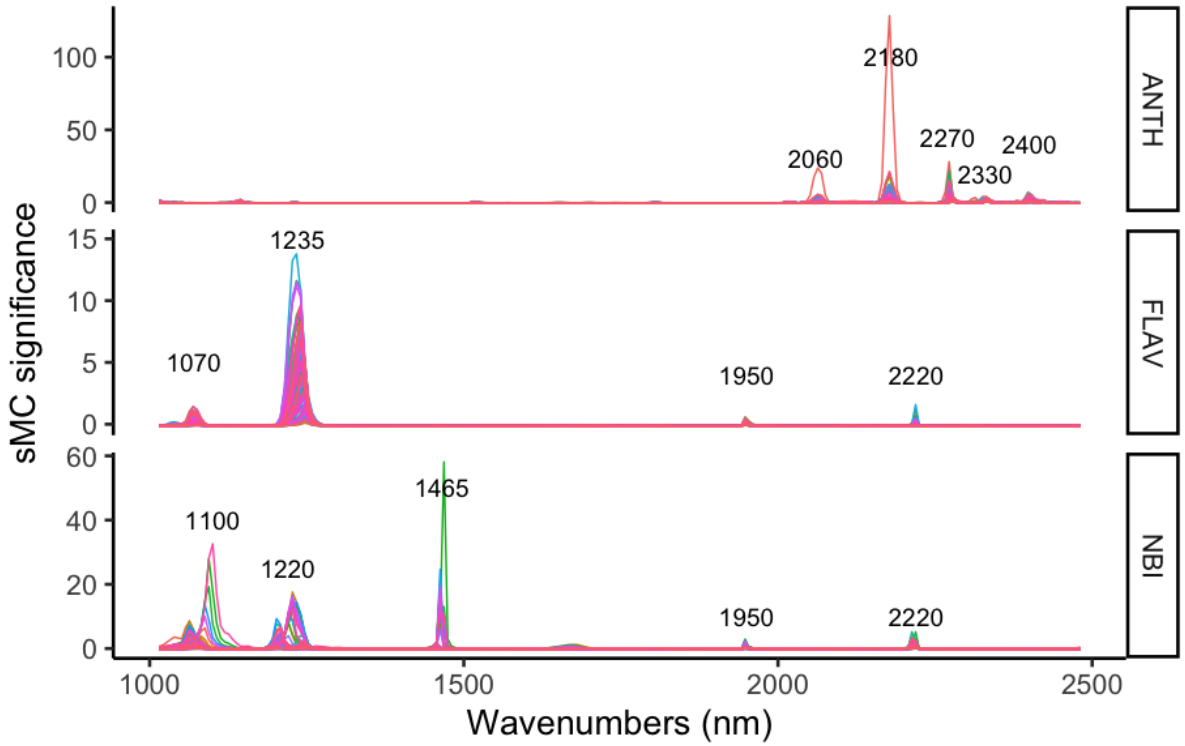


Figure 7 Spectra influence in ANTH, FLAV and NBI models that randomly being conducted 200 times; each line means one time of modelling with sMC variable selection.

Heritability, genetic and phenotypic correlation among traits

Table 1 shows the correlation (genetic and phenotypic) and heritability of three traits. Leaf ANTH produces the highest heritability of 0.78, followed by FLAV and NBI with h^2 of 0.58 and 0.61 respectively. There has no significant genetic and phenotypic correlation between

ANTH, FLAV and NBI. FLAV was found to have the highest positive genetic correlation with ANTH of a value of 0.36.

Table 1 The heritability, genetic (above diagonal) and phenotypic correlation (below diagonal) between ANTH, FLAV and NBI traits with standard error between parentheses.

Traits	ANTH	FLAV	NBI	h^2
ANTH		*0.36 (0.01)*	*0.11 (0.02)*	0.78 (0.10)
FLAV	0.16 (0.03)		*0.09 (0.01)*	0.58 (0.11)
NBI	0.09 (0.01)	0. 12 (0.01)		0.61 (0.08)

Family selection

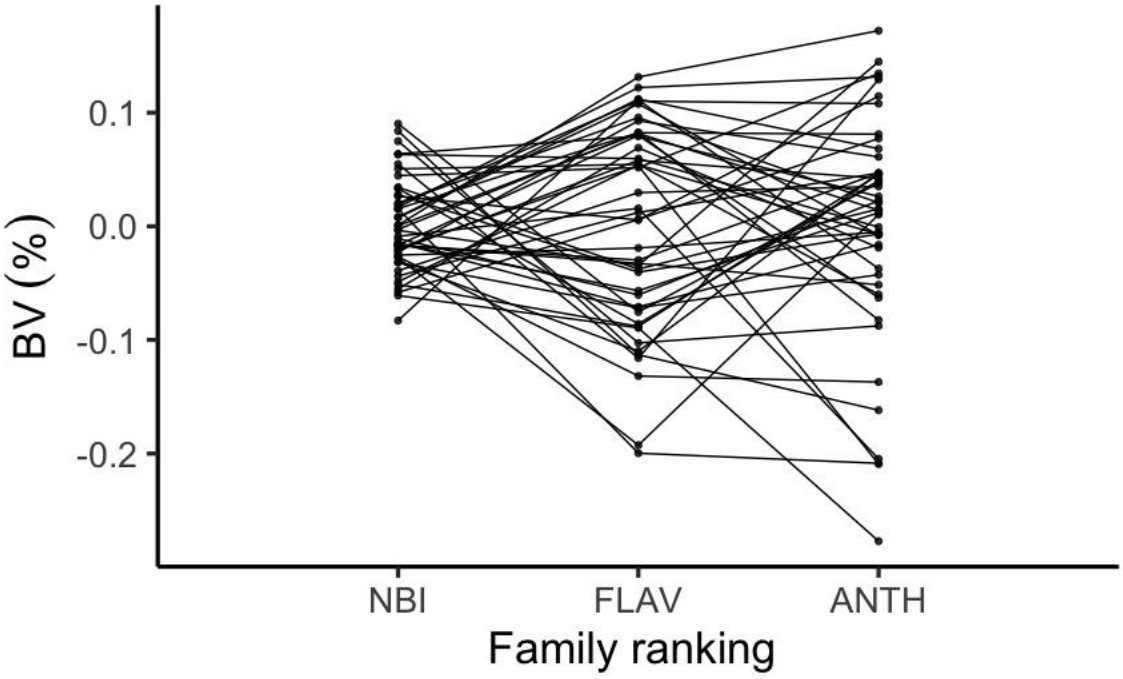


Figure 8 Family ranking for ANTH, FLAV and NBI content in *Sassafras tzumu* at age 2. Family values are expressed as deviation from each trait mean. BV: Breeding values.

The best models of ANTH, FLAV and NBI were applied to predict the remaining 1000 trees spectra. In total ,1500 trees of 50 families were selected for breeding analysis. Figure 8 shows

the distribution of three leaf traits in the ranking of breeding value from 50 families. The ranking of three leaf traits in different families is inconsistent as well as a part of families consistently displaying in the breeding value, which explains that it is feasible to make a family selection of ANTH, FLAV and NBI at the same time through genetic selections.

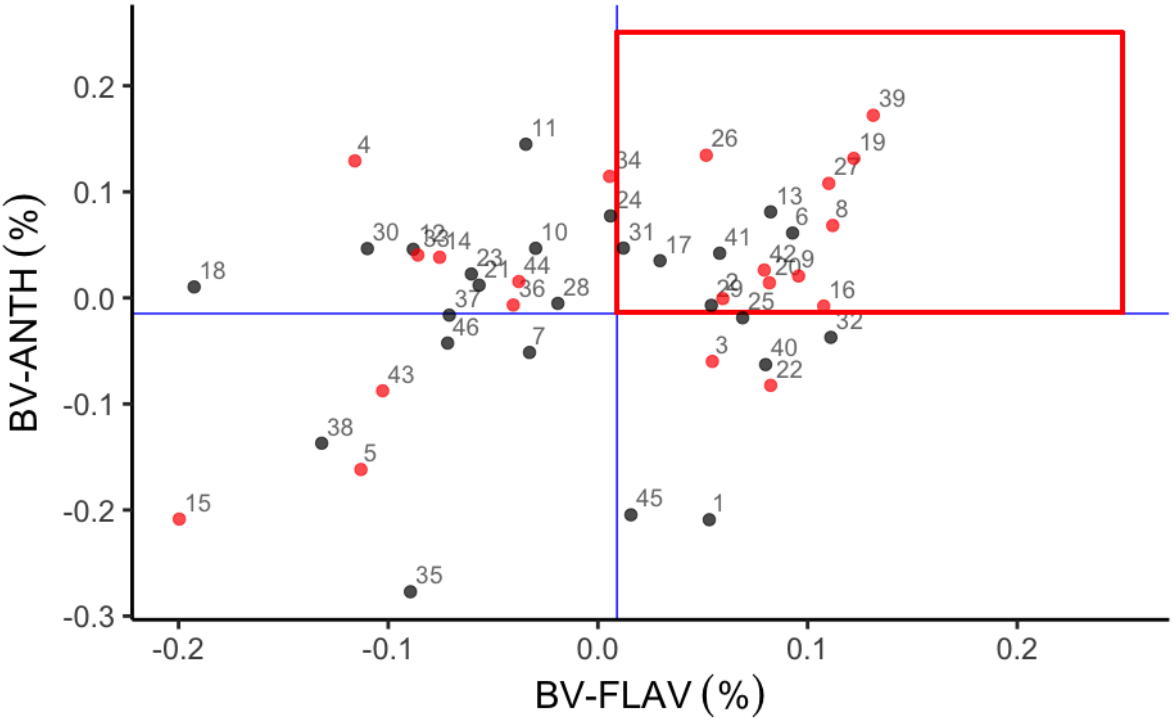


Figure 9 Relationship between ANTH, FLAV and NBI content breeding values of *Sassafras tzumu* families at age 2. BV-ANTH: breeding value of ANTH; BV-FLAV: breeding value of FLAV; the blue solid line: the mean value of each trait breeding value; red square: the region that most interesting.

Figure 9 demonstrates the breeding value distribution of 50 families of three leaf traits. The blue solid lines represent the average of ANTH and FLAV respectively. The families with a higher NBI breeding value than its mean are shown in red, and below the mean is in black. 16 families have high FLAV and ANTH breeding value. 10 families with a high breeding value will be selected If NBI breeding values are required to be above mean. These families can be further taken as genetic family materials for second-generation breeding.

Discussion

The health of tree growth is dictated by main factors, such as soil, nutrients, environment, genetic and so on. N is a key role of nutrient which highly influences the tree growth.

The internal N cycling in trees (Swarts, 2016) is a hot topic in numerous study (Millard and Neilsen, 1989; García-Sánchez *et al.*, 2017; Li and Coleman, 2019). However, the measurement of N concentration limits the access to the further study. In this study, the filed-base reflectance spectroscopy is proved to be a reliable and accurate method to characterize the plant nutrition properties in fresh leaves. The SNV + 2nd derivative spectra for ANTH, and 1st derivative spectra for FLAV and NBI have been identified to increase the model accuracy when calibrating the PLSR prediction models. Incorporate with spectra variable selection, the model accuracy is significantly improved with less variables for the prediction of leaf nutrition traits. Our model offered a promising and reliable result for predicting the FLAV content in fresh leaf (R^2_{Val} =0.58, $RMSE_{Val}$ = 0.11%), which was lower than the result reported for *fresh Ginkgo biloba* leaf in different colors (R^2_{CV} = 0.82 and RMSE = 2.62 %). The variability lessened by small range of NBI value lead to an inefficient prediction. (Blanco and Villarroya, 2002).

Conversely, our result of the prediction of ANTH content illustrates a higher accuracy then the other two leaf traits, with a mean R^2_{Val} of 0.72 (range: 0.69-0.75) and a mean of $RMSE_{Val}$ = 0.09 % (range: 0.09-0.10 %). Similar result was discovered in wine grapes by NIR hyperspectral imaging and PLSR model, which gave R^2 of 0.84 and RMSEP of 0.013% for estimating ANTH content.

A robust statistical methodology for model calibration which was first conducted by Couture *et al.* (2016) was carried out to predict plant leaf secondary metabolites with reflectance spectroscopy. It was being run 200 randomized simulations for calibrating the models to provide an estimation of the model uncertainty and overall stability (Figures 1-7). It is similar to our previous study which takes use of filed spectroscopy to predict the leaf colour and

chlorophyll content (Li *et al.*, 2019). Random sampling (Quentin *et al.*, 2017) and Kennard-Stone sampling algorithm (Li *et al.*, 2018b) in other studies, which sample only once for model calibration, may cause instability for model prediction. Thus, we highly recommend to use this methodology for model calibration and validation on NIR analysis.

The NIR spectra involves not only the favourable information but noise and irreverent information which will encumbrance the model accuracy of prediction. Therefore, variable selection is regarded as an efficient way to find out the most important wavelengths which contributes the minimum error for model calibration and helps to reduce the model processing time For spectral models. Variables in the spectrum play a key role in the predictive accuracy of the model. The spectral information is extensive along with the relevant and irrelevant information, both of which will overlap to interfere the model construction of the useful information and the PLSR model with a specific trait (Workman Jr and Weyer, 2012). Thus, it is vital to screen important variables for spectral information. In this study, four variable selection methods were compared to pick the best variable selection method. It shows that the sMC-PLSR model efficiently identified the key wavelengths and enables us to select a small set of variables to yield a promising and robust calibrated model for the prediction of ANTH, FLAV and NBI. Our results support the research announced by Li and Altaner (2018), who successfully took the sMC variable selection method to improve the accuracy of an NIR calibration model to predict concentrations in extracts of heartwood of *Eucalyptus bosistoana* trees, and Li *et al.* (2019) who found that sMC selection algorithm held the advantage of finding the most relevant variables for the prediction of leaf chlorophyll content and colour parameters. Some studies also states that significance multivariate correlation (sMC) (Tran *et al.*, 2014) is a positive algorithm to remove confounding effects from NIR calibrations (Wijewardane *et al.*, 2016).

Several important variables which are related to the ANTH, FLAV, and NBI have been selected similarly in each model, including the range at 2060，2180，2270，2330 and 2440 nm for ANTH, 1070, 1235, 1950, 2220 nm for FLAV, and 1100, 1220, 1465, 1950, 2220 nm for NBI respectively. As reported by Ramirez *et al.* (2015), the regions around 2060, 2180, 2270, 2330 and 2440 nm are mostly associated with O–H and C–H stretching vibrations as well as the starch and sugar (Decruyenaere *et al.*, 2012). However, in our study, these regions have been ignored. The regions around 1070, 1100, 1220, 1235 nm are mainly assigned to the 1st overtones of C–H combination bands and 1st and 2nd overtones of O–H and N–H stretching vibrations, while the bands around 1465 nm are mostly related to the 1st overtones of O–H stretching vibration, both of which are associated with starch and protein (Curran, 1989; Kokaly, 2001; De Bei *et al.*, 2017). In NIR spectra, water has a wide absorbance region which is a major influence on the other chemical information because of spectra overlap. In our study, the band around 1950 nm related to the water has less contribution to the FLAV and NBI model but no influence on the ANTH model. It probably influences the accuracy of model for the prediction of FLAV and NBI. Correlational study was found by Min *et al.* (2006), who stressed that the regions of 1910 and 1938 nm highly related to water might have a strong impact on the N concentration prediction.

Trees N internal cycling is considered as one of the major ecology factors for tree growth and is an augment for the tree uptake of soil N (Millard, 1989). In addition, it also helps to understand numerous aspects of plant ecology, for instance, to evaluate the effect of the N storage and remobilization in different part tissues of trees in relation to current demands for growth (Schneider *et al.*, 1996), to find out the role of N on growth stress, the relationship with N deposition in forest (Gundersen, 1991; Gundersen *et al.*, 1998) and the relationship with dynamics of carbon recourse in trees (Villar-Salvador *et al.*, 2015; Han and Kabeya, 2017). Our fast and accurate measurement of N index, including ANTH, FLAV and NBI traits of trees

with NIR spectroscopy provides an advanced way for the study of N internal cycling and allows to quickly measure large number of samples.

In this study, we continue to use the coefficients of 1/2.5 for the calculation of heritability of ANTH, FLAV and NBI traits based on our previous study to avoid the assembling of half-siblings and inbreeding effects. (Li *et al.*, 2019). The moderate heritability of ANTH, FLAV and NBI was found, with the value of h^2 ranging from 0.61 to 0.78. The leaf ANTH heritability of 0.78 in our study is similar to the result found by Yihu *et al.* (2009) who figured out the anthocyanin content heritability ranging from 0.79 to 0.91 in leaves of chili pepper higher than 0.29 reported in the leaf of Aspen (*Populus tremula* L.) (Robinson *et al.*, 2012). For FLAV, a significant high rang of heritability from 0.94 to 0.99 was reported in the leave of Ginkgo Trees (Zhang *et al.*, 2017) which was much higher than our study (h^2= 0.58). It indicates that genetic control capacity is different between species even the same traits. Our study proves there is also a potential for the selection for NBI traits in breeding programs while with less study on the estimation of NBI heritability.

The consistence of families ranking of ANTH, FLAV and NBI indicates that the selection for a good leaf nutrition tree is workable, and the selection of qualified nutrition plant is supposed to involve multiple traits, which will afford a stable inheritance.

Conclusion

In conclusion, NIR spectroscopy is potentially taken to estimate the nutrition related traits by fresh leaf. Our reasonable high accuracy of ANTH and FLAV prediction and moderate accuracy of NBI prediction provide an alternative way for the N index traits and open a door to the efficient analysis of the internal N cycling in trees. The pre-processing method and variable selection much influence the performance of model prediction. Our study found that by using of 1st and SNV+ 2nd derivative spectra processing method and sMC variable selection algorithm, the PLSR models have been highly improved. In addition, the repeated spectral

statistical methodology that we applied provided an efficient way to deal with variation in calibration data and generate information on the response of plant nutrition traits with NIR spectra. NIR model serves as an efficient tool for the estimation of genetic parameters and breeding selection in high throughput way to improve the leaf traits quality.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

Not applicable.

Competing interests

The authors declare that there is no conflict of interest.

Funding

This work was funded by National Natural Science Foundation of China (31901323) and Zhejiang Science and Technology Major Program on Agricultural (Forestry) New Variety Breeding (2016C02056-10).

Authors' contributions

Yanjie Li designed the study, conducted the experiment, analysed the data and wrote the manuscript. Wenjian Liu and Zifeng Tan conducted lab experiments, Jun Liu and Jingmin Jiang supervised the experiments at all stages and reviewed the manuscript. All authors have read and approved the final manuscript.

Acknowledgements

Not applicable.

Authors' information

Not applicable.

Reference:

Asner, G.P., et al., 2011. Spectroscopy of canopy chemicals in humid tropical forests. Remote. Sens. Environ. 115, 3587-3598. http://dx.doi.org/10.1016/j.rse.2011.08.020

Bates, D., et al., 2015. Package 'lme4'. Convergence 12, 2.

Blanco, M., Villarroya, I.N.I.R., 2002. NIR spectroscopy: a rapid-response analytical tool. TrAC, Trends. Anal. Chem. 21, 240-250. http://dx.doi.org/10.1016/S0165-9936(02)00404-1

Bokobza, L., 2002. Origin of near-infrared absorption bands. Wiley-VCH. http://dx.doi.org/10.1002/9783527612666.ch02

Bolster, K.L., et al., 1996. Determination of carbon fraction and nitrogen concentration in tree foliage by near infrared reflectances: A comparison of statistical methods. Can. J. Forest. Res.26, 590-600. http://dx.doi.org/10.1139/x26-068

Borges, D.L.G., Holcombe, J.A., 2017. Graphite Furnace Atomic Absorption Spectrometry. Encyclopedia of Analytical Chemistry, 1-20. http://dx.doi.org/10.1002/9780470027318.a5108.pub3

Caliari, Í.P., et al., 2017. Estimation of cellulose crystallinity of sugarcane biomass using near infrared spectroscopy and multivariate analysis methods. Carbohydr. Polym. 158, 20-28. http://dx.doi.org/10.1016/j.carbpol.2016.12.005

Cánovas, F.M., et al., 2018. Nitrogen metabolism and biomass production in forest trees. Front. Plant. Sci. 9, 1449. http://dx.doi.org/10.3389/fpls.2018.01449

459 Couture, J.J., et al., 2016. Spectroscopic determination of ecologically relevant plant secondary
460 metabolites. Methods. Ecol. Evol. 7, 1402-1412. http://dx.doi.org/10.1111/2041-210X.12596
461 Croft, H., Chen, J., 2017. Leaf pigment content. Reference Module in Earth Systems and
462 Environmental Sciences. Oxford: Elsevier Inc, 1-22.
463 Curran, P.J., 1989. Remote sensing of foliar chemistry. Remote. Sens. Environ. 30, 271-278.
464 http://dx.doi.org/10.1016/0034-4257(89)90069-2
465 Cyr, D., et al., 1990. Seasonal dynamics of carbohydrate and nitrogenous components in the
466 roots of perennial weeds. Plant. Cell Environ.13, 359-365. http://dx.doi.org/10.1111/j.1365-
467 3040.1990.tb02139.x
468 De Bei, R., et al., 2017. Rapid measurement of total non-structural carbohydrate concentration
469 in grapevine trunk and leaf tissues using near infrared spectroscopy. Comput. Electron. Agric.
470 136, 176-183. http://dx.doi.org/10.1016/j.compag.2017.03.007
471 Decruyenaere, V., et al., 2012. Development of near-infrared spectroscopy calibrations to
472 quantify starch and soluble sugar content in the roots of *Rumex obtusifolius*. Weed. Res. 52, 1-
473 5. http://dx.doi.org/10.1111/j.1365-3180.2011.00895.x
474 Eom, S.-H., et al., 2007. Far infrared ray irradiation stimulates antioxidant activity in *Vitis*
475 *flexuosa* THUNB. berries. Korean. J. Medicinal. Crop. Sci. 15, 319-323.
476 Famula, R.A., et al., 2019. Association genetics of carbon isotope discrimination and leaf
477 morphology in a breeding population of *Juglans regia* L. Tree. Genet. Genom. 15, 6.
478 http://dx.doi.org/10.1007/s11295-018-1307-4
479 Fernández, J.L., et al., 2019. Determination of the Lignocellulosic Components of Olive Tree
480 Pruning Biomass by Near Infrared Spectroscopy. Energies. 12, 2497.
481 http://dx.doi.org/10.3390/en12132497
482 Forina, M., et al., 1999. Iterative predictor weighting (IPW) PLS: a technique for the
483 elimination of useless predictors in regression problems. J. Chemom. 13, 165-184.

Forina, M., et al., 2015. Artificial nose, NIR and UV–visible spectroscopy for the characterisation of the PDO Chianti Classico olive oil. Talanta. 144, 1070-1078. http://dx.doi.org/10.1016/j.talanta.2015.07.067

García-Sánchez, F., et al., 2017. 7 Irrigation and Fertilization Management in Lime Trees. The Lime: Botany, Production and Uses, 91. http://dx.doi.org/10.1079/9781780647845.0091

Garriga, M., et al., 2014. Chlorophyll, anthocyanin, and gas exchange changes assessed by spectroradiometry in *Fragaria chiloensis* under salt stress. J. Integr. Plant. Biol. 56, 505-515.

Gebreselassie, M.N., et al., 2017. Near-infrared spectroscopy enables the genetic analysis of chemical properties in a large set of wood samples from *Populus nigra* (L.) natural populations. Ind. Crops. Prod. 107, 159-171. http://dx.doi.org/10.1016/j.indcrop.2017.05.013

Guillemain, A., et al., 2017. Performance of NIR handheld spectrometers for the detection of counterfeit tablets. Talanta. 165, 632-640. http://dx.doi.org/10.1016/j.talanta.2016.12.063

Gundersen, P., 1991. Nitrogen deposition and the forest nitrogen cycle: role of denitrification. For. Ecol. Manage. 44, 15-28. http://dx.doi.org/10.1016/0378-1127(91)90194-Z

Gundersen, P., et al., 1998. Impact of nitrogen deposition on nitrogen cycling in forests: a synthesis of NITREX data. For. Ecol. Manage. 101, 37-55. http://dx.doi.org/10.1016/S0378-1127(97)00124-2

Han, Q., Kabeya, D., 2017. Recent developments in understanding mast seeding in relation to dynamics of carbon and nitrogen resources in temperate trees. Ecol. Res. 32, 771-778. http://dx.doi.org/10.1007/s11284-017-1494-8

Hardin, J.A., et al., 2012. In situ measurement of pecan leaf nitrogen concentration using a chlorophyll meter and vis-near infrared multispectral camera. HortScience. 47, 955-960. http://dx.doi.org/10.21273/HORTSCI.47.7.955

507 Hu, X.-F., et al., 2019. The effects of simulated acid rain on internal nutrient cycling and the
508 ratios of Mg, Al, Ca, N, and P in tea plants of a subtropical plantation. Environ. Monit. Assess.
509 191, 99. http://dx.doi.org/10.1007/s10661-019-7248-z
510 Inagaki, T., et al., 2018. Determination of physical and chemical properties and degradation of
511 archeological Japanese cypress wood from the Tohyamago area using near-infrared
512 spectroscopy. J. Wood. Sci. 64, 347-355. http://dx.doi.org/10.1007/s10086-018-1718-8
513 Jiang aiping, et al., 2016. Relationships of Leaf Color and Pigment and Nutrient Elements in
514 Senescing Leaves of *Sassafras tsumu*. For. Res. 29, 362.
515 Jin, X., et al., 2017. Determination of hemicellulose, cellulose and lignin content using visible
516 and near infrared spectroscopy in *Miscanthus sinensis*. Bioresour. Technol. 241, 603-609.
517 http://dx.doi.org/10.1016/j.biortech.2017.05.047
518 Kokaly, R.F., 2001. Investigating a physical basis for spectroscopic estimates of leaf nitrogen
519 concentration. Remote. Sens. Environ. 75, 153-161. http://dx.doi.org/10.1016/S0034-
520 4257(00)00163-2
521 Li, G., Coleman, G.D., 2019. Nitrogen storage and cycling in trees. In, Advances in Botanical
522 Research. Elsevier, pp. 127-155.
523 Li, Y., Altaner, C., 2018. Predicting extractives content of *Eucalyptus bosistoana* F. Muell.
524 Heartwood from stem cores by near infrared spectroscopy. Spectrochim. Acta. A Mol. Biomol.
525 Spectrosc. 198, 78-87. http://dx.doi.org/10.1016/j.saa.2018.02.068
526 Li, Y., et al., 2018a. Genetic variation in heartwood properties and growth traits of *Eucalyptus*
527 *bosistoana*. Eur. J. For. Res. 137, 565-572. http://dx.doi.org/10.1007/s10342-018-1125-0
528 Li, Y., et al., 2019. Spectroscopic determination of leaf chlorophyll content and color for
529 genetic selection on *Sassafras tzumu*. Plant. Methods. 15, 73.
530 http://dx.doi.org/10.1186/s13007-019-0458-0

Li, Z., et al., 2018b. Identification of oil, sugar and crude fiber during tobacco (*Nicotiana tabacum* L.) seed development based on near infrared spectroscopy. Biomass. Bioenergy. 111, 39-45. http://dx.doi.org/10.1016/j.biombioe.2018.01.017

Liang, L., et al., 2020. Prediction of holocellulose and lignin content of pulp wood feedstock using near infrared spectroscopy and variable selection. Spectrochim. Acta. A Mol. Biomol. Spectrosc. 225, 117515. http://dx.doi.org/10.1016/j.saa.2019.117515

Liu, X., et al., 2019. Comparison of prediction power of three multivariate calibrations for estimation of leaf anthocyanin content with visible spectroscopy in *Prunus cerasifera*. PeerJ. 7, e7997. http://dx.doi.org/10.7717/peerj.7997

Malagoli, P., et al., 2005. Dynamics of nitrogen uptake and mobilization in field-grown winter oilseed rape (*Brassica napus*) from stem extension to harvest: I. Global N flows between vegetative and reproductive tissues in relation to leaf fall and their residual N. Ann. Bot. 95, 853-861. http://dx.doi.org/10.1093/aob/mci091

Malegori, C., et al., 2017. Comparing the analytical performances of Micro-NIR and FT-NIR spectrometers in the evaluation of acerola fruit quality, using PLS and SVM regression algorithms. Talanta. 165, 112-116. http://dx.doi.org/10.1016/j.talanta.2016.12.035

Mancini, M., et al., 2018. Near infrared spectroscopy for the discrimination between different residues of the wood processing industry in the pellet sector. Fuel. 217, 650-655. http://dx.doi.org/10.1016/j.fuel.2018.01.008

Mehmood, T., et al., 2012. A review of variable selection methods in partial least squares regression. Chemometr. Intell. Lab. Syst. 118, 62-69. http://dx.doi.org/10.1016/j.chemolab.2012.07.010

Mevik, B., et al., 2015. Partial Least Squares and Principal Component Regression. R package version 2.5-0. In.

555 Millard, P., 1989. Effect of nitrogen supply on growth and internal nitrogen cycling within
556 deciduous trees. In, Annales des Sciences forestières. EDP Sciences, pp. 666s-668s.
557 Millard, P., 1996. Ecophysiology of the internal cycling of nitrogen for tree growth. J. Plant.
558 Nutr. Soil. Sci. 159, 1-10. http://dx.doi.org/10.1002/jpln.1996.3581590102
559 Millard, P., Neilsen, G., 1989. The influence of nitrogen supply on the uptake and
560 remobilization of stored N for the seasonal growth of apple trees. Ann. Bot. 63, 301-309.
561 http://dx.doi.org/10.1093/oxfordjournals.aob.a087746
562 Millard, P., Proe, M.F., 1993. Nitrogen uptake, partitioning and internal cycling in *Picea*
563 *sitchensis* (Bong.) Carr. as influenced by nitrogen supply. New. Phytol. 125, 113-119.
564 http://dx.doi.org/10.1111/j.1469-8137.1993.tb03869.x
565 Min, M., et al., 2006. Nondestructive detection of nitrogen in Chinese cabbage leaves using
566 VIS–NIR spectroscopy. HortScience. 41, 162-166.
567 http://dx.doi.org/10.21273/HORTSCI.41.1.162
568 Moreau, B., et al., 2004. Estimating leaf nitrogen of eastern cottonwood trees with a
569 chlorophyll meter. Gen Tech Rep SRS-71. US Department of Agriculture, Forest Service,
570 Southern Research Station, Asheville, NC, 487-491.
571 Naing, A.H., et al., 2017. Overexpression of snapdragon Delila (Del) gene in tobacco enhances
572 anthocyanin accumulation and abiotic stress tolerance. BMC. Plant. Biol. 17, 65.
573 http://dx.doi.org/10.1186/s12870-017-1015-5
574 Nambiar, E.K.S., Fife, D.N., 1991. Nutrient retranslocation in temperate conifers. Tree. Physiol.
575 9, 185-207. http://dx.doi.org/10.1093/treephys/9.1-2.185
576 Padilla, F.M., et al., 2014. Evaluation of optical sensor measurements of canopy reflectance
577 and of leaf flavonols and chlorophyll contents to assess crop nitrogen status of muskmelon.
578 Eur. J. Agron. 58, 39-52. http://dx.doi.org/10.1016/j.eja.2014.04.006

579 Park, S., et al., 2018. Rapid Prediction of the Chemical Information of Wood Powder from
580 Softwood Species Using Near-Infrared Spectroscopy. Bioresources. 13, 2440-2451.
581 http://dx.doi.org/10.15376/biores.13.2.2440-2451
582 Paul, V., et al., 2017. Estimation of chlorophylls/photosynthetic pigments–their stability is an
583 Indicator of crop plant tolerance to abiotic stresses. Manual of ICAR Sponsored Training
584 Programme for Technical Staff of ICAR Institutes on "Physiological Techniques to Analyze
585 the Impact of Climate Change on Crop Plants", 8.
586 Payne, R.J., et al., 2017. Nitrogen deposition and plant biodiversity: past, present, and future.
587 Front. Ecol. Environ. 15, 431-436. http://dx.doi.org/10.1002/fee.1528
588 Press, W.H., Teukolsky, S.A., 1990. Savitzky-Golay smoothing filters. Comput. Phys. 4, 669-
589 672. http://dx.doi.org/10.1063/1.4822961
590 Quentin, A.G., et al., 2017. Application of near-infrared spectroscopy for estimation of non-
591 structural carbohydrates in foliar samples of *Eucalyptus globulus* Labilladière. Tree. Physiol.
592 37, 131-141.
593 R Core Team, 2017. R: A Language and Environment for Statistical Computing. In, R
594 Foundation for Statistical Computing, Vienna, Austria.
595 Ramirez, J.A., et al., 2015. Near-infrared spectroscopy (NIRS) predicts non-structural
596 carbohydrate concentrations in different tissue types of a broad range of tree species. Methods.
597 Ecol. Evol. 6, 1018-1025. http://dx.doi.org/10.1111/2041-210X.12391
598 Rivero-Villar, A., et al., 2018. Differences in nitrogen cycling between tropical dry forests with
599 contrasting precipitation revealed by stable isotopes of nitrogen in plants and soils. Biotropica.
600 50, 859-867. http://dx.doi.org/10.1111/btp.12612
601 Robinson, K.M., et al., 2012. Genetic variation in functional traits influences arthropod
602 community composition in aspen (*Populus tremula* L.). PLoS One 7.
603 http://dx.doi.org/10.1371/journal.pone.0037679

604 Schneider, S., et al., 1996. Soluble N compounds in trees exposed to high loads of N: a
605 comparison of spruce (*Picea abies*) and beech (*Fagus sylvatica*) grown under field conditions.
606 New. Phytol. 134, 103-114. http://dx.doi.org/10.1111/j.1469-8137.1996.tb01150.x
607 Siesler, H.W., et al., 2008. Near-infrared spectroscopy: principles, instruments, applications.
608 John Wiley & Sons. http://dx.doi.org/10.1002/9783527612666
609 Sinclair, T., Horie, T., 1989. Leaf nitrogen, photosynthesis, and crop radiation use efficiency:
610 a review. Crop. Sci. 29, 90-98.
611 http://dx.doi.org/10.2135/cropsci1989.0011183X002900010023x
612 Srinivasan, A., 2006. Handbook of precision agriculture: principles and applications. CRC
613 press. http://dx.doi.org/10.1201/9781482277968
614 Stevens, A., Ramirez–Lopez, L., 2014. An introduction to the prospectr package. R Package
615 Vignette, Report No.: R Package Version 0.1 3.
616 Strack, D., Wray, V., 1989. Anthocyanins. In, Methods in plant biochemistry. Elsevier, pp.
617 325-356.
618 Swarts, N., 2016. Precision fertigation for improved apple orchard productivity. In, 2016
619 Future Orchards National Tour, pp. 1-25.
620 Tran, T.N., et al., 2014. Interpretation of variable importance in partial least squares with
621 significance multivariate correlation (sMC). Chemometr. Intell. Lab. Syst. 138, 153-160.
622 http://dx.doi.org/10.1016/j.chemolab.2014.08.005
623 Tremblay, N., et al., 2012. Sensing crop nitrogen status with fluorescence indicators. A review.
624 Agron. Sustain. Dev. 32, 451-464. http://dx.doi.org/10.1007/s13593-011-0041-1
625 Trojak, M., Skowron, E., 2017. Role of anthocyanins in high-light stress response. World. Sci.
626 News. 81, 150-168.

627 Villar-Salvador, P., et al., 2015. The role of stored carbohydrates and nitrogen in the growth
628 and stress tolerance of planted forest trees. New. For. 46, 813-839.
629 http://dx.doi.org/10.1007/s11056-015-9499-z
630 Wellburn, A.R., 1994. The spectral determination of chlorophylls a and b, as well as total
631 carotenoids, using various solvents with spectrophotometers of different resolution. J. Plant.
632 Physiol. 144, 307-313. http://dx.doi.org/10.1016/S0176-1617(11)81192-2
633 Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer.
634 Wijewardane, N.K., et al., 2016. Moisture insensitive prediction of soil properties from VNIR
635 reflectance spectra based on external parameter orthogonalization. Geoderma. 267, 92-101.
636 http://dx.doi.org/10.1016/j.geoderma.2015.12.014
637 Wold, S., et al., 2001. PLS-regression: a basic tool of chemometrics. Chemometr. Intell. Lab.
638 Syst. 58, 109-130. http://dx.doi.org/10.1016/S0169-7439(01)00155-1
639 Wood, C., et al., 1993. Relationships between chlorophyll meter readings and leaf chlorophyll
640 concentration, N status, and crop yield: a review. In, Proceedings of the Agronomy Society of
641 New Zealand, pp. 1-9.
642 Workman Jr, J., Weyer, L., 2012. Practical guide and spectral atlas for interpretive near-
643 infrared spectroscopy. CRC press. http://dx.doi.org/10.1201/b11894
644 Yang, S., et al., 2018. Classification of the hot air heat treatment degree of larch wood using a
645 multivariate analysis of near-infrared spectroscopy. J. Wood. Sci. 64, 220-225.
646 http://dx.doi.org/10.1007/s10086-018-1706-z
647 Yihu, S., et al., 2009. Inheritance of anthocyanin contents in leaves of chili pepper. J. Nanjing.
648 Agricult. Univ.
649 Zhang, Y., et al., 2017. The Contents of Terpene Trilactone and Flavonoid in Leaves of
650 Seedlings from Ancient Female *Ginkgo* Trees in China. Hortic.Plant. J.3, 165-171.
651 http://dx.doi.org/10.1016/j.hpj.2017.06.002

652 Zhao, P., Cao, J., 2016. Wood species identification using spectral reflectance feature and
653 optimal illumination radian design. J. For. Res. 27, 219-224. http://dx.doi.org/10.1007/s11676-
654 015-0171-4
655