Intraspecific genetic variation of a Fagus sylvatica population in a temperate forest derived from airborne imaging spectroscopy time series

Czyz, Ewa A ; Guillén Escribà, Carla ; Wulf, Hendrik ; Tedder, Andrew ; Schuman, Meredith Christine ; Schneider, Fabian D ; Schaepman, Michael E

Abstract: The growing pace of environmental change has increased the need for large-scale monitoring of biodiversity. Declining intraspecific genetic variation is likely a critical factor in biodiversity loss, but is especially difficult to monitor: assessments of genetic variation are commonly based on measuring allele pools, which requires sampling of individuals and extensive sample processing, limiting spatial coverage. Alternatively, imaging spectroscopy data from remote platforms may hold the potential to reveal genetic structure of populations. In this study, we investigated how differences detected in an airborne imaging spectroscopy time series correspond to genetic variation within a population of Fagus sylvatica under natural conditions. We used multi-annual APEX (Airborne Prism Experiment) imaging spectrometer data from a temperate forest located in the Swiss midlands (Laegern, 47°28'N, 8°21'E), along with microsatellite data from F. sylvatica individuals collected at the site. We identified variation in foliar reflectance independent of annual and seasonal changes which we hypothesize is more likely to correspond to stable genetic differences. We established a direct connection between the spectroscopy and genetics data by using partial least squares (PLS) regression to predict the probability of belonging to a genetic cluster from spectral data. We achieved the best genetic structure prediction by using derivatives of reflectance and a subset of wavebands rather than full-analyzed spectra. Our model indicates that spectral regions related to leaf water content, phenols, pigments, and wax composition contribute most to the ability of this approach to predict genetic structure of F. sylvatica population in natural conditions. This study advances the use of airborne imaging spectroscopy to assess tree genetic diversity at canopy level under natural conditions, which could overcome current spatiotemporal limitations on monitoring, understanding, and preventing genetic biodiversity loss imposed by requirements for extensive in situ sampling.

Czyz, Ewa A; Guillén Escribà, Carla; Wulf, Hendrik; Tedder, Andrew; Schuman, Meredith Christine; Schneider, Fabian D; Schaepman, Michael E (2020). Intraspecific genetic variation of a Fagus sylvatica population in a temperate forest derived from airborne imaging spectroscopy time series. Ecology and Evolution, 10(14):7419-7430.

DOI: https://doi.org/10.1002/ece3.6469

Ecology and Evolution
Open Access

WILEY

Intraspecific genetic variation of a *Fagus sylvatica* population in a temperate forest derived from airborne imaging spectroscopy time series

Ewa A. Czyż[1] | Carla Guillén Escribà[1] | Hendrik Wulf[1] | Andrew Tedder[2,3] | Meredith C. Schuman[1,4] | Fabian D. Schneider[1,5] | Michael E. Schaepman[1]

[1]Remote Sensing Laboratories, Department of Geography, University of Zürich, Zürich, Switzerland

[2]Department of Evolutionary Biology and Environmental Studies, University of Zürich, Zürich, Switzerland

[3]School of Chemistry and Biosciences, Faculty of Life Sciences, University of Bradford, Bradford, UK

[4]Department of Chemistry, University of Zürich, Zürich, Switzerland

[5]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

Correspondence
Ewa A. Czyż, Remote Sensing Laboratories, Department of Geography, University of Zürich, Zürich, Switzerland.
Email: ewa.czyz@geo.uzh.ch

Abstract

1. The growing pace of environmental change has increased the need for large-scale monitoring of biodiversity. Declining intraspecific genetic variation is likely a critical factor in biodiversity loss, but is especially difficult to monitor: assessments of genetic variation are commonly based on measuring allele pools, which requires sampling of individuals and extensive sample processing, limiting spatial coverage. Alternatively, imaging spectroscopy data from remote platforms may hold the potential to reveal genetic structure of populations. In this study, we investigated how differences detected in an airborne imaging spectroscopy time series correspond to genetic variation within a population of *Fagus sylvatica* under natural conditions.

2. We used multi-annual APEX (Airborne Prism Experiment) imaging spectrometer data from a temperate forest located in the Swiss midlands (Laegern, 47°28'N, 8°21'E), along with microsatellite data from *F. sylvatica* individuals collected at the site. We identified variation in foliar reflectance independent of annual and seasonal changes which we hypothesize is more likely to correspond to stable genetic differences. We established a direct connection between the spectroscopy and genetics data by using partial least squares (PLS) regression to predict the probability of belonging to a genetic cluster from spectral data.

3. We achieved the best genetic structure prediction by using derivatives of reflectance and a subset of wavebands rather than full-analyzed spectra. Our model indicates that spectral regions related to leaf water content, phenols, pigments, and wax composition contribute most to the ability of this approach to predict genetic structure of *F. sylvatica* population in natural conditions.

4. This study advances the use of airborne imaging spectroscopy to assess tree genetic diversity at canopy level under natural conditions, which could overcome current spatiotemporal limitations on monitoring, understanding, and preventing genetic biodiversity loss imposed by requirements for extensive in situ sampling.

1 | INTRODUCTION

It has long been recognized that declining genetic variation within species is a key factor in biodiversity loss (Wilson & Peter, 1988). Since then several studies have acknowledged that both inter- and intraspecific genetic variation have an important influence on ecosystem structure and functioning (Bolnick et al., 2011; Des Roches et al., 2018; Hughes, Inouye, Johnson, Underwood, & Vellend, 2008). A reduction in the genetic variability of a population increases its susceptibility to diseases (Schmid, 1994), limits its evolutionary potential and reduces the fitness of the next generation (Ellstrand & Antonovics, 1985). Therefore, maintaining genetically variable populations, with their wider potential range of adaptive responses, can be important for conserving biodiversity under changing environmental conditions (Gienapp, Teplitsky, Alho, Mills, & Merilä, 2008; Szathmary, Jordán, & Pál, 2001) and more frequent stochastic climatic events (Hartmann, 2013).

Considering the rapid pace of global climate change and habitat degradation in comparison with timescales of evolutionary processes, the ability of existing populations to adapt to the changes is more important than the potential emergence of new variants (Frankham, 2010). Both, the short-term evolutionary potential and the potential phenotypic plasticity of a population are positively correlated with its allelic variation (Gratani, 2014). Therefore, the number and frequency of alleles changing in time and space is considered to be a suitable measurement of change in genetic diversity (Hoban et al., 2014).

Measurements of the allele pool, together with DNA and RNA sequencing-based techniques (Bruford, Davies, Dulloo, Faith, & Walters, 2017; Yamasaki et al., 2017), provide direct estimates of population genetic composition, which is defined as one of the six essential biodiversity variables (EBV) for monitoring worldwide biodiversity status (Pereira, 2013). These techniques require physical sampling of individuals combined with extraction and analysis of samples (Davies et al., 2012) and are costly and time-consuming. The resulting measurements usually lack the spatial and temporal extent relevant for biodiversity monitoring. Whereas, large-scale monitoring is essential for understanding the drivers of biodiversity change arising from key global change at various spatial and temporal scales (Jetz et al., 2016). Continuous temporal, spatial, and spectral data derived from remote sensing platforms have the potential to overcome scale-induced limitations and are therefore receiving increasing attention for achieving global biodiversity assessments (Navarro et al., 2017; O'Connor et al., 2015; Skidmore et al., 2015; Turner, 2014).

Both passive and active remote sensing technologies have been used to estimate functional, taxonomic, and phylogenetic diversity of plants in a variety of ecosystems. For example, Schneider et al. (2017) evaluated functional diversity in a temperate forest based on vegetation traits interpreted from imaging spectroscopy and light detection and ranging (LiDAR) data. Taxonomic diversity assessment of other sites has been conducted using both trait- and spectra-based approaches (e.g., Asner & Martin, 2009; Martin, Newman, Aber, & Congalton, 1998). However, very few remote sensing studies to date have provided within-species genetic diversity measures at a canopy level under natural conditions.

Like different species, genetically different tree individuals can express different morphological and physiological traits, which shape their reflectance features. These differences are likely to be less pronounced at lower taxonomic ranks than at the species level (Hulshof & Swenson, 2010). Therefore, the recognition of individuals of different genotypes based only on spectral information gained in nonexperimental conditions is limited and not commonly attempted. However, the increasing use of remote sensing data in ecological assessments has led to recognition of the potential for linking genetic with spectral variation. Cavender-Bares et al. (2016) correlated spectral data from on-leaf measurements with genetic clusters as well as species divisions for several species of oak (*Quercus*); and Schweiger et al. (2018) correlated phylogenetic relationships among grassland species with spectral information from both on-leaf and remote (tram-based) measurements. A remote sensing study by Madritch et al. (2014) was able to correlate foliar reflectance with genotype for quaking aspen (*Populus tremuloides*) clones under natural conditions, and furthermore revealed variation in below ground processes. Recent reviews by Bush et al. (2017) and Yamasaki et al. (2017) discussed the potential of remote sensing to reveal genetic composition in natural habitats.

Given the need for novel approaches to detect intraspecific genetic variation and the potential of using imaging spectroscopy data (Geijzendorffer et al., 2016; Navarro et al., 2017; Vihervaara et al., 2017), we attempt to identify a direct connection between the spectral and genetic information from individual trees within a temperate forest. We aim to demonstrate that genotype-specific phenotypic features can be detected in spectral reflectance data acquired under natural conditions and at the canopy level. Furthermore, we emphasize that analyzing this spectral information could represent a time- and cost-efficient tool to reveal the genetic composition of forests, repeatedly and at large spatial scales.

We base our study on the hypothesis that the foliar reflectance changes on an annual and seasonal basis, which contrasts the expression of stable, genotype-specific phenotypic features of individual trees that are maintained over years. Based on that, we expect to identify links between remotely sensed predictors (spectral bands) and genetic structure (membership probability to the genetic clusters) which are maintained over multiple years. To test this hypothesis, we use multi-annual airborne imaging spectroscopy data from

a temperate forest in Switzerland along with genetic information derived from microsatellite analyses of individual trees at the study site. Our approach is based on the conditions that high-fidelity spectral measurements are available (Schaepman et al., 2009, 2015), with residual measurement noise (Hueni, Damm, Kneubuehler, Schlapfer, & Schaepman, 2017) lower than the expected genetic variation to be detected; and further, that individual tree crowns (here: dominant *F. sylvatica* trees) can be detected using the spatial resolution of the imaging spectrometer, and that a minimum of 2–3 sunlit crown pixels for each individual can be identified in the airborne data. We established the link between spectral and genetic data by using partial least square (PLS) regression to assess the explanatory power of distinct wavelength regions (between 372 and 2,540 nm) within the solar radiation reflected from the tree canopy. By combining interdisciplinary approaches, we take a step toward using the potential of imaging spectrometry for spatiotemporal biodiversity mapping of genetic variation within species.

2 | MATERIALS AND METHODS

2.1 | Study area

The study area covers 12.6 ha of seminatural temperate mixed forest located on the Laegern mountain on the northern boundary of the Swiss Plateau (47°28'N, 8°21'E) (Figure 1). The climate of this region is characterized by a mean annual temperature of 7.4°C and mean annual precipitation of 1,000 mm (Etzold et al., 2011). Our study site is located on an up-to-60°-steep south-facing slope with an elevation range of 620 to 810 m a.s.l (Guillén Escribà et al.,

in press). According to the United Nations Environment World Conservation Monitoring Centre (UNEP-WCMC), the vegetation cover is classified as Temperate Deciduous Broadleaf Forest, with 13 tree species consisting of 3 conifers and 10 angiosperms; the European common beech (*Fagus sylvatica)* is the dominant species. *F. sylvatica* trees are wind-pollinated, monoecious plants. Tree age spans between 53 and 185 years with a mean height of 30.6 m and a diameter at breast height of up to 150 cm (Eugster et al., 2007). This composition creates a complex vertical structure of the mainly closed canopy (Schneider et al., 2017). The study area is located in an unmanaged part of the forest and has been a forest ecosystem research site for the last four decades (Kloeti, Keller, & Guecheva, 1989). Individual trees in the study area have been reconstructed in 3D using ground and airborne laser scanning, and modeled using 3D radiative transfer models (Schneider et al., 2014; Schneider et al. 2017), allowing to model and validate airborne data with high accuracy.

2.2 | Genetic data

Microsatellite genotyping was conducted on genomic DNA samples from 77 dominant *F. sylvatica* individuals located in the study area. The individuals were georeferenced using a tachymeter in April 2013. The tachymeter measurements were tied to reference points of the Swiss national grid (LV95). We used a polygon traverse and additional fixed points to measure all tree locations in the challenging terrain (Leiterer, Furrer, Schaepman, & Morsdorf, 2015). Tree crowns and tree trunks were mapped separately and linked to each other using photogrammetric approaches, due to slope geometries
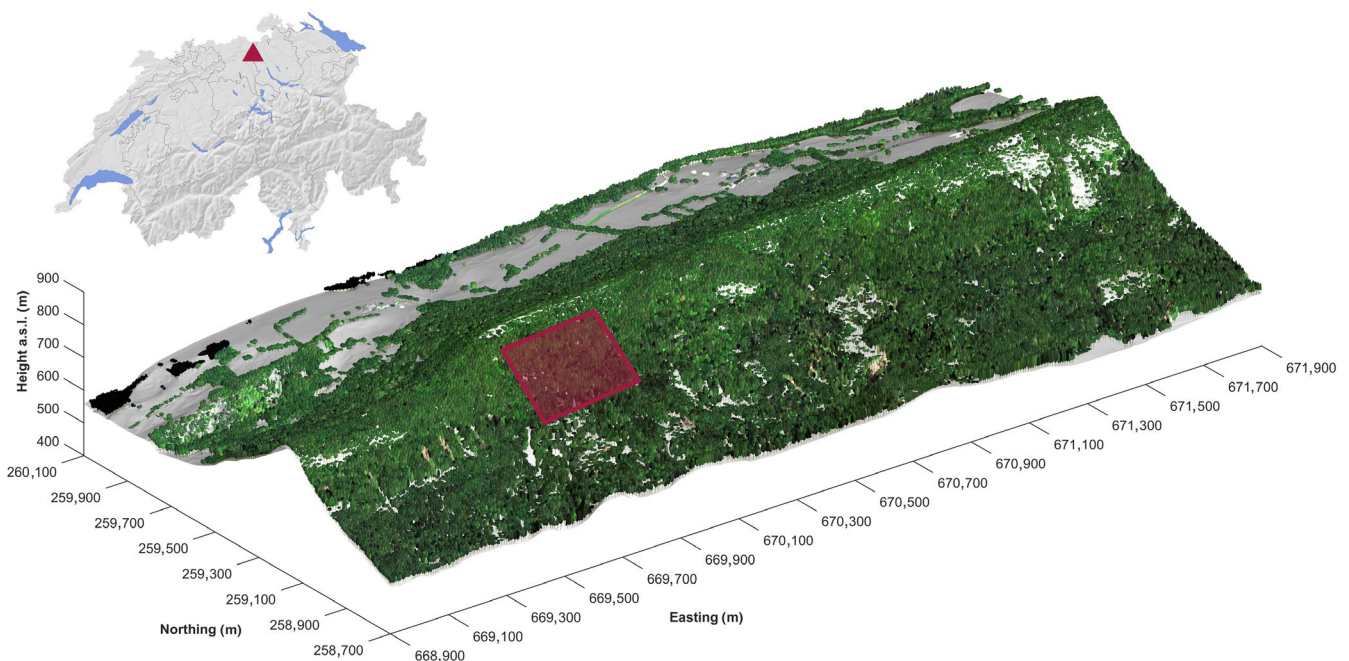
FIGURE 1 Location of the study site (rectangle). The temperate forest is located in northern Switzerland (triangle) on a portion of the south-facing slope of the Laegern mountain

distorting nadir projection of crowns to trunks (Guillén Escribà et al. in press; Torabzadeh, Leiterer, Hueni, Schaepman, & Morsdorf, 2019). Genomic DNA was extracted from leaf disks (diameter 1.15 cm) sampled from each tree in September 2013. The collected material was stored on silica gel. The DNA from the sampled material was extracted using the cetyl trimethylammonium bromide (CTAB) method following the procedure of Doyle and Doyle (1990). From the extracted DNA, five highly variable microsatellite loci (FS1-03, FS1-15, FS3-04, FS4-46, FCM5; Pastorelli, et al., 2003) were amplified using Polymerase Chain Reaction (PCR). To assess the degree of polymorphism at each microsatellite locus, capillary electrophoresis was performed on an ABI-3720 sequencer (Thermo Fisher Scientific, UK) and GeneMapper software was used to determine the length of the analyzed microsatellites for each sampled tree (Table S1). Population structure was inferred using a model-based Bayesian clustering approach implemented in the software TESS2 (Durand, Chen, & François, 2009). Twenty independent runs were performed for K, genetic clusters; the total number of sweeps and burns was 1,200 and 200, respectively, and the degree of trend was set to linear. Based on an average cross-entropy of the admixture model, we determined five genetic clusters, to which each tree was assigned its membership probability (Figure 2).

2.3 | Optical data

The spectral dataset contains seven acquisitions of the Airborne Prism Experiment (APEX) Airborne Imaging Spectrometer (AIS) (Schaepman et al., 2015) acquired between 2009 and 2016. To compare trees from similar development stages, images from the range of 300 – 700 cumulative growing degree days (CGDD) (corresponding to June - July) within each year were selected (Table S2). The raw APEX data preprocessing chain including calibration and correction for spectral shifts with smile effect were performed in the APEX Processing and Archiving Facility and the ATCOR smile module, respectively (Hueni et al., 2012; Hueni et al., 2009; Richter, Schlapfer, & Muller, 2011). Calibrated and corrected radiance data were atmospherically corrected to surface reflectances in ATCOR (Hueni et al., 2017; Schläpfer & Richter, 2002) resulting in imaging spectroscopy datasets of 284 spectral bands, each in the range of 372–2540 nm with 2 m spatial resolution. The wavebands most affected by atmospheric absorption features were interpolated due to high noise level. Each dataset was vicariously calibrated with repeated field spectroradiometer (ASD FieldSpec 4, Boulder, CO, USA) measurements of ground targets to ensure consistent data quality standards and cross-compatibility of the datasets.

We geometrically coregistered the multi-temporal APEX datasets based on a spectrally distinct landmark (i.e., Laegern flux tower). As multitemporal pixels never align perfectly, we used the spectral angle mapper (SAM) method (Kruse et al., 1993) to fill a canopy projected rectangular grid with residual abundances of the SAM to generate reflectance data within one sampling grid. Furthermore, we masked shadows as they created high spatial heterogeneity, and are affected by a lower signal-to-noise ratio in high spatial resolution data (Nagendra & Rocchini, 2008; Stickler & Southworth, 2008). Thus, the pixels with a cumulative reflectance less than the 30th percentile of the entire dataset were excluded from the analysis. We homogenized the tree-specific spectral signature by adopting object-based rather than pixel-based analyses (Karl & Maurer, 2010). The object was defined as a tree crown for each sampled *F. sylvatica* individual; individuals were identified using images derived from light detection and ranging (LiDAR) and high-resolution drone measurements resulting in a crown map of the site (Torabzadeh et al., 2019; Guillén-Escribà et al. in press). The spectral signature for each *F. sylvatica* individual was a mean of sunlit pixels' reflectance from the delineated crown. On average, 17 ± 9 pixels (mean ± standard deviation) were averaged per crown.

Using the annual APEX imagery, we based our analysis on the following three datasets: (a) yearly averaged reflectance data per tree crown, (b) z-score of "a," and (c) the 1st derivative of "b." The z-score of the mean tree crown reflectance was calculated separately for each spectral band and independently for each year with the formula: $(x_l − \mu_l)/\sigma_1$, where x is a value of a single spectral band from each *F. sylvatica* individual, μ and σ are respectively mean and standard deviation values of a single spectral band from the whole analyzed *F. sylvatica* population and l stands for spectral band. Z-score of reflectance (dataset b) and their 1st derivative (dataset c) were included in the analysis to reduce the impact of multi-temporal variation in the reflectance magnitudes and simultaneously emphasize the relative differences between reflectance and absorption/transmittance influenced by structure, water content, and organic compounds (Huesca, García, Roth, Casas, & Ustin, 2016).
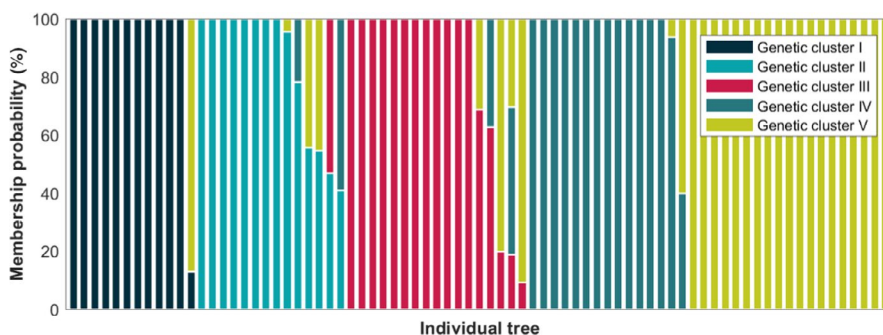
FIGURE 2 Membership probabilities for 77 sampled *Fagus sylvatica* individuals in five genetic clusters as determined by microsatellite analyses

2.4 | Statistical analysis

We investigated the relationship between imaging spectroscopy and genetic information by using partial least squares (PLS) regression. This method is commonly used in chemometrics (Wold, 2001) and more generally to analyze datasets that are highly collinear and have a high ratio of independent variables to observations (Wold, Ruhe, Wold, & Dunn, 1984). In comparison with multiple linear regression (MLR), the latent variables of PLS regression are generated not only based on the best explanation of dependent and independent variables, but also with respect to the relationship between them (Wold, Eriksson, Trygg, & Kettaneh, 2004). The method is used in analyzing spectral data for classification and prediction purposes (Cavender-Bares et al., 2016; Lee et al., 2016; Peerbhay, Mutanga, & Ismail, 2013; Singh, Serbin, McNeil, Kingdon, & Townsend, 2015). In this study, for each sampled *F. sylvatica* individual, we predicted the probability of membership in one of five recognized genetic clusters rather than assigning membership, because the total genetic distance among trees in this population is relatively small. Thus, we conducted a PLS regression rather than PLS discriminant analysis (PLS-DA).

For each *F. sylvatica* individual, we predicted membership probability to five identified genetic clusters by using the spectral fingerprint (284 spectral bands) as predictors. Due to the relationship between genetic clusters of the analyzed population, we used one model to predict membership probability in all five genetic clusters, rather than constructing separate models for each of the genetic clusters. We predicted genetic information of each analyzed tree with leave-one-out cross-validation and adjusted the number of components based on the PRESS statistic (Chen, Hong, Harris, & Sharkey, 2004).

Firstly, we developed PLS models from all spectral predictors for each year separately and we calculated the variable importance of projection (VIP), by waveband. The VIP score corresponds to the contribution of each waveband in the model prediction and is dependent on variance of wavebands and the variance of genetic structure of the trees, as well as the variances between the two (Wold, 2001). We averaged derived VIP scores over 7 years.

Subsequently, we segregated all the wavebands by their importance in genetic cluster prediction. To do so, we identified local maxima of the averaged VIP scores over years and ranked them in descending order based on their prominences. We refer to prominences as the measure of how much the local maxima stands out due to its intrinsic magnitude and its location relative to other local maxima. The threshold for local maxima identification was set to 10^{-5} minimum vertical distance. We estimated the prominences using the *findpeaks* function in MATLAB (MATLAB ver. R2017b). The VIP scores that were not identified as local maxima were assumed to have lower importance than any local maxima and were assigned a ranking based on their absolute magnitude.

Afterward, we generated N models out of n spectral predictors selected based on the nth most prominent VIP scores, where n is the number of spectral variables included for each model, increasing from 1 to the total number of predictors (N). We calculated the root-mean-square error (RMSE) for each model. We developed individual models for each year separately in order to retain intra-annual variation. We then averaged the RMSE over all seven years, which we expect should reduce the influence of phenotypic plasticity, and thus increase the influence of stable genetic variation on differences calculated from spectra.

We conducted the procedure for each of the signal transformations (datasets a, b, c under "optical data") separately and used the *plsregress* function in MATLAB (MATLAB ver. R2017b) for model development.

We expect that the model with the lowest RMSE is constructed based on spectral information that is the most relevant to the genotype-specific phenotypic features maintained throughout the seven years.
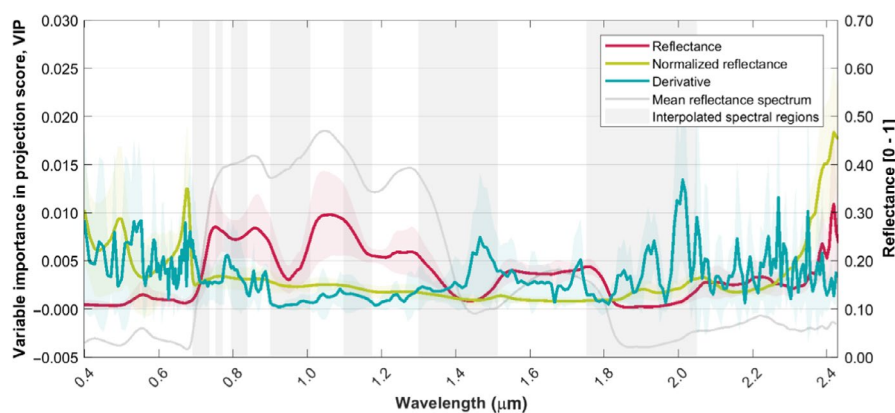
FIGURE 3 VIP scores of PLS models generated to predict genetic structure of *Fagus sylvatica* population in a temperate forest from airborne imaging spectroscopy data. The red, green, and blue solid lines represent the multi-year mean scores achieved from analyses of reflectance (reflectance), z-score of reflectance (normalized reflectance) and a 1st derivative of z-score of reflectance (derivative) signal transformations, respectively. The shaded areas represent the standard deviation of the scores from seven years of data acquired between 2009 and 2016. The gray bars indicate bands that have been interpolated during data processing due to high noise levels in atmospheric absorption bands, and the gray solid line represents the mean reflectance from the analyzed trees
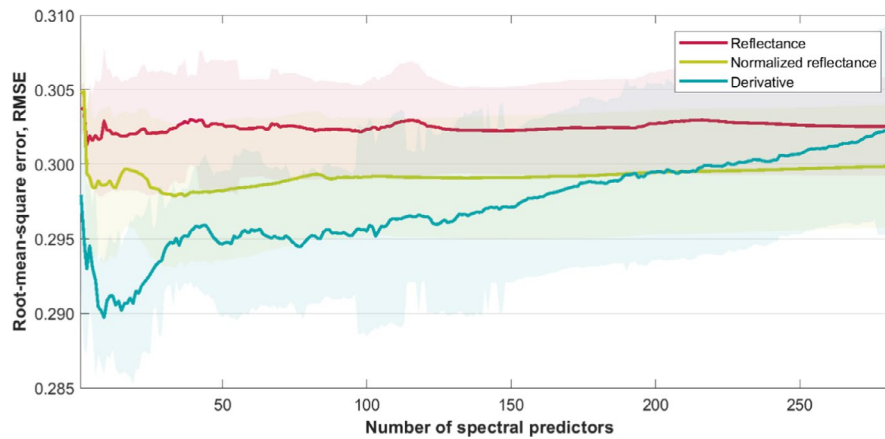
FIGURE 4 Root-mean-square error (RMSE) of genetic structure prediction from PLS models generated from an increasing number of most prominent (defined under "statistical analysis") spectral predictors derived based on the VIP score of the models. The red, green, and blue solid lines represent multi-year mean RMSE achieved from analyses made on reflectance (reflectance), z-score of reflectance (normalized reflectance), and a 1st derivative of z-score of reflectance (derivative) signal transformations, respectively. The shaded areas represent the standard deviation of RMSE calculations derived from data acquired between 2009 and 2016

3 | RESULTS

3.1 | VIP scores

The VIP score reveals the relevant contribution of different wavelength regions for predicting the genetic structure of sampled *F. sylvatica* individuals over the full-analyzed solar spectrum (Figure 3).

In the reflectance dataset (dataset a, under "optical data"), the near-infrared (0.75–1.4 µm) region of the spectrum is the most influential on predicting genetic structure. Over the full-analyzed reflectance spectrum, the VIP score is positively correlated to the absolute reflectance (Pearson coefficient: .84, *p* < .001). Wavelengths with high VIP scores in this dataset are also characterized by higher standard deviation of interannual data than those with lower reflectance (Pearson coefficient: .86, *p* < .001).

In the normalized reflectance dataset (dataset b), the highest VIP score in predicting the genetic structure was identified for spectral features around 0.48, 0.70, and 2.40 µm. In contrast to the reflectance- and derivative-based analyses, the VIP scores do not vary significantly between 0.70 and 2.30 µm and are not strongly correlated to the absolute reflectance (Pearson's coefficient: −.29, *p* < .001).

In the derivative dataset (dataset c), the spectral regions that show a closer relation to the genetic structure of analyzed trees are located at 0.55, 0.68, 1.45, 2.00, and 2.27 µm. In comparison with the VIP scores derived from reflectance and normalized reflectance datasets, the derivative dataset shows significant variation in VIP scores throughout the solar spectrum and exhibits a moderate negative correlation to the absolute reflectance (Pearson's coefficient: −.50, *p* < .001).

### 3.2 | Root-mean-square error in prediction of the genetic structure

Prediction of genetic structure from the full-analyzed spectrum resulted in higher mean RMSE than the analyses on a subset of spectral bands in all the signal transformation datasets (Figure 4). The RMSE of membership probability to the five detected genetic clusters averaged over years, trees, and genetic clusters over the full-analyzed spectrum were 0.3025, 0.2998, and 0.3024 for reflectance (datasets a)-, normalized reflectance (datasets b)-, and derivative (datasets c)-based analyses, respectively.

In the reflectance-based analyses, the use of three spectral bands with the highest prominence of the VIP scores (under "statistical analysis") resulted in RMSE of genetic structure prediction of 0.3013. Compared with the full spectrum-based analyses, the prediction improvement is of 0.4%. Likewise, the use of 37 spectral bands in the normalized reflectance dataset (RMSE: 0.2978) reduced RMSE by 0.7% in comparison with the model prediction based on all spectral bands. The best membership prediction was achieved for a subset of nine predictors derived from the derivative-based signal (RMSE: 0.2897). The prediction of the model constructed from those nine variables resulted in an improvement of 4.2% in comparison with the model based on the full spectrum. In the reflectance and normalized reflectance datasets, the use of the single most prominent band did not improve the prediction. However, in the derivative dataset, the use of the single most informative predictor selected based on the VIP score, rather than the full spectrum, resulted in a 1.5% improvement in genetic structure prediction (RMSE: 0.2979). The spectral indicator subsets of the datasets that result in the best-performing models are presented in Figure 5.

Prediction of genetic structure does not change significantly with the use of more than 120 and 90 spectral bands in reflectance and normalized reflectance datasets, respectively. In contrast, the addition of further information (i.e., additional spectral variables) in the derivative dataset has an influence on the PLS model performance. In addition to the spectral predictors that construct the models resulting in lowest RMSE, incorporation of predictors derived from 1.74, 2.24, and 2.33 µm wavelengths in the derivative and 1.55 µm wavelength in the reflectance dataset substantially reduce the RMSE of the genetic structure prediction (not shown).
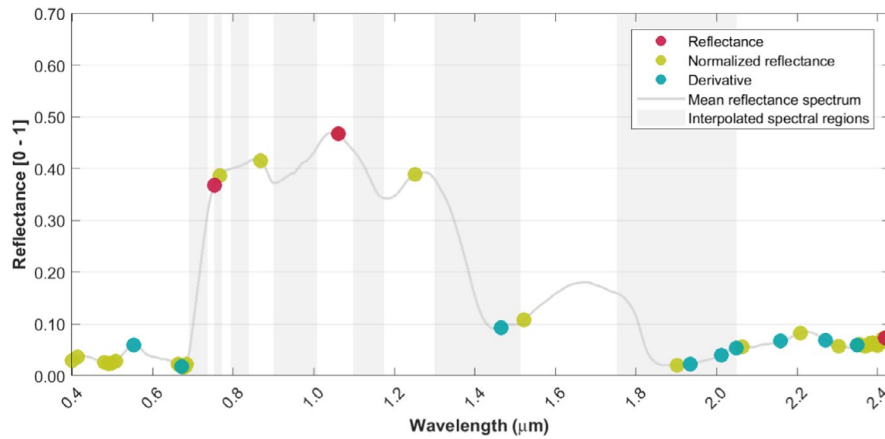
FIGURE 5 Spectral predictors selected based on the highest prominence of VIP scores (defined under "statistical analysis"), from which the genetic structure was predicted with the highest accuracy. The red, green, and blue refers to reflectance (reflectance), *z*-score of reflectance (normalized reflectance), and a 1st derivative of *z*-score of reflectance (derivative) signal transformations, respectively. The gray bars indicate bands that have been interpolated during data processing due to high noise levels in atmospheric absorption bands, and the gray solid line represents the mean reflectance from the analyzed trees

4 | DISCUSSION

4.1 | Spectral subsets

Our results show that the use of spectral subsets rather than the full-analyzed spectrum results in a lower error in the prediction of genetic structure.

In contrast to spectroscopic measurements made in fully controlled environments, aerial measurements under natural conditions are influenced by many factors other than the genotype-specific features. The atmospheric influence (Richter & Schläpfer, 2002) as well as phenotypic plasticity of the trees could mask genetic variation in an electromagnetic signal acquired from airborne spectroscopy, if not correctly interpreted. Reduction of nongenetically relevant variation would result in better prediction of genetic structure. Accordingly, the use of 1%, 13%, and 3% of spectral information from the reflectance, normalized reflectance, and derivative datasets, respectively, yields the best prediction of the genetic structure in our study. Furthermore, accuracy in predicting genetic structure does not change substantially after incorporating more than 42% of reflectance- and 32% of normalized reflectance-derived variables.

Despite the relatively small differences in the prediction of genetic structure for *F. sylvatica* from different spectral subsets, we present evidence that the electromagnetic spectrum of the canopy surface contains information on intraspecific genetic diversity. This conclusion is supported by the analyses of different wavelength regions having varying performance in prediction of genetic structure. The improved performance of the spectral subset-based approach is consistent with a study by Cavender-Bares et al. (2016), in which the bands with highest VIP scores were selected for analyses. In our study, we additionally show that the results could be improved by choosing not only the highest but also the most prominent VIP scores (defined under "statistical analysis") derived over the spectrum (Figure S1).

Our method also reveals the spectral regions that are most appropriate for genetic structure prediction, reflecting genotype-specific phenotypic responses, which are most consistent across years. Those regions vary in each signal dataset. In all the datasets, the most influential wavelengths are those in the short-wavelength infrared (1.40–2.54 μm) and specifically those influenced by tissue water content (Gausman & Weidner, 1985; Tucker, 1980). The analyses may also indicate the importance of wavelengths connected with C-H bond absorption of phenolic compounds located in this region of the spectrum (Kokaly, Asner, Ollinger, Martin, & Wessman, 2009). Additionally, in the normalized and derivative datasets, the visible region of the spectrum (0.37–0.75 μm), which is influenced by pigment amount and composition (Ustin et al., 2009) as well as epicuticular wax content (Petibon F., unpublished data), is among the most informative about genetic structure.

Our results suggest that there is intraspecific variation in water balance as well as the composition and content of phenols, pigments, and waxes for the analyzed population of *F. sylvatica*, which can be elucidated remotely using aerial imaging spectroscopy. This outcome is consistent with genetic and physiological studies. *F. sylvatica* individuals originating from different populations have provenance-specific genetic backgrounds (Demesure, Comps, & Petit, 1996) and differences in water management (Peuke, Schraml, Hartung, & Rennenberg, 2002). Further, phenolic compounds are related to many physiological reactions of plants, including protection against UV radiation (Close & McArthur, 2002), microbial (Scalbert, 1991), fungal (Telles, Kupski, & Furlong, 2017), and herbivorous (War et al., 2012) attackers, as well as pollution (Pasqualini et al., 2003) and climatic responses (Stark, 2015). The variation in the abundance of phenols is in part under genetic control (Pereira, 2007) and is already used for remote taxonomic identification of trees (Asner, Martin, & Suhaili, 2012), and within-species assessments. Additionally, pigments influence photosynthetic performance (Flexas, Ribas-Carbó,, Diaz-Espejo,

Galmes, & Medrano, 2008; Flexas et al., 2012). This performance is also related to the physiological adaptation of the organism and therefore is subject to purifying selection (Arntz & Delph, 2001). Similarly, epicuticular wax content responds to environmental conditions (e.g., Schreiber, Kirsch, & Riederer, 1996) and may be genetically constrained. Correlation of the spectral features identified as being most informative about genetic structure with variation in abiotic and biotic factors may be a first step to elucidate the influence of genotype by environment interactions.

It should be noted that, due to the uncertainty of spectral measurements below 0.45 μm and above 2.20 μm, and in the spectral regions where the signal was interpolated during data processing due to high noise levels in atmospheric absorption bands (indicated by gray shading in figures), the outcomes related to those specific wavelengths should be interpreted with caution.

4.2 | Signal transformation

The difference in outcome from analyses made on various forms of the electromagnetic signal indicate the importance of signal transformation in multi-temporal data analyses. Our analyses indicate that the transformation of the spectrum may be particularly important in analyzing fine-scale characteristics like intraspecific genetic variation.

The strong correlation of wavelength importance and the magnitude of the nontransformed signal could be caused by statistical limitations of the PLS regression method. The importance of specific wavelength regions might be assigned by overbalance of variation within one wavelength over fine genetically relevant information recorded at that particular wavelength. This could be supported by the observed reduction of importance of wavelengths with high magnitude in the normalized reflectance dataset, so in the signal where interannual variation caused by system instability is reduced. However, normalization may remove many sources of spectral variation, including the variation among spectral features of genetic clusters. This problem is remedied in the derivative dataset, where differences among individual measurements are emphasized and the intensity is not a major source of variation across acquisitions. Additionally, each predicting variable in the derivative dataset is an outcome of two neighboring spectral bands and thus represents more information, but also information that is averaged over a wider spectral window. The reduction of spectral resolution from high spectral resolution data results in higher signal-to-noise ratios (Karl & Maurer, 2010; Nagendra & Rocchini, 2008; Stickler & Southworth, 2008) and therefore may expose the spectral features for which genotype-specific variation is greater than variation of the system acquisition.

Indeed, the derivative dataset over multi-year analyses performed best in predicting genetic structure of the *F. sylvatica* study population, and thus, we expect this dataset conveys relatively the most information about intraspecific genetic variation derived from time series.

4.3 | Limitations and outlook

It should be emphasized that the approach we used takes advantage of relatively high temporal, spectral and spatial resolution of available airborne data.

Using a dataset covering seven years, we have been able to reduce the influence of phenotypic responses related to interannual and phenological variability and focus on differences which remain stable across time and may be genotype-specific. However, there are also likely to be genotype-specific responses which are more distinct on an intra-annual or intradaily basis. For example, investigations over the phenological cycle of plant communities (e.g., Merton, 1998) and plant species (e.g., Somers & Asner, 2014) improved genetic cluster classification. Beyond the differences in phenological responses, physiological responses over the day could be indicative of genetic differences in a population (Gallé & Feller, 2007). Therefore, we expect that multi-seasonal and daily resolved analyses may improve the prediction of intraspecific genetic structure from spectral information.

Using other datasets lacking spectrally detailed information may not be sufficient to resolve genetic variation. At the same time, the spectral resolution we used is most certainly not sufficient to identify all genetically relevant spectral features. Next steps might be to use finer genetic resolution and to investigate various populations, anticipating a wider genetic pool that potentially expresses a larger distribution of phenotypic features accessible with the spectral resolution of current airborne sensors. Additionally, incorporation of other remotely sensed data (e.g., LiDAR (Torabzadeh et al., 2019; Valbuena et al. 2020), thermal (Ullah, Schlerf, Skidmore, & Hecker, 2012), or solar-induced chlorophyll fluorescence (SIF) (Keller et al., 2019)) can provide means to detect additional phenotypic features indicative of genetic differences.

The 2 m spatial resolution of the airborne data and the prior mapping of the Laegern experimental forest using LiDAR data allowed us to work on canopy level in natural conditions. By incorporating structural information from ground-based LiDAR data (Morsdorf, Kükenbrink, Schneider, Abegg, & Schaepman, 2018; Schneider, Kükenbrink, Schaepman, Schimel, & Morsdorf, 2019) as well as airborne laser scanning (Kaartinen, 2012; Schneider et al., 2014; Wang, 2016), we have been able to delineate species crowns of the area and link spectral information to *F. sylvatica* individuals. Hence, this methodological approach is expandable to other species, where individuals have a certain probability of belonging to a genetic cluster within the detection sensitivity of the imaging instrument used. RS-based studies of genetic composition, where spatial resolution is not sufficient to recognize individual trees, might be limited to cases where the genetics of the area are highly homogenous (c.f., Madritch et al., 2014).

It should be noted that our results represent a case study limited to location and species. The available genetic information of the population was not sufficiently large to perform a spatial classification to distinct genetic clusters with high confidence using the selected genetic markers. However, we relate differences among

spectral fingerprints of individuals in a population of *F. sylvatica* to genetic differences among those individuals, and we identify spectral regions and signal transformation informative about genetic structure. Using a dataset with relatively high temporal, spectral, and spatial information, we demonstrate the potential of linking spectral and genetic information. This may represent a step toward developing universal models to identify genetic variation from spectral data.

5 | CONCLUSIONS

This study highlights the potential for multi-temporal imaging spectroscopy data to detect intraspecific genetic variation of trees in temperate forests. We investigated the use of derivative-based analyses and PLS-based methods to overcome discrimination challenges caused by multi-factorial influences on the spectral canopy signature acquired under natural conditions and at various time steps. We demonstrate that, using this approach, intraspecific genetic diversity can best be assessed using spectral subsets, rather than the full spectrum influenced by various sources. Moreover, our method successfully detected the spectral regions most indicative for genetic structure prediction, without introducing prior knowledge.

In this study, we contribute to resources available for further studies focusing on genetic diversity using remote sensing techniques. Accordingly, our results suggest that this kind of analysis, where the genetic resolution is low enough for remote detection and high enough for practical purposes, is a promising tool for tracing the landscape of genetic variation (Rocchini et al., 2010) in a direct, efficient and globally consistent way.

CONFLICT OF INTEREST
The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS
Ewa A. Czyż: Conceptualization (equal); Data curation (lead); Formal analysis (lead); Investigation (lead); Methodology (lead); Visualization (lead); Writing-original draft (lead); Writing-review & editing (lead). **Carla Guillén Escribà:** Conceptualization (equal); Data curation (lead); Investigation (equal); Methodology (equal); Supervision (equal); Writing-original draft (equal). **Hendrik Wulf:** Conceptualization (equal); Data curation (lead); Investigation (equal); Methodology (equal); Supervision (equal); Writing-original draft (equal); Writing-review & editing (equal). **Andrew Tedder:** Data curation (equal); Formal analysis (supporting); Investigation (supporting); Methodology (supporting); Writing-review & editing (supporting). **Meredith C. Schuman:** Writing-original draft (equal); Writing-review & editing (equal). **Fabian D. Schneider:** Data curation (equal); Writing-review & editing (equal). **Michael E. Schaepman:** Conceptualization (lead); Project administration (lead); Resources (lead); Supervision (equal); Writing-original draft (equal).

ORCID
Ewa A. Czyż https://orcid.org/0000-0002-4505-0701
Carla Guillén Escribà https://orcid.org/0000-0002-4075-7177
Andrew Tedder https://orcid.org/0000-0002-7378-4673
Michael E. Schaepman https://orcid.org/0000-0002-9627-9565

REFERENCES
Arntz, M. A., & Delph, L. F. D. (2001). Pattern and process: Evidence for the evolution of photosynthetic traits in natural populations. *Oecologia*, *127*(4), 455–467. https://doi.org/10.1007/s004420100650

Asner, G. P., & Martin, R. E. (2009). Airborne spectranomics: Mapping canopy chemical and taxonomic diversity in tropical forests. *Frontiers in Ecology and the Environment*, *7*(5), 269–276. https://doi.org/10.1890/070152

Asner, G. P., Martin, R. E., & Suhaili, A. B. (2012). Sources of canopy chemical and spectral diversity in lowland Bornean forest. *Ecosystems*, *15*(3), 504–517. https://doi.org/10.1007/s10021-012-9526-2

Bolnick, D. I., Amarasekare, P., Araújo, M. S., Bürger, R., Levine, J. M., Novak, M., … Vasseur, D. A. (2011). Why intraspecific trait variation matters in community ecology. *Trends in Ecology & Evolution*, *26*(4), 183–192. https://doi.org/10.1016/j.tree.2011.01.009

Bruford, M. W., Davies, N., Dulloo, M. E., Faith, D. P., & Walters, M. (2017). Monitoring changes in genetic diversity. In *The GEO handbook on biodiversity observation networks* (pp. 107–128). Cham, Switzerland: Springer.

Bush, A., Sollmann, R., Wilting, A., Bohmann, K., Cole, B., Balzter, H., … Yu, D. W. (2017). Connecting Earth observation to high-throughput biodiversity data. *Nature Ecology & Evolution*, *1*(7), 176. https://doi.org/10.1038/s41559-017-0176

Cavender-Bares, J., Meireles, J., Couture, J., Kaproth, M., Kingdon, C., Singh, A., … Townsend, P. (2016). Associations of leaf spectra with genetic and phylogenetic variation in oaks: Prospects for remote detection of biodiversity. *Remote Sensing*, *8*(3), 221. https://doi.org/10.3390/rs8030221

Chen, S., Hong, X., Harris, C. J., & Sharkey, P. M. (2004). Sparse modeling using orthogonal forward regression with PRESS statistic and regularization. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *34*(2), 898–911. https://doi.org/10.1109/TSMCB.2003.817107

Close, D. C., & McArthur, C. (2002). Rethinking the role of many plant phenolics–protection from photodamage not herbivores? *Oikos*, *99*(1), 166–172. https://doi.org/10.1034/j.1600-0706.2002.990117.x

Davies, N., Meyer, C., Gilbert, J. A., Amaral-Zettler, L., Deck, J., Bicak, M., … Field, D. (2012). A call for an international network of

genomic observatories (GOs). *GigaScience*, *1*(1), 5. https://doi.org/10.1186/2047-217X-1-5

Demesure, B., Comps, B., & Petit, R. J. (1996). Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. *Evolution*, *50*(6), 2515–2520. https://doi.org/10.1111/j.1558-5646.1996.tb03638.x

Des Roches, S., Post, D. M., Turley, N. E., Bailey, J. K., Hendry, A. P., Kinnison, M. T., ... Palkovacs, E. P. (2018). The ecological importance of intraspecific variation. *Nature Ecology & Evolution*, *2*(1), 57. https://doi.org/10.1038/s41559-017-0402-5

Doyle, J. J., & Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. *Focus*, *12*(13), 39–40.

Durand, E., Chen, C., & François, O. (2009). *Tess version 2.3—reference manual*, August 2009. Available at: memberstimc.imag.fr/Olivier.Francois/tess.html, 1, 30.

Ellstrand, N. C., & Antonovics, J. (1985). Experimental studies of the evolutionary significance of sexual reproduction II. A test of the density-dependent selection hypothesis. *Evolution*, *39*(3), 657–666. https://doi.org/10.1111/j.1558-5646.1985.tb00402.x

Etzold, S., Ruehr, N. K., Zweifel, R., Dobbertin, M., Zingg, A., Pluess, P., ... Buchmann, N. (2011). The carbon balance of two contrasting mountain forest ecosystems in Switzerland: Similar annual trends, but seasonal differences. *Ecosystems*, *14*(8), 1289–1309. https://doi.org/10.1007/s10021-011-9481-3

Eugster, W., Zeyer, K., Zeeman, M., Michna, P., Zingg, A., Buchmann, N., & Emmenegger, L. (2007). Methodical study of nitrous oxide eddy covariance measurements using quantum cascade laser spectrometery over a Swiss forest. *Biogeosciences*, *4*(5), 927–939. https://doi.org/10.5194/bg-4-927-2007

Flexas, J., Ribas-Carbó, M., Diaz-Espejo, A., Galmes, J., & Medrano, H. (2008). Mesophyll conductance to CO_2: Current knowledge and future prospects. *Plant, Cell & Environment*, *31*(5), 602–621. https://doi.org/10.1016/j.plantsci.2012.05.009

Flexas, J., Barbour, M. M., Brendel, O., Cabrera, H. M., Carriquí, M., Díaz-Espejo, A., ... Warren, C. R. (2012). Mesophyll diffusion conductance to CO_2: An unappreciated central player in photosynthesis. *Plant Science*, *193*, 70–84. https://doi.org/10.1016/j.plantsci.2012.05.009

Frankham, R. (2010). Challenges and opportunities of genetic approaches to biological conservation. *Biological Conservation*, *143*(9), 1919–1927. https://doi.org/10.1016/j.biocon.2010.05.011

Gallé, A., & Feller, U. (2007). Changes of photosynthetic traits in beech saplings (*Fagus sylvatica*) under severe drought stress and during recovery. *Physiologia Plantarum*, *131*(3), 412–421. https://doi.org/10.1111/j.1399-3054.2007.00972.x

Gausman, H. W., & Weidner, V. R. (1985). Plant leaf optical properties in visible and near-infrared light. *Applied Optics*, *24*, 3305.

Geijzendorffer, I. R., Regan, E. C., Pereira, H. M., Brotons, L., Brummitt, N., Gavish, Y., ... Walters, M. (2016). Bridging the gap between biodiversity data and policy reporting needs: An Essential Biodiversity Variables perspective. *Journal of Applied Ecology*, *53*(5), 1341–1350. https://doi.org/10.1111/1365-2664.12417

Gienapp, P., Teplitsky, C., Alho, J. S., Mills, J. A., & Merilä, J. (2008). Climate change and evolution: Disentangling environmental and genetic responses. *Molecular Ecology*, *17*(1), 167–178. https://doi.org/10.1111/j.1365-294X.2007.03413.x

Gratani, L. (2014). Plant phenotypic plasticity in response to environmental factors. *Advances in Botany*, *2014*, 1–17. https://doi.org/10.1155/2014/208747

Hartmann, D. L., Tank, A. M. K., Rusticucci, M., Alexander, L. V., Brönnimann, S., Charabi, Y. A. R., ... & Soden, B. J. (2013). Observations: atmosphere and surface. In: *Climate change 2013 the physical science basis: Working group I contribution to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. (pp. 159 –254).

Hoban, S., Arntzen, J. A., Bruford, M. W., Godoy, J. A., Rus Hoelzel, A., Segelbacher, G., ... Bertorelle, G. (2014). Comparative evaluation of potential indicators and temporal sampling protocols for monitoring genetic erosion. *Evolutionary Applications*, *7*(9), 984–998. https://doi.org/10.1111/eva.12197

Hueni, A., Sterckx, S., Jehle, M., D'Odorico, P., Vreys, K., Bomans, B., ... & Schaepman, M. (2012). Operational status of APEX and characteristics of the APEX open science data set. In *Geoscience and remote sensing symposium (IGARSS), 2012 IEEE International* (pp. 5009–5012). IEEE.

Hueni, A., Biesemans, J., Meuleman, K., Dell'Endice, F., Schlapfer, D., Odermatt, D., ... Itten, K. I. (2009). Structure, components, and interfaces of the airborne prism experiment (APEX) processing and archiving facility. *IEEE Transactions on Geoscience and Remote Sensing*, *47*(1), 29–43. https://doi.org/10.1109/TGRS.2008.2005828

Hueni, A., Damm, A., Kneubuehler, M., Schlapfer, D., & Schaepman, M. E. (2017). Field and Airborne Spectroscopy Cross Validation—Some Considerations. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(3), 1117–1135. https://doi.org/10.1109/JSTARS.2016.2593984

Huesca, M., García, M., Roth, K. L., Casas, A., & Ustin, S. L. (2016). Canopy structural attributes derived from AVIRIS imaging spectroscopy data in a mixed broadleaf/conifer forest. *Remote Sensing of Environment*, *182*, 208–226. https://doi.org/10.1016/j.rse.2016.04.020

Hughes, A. R., Inouye, B. D., Johnson, M. T. J., Underwood, N., & Vellend, M. (2008). Ecological consequences of genetic diversity. *Ecology Letters*, *11*(6), 609–623. https://doi.org/10.1111/j.1461-0248.2008.01179.x

Hulshof, C. M., & Swenson, N. G. (2010). Variation in leaf functional trait values within and across individuals and species: An example from a Costa Rican dry forest. *Functional Ecology*, *24*(1), 217–223. https://doi.org/10.1111/j.1365-2435.2009.01614.x

Jetz, W., Cavender-Bares, J., Pavlick, R., Schimel, D., Davis, F. W., Asner, G. P., ... Ustin, S. L. (2016). Monitoring plant functional diversity from space. *Nature Plants*, *2*(3), 1–5. https://doi.org/10.1038/nplants.2016.24

Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., ... Wu, J.-C. (2012). An international comparison of individual tree detection and extraction using airborne laser scanning. *Remote Sensing*, *4*(4), 950–974. https://doi.org/10.3390/rs4040950

Karl, J. W., & Maurer, B. A. (2010). Multivariate correlations between imagery and field measurements across scales: Comparing pixel aggregation and image segmentation. *Landscape Ecology*, *25*(4), 591–605. https://doi.org/10.1007/s10980-009-9439-4

Keller, B., Matsubara, S., Rascher, U., Pieruschka, R., Steier, A., Kraska, T., & Muller, O. (2019). Genotype specific photosynthesis x environment interactions captured by automated fluorescence canopy scans over two fluctuating growing seasons. *Frontiers in Plant Science*, *10*, 1482. https://doi.org/10.3389/fpls.2019.01482

Kloeti, P., Keller, H. M., & Guecheva, M. (1989). Effects of forest canopy on throughfall precipitation chemistry. In *Baltimore Symposium, Baltimore* (Vol. *179*, pp. 203–209).

Kokaly, R. F., Asner, G. P., Ollinger, S. V., Martin, M. E., & Wessman, C. A. (2009). Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. *Remote Sensing of Environment*, *113*, 78–91. https://doi.org/10.1016/j.rse.2008.10.018

Kruse, F. A., Lefkoff, A. B., Boardman, J. W., Heidebrecht, K. B., Shapiro, A. T., Barloon, P. J., & Goetz, A. (1993). The Spectral Image Processing System (SIPS) – Interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, *44*, 145–163. https://doi.org/10.1016/0034-4257(93)90013-N

Lee, J., Cai, X., Lellmann, J., Dalponte, M., Malhi, Y., Butt, N., ... Coomes, D. A. (2016). Individual tree species classification from airborne multisensor imagery using robust PCA. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(6), 2554–2567. https://doi.org/10.1109/JSTARS.2016.2569408

Leiterer, R., Furrer, R., Schaepman, M. E., & Morsdorf, F. (2015). Forest canopy-structure characterization: A data-driven approach. *Forest Ecology and Management*, *358*, 48–61. https://doi.org/10.1016/j.foreco.2015.09.003

Madritch, M. D., Kingdon, C. C., Singh, A., Mock, K. E., Lindroth, R. L., & Townsend, P. A. (2014). Imaging spectroscopy links aspen genotype with below-ground processes at landscape scales. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1643), 20130194. https://doi.org/10.1098/rstb.2013.0194

Martin, M. E., Newman, S. D., Aber, J. D., & Congalton, R. G. (1998). Determining forest species composition using high spectral resolution remote sensing data. *Remote Sensing of Environment*, *65*(3), 249–254. https://doi.org/10.1016/S0034-4257(98)00035-2

Merton, R. N. (1998, January). Monitoring community hysteresis using spectral shift analysis and the red-edge vegetation stress index. In: *Proceedings of the Seventh Annual JPL Airborne Earth Science Workshop*, 12–16.

Morsdorf, F., Kükenbrink, D., Schneider, F. D., Abegg, M., & Schaepman, M. E. (2018). Close-range laser scanning in forests - towards physically based semantics across scales. *Interface Focus*, *8*(2), 20170046. https://doi.org/10.1098/rsfs.2017.0046

Nagendra, H., & Rocchini, D. (2008). High resolution satellite imagery for tropical biodiversity studies: The devil is in the detail. *Biodiversity and Conservation*, *17*(14), 3431. https://doi.org/10.1007/s10531-008-9479-0

Navarro, L. M., Fernández, N., Guerra, C., Guralnick, R., Kissling, W. D., Londoño, M. C., … Pereira, H. M. (2017). Monitoring biodiversity change through effective global coordination. *Current Opinion in Environmental Sustainability*, *29*, 158–169. https://doi.org/10.1016/j.cosust.2018.02.005

O'Connor, B., Secades, C., Penner, J., Sonnenschein, R., Skidmore, A., Burgess, N. D., & Hutton, J. M. (2015). Earth observation as a tool for tracking progress towards the Aichi Biodiversity Targets. *Remote Sensing in Ecology and Conservation*, *1*(1), 19–28. https://doi.org/10.1002/rse2.4

Pasqualini, V., Robles, C., Garzino, S., Greff, S., Bousquet-Mélou, A., & Bonin, G. (2003). Phenolic compounds content in *Pinus halepensis* Mill. needles: A bioindicator of air pollution. *Chemosphere*, *52*(1), 239–248. https://doi.org/10.1016/S0045-6535(03)00268-6

Pastorelli, R., Smulders, M. J. M., Van't Westende, W. P. C., Vosman, B., Giannini, R., Vettori, C., & Vendramin, G. G. (2003). Characterization of microsatellite markers in *Fagus sylvatica* L. and *Fagus orientalis* Lipsky. *Molecular Ecology Resources*, *3*(1), 76–78. https://doi.org/10.1046/j.1471-8286.2003.00355.x

Peerbhay, K. Y., Mutanga, O., & Ismail, R. (2013). Commercial tree species discrimination using airborne AISA Eagle hyperspectral imagery and partial least squares discriminant analysis (PLS-DA) in KwaZulu–Natal, South Africa. *ISPRS Journal of Photogrammetry and Remote Sensing*, *79*, 19–28. https://doi.org/10.1016/j.isprsjprs.2013.01.013

Pereira, H. M., Ferrier, S., Walters, M., Geller, G. N., Jongman, R. H. G., Scholes, R. J., … & Coops, N. C. (2013). Essential biodiversity variables. *Science*, *339*(6117), 277–278. http://doi.org/10.1126/science.1229931

Pereira, J. A., Oliveira, I., Sousa, A., Valentão, P., Andrade, P. B., Ferreira, I. C., … & Estevinho, L. (2007). Walnut (*Juglans regia* L.) leaves: Phenolic compounds, antibacterial activity and antioxidant potential of different cultivars. *Food and Chemical Toxicology*, *45*(11), 2287–2295. https://doi.org/10.1016/j.fct.2007.06.004

Peuke, A. D., Schraml, C., Hartung, W., & Rennenberg, H. (2002). Identification of drought-sensitive beech ecotypes by physiological parameters. *New Phytologist*, *154*(2), 373–387. https://doi.org/10.1046/j.1469-8137.2002.00400.x

Richter, R., & Schläpfer, D. (2002). Geo-atmospheric processing of airborne imaging spectrometry data. Part 2: atmospheric/topographic correction. *International Journal of Remote Sensing*, *23*(13), 2631–2649.

Richter, R., Schlapfer, D., & Muller, A. (2011). Operational atmospheric correction for imaging spectrometers accounting for the smile effect. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(5), 1772–1780. https://doi.org/10.1109/TGRS.2010.2089799

Rocchini, D., Balkenhol, N., Carter, G. A., Foody, G. M., Gillespie, T. W., He, K. S., … Neteler, M. (2010). Remotely sensed spectral heterogeneity as a proxy of species diversity: Recent advances and open challenges. *Ecological Informatics*, *5*(5), 318–329. https://doi.org/10.1016/j.ecoinf.2010.06.001

Scalbert, A. (1991). Antimicrobial properties of tannins. *Phytochemistry*, *30*(12), 3875–3883. https://doi.org/10.1016/0031-9422(91)83426-L

Schaepman, M. E., Jehle, M., Hueni, A., D'Odorico, P., Damm, A., Weyermann, J., … Itten, K. I. (2015). Advanced radiometry measurements and Earth science applications with the Airborne Prism Experiment (APEX). *Remote Sensing of Environment*, *158*, 207–219. https://doi.org/10.1016/j.rse.2014.11.014

Schaepman, M. E., Ustin, S. L., Plaza, A. J., Painter, T. H., Verrelst, J., & Liang, S. (2009). Earth system science related imaging spectroscopy–An assessment. *Remote Sensing of Environment*, *113*, S123–S137. https://doi.org/10.1016/j.rse.2009.03.001

Schmid, B. (1994). Effects of genetic diversity in experimental stands of *Solidago altissima*–evidence for the potential role of pathogens as selective agents in plant populations. *Journal of Ecology*, 165–175. https://doi.org/10.2307/2261395

Schneider, F. D., Kükenbrink, D., Schaepman, M. E., Schimel, D. S., & Morsdorf, F. (2019). Quantifying 3D structure and occlusion in dense tropical and temperate forests using close-range LiDAR. *Agricultural and Forest Meteorology*, *268*, 249–257. https://doi.org/10.1016/j.agrformet.2019.01.033

Schneider, F. D., Leiterer, R., Morsdorf, F., Gastellu-Etchegorry, J.-P., Lauret, N., Pfeifer, N., & Schaepman, M. E. (2014). Simulating imaging spectrometer data: 3D forest modeling based on LiDAR and in situ data. *Remote Sensing of Environment*, *152*, 235–250. https://doi.org/10.1016/j.rse.2014.06.015

Schreiber, L., Kirsch, T., & Riederer, M. (1996). Transport properties of cuticular waxes of Fagus sylvatica L. and Picea abies (L.) Karst.: estimation of size selectivity and tortuosity from diffusion coefficients of aliphatic molecules. *Planta*, *198*(1), 104–109.

Schneider, F. D., Morsdorf, F., Schmid, B., Petchey, O. L., Hueni, A., Schimel, D. S., & Schaepman, M. E. (2017). Mapping functional diversity from remotely sensed morphological and physiological forest traits. *Nature Communications*, *8*(1), 1441. https://doi.org/10.1038/s41467-017-01530-3

Schweiger, A. K., Cavender-Bares, J., Townsend, P. A., Hobbie, S. E., Madritch, M. D., Wang, R., … Gamon, J. A. (2018). Plant spectral diversity integrates functional and phylogenetic components of biodiversity and predicts ecosystem function. *Nature Ecology & Evolution*, *2*(6), 976–982. https://doi.org/10.1038/s41559-018-0551-1

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*(8), 2180–2197.

Skidmore, A. K., Pettorelli, N., Coops, N. C., Geller, G. N., Hansen, M., Lucas, R., … Wegmann, M. (2015). Environmental science: Agree on biodiversity metrics to track from space. *Nature*, *523*, 403–405. https://doi.org/10.1038/523403a

Somers, B., & Asner, G. P. (2014). Tree species mapping in tropical forests using multi-temporal imaging spectroscopy: Wavelength adaptive spectral mixture analysis. *International Journal of Applied Earth Observation and Geoinformation*, *31*, 57–66. https://doi.org/10.1016/j.jag.2014.02.006

Stark, S., Väisänen, M., Ylänne, H., Julkunen-Tiitto, R., & Martz, F. (2015). Decreased phenolic defence in dwarf birch (*Betula nana*) after warming in subarctic tundra. *Polar Biology*, *38*(12), 1993–2005. https://doi.org/10.1007/s00300-015-1758-0

Stickler, C. M., & Southworth, J. (2008). Application of multi-scale spatial and spectral analysis for predicting primate occurrence and habitat associations in Kibale National Park, Uganda. *Remote Sensing of Environment*, *112*(5), 2170–2186. https://doi.org/10.1016/j.rse.2007.10.013

Szathmáry, E., Jordán, F., & Pál, C. (2001). Can genes explain biological complexity? *Science*, *292*(5520), 1315–1316. https://doi.org/10.1126/science.1060852

Telles, A. C., Kupski, L., & Furlong, E. B. (2017). Phenolic compound in beans as protection against mycotoxins. *Food Chemistry*, *214*, 293–299. https://doi.org/10.1016/j.foodchem.2016.07.079

Torabzadeh, H., Leiterer, R., Hueni, A., Schaepman, M. E., & Morsdorf, F. (2019). Tree species classification in a temperate mixed forest using a combination of imaging spectroscopy and airborne laser scanning. *Agricultural and Forest Meteorology*, *279*, 107744. https://doi.org/10.1016/j.agrformet.2019.107744

Tucker, C. J. (1980). Remote sensing of leaf water content in the near infrared. *Remote Sensing of Environment*, *10*(1), 23–32. https://doi.org/10.1016/0034-4257(80)90096-6

Turner, W. (2014). Sensing biodiversity. *Science*, *346*(6207), 301–302. https://doi.org/10.1126/science.1256014

Ullah, S., Schlerf, M., Skidmore, A. K., & Hecker, C. (2012). Identifying plant species using mid-wave infrared (2.5–6 μm) and thermal infrared (8–14 μm) emissivity spectra. *Remote Sensing of Environment*, *118*, 95–102. https://doi.org/10.1016/j.rse.2011.11.008

Ustin, S. L., Gitelson, A. A., Jacquemoud, S., Schaepman, M., Asner, G. P., Gamon, J. A., & Zarco-Tejada, P. (2009). Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. *Remote Sensing of Environment*, *113*, 67–77. https://doi.org/10.1016/j.rse.2008.10.019

Valbuena, R., O'Connor, B., Zellweger, F., Simonson, W., Vihervaara, P., Maltamo, M., … Chirici, G. (2020). Standardizing Ecosystem Morphological Traits from 3D Information Sources. *Trends in Ecology & Evolution.*, In press.

Vihervaara, P., Auvinen, A.-P., Mononen, L., Törmä, M., Ahlroth, P., Anttila, S., … Virkkala, R. (2017). How essential biodiversity variables and remote sensing can help national biodiversity monitoring. *Global Ecology and Conservation*, *10*, 43–59. https://doi.org/10.1016/j.gecco.2017.01.007

Wang, Y., Hyyppa, J., Liang, X., Kaartinen, H., Yu, X., Lindberg, E., … Alho, P. (2016). International benchmarking of the individual tree detection methods for modeling 3-D canopy structure for silviculture and forest ecology using airborne laser scanning. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(9), 5011–5027. https://doi.org/10.1109/TGRS.2016.2543225

War, A. R., Paulraj, M. G., Ahmad, T., Buhroo, A. A., Hussain, B., Ignacimuthu, S., & Sharma, H. C. (2012). Mechanisms of plant defense against insect herbivores. *Plant Signaling & Behavior*, *7*(10), 1306–1320. https://doi.org/10.4161/psb.21663

Wilson, E. O., & Peter, F. M. (1988). *The loss of diversity causes and consequences. Biodiversity*. Washington, DC: National Academies Press (US).

Wold, S., Ruhe, A., Wold, H., & Dunn, III, W. J. (1984). The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM Journal on Scientific and Statistical Computing*, *5*(3), 735–743. https://doi.org/10.1137/0905052

Wold, S., Eriksson, L., Trygg, J., & Kettaneh, N. (2004). *The PLS method–partial least squares projections to latent structures–and its applications in industrial RDP (research, development, and production)*. Unea University.

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, *58*(2), 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1

Yamasaki, E., Altermatt, F., Cavender-Bares, J., Schuman, M. C., Zuppinger-Dingley, D., Garonna, I., … Shimizu, K. K. (2017). Genomics meets remote sensing in global change studies: Monitoring and predicting phenology, evolution and biodiversity. *Current Opinion in Environmental Sustainability*, *29*, 177–186. https://doi.org/10.1016/j.cosust.2018.03.005

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.