Remote Sensing Series 91

Ewa A. Czyż

Remotely Sensing Ecological Genomics

Remote Sensing Laboratories
Department of Geography
University of Zurich, 2023

Remote Sensing Series 91

Ewa A. Czyż

Remotely Sensing Ecological Genomics

Remote Sensing Laboratories
Department of Geography
University of Zurich, 2023

Front page: reflected and transmitted light from canopies of *Fagus sylvatica* L. (European beech) in the Lägern forest, Switzerland. (Photo: F. D. Schneider).

Czyż, Ewa A.

Remotely Sensing Ecological Genomics.

Remote Sensing Series, Vol. 91
Remote Sensing Laboratories, Department of Geography, University of Zurich
Switzerland, 2023

ISBN: 978-3-906894-29-4

This work was approved as a PhD thesis by the Faculty of Science of the University of Zurich in the spring semester 2023. Doctorate committee: Prof. Dr. Michael E. Schaepman (dissertation supervisor), Prof. Dr. Meredith C. Schuman (chair), Prof. Dr. Maria J. Santos, PD Dr. Felix Morsdorf, and Prof. Dr. Karl Gademann.

Remotely Sensing Ecological Genomics

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)
vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich

von
Ewa A. Czyż
aus
Polen

Promotionskommission
Prof. Dr. Michael E. Schaepman
(Leitung der Dissertation)
Prof. Dr. Meredith C. Schuman
(Vorsitz der Dissertation)
Prof. Dr. Maria J. Santos
PD Dr. Felix Morsdorf
Prof. Dr. Karl Gademann

Zürich, 2023

Abstract

Solar radiation is the prime energy source on Earth. It reaches any object in the form of electromagnetic radiation that may be absorbed, transmitted or reflected. The magnitude of these optical processes depends on the optical properties of each object, which in the case of plants relate to their biochemical and structural traits. These plant phenotypic traits result from gene expression underpinned by an individual's genotype constrained by phylogeny, the environment the individual is exposed to, and the interaction between genotype and the environment. Remote observations of plant phenotypes across space and time may thus hold information about the composition and structure of genetic variation, if a link between spectral and genetic information can be established.

This dissertation encompasses studies linking information derived from imaging spectrometer acquisitions under natural conditions with *in situ* collected information about genetic variation within a tree species, the European beech *Fagus sylvatica*. It presents the correlation between spectral and genetic information by sequentially expanding temporal, spatial and genetic aspects, and simultaneously accounting for environmental contexts that impact gene expression. By evaluating spectral-genetic similarities across decadal airborne imaging spectrometer acquisitions and accounting for spectral phenotypes and whole-genome sequences of tree individuals from across the species range, the studies provide a proof that observed reflectance spectra hold information about genetic variation within the species. Further, by accounting on uncertainties of spectral measurements and deriving genetic structure of the most abundant tree species in Europe, the dissertation advances the current remote sensing approaches and the knowledge on intraspecific genetic variation. The studies focus particularly on the genetic relatedness between the trees of the test species, whereas the acquired data may allow to establish direct associations between genes and spectral features. The methods used may be expanded to other tree species or applied to spectral data acquired by upcoming spaceborne imaging spectrometers, which overcome current spatiotemporal limitations of data collection, and demonstrate further paths towards the association of genetic variation with variation in spectral phenotypes.

The thesis presents the potential of spectral derivation of intraspecific genetic variation within tree species and discusses associated limitations induced by spectral, temporal, spatial and genetic scopes of analysis. This sets a stage towards establishing a means of remote observations of spectral signatures to contribute to monitoring biological variation at the fundamental genetic level, which correlates with ecosystem performance and is an insurance mechanism for populations to adapt to global change.

Zusammenfassung

Die Sonneneinstrahlung ist die wichtigste Energiequelle der Erde. Sie erreicht die Oberfläche in Form von elektromagnetischer Strahlung, welche von Objekten absorbiert, transmittiert oder reflektiert werden kann. Der Ablauf dieser optischen Prozesse hängt von den Eigenschaften des jeweiligen Objekts ab, die im Fall von Pflanzen mit ihren individuellen biochemischen und strukturellen Eigenschaften zusammenhängen. Diese Eigenschaften werden durch den Genotyp der Individuen vorgegeben, der die genetische Ausprägung und damit den Phänotyp, der beobachtet werden kann, bestimmt. Wird eine Verbindung zwischen spektraler und genetischer Information hergestellt, können Raum und Zeit übergreifende fernerkundliche Beobachtungen von Pflanzenphänotypen zum Verständnis über die genetische Zusammensetzung und Struktur einer Pflanzenart beitragen.

Die vorliegende Dissertation befasst sich mit der Verknüpfung von spektraler Information von Individuen derselben Baumart, der Buche *Fagus sylvatica*, unter natürlichen Bedingungen mit genetischer Information aus DNA-Proben von Blättern. Aus Ganzgenomsequenzen und Feldspektrometermessungen aus dem gesamten Verbreitungsgebiet von *Fagus sylvatica* sowie langzeitlicher Datenerhebung mit luftgestützen Spektrometern, stelle ich die Zusammenhänge zwischen spektralen und genetischen Informationen dar. Dazu beleuchte ich sequenziell die zeitlichen, räumlichen und genetischen Aspekte und berücksichtige dabei den Umweltkontext, der die Genexpression beeinflusst. Durch das Einbeziehen der Unsicherheiten von Spektralmessungen und die Beschreibung der genetischen Struktur der am weitesten verbreiteten Baumart in Europa trägt diese Arbeit massgeblich zu den derzeitigen Fernerkundungsansätzen und dem Wissen über die intraspezifische genetische Variation bei. Die Arbeit befasst sich mit der genetische Ähnlichkeit innerhalb der Art, die gewonnenen Daten können aber auch ermöglichen, direkte Assoziationen zwischen Genen und spektralen Signaturen herzustellen. Zukünftig können die gezeigten Methoden auf andere Baumarten ausgeweitet oder auf Spektraldaten angewandt werden, die von weltraumgestützten Spektrometern erfasst werden, welche die aktuellen räumlichen und zeitlichen Beschränkungen der Datenerfassung überwinden.

Die vorgestellte Arbeit zeigt das Potenzial der Erfassung innerartlicher genetischer Variation über spektrale Eigenschaften. Sie erörtert die derzeitigen Limitierungen dieses Ansatzes, welche die spektralen, zeitlichen, räumlichen und genetischen Aspekte der Datengewinnung betreffen. Dies schafft die Voraussetzungen für die Einrichtung von Fernbeobachtungen spektraler Signaturen zur Überwachung biologischer Variation auf ihrer grundlegenden genetischen Ebene, die eine Lebensversicherung für Organismen bei der Anpassung an den globalen Wandel darstellt.

Acknowledgements

This dissertation is an outcome of support, collaboration, and enthusiasm. I would like to thank Michael for entrusting me with the topic and sharing perspectives on how the beauty of science aligns with the bigger picture. I express my gratitude to Merry, who patiently guided the work across the fields, and to Bernhard who always reminds about the wonderful merits of the research in all its aspects. Further thanks go to Carla and Hendrik, who accompanied me with curiosity and maintained the direction in the first steps through the discovery of the topic. Moreover, my thanks go to Fabian, Maarten, and Anna for their openness to exchanging countless thoughts on remote sensing, genetics, and everything in between.

This work accompanied the adventure that I gladly shared with colleagues, without whom the memories collected in the European woods would not have been so wonderful. Especially Marylaure, Cheng, Nemish, Isa, Abo, and Helena – thanks a lot for each and every moment we spent together. Further gratitude goes to colleagues thanks to whom every single day at work was better: Alex, Aline, Aman, Andy, Bruno, Carmen, Chengxiu, Christian, Christoph, Claudia, Dabwiso, Daniel H., Daniel K., David, Devin, Diego, Dominic, Elias, Emiliano, Eugénie, Fanny, Felix, Giulia, Gillian, Hanneke, Jasmin, Jakob, Jen, Joan, Jonas, Julia, Julian, Kathrin, Kien, Kimberley, Leon, Lidia, Maria, Marius, Mary Ann, Mathias, Michi, Mike, Nargiz, Nathalie, Nicole, Oli, Peter, Reik, Remika, Rifat, Sergio, Simon, Sofia, Ting, Tiziana, Valantina, Veronika, Vladimir, and Zhaoju. And yet, none of these thanks would ever have been possible if Sandra and Rita would not have taken care of us! I also acknowledge the University Research Priority Program on Global Change and Biodiversity, members of the Department of Geography, and committee members of this thesis for facilitating the great exchange of knowledge.

Finally, I thank my friends and family for the encouragement necessary to embark on and complete this dissertation. I acknowledge your never-ending support. Family, you know better than anyone how important you were to keeping me on course and navigating me back from inherent ups and downs – with no excuses, you helped me keep my head up and stay on stable ground. I acknowledge that each and everyone of you were the ones who taught me the beauty and appreciation of nature, the greatest motivation for this study.

Contents

List of Abbreviations

AFLP	Amplified Fragment Length Polymorphism
AIC	Akaike Information Criterion
AIS	Airborne Imaging Spectrometer
ANOVA	analysis of variance
APEX	Airborne Prism Experiment
ASAS	Advanced Solidstate Array Spectroradiometer
ATCOR	Atmospheric/Topographic Correction for Satellite Imagery
AVIRIS	Airborne Visible/Infrared Imaging Spectrometer
AVIRIS-NG	Airborne Visible/Infrared Imaging Spectrometer-Next Generation
CGDD	Cumulative Growing Degree Days
CGVPDD	Cumulative Growing Vapor Pressure Deficit Days
CHIME	Copernicus Hyperspectral Imaging Mission for the Environment
CTAB	cetyl trimethylammonium bromide
DESIS	DLR Earth Sensing Imaging Spectrometer
DLR	Das Deutsche Zentrum für Luft- und Raumfahrt
DLSF	Day of the Last Spring Frost
DOY	Day Of the Year
EBV	Essential Biodiversity Variables
EMIT	Earth Surface Mineral Dust Source Investigation
EnMAP	Environmental Mapping and Analysis Program
EO-1	Earth Observing One
ESA	European Space Agency
FIR	far infrared
FLI/PMI	Fluorescence Line Imager/Programmable Multiband Imager
FWHM	Full Width at Half Maximum
GPS	Global Positioning System
HISUI	Hyperspectral Imager Suite
ISS	International Space Station
JCGM	Joint Committee for Guides in Metrology
JPL	Jet Propulsion Laboratory
LiDAR	Light Detection and Ranging
LV95	Swiss national grid
LYCWB	Last Year Climatic Water Balance
MLR	multiple linear regression
MV	microwave
NASA	National Aeronautics and Space Administration

NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NIR	near infrared
PCO	Principal Coordinate
PCoA	Principal coordinate analysis
PCR	Polymerase Chain Reaction
PLS	partial least-squares
PLS-DA	PLS discriminant analysis
PARGE	Parametric Geocoding and Orthorectification for Airborne Optical Scanner Data
PRESS	Predicted Residual Error Sum of Squares
PRI	Photochemical Reflectance Index
PRISMA	Precursore Iperspettrale della Missione Applicativa
RAPD	Random Amplified Polymorphic DNA
RDA	redundancy analysis
RMSE	Root-Mean-Square Error
SAM	Spectral Angle Mapper
SBG	Surface Biology and Geology
SIF	Solar Induced Fluorescence
SIS	Shuttle Imaging Spectrometer
SNPs	Single-Nucleotide Polymorphisms
SRF	Spectral Response Function
SWIR	short-wave infrared
TMP	temperature
UNEP	United Nations Environment Programme
URPP GCB	University of Zürich Research Priority Program on Global Change and Biodiversity
UV	ultraviolet
UZH	University of Zürich
WCMC	World Conservation Monitoring Centre
VIF	Variance Inflation Factor
VIP	Variable Importance of Projection
VIS	visible
VPD	Vapor Pressure Deficit

Chapter 1

Introduction

Remote observations of genetically underpinned plant traits hold the potential to monitor biodiversity on a global scale (e.g. Navarro et al., 2017; O'Connor et al., 2015; Pettorelli et al., 2014; Skidmore et al., 2015; Turner, 2014). Current airborne and spaceborne imaging spectrometers provide arguments towards repeated acquisitions of plant phenotypes under natural conditions across spatial and temporal ranges. These plant phenotypes are the result of interactions between environmental conditions and plant genotypes. Current advances in genotyping have improved the derivation of genetic variations between individuals within classified species. The continuous characterization of plant individuals within species acknowledges the importance of broad genetic pools to maintain ecosystem functions under current conditions (Crutsinger et al., 2006) and potentially allow populations to adapt to global change (Des Roches et al., 2021). By linking spectral observations under natural conditions with *in situ* collected information about intraspecific genetic variation of the species *Fagus sylvatica* L. (European beech), the thesis develops and presents a proof of concept of how spectral information may contribute to monitoring genetic diversity (Fig. 1.1), and thus helps to mitigate risks resulting from the current crisis of biodiversity loss (IPBES, 2019).

FIGURE 1.1: Genotype and environment define the expression of plant traits, which in turn affect their spectral phenotypes. The observations of plants' spectral phenotypes across wide spectral, spatial, temporal and genetic resolutions and ranges hold the potential to unravel the composition and distribution of intraspecific genetic variation.

1.1 Plant spectroscopy

Spectroscopy is the study of the interaction of electromagnetic radiation with matter as a function of wavelength. It allows for non-destructive characterization of matter properties through quantification of reflection, absorption or transmission of radiation in many contiguous spectral bands (Rosencwaig, 1980). In the case of plants, the proportion of these optical processes depends on genetically underpinned biochemical and architectural characteristics of leaves and canopies.

The radiation incoming to a leaf may be reflected (Fresnel, 1866) or refracted (Snell's Law) depending on its incidence angle, its polarization, and the refractive index of the matter. The radiation directly reflected from the leaf surface emerges from single scattering (e.g. Grant et al., 1993) and relates to the roughness of the leaf surface (e.g. Vanderbilt et al., 1985). Part of the radiation that is entering into a leaf by refraction may be attenuated by multiple scattering on air and cell interfaces (Kumar and Silva, 1973) and may be diffusely reflected, or transmitted through the leaf (Breece and Holmes, 1971). These single and multiple scattering behaviors of radiation depend on structural properties of leaves, whereas biochemistry affects their absorptive properties. Absorption takes place when radiation of a particular wavelength, thus carrying particular energy, encounters photoactive molecules sensitive at the same energy levels. The high energies of wavelength in the visible (VIS) part of the solar radiation trigger electronic transitions in molecules (Burns and Burns, 1993). Electromagnetic radiation in longer wavelengths of near infrared (NIR) and short-wave infrared (SWIR) can be absorbed by the vibrational changes of molecules. Their rotation can also trigger absorption of energy of wavelengths in the far infrared (FIR) and the microwave (MV) spectral ranges. Both scattering and absorptive properties influence optical properties of plants, which can be estimated as their reflectance spectra measured across the solar spectral range.

While quantification of the optical properties of organic tissues beyond visible light was conducted in laboratories from the 1930's onwards (Stair and Coblentz, 1933), the development of devices sensitive to wider spectral ranges (Wright, 1941) and 2D array-sensor technologies for the detection of scattered radiation in the early 1980's allowed for the development of imaging spectrometers towards broader ranges of wavelengths and spatially resolved quantification of Earth surface properties (Boardman, 1990).

1.2 Imaging spectrometers for Earth observation

The development of technologies allowing for spectrally and spatially resolved acquisitions made observations of spectral features diagnostic of plant traits from air and space feasible (Goetz et al., 1985).

Imaging spectrometers carried on airborne platforms allowed first remote measurements of Earth surface properties derived from the entire spectral range of reflected solar energy. The Airborne Imaging Spectrometer (AIS) developed in the early 1980's was the first instrument that allowed the detection of reflected solar radiation from the Earth's surface (Vane et al., 1984). The extension of spectral coverage from VIS to NIR and SWIR (Goetz et al., 1983) and improvements to instrument subsystems (Bailey, 1987; Chrisp et al., 1987; Macenka and Chrisp, 1987) resulted in the design of the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). AVIRIS along with other airborne imaging spectrometers, such as Fluorescence Line Imager/Programmable Multiband Imager (FLI/PMI) (Miller et al., 1987), Advanced Solidstate Array Spectroradiometer (ASAS) (Huegel, 1987), and the planned Shuttle Imaging Spectrometer (SIS) (Wellman et al., 1983) were leading the path towards the launch of spaceborne imaging spectrometers. It was not until 2000 when Hyperion — the first spaceborne imaging spectrometer — was launched to acquire spectral properties of the Earth's surface from the Earth Observing One (EO-1) platform. Until now, spaceborne imaging spectrometers represent exploratory missions with limited acquisition capabilities such as the currently operational Precursore Iperspettrale della Missione Applicativa (PRISMA) and Environmental Mapping and Analysis Program (EnMAP) systems on their own platforms and the DLR Earth Sensing Imaging Spectrometer (DESIS), the Hyperspectral Imager Suite (HISUI), and the Earth Surface Mineral Dust Source Investigation (EMIT) systems mounted on the International Space Station (ISS). These imaging spectrometers are demonstrations towards spatially and temporally continuous measurements by the planned Earth-orbiting imaging spectrometers, as for example, Surface Biology and Geology (SBG) of the National Aeronautics and Space Administration National Aeronautics and Space Administration (NASA) (Cawse-Nicholson et al., 2021), the Copernicus Hyperspectral Imaging Mission for the Environment (CHIME) of the European Space Agency (ESA) (Rast et al., 2021), and their constellations.

As the development of spaceborne imaging spectrometers is progressing, their spectral,

spatial and temporal requirements of acquired data are being defined based on investigations of data acquired by current airborne imaging spectrometers.

1.3 Acquisition and processing of airborne imaging spectrometer data

Airborne imaging spectrometers provide the ability to assess Earth surface properties on a landscape scale. The accurate derivation of the surface properties depends on the technical sensor characteristics as well as on the data processing chain, from acquired digital numbers representing the electromagnetic radiation to reflectance spectra holding information about optical properties of the imaged surface.

Solar radiation reflected from the Earth's surface towards the imaging spectrometer is consecutively captured by the spectrometer, split in wavelength domains by a spectral disperser, and focused on a detector array. The properties of the components of the electro-optical chain, such as the sensitivity of the detector and the uniformity of the spectral dispersion, define the precision of the transformation of the at-sensor radiance to electrical signals. The electrical signal is further converted to digital numbers by an analog-to-digital converter, then to radiance units, and aligned to spectral sampling (Vane et al., 1987) based on the radiometric and spectral calibration of the imaging spectrometer. Spectrally aligned radiance data acquired over the time of the flight are georectified or geoindexed to link pixels from imaging space into object space (Schlapfer et al., 2004). Pixels in radiance represent spatially resolved information about optical properties of the imaged surface and the atmospheric pathway between the surface and the sensor. Georectified or geoindexed radiance data are further atmospherically corrected based on the acquisition and illumination geometry and the composition of the atmosphere (Conel et al., 1988; Gao and Goetz, 1990) to derive reflectance properties of the imaged surface only. Currently, efforts are being taken toward quantification and propagation schemes of the effect of the stability of the sensor system over time, the uncertainty of the calibration, and the accuracy of the data processing chain, all contributing to uncertainties of spatially resolved reflectance data.

Radiometric, spectral and spatial resolutions and the fidelity of reflectance data (Chrien et al., 1990; Green, 1990) govern the capabilities for remotely deriving plant traits at the landscape level.

1.4 Plant traits derived from airborne imaging spectrometer data

Remotely acquired reflectance spectra of plants hold information about their optical properties, which relate to multiple biochemical and structural plant traits. The plant traits and their distribution over the landscape affect ecosystem functions and ecosystem dynamics on various spatial scales.

Plant biochemical traits may be derived from the magnitude of absorptions of electromagnetic radiation indicative of the quantity and quality of a plant's photoactive molecules. The electronic transitions in molecules triggered by light in the visible part of the spectrum (VIS) underlay, for example, the light absorption by plant pigments (Moss and Loomis, 1952). In contrast, electromagnetic radiation of longer wavelengths, such as near and short-wave infrared (NIR and SWIR), can be absorbed by the vibrational changes of molecules such as water, lignin, cellulose, and leaf proteins (Peterson et al., 1988). These plant biochemical properties may be derived from spectral indexes that account for particular spectral features, from empirical models that account for the whole solar spectral range (Singh et al., 2015), or from physical modeling (Jacquemoud and Baret, 1990). The width of spectral absorption features and the total magnitude of reflections are modified by structural traits of plants, which affect radiation scattering at the leaf and the canopy level (Allen et al., 1970; Gastellu-Etchegorry et al., 1995; Guyot et al., 1992; Kupiec and Curran, 1995). The changes in reflection magnitude also highly depend on the surface anisotropy (Vögtli et al., 2020) and the acquisition geometry (Schaepman-Strub et al., 2006; Zarco-Tejada and Ustin, 2001), which affect the remote derivation of plant properties and their distribution. Spatial and temporal distribution of both biochemical and structural plant traits (Peterson et al., 1988; Wessman, 1987) drive scale-dependent ecological processes (Goetz and Herring, 1989). Temporally repeated acquisitions of trait distributions may contribute to the monitoring of vegetation health (Herrmann et al., 1988; Rock et al., 1988; Ustin et al., 1998), calibration of ecological models (M. E. Martin and Aber, 1997; Zarco-Tejada and Miller, 1999), and understanding of ecological function (Asner et al., 2017). As airborne campaigns are often restricted by the tradeoff between spatial and temporal coverage, the current efforts towards dense multi-temporal acquisitions will further enhance our understanding of repeated acquisitions by upcoming spaceborne imaging spectrometers and the derived temporal dynamics of plant traits (Chlus and Townsend, 2022) from local to global scales.

Derivations of remotely observable plant traits across space and their dynamics across multiple time scales under natural conditions may be diagnostic of genetic variation within tree species, which is underpinned by evolutionary processes as described in the following section (Fig. 1.2).

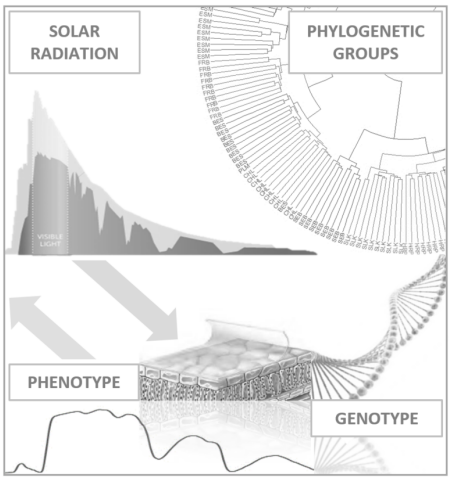

FIGURE 1.2: The solar radiation interacts with plant tissues, whose phenotypic traits, such as morphological structure and biochemical composition are constrained by individual's genotype. The observable spectral phenotypes may be indicative of genetic variation among phylogenetic groups and among individuals within a species.

1.5 Evolutionarily underpinned plant traits and vegetation mapping

Observable plant traits are determined by a plant's genotype and the influences of the environment to which the plant is exposed. The interaction between genotype and the environment results in the plant phenotype, and genetic and environmental variation among plants is reflected in the phenotypic variation of plant traits (Schmid, 1992; Smith et al., 1967).

During the life span of a plant, it acclimates to its environment and adjusts its phenotype to enhance fitness and survival (Schmid, 1992). On longer, evolutionary time scales, plant populations adapt to their environment through heritable changes in genotypes, which may be passed down through generations. Adaptation to similar abiotic and biotic conditions can drive convergent evolution of distant phylogenetic groups, whereas adaptation to different abiotic and biotic conditions can drive divergent evolution that can cause trait diversification between closely related phylogenetic groups or even within species. Diversification of traits through evolution allows different species or genotypes to fill particular ecological niches and utilize different resources across space and time (Vellend and Geber, 2005). Adaptation through genetic change may be phylogenetically constrained, meaning that not every genotype may be acquired via evolution by a particular phylogenetic

group, at least not on ecologically and evolutionarily relevant time-scales. Furthermore, the phenotypic expression of traits is genetically constrained, meaning that a particular trait is expressed for a particular genotype under a particular environment. Under these conditions, when individuals grow in the same environment at the same time, more genetically similar individuals will express more similar phenotypic traits, whereas genetically distant individuals will express more different phenotypic traits. Thus, genetically different individuals become distinguishable based on their phenotypic traits under the same environmental conditions, or genetically different individuals may be distinguishable based on variation in their phenotypic traits separated from the variation in environmental conditions to which they are exposed. Based on this, simultaneous acquisitions of multiple plant phenotypic traits by imaging spectrometers across space and time may allow the mapping of genetic groups (Meireles et al., 2020; Schweiger et al., 2018), as long as observed phenotypic traits are discriminative of genetic groups at the available spectral, spatial and temporal scales. Earlier work has shown that plant traits derived from airborne imaging spectrometer data can be indicative of plant communities (M. Martin et al., 1998; Rock, 1983), species assemblages (D. A. Roberts et al., 1998; D. Roberts et al., 1993), functionally distinct species (Asner et al., 2008; Gamon et al., 1993; Ustin et al., 2004), and particular species (Borregaard et al., 2000; Zarco-Tejada and Miller, 1999). So far, few studies have used remote sensing to assess genetic variation within species under natural conditions (Madritch et al., 2014), presumably because within-species phenotypic variation tends to be smaller than between-species phenotypic variation (Cavender-Bares et al., 2016; Guillén-Escribà et al., 2021), and because the impact of environmental variations on variation in phenotypic traits may obscure variation in phenotypic traits resulting from within-species genetic variation. Linkages between genetic variation and variation in spectral phenotypes have been shown for laboratory conditions (Matsuda et al., 2012) and for experimental designs (Cavender-Bares et al., 2016). Relating phenotypic variation to genetic variation is becoming more useful on broader spatial extents as individuals within one species may cover a wider range of environmental and genetic variation, exposing a wider range of phenotypes (Castro-Esau et al., 2004; Cochrane, 2000; Roth et al., 2015), and thus the derived correlation between genetic and phenotypic variation may be more generic and potentially more transferable across space and time.

If observed within-species variation resulting from genetic variation could be separated from within-species variation resulting from environmental variation, the remote observations would facilitate mapping of genetic composition and its distribution across space and

time. Such assessments are relevant because this fine-scale intraspecific genetic variation underpins global biological variation.

1.6 Intraspecific genetic variation across space and time

Genetic variation within species represents a fundamental level of biological variation, which underpins species and ecosystem diversity, ecological processes, and the potential of populations to adapt to environmental change.

The exposure of plant populations to changing environments and their demographic history drives genetic diversification within species via evolutionary processes. Such processes include selection, drift, mutation, and gene flow, which directly affect changes in allele (variant of a gene) frequencies (Laland, 2015; Wilson, Bossert, et al., 1971). Limited genetic diversity may limit phenotypic diversity, thus affecting the growth, reproduction, and survival of individuals under environmental pressure, with selection of the best suitable genotypes (Irgens-Moller, 1958). While selection acts on adaptive plant traits, genetic drift is a neutral transition in allele frequencies, which together with genetic mutations drive genetic variations with no direction towards particular phenotypes (Kimura, 1968). Mutation is the only evolutionary process that beyond changing allele frequencies can also create new genes. Adaptive genetic variation results in particular distributions of expressed plant traits, and impacts functions on intraspecific, interspecific and ecosystem levels at local to global scales (Agashe, 2009; Reusch et al., 2005). Neutral genetic variation can change into adaptive genetic variation if conditions change and therefore contributes to the evolutionary potential of a population or species to adapt to environmental change (Frankham, 2010).

Adaptive and neutral genetic variation may further affect the temporal turnover of genetic diversity; therefore, monitoring and understanding drivers of genetic variation across space and time is especially important in the light of global change (Carroll et al., 2014; Des Roches et al., 2021; Thomassen et al., 2011). Yet, despite the urgency, there are no global schemes for monitoring intraspecific genetic diversity.

1.7 Monitoring of genetic variation with imaging spectrometers

Intraspecific genetic diversity refers to differences in the genotypic composition between individuals of a single species and may be measured from differences among genomes of individuals of that species (Bruford et al., 2017). Given the methodological, spatial and temporal restrictions of genotyping, genetic variation is challenging to monitor.

Before the development of genotyping methods, physiological and morphological traits were used to measure variations among individuals. These phenotypic responses depend on expression of particular genes at particular ontogenetic stages in particular environments and do not reflect the total of all genetic differences between two individuals. The development of molecular techniques allowed for the estimation of genetic diversity between individuals based on gene expression products such as particular enzymes or based on DNA sequences directly. Differences among isozymes, which represent enzymes serving the same function but which may differ in the genetic sequences between organisms, were the first approach to derive genetic variation using molecular techniques (Hunter and Markert, 1957; Lewontin and Hubby, 1966). The development of the Polymerase Chain Reaction (PCR) in the 1980's (Mullis et al., 1992) allowed for direct DNA sequencing techniques, which subsequently utilized short reads of repeated DNA known as microsatellites (Hughes and Queller, 1993), Random Amplified Polymorphic DNA (RAPD) fragments (Williams et al., 1990), and Amplified Fragment Length Polymorphism (AFLP) (Vos et al., 1995). Molecular techniques advanced towards whole-genome sequencing of plant species in 2000 (The Arabidopsis Genome Initiative 2000). Whole-genome sequences, as well as genomic sequences of selected DNA fragments, require physical sampling and, for the current stage of innovation, usually laboratory work. The necessary time and resources limit the feasibility of mapping genetic diversity across space and time via genotyping of individual plants. Remote sensing with its capacity for spatially wide and temporally repeated acquisitions may contribute to biodiversity monitoring (Turner, 2014). Imaging spectrometers acquiring spectrally dense information may provide information on integrative phenotypic responses extensively across spatially and temporally variable conditions. As each phenotypic response is a result of environment interaction with plant genotypes, assessing phenotypic variation over large ranges of genotypes and environmental conditions may allow for the establishing of links between remotely acquired information and

genetic variation (Yamasaki et al., 2017). Finding associations between genetic and spectral variation may then help to predict genetic composition and its distribution. Moreover, associations of genome-wide sequences with spectral and environmental variation, supported by common garden and ultimately transplant experiments, may facilitate gene recognition underpinning variation in spectral phenotypes and the recognition of spectral features by genes of known function. Such associations would advance the understanding of selective processes that underlie genetic adaptations across wide spatial scales.
Contributions by remote observation to the mapping of genetic variation and its spatial structure, and gene variant distribution and its adaptation across space, may improve current strategies for maintaining genetic diversity. It could help in the recognition of genetically rich or distinct areas or areas where genetic variation is changing, enhance sensitivity of global biodiversity indicators (Scholes et al., 2008) to early warning signs (Lindenmayer et al., 2012), improve our understanding of adaptive capacity underpinned by genes, and further introduce adaptive management strategies for the conservation and restoration of biological diversity at local to global scales (Sarkar, 2002; Sutherland et al., 2004).

1.8 Structure and aim of the dissertation

This thesis represents interdisciplinary research at the interface of remote sensing and genetics. It shows empirical links between spectral data derived from spectroradiometers and genetic data derived from *in situ* collected plant material. The presented approaches are based on observations conducted under natural conditions and thus represent the complexity of plant interactions with their temporally and spatially heterogeneous environment. The overarching goal of the thesis is to provide a proof of concept on how remote observations may contribute to the monitoring of biodiversity at the intraspecific genetic level.

The work particularly focuses on the tree species *Fagus sylvatica* L., the European beech, and the correlation of its spectral variation with genetic variation. Consequently, the studies expand the potential of imaging spectrometer data across temporal scales and the natural range of species, and exploits genetic information from short genetic markers to whole-genome sequences (Fig. 1.3). The expansion of the temporal, spatial and genetic components provides form to the structure of the thesis:

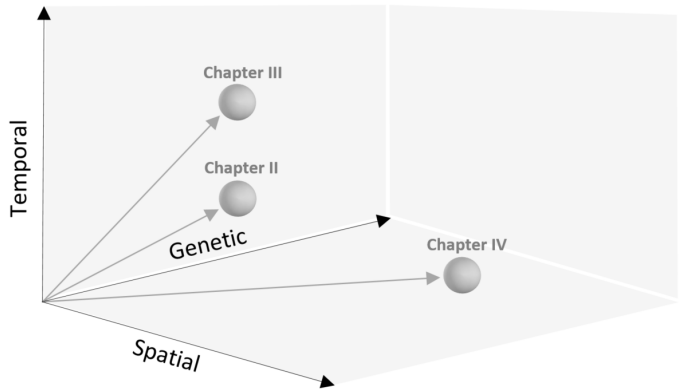

FIGURE 1.3: This thesis evaluates the correlation of spectral with genetic variation. It subsequently explores temporal, spatial and genetic ranges and resolutions and discusses the current limitations and further potential of imaging spectrometers to monitor genetic variation.

Chapter 1 An introduction that provides a background to the studies with emphasis on information about vegetation traits derived from imaging spectrometer data and the relevance of intraspecific genetic variation for biodiversity assessments.

Chapter 2 Presents a population-wide study unravelling links between spectral information derived from an airborne imaging spectrometer data and genetic groups of *F. sylvatica* individuals derived from microsatellite data. The study takes advantage of multitemporal acquisitions conducted at the peak of the growing season over multiple years (Czyż et al., 2020).

Chapter 3 Expands the temporal aspect to decade-long airborne acquisitions over the growing seasons and considers the genetic representation of individuals as their relatedness derived from microsatellite data. Further, the study incorporates intrinsic uncertainties of spectral measurements and addresses ecophysiological processes explaining spectral-genetic similarities on temporal ranges from days to years (Czyż et al., 2023).

Chapter 4 Explores the correlation between the variation of leaf-level spectral reflectance of *F. sylvatica* populations from across the species range measured by a contact probe attached to a spectroradiometer, and the species genetic structure derived from short-read whole-genome sequences. The study addresses the potential of environmental and spectral data to predict genetic variation of a species across its natural range (Czyż et al. in preparation).

Chapter 5 Discusses contributions of the studies in respect to knowledge gathered in the context of spectral, temporal, spatial and genetic resolutions and ranges. Further, it presents limitations of the studies and proposes possible future

directions towards understanding how remote sensing techniques could best contribute to the monitoring of biodiversity, in particular genetic diversity within tree species.

The studies improve understanding of the benefits and caveats of spectral information in relation to contributing to not yet established worldwide schemes for genetic diversity monitoring, and adds to a limited number of studies focusing on remotely sensed intraspecific variation under natural conditions (Guillén-Escribà et al., 2021; Madritch et al., 2014). Further, the studies step ahead to describe the benefits and caveats associated with dense multi-temporal spectral acquisitions, which are so far not extensively used (but see Chlus and Townsend, 2022) but in the future will presumably be widely available (Cawse-Nicholson et al., 2021). By simultaneously accounting for uncertainties in spectral measurements, the studies potentially help to set new standards for the delivery and propagation of uncertainties of acquired spectral information. Finally, in using advances in current technologies on genome-wide sequences, the studies builds on the knowledge on genetic structure of ecologically and economically important species strongly affected by current environmental pressure (Pfenninger et al., 2020; Pluess and Weber, 2012), and discusses a potential path for further contribution of spectral data for understanding genetic variation under global change.

References

Agashe, D. (2009). The stabilizing effect of intraspecific genetic variation on population dynamics in novel and ancestral habitats. *The American Naturalist*, *174*(2), 255–267.

Allen, W. A., Gausman, H., Richardson, A., & Wiegand, C. (1970). Mean effective optical constants of thirteen kinds of plant leaves. *Applied Optics*, *9*(11), 2573–2577.

Asner, G. P., Jones, M. O., Martin, R. E., Knapp, D. E., & Hughes, R. F. (2008). Remote sensing of native and invasive species in hawaiian forests. *Remote sensing of Environment*, *112*(5), 1912–1926.

Asner, G. P., Martin, R. E., Knapp, D., Tupayachi, R., Anderson, C., Sinca, F., Vaughn, N., & Llactayo, W. (2017). Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. *Science*, *355*(6323), 385–389.

Bailey, G. C. (1987). Visible and infrared linear detector arrays for the airborne visible/infrared imaging spectrometer (aviris). *Imaging spectroscopy II*, *834*, 50–54.

Boardman, J. W. (1990). Inversion of high spectral resolution data. *Imaging Spectroscopy of the Terrestrial Environment*, *1298*, 222–233.

Borregaard, T., Nielsen, H., Nørgaard, L., & Have, H. (2000). Crop–weed discrimination by line imaging spectroscopy. *Journal of agricultural engineering research*, *75*(4), 389–400.

Breece, H. T., & Holmes, R. A. (1971). Bidirectional scattering characteristics of healthy green soybean and corn leaves in vivo. *Applied Optics*, *10*(1), 119–127.

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

Burns, R. G., & Burns, R. G. (1993). *Mineralogical applications of crystal field theory*. Cambridge university press.

Carroll, S. P., Jørgensen, P. S., Kinnison, M. T., Bergstrom, C. T., Denison, R. F., Gluckman, P., Smith, T. B., Strauss, S. Y., & Tabashnik, B. E. (2014). Applying evolutionary biology to address global challenges. *science*, *346*(6207), 1245993.

Castro-Esau, K. L., Sánchez-Azofeifa, G., & Caelli, T. (2004). Discrimination of lianas and trees with leaf-level hyperspectral data. *Remote Sensing of Environment*, *90*(3), 353–372.

Cavender-Bares, J., Meireles, J. E., Couture, J. J., Kaproth, M. A., Kingdon, C. C., Singh, A., Serbin, S. P., Center, A., Zuniga, E., Pilz, G., et al. (2016). Associations of leaf spectra with genetic and phylogenetic variation in oaks: Prospects for remote detection of biodiversity. *Remote Sensing*, *8*(3), 221.

Cawse-Nicholson, K., Townsend, P. A., Schimel, D., Assiri, A. M., Blake, P. L., Buongiorno, M. F., Campbell, P., Carmon, N., Casey, K. A., Correa-Pabón, R. E., et al. (2021). Nasa's surface biology and geology designated observable: A perspective on surface imaging algorithms. *Remote Sensing of Environment*, *257*, 112349.

Chlus, A., & Townsend, P. A. (2022). Characterizing seasonal variation in foliar biochemistry with airborne imaging spectroscopy. *Remote Sensing of Environment*, *275*, 113023.

Chrien, T. G., Green, R. O., & Eastwood, M. L. (1990). Accuracy of the spectral and radiometric laboratory calibration of the airborne visible/infrared imaging spectrometer. *Imaging Spectroscopy of the Terrestrial Environment*, *1298*, 37–49.

Chrisp, M. P., Chrien, T., & Steimle, L. (1987). Aviris foreoptics, fiber optics and on-board calibrator. *Imaging spectroscopy II*, *834*, 44–49.

Cochrane, M. (2000). Using vegetation reflectance variability for species level classification of hyperspectral data. *International journal of remote sensing*, *21*(10), 2075–2087.

Conel, J., Adams, S., Alley, R., Hoover, G., & Schultz, S. (1988). Ais radiometry and the problem of contamination from mixed spectral orders. *Remote Sensing of Environment*, *24*(1), 179–200.

Crutsinger, G. M., Collins, M. D., Fordyce, J. A., Gompert, Z., Nice, C. C., & Sanders, N. J. (2006). Plant genotypic diversity predicts community structure and governs an ecosystem process. *science*, *313*(5789), 966–968.

Czyż, E. A., Guillén Escribà, C., Wulf, H., Tedder, A., Schuman, M. C., Schneider, F. D., & Schaepman, M. 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.

Czyż, E. A., Schmid, B., Hueni, A., Eppinga, M. B., Schuman, M. C., Schneider, F. D., Guillén-Escribà, C., & Schaepman, M. E. (2023). Genetic constraints on temporal variation of airborne reflectance spectra and their uncertainties over a temperate forest. *Remote Sensing of Environment*, *284*, 113338.

Des Roches, S., Pendleton, L. H., Shapiro, B., & Palkovacs, E. P. (2021). Conserving intraspecific variation for nature's contributions to people. *Nature Ecology & Evolution*, *5*(5), 574–582.

Frankham, R. (2010). Challenges and opportunities of genetic approaches to biological conservation. *Biological conservation*, *143*(9), 1919–1927.

Fresnel, A. (1866). Œuvres complètes d'augustin fresnel, publiées par mm. *Henri de Senarmont, Verdet et Léonor Fresnel, Paris, Imprimerie impériale*, *3*.

Gamon, J. A., Field, C. B., Roberts, D. A., Ustin, S. L., & Valentini, R. (1993). Functional patterns in an annual grassland during an aviris overflight. *Remote Sensing of Environment*, *44*(2-3), 239–253.

Gao, B.-C., & Goetz, A. F. (1990). Determination of total column water vapor in the atmosphere at high spatial resolution from aviris data using spectral curve fitting and band ratioing techniques. *Imaging Spectroscopy of the Terrestrial Environment*, *1298*, 138–149.

Gastellu-Etchegorry, J., Zagolski, F., Mougtn, E., Marty, G., & Giordano, G. (1995). An assessment of canopy chemistry with aviris—a case study in the landes forest, south-west france. *Remote Sensing*, *16*(3), 487–501.

Goetz, A. F., & Herring, M. (1989). The high resolution imaging spectrometer (hiris) for eos. *IEEE Transactions on Geoscience and Remote Sensing*, *27*(2), 136–144.

Goetz, A. F., Rock, B. N., & Rowan, L. C. (1983). Remote sensing for exploration; an overview. *Economic Geology*, *78*(4), 573–590.

Goetz, A. F., Vane, G., Solomon, J. E., & Rock, B. N. (1985). Imaging spectrometry for earth remote sensing. *science*, *228*(4704), 1147–1153.

Grant, L., Daughtry, C., & Vanderbilt, V. (1993). Polarized and specular reflectance variation with leaf surface features. *Physiologia Plantarum*, *88*(1), 1–9.

Green, R. O. (1990). Retrieval of reflectance from calibrated radiance imagery measured by the airborne visible/infrared imaging spectrometer (aviris) for lithological mapping of clark mountains, california. *Annual JPL Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Workshop*, *2*, 90–54.

Guillén-Escribà, C., Schneider, F. D., Schmid, B., Tedder, A., Morsdorf, F., Furrer, R., Hueni, A., Niklaus, P. A., & Schaepman, M. E. (2021). Remotely sensed between-individual functional trait variation in a temperate forest. *Ecology and Evolution*, *11*(16), 10834–10867.

Guyot, G., Baret, F., & Jacquemoud, S. (1992). *Imaging spectroscopy for vegetation studies* (Vol. 2). Kluwer Academic Publishers: Norwell, MA, USA.

Herrmann, K., Rock, B. N., Ammer, U., & Paley, H. N. (1988). Preliminary assessment of airborne imaging spectrometer and airborne thematic mapper data acquired for forest decline areas in the federal republic of germany. *Remote sensing of environment*, *24*(1), 129–149.

Huegel, F. G. (1987). Advanced solidstate array spectroradiometer: Sensor and calibration improvements. *Imaging spectroscopy II*, *834*, 12–15.

Hughes, C., & Queller, D. C. (1993). Detection of highly polymorphic microsatellite loci in a species with little allozyme polymorphism. *Molecular Ecology*, *2*(3), 131–137.

Hunter, R. L., & Markert, C. (1957). Histochemical demonstration of enzymes separated by zone electrophoresis in starch gels. *Science*, *125*(3261), 1294–1295.

IPBES. (2019). *Global assessment report of the intergovernmental science-policy platform on biodiversity and ecosystem services* (E. Brondizio, J. Settele, S. Díaz, & H. T. Ngo, Eds.). https://doi.org/10.5281/zenodo.3831673

Irgens-Moller, H. (1958). Genetic variations in length of active growth period among races of douglas fir, pseudotsuga menziesii (mirb.) franco.

Jacquemoud, S., & Baret, F. (1990). Prospect: A model of leaf optical properties spectra. *Remote sensing of environment*, *34*(2), 75–91.

Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature*, *217*, 624–626.

Kumar, R., & Silva, L. (1973). Light ray tracing through a leaf cross section. *Applied Optics*, *12*(12), 2950–2954.

Kupiec, J., & Curran, P. (1995). Decoupling effects of the canopy and foliar biochemicals in aviris spectra. *International Journal of Remote Sensing*, *16*(9), 1731–1739.

Laland, K. N. (2015). On evolutionary causes and evolutionary processes. *Behavioural Processes*, *117*, 97–104.

Lewontin, R. C., & Hubby, J. L. (1966). A molecular approach to the study of genic heterozygosity in natural populations. ii. amount of variation and degree of heterozygosity in natural populations of drosophila pseudoobscura. *Genetics*, *54*(2), 595.

Lindenmayer, D. B., Gibbons, P., Bourke, M., Burgman, M., Dickman, C. R., Ferrier, S., Fitzsimons, J., Freudenberger, D., Garnett, S. T., Groves, C., et al. (2012). Improving biodiversity monitoring. *Austral Ecology*, *37*(3), 285–294.

Macenka, S. A., & Chrisp, M. P. (1987). Airborne visible/infrared imaging spectrometer (aviris) spectrometer design and performance. *Imaging spectroscopy II*, *834*, 32–43.

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.

Martin, M. E., & Aber, J. D. (1997). High spectral resolution remote sensing of forest canopy lignin, nitrogen, and ecosystem processes. *Ecological applications*, *7*(2), 431–443.

Martin, M., 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.

Matsuda, O., Tanaka, A., Fujita, T., & Iba, K. (2012). Hyperspectral imaging techniques for rapid identification of arabidopsis mutants with altered leaf pigment status. *Plant and Cell Physiology*, *53*(6), 1154–1170.

Meireles, J. E., Cavender-Bares, J., Townsend, P. A., Ustin, S., Gamon, J. A., Schweiger, A. K., Schaepman, M. E., Asner, G. P., Martin, R. E., Singh, A., et al. (2020). Leaf reflectance spectra capture the evolutionary history of seed plants. *New Phytologist*, *228*(2), 485–493.

Miller, J., Hare, E., Hollinger, A., & Sturgeon, D. (1987). Imaging spectrometry as a tool for botanical mapping. *Imaging spectroscopy II*, *834*, 108–113.

Moss, R., & Loomis, W. (1952). Absorption spectra of leaves. i. the visible spectrum. *Plant Physiology*, *27*(2), 370.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., Erlich, H., et al. (1992). Specific enzymatic amplification of dna in vitro: The polymerase chain reaction. *Biotechnology Series*, 17–17.

Navarro, L. M., Fernandez, N., Guerra, C., Guralnick, R., Kissling, W. D., Londono, M. C., Muller-Karger, F., Turak, E., Balvanera, P., Costello, M. J., et al. (2017). Monitoring biodiversity change through effective global coordination. *Current opinion in environmental sustainability*, *29*, 158–169.

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.

Peterson, D. L., Aber, J. D., Matson, P. A., Card, D. H., Swanberg, N., Wessman, C., & Spanner, M. (1988). Remote sensing of forest canopy and leaf biochemical contents. *Remote Sensing of Environment*, *24*(1), 85–108.

Pettorelli, N., Safi, K., & Turner, W. (2014). Satellite remote sensing, biodiversity research and conservation of the future.

Pfenninger, M., Reuss, F., Kiebler, A., Schönnenbeck, P., Caliendo, C., Gerber, S., Cocchiararo, B., Reuter, S., Blüthgen, N., Mody, K., et al. (2020). Genomic basis of drought resistance in fagus sylvatica. *bioRxiv*, 2020–12.

Pluess, A. R., & Weber, P. (2012). Drought-adaptation potential in fagus sylvatica: Linking moisture availability with genetic diversity and dendrochronology. *PloS one*, *7*(3), e33636.

Rast, M., Nieke, J., Adams, J., Isola, C., & Gascon, F. (2021). Copernicus hyperspectral imaging mission for the environment (chime). *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 108–111.

Reusch, T. B., Ehlers, A., Hämmerli, A., & Worm, B. (2005). Ecosystem recovery after climatic extremes enhanced by genotypic diversity. *Proceedings of the National Academy of Sciences*, *102*(8), 2826–2831.

Roberts, D. A., Gardner, M., Church, R., Ustin, S., Scheer, G., & Green, R. (1998). Mapping chaparral in the santa monica mountains using multiple endmember spectral mixture models. *Remote sensing of environment*, *65*(3), 267–279.

Roberts, D., Green, R., Sabol, D., & Adams, J. (1993). Temporal changes in endmember abundances, liquid water and water vapor over vegetation at jasper ridge.

Rock, B. (1983). Preliminary airborne imaging spectrometer vegetation data. *1983 International Geoscience and Remote Sensing Symposium (IGARSS"83)*.

Rock, B., Hoshizaki, T., & Miller, J. (1988). Comparison of in situ and airborne spectral measurements of the blue shift associated with forest decline. *Remote Sensing of Environment*, *24*(1), 109–127.

Rosencwaig, A. (1980). Photoacoustic spectroscopy. *Annual review of biophysics and bioengineering*, *9*(1), 31–54.

Roth, K. L., Roberts, D. A., Dennison, P. E., Alonzo, M., Peterson, S. H., & Beland, M. (2015). Differentiating plant species within and across diverse ecosystems with imaging spectroscopy. *Remote Sensing of Environment*, *167*, 135–151.

Sarkar, S. (2002). Defining "biodiversity"; assessing biodiversity. *The Monist*, *85*(1), 131–155.

Schaepman-Strub, G., Schaepman, M. E., Painter, T. H., Dangel, S., & Martonchik, J. V. (2006). Reflectance quantities in optical remote sensing—definitions and case studies. *Remote sensing of environment*, *103*(1), 27–42.

Schlapfer, D. R., Kaiser, J. W., Brazile, J., Schaepman, M. E., & Itten, K. I. (2004). Calibration concept for potential optical aberrations of the apex pushbroom imaging spectrometer. *Sensors, Systems, and Next-Generation Satellites VII*, *5234*, 221–231.

Schmid, B. (1992). Phenotypic variation in plants. *Evolutionary trends in plants*, *6*(1), 45–60.

Scholes, R. J., Mace, G. M., Turner, W., Geller, G. N., Jürgens, N., Larigauderie, A., Muchoney, D., Walther, B. A., & Mooney, H. (2008). Toward a global biodiversity observing system.

Schweiger, A. K., Cavender-Bares, J., Townsend, P. A., Hobbie, S. E., Madritch, M. D., Wang, R., Tilman, D., & 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.

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., Mücher, C. A., O'Connor, B., Paganini, M., Pereira, H. M., et al. (2015). Environmental science: Agree on biodiversity metrics to track from space. *Nature*, *523*(7561), 403–405.

Smith, S., et al. (1967). Experimental and natural hybrids in north america. typha (typhaceae). *American Midland Naturalist*, *78*(2), 295–97.

Stair, R., & Coblentz, W. (1933). Infrared absorption spectra of some plant pigments. *J Res*, *11*, 703–711.

Sutherland, W. J., Pullin, A. S., Dolman, P. M., & Knight, T. M. (2004). The need for evidence-based conservation. *Trends in ecology & evolution*, *19*(6), 305–308.

Thomassen, H. A., Fuller, T., Buermann, W., Mila, B., Kieswetter, C. M., Jarrın-V, P., Cameron, S. E., Mason, E., Schweizer, R., Schlunegger, J., et al. (2011). Mapping evolutionary process: A multi-taxa approach to conservation prioritization. *Evolutionary applications*, *4*(2), 397–413.

Turner, W. (2014). Sensing biodiversity. *Science*, *346*(6207), 301–302.

Ustin, S. L., Roberts, D. A., Gamon, J. A., Asner, G. P., & Green, R. O. (2004). Using imaging spectroscopy to study ecosystem processes and properties. *BioScience*, *54*(6), 523–534.

Ustin, S. L., Roberts, D. A., Pinzon, J., Jacquemoud, S., Gardner, M., Scheer, G., Castaneda, C. M., & Palacios-Orueta, A. (1998). Estimating canopy water content of chaparral shrubs using optical methods. *Remote Sensing of Environment*, *65*(3), 280–291.

Vanderbilt, V. C., Grant, L., & Daughtry, C. (1985). Polarization of light scattered by vegetation. *Proceedings of the IEEE*, *73*(6), 1012–1024.

Vane, G., Chrien, T. G., Miller, E. A., & Reimer, J. H. (1987). Spectral and radiometric calibration of the airborne visible/infrared imaging spectrometer. *Imaging spectroscopy II*, *834*, 91–107.

Vane, G., Goetz, A. F., & Wellman, J. B. (1984). Airborne imaging spectrometer: A new tool for remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, (6), 546–549.

Vellend, M., & Geber, M. A. (2005). Connections between species diversity and genetic diversity. *Ecology letters*, *8*(7), 767–781.

Vögtli, M., Schläpfer, D., Richter, R., Hueni, A., Schaepman, M. E., & Kneubühler, M. (2020). About the transferability of topographic correction methods from spaceborne to airborne optical data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 1348–1362.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., et al. (1995). A new technique for dna fingerprinting. *Nucleic acids research*, *23*(21), 4407–4414.

Wellman, J., Goetz, A., Herring, M., & Vane, G. (1983). An imaging spectrometer experiment for the shuttle. *1983 International Geoscience and Remote Sensing Symposium (IGARSS"83)*.

Wessman, C. A. (1987). *Estimating key forest ecosystem parameters through remote sensing (wisconsin)*. The University of Wisconsin-Madison.

Williams, J. G., Kubelik, A. R., Livak, K. J., Rafalski, J. A., & Tingey, S. V. (1990). Dna polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic acids research*, *18*(22), 6531–6535.

Wilson, E. O., Bossert, W. H., et al. (1971). *A primer of population biology* (Vol. 3). Sinauer Associates Sunderland, MA.

Wright, N. (1941). Application of infrared spectroscopy to industrial research. *Industrial & Engineering Chemistry Analytical Edition*, *13*(1), 1–8.

Yamasaki, E., Altermatt, F., Cavender-Bares, J., Schuman, M. C., Zuppinger-Dingley, D., Garonna, I., Schneider, F. D., Guillén-Escribà, C., van Moorsel, S. J., Hahl, T., et al. (2017). Genomics meets remote sensing in global change studies: Monitoring and predicting phenology, evolution and biodiversity. *Current opinion in environmental sustainability*, *29*, 177–186.

Zarco-Tejada, P. J., & Miller, J. R. (1999). Land cover mapping at boreas using red edge spectral parameters from casi imagery. *Journal of Geophysical Research: Atmospheres*, *104*(D22), 27921–27933.

Zarco-Tejada, P. J., & Ustin, S. (2001). Modeling canopy water content for carbon estimates from modis data at land eos validation sites. *IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No. 01CH37217)*, *1*, 342–344.

Chapter 2

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

Ewa A. Czyż, Carla Guillén Escribà, Hendrik Wulf, Andrew Tedder, Meredith C. Schuman, Fabian D. Schneider, Michael E. Schaepman

This chapter is based on the peer-reviewed article:
Intraspecific genetic structure of a Fagus sylvatica population in a temperate forest derived from airborne imaging spectroscopy time series.
Ecology and evolution, 2020, 10(14), 7419-7430.
DOI: 10.1002/ece3.6469.
It is reprinted as the final submitted manuscript and has been modified to fit the layout of this thesis.

Author contributions. Conceptualization: EAC, CGE, HW, MES. Methodology: EAC, CGE, HW, AT. Formal analysis: EAC, AT. Investigation: EAC, CGE, HW, AT. Resources: MES. Data Curation: EAC, CGE, HW, AT, FDS. Writing - Original Draft: EAC, CGE, HW, MCS, MES. Writing - Review and Editing: EAC, HW, AT, MCS, FDS. Visualization: EAC. Supervision: CGE, HW, MES. Project administration: MES.

Detailed author contributions

EA Czyż conceptualized the study together with ME Schaepman, C Guillén Escribà, and H Wulf. ME Schaepman obtained funding. C Guillén Escribà, ME Schaepman, and H Wulf supervised the work. Nuclear microsatellite analyses were conducted by A Tedder with support from C Guillén Escribà. C Guillén Escribà derived the georeferenced tree crown map supported by FD Schneider. Airborne imaging spectrometer data were collected with the Airborne Prism Experiment (APEX) during yearly campaigns administrated by ME Schaepman and airborne data were curated by EA Czyż with H Wulf, C Guillén Escribà, and FD Schneider. EA Czyż conducted the further curation of genetic and spectral data and led the development of the methodology to link spectral and genetic data with support from C Guillén Escribà, H Wulf and A Tedder. EA Czyż conducted the main analysis presented in this manuscript. The original draft of the manuscript with figures was written by EA Czyż with support from H Wulf, C Guillén Escribà, ME Schaepman, and MC Schuman. The revision of the manuscript was led by EA Czyż with support from MC Schuman, H Wulf, A Tedder and FD Schneider. Additional contributions not constituting authorship are acknowledged for the support of field sampling, and generation of APEX data products.

2.1 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 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.

2.2 Introduction

It has long been recognized that declining genetic variation within species is a key factor in biodiversity loss (Wilson et al., 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 et al., 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 and 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 et al., 2008; Szathmáry et al., 2001) and more frequent stochastic climatic events.

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 et al., 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 (H. M. Pereira et al., 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 and Martin, 2009; Martin et al., 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 and 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., 2015; Schaepman et al., 2009), with residual measurement noise (Hueni et al., 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 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.

2.3 Materials and Methods

2.3.1 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) (Fig. 2.1). 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., 2021). According to the United Nations Environment Programme (UNEP)-World Conservation Monitoring Centre (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 et al., 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.

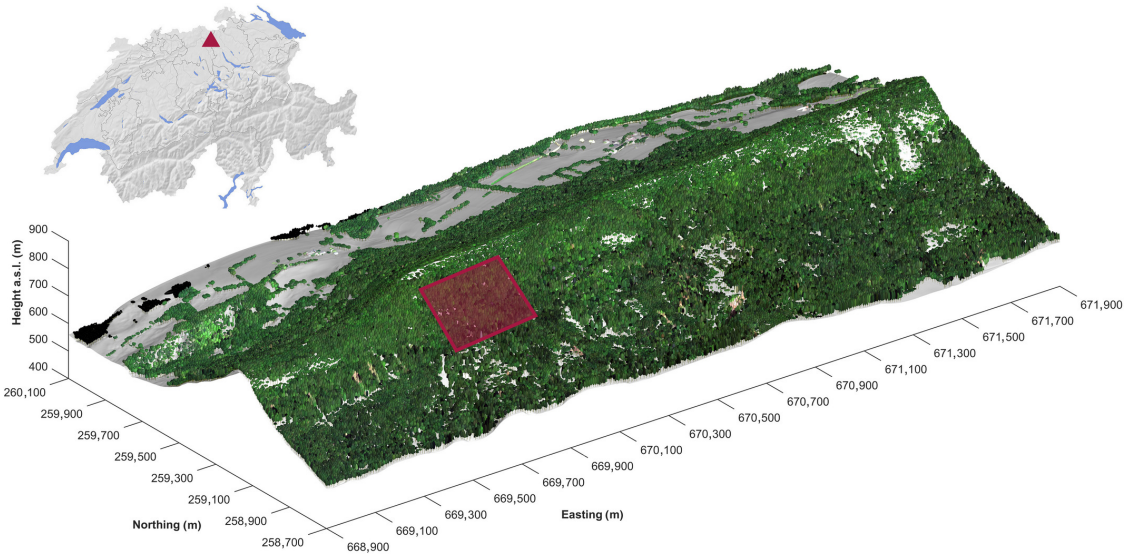

FIGURE 2.1: Location of the study site (rectangle). The temperate forest is located in northern Switzerland (triangle) on a portion of the south-facing slope of the Laegern mountain

2.3.2 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 et al., 2015). Tree crowns and tree trunks were mapped separately and linked to each other using photogrammetric approaches, due to slope geometries distorting nadir projection of crowns to trunks (Guillén-Escribà et al., 2021; Torabzadeh et al., 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, 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 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 (S.2.1). Population structure was inferred using a model-based Bayesian clustering approach implemented in the software *TESS2* (Durand et al., 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

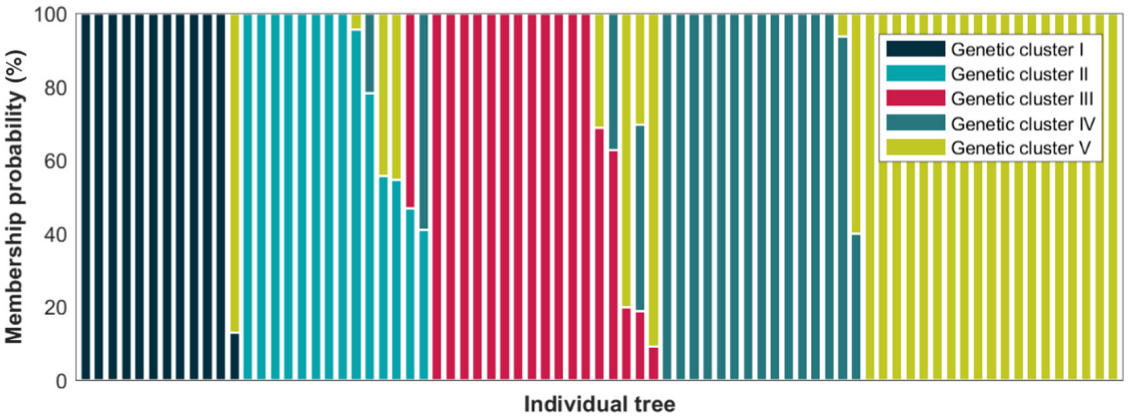

FIGURE 2.2: Membership probabilities for 77 sampled *Fagus sylvatica* individuals in five genetic clusters as determined by microsatellite analyses.

on an average cross-entropy of the admixture model, we determined five genetic clusters, to which each tree was assigned its membership probability (Fig. 2.2).

2.3.3 Optical Data

The spectral dataset contains seven acquisitions of the APEX 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 (S.2.2). 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 Atmospheric/Topographic Correction for Satellite Imagery (ATCOR) smile module, respectively (Hueni et al., 2008; Hueni et al., 2012; Richter et al., 2010). Calibrated and corrected radiance data were atmospherically corrected to surface reflectances in ATCOR (Hueni et al., 2017; Schläpfer and 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 and Rocchini, 2008; Stickler and Southworth, 2008). Thus, the pixels with a cumulative reflectance less than the 30^{th} 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 and Maurer, 2010). The object was defined as a tree crown for each sampled *F. sylvatica* individual; individuals were identified using images derived from LiDAR and high-resolution drone measurements resulting in a crown map of the site (Guillén-Escribà et al., 2021; Torabzadeh et al., 2019). 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 1^{st} 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: $(x_l - \mu_l)/\sigma_l$, 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 1^{st} 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 et al., 2016).

2.3.4 Statistical analysis

We investigated the relationship between imaging spectroscopy and genetic information by using PLS regression. This method is commonly used in chemometrics (Wold et al., 2001) and more generally to analyze datasets that are highly collinear and have a high ratio of independent variables to observations (Wold et al., 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 et al., 2004). The method is used in analyzing spectral data for classification and prediction purposes (Cavender-Bares et al., 2016; Lee et al., 2016; Peerbhay et al., 2013; Singh et al., 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 Predicted Residual Error Sum of Squares (PRESS) statistic (Chen et al., 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 et al., 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 n^{th} 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 2.3.3) 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.

2.4 Results

2.4.1 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 (Fig. 2.3).

In the reflectance dataset (dataset a, under 2.3.3), 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, $p < .001$).

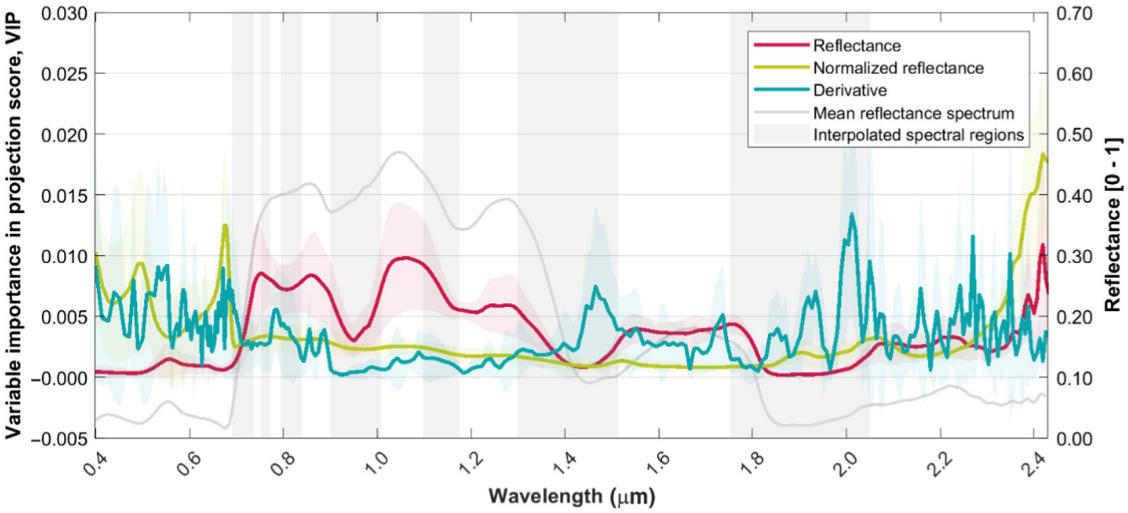

FIGURE 2.3: VIP scores of PLS models generated to predict genetic structure of *Fagus sylvatica* population in a temperate forest from airborne imaging spectroscopy data. The red, green, and blue solid lines represent the multi-year mean scores achieved from analyses of reflectance (reflectance), z-score 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.

2.4.2 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 (Fig. 2.4). 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 2.3.4) 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

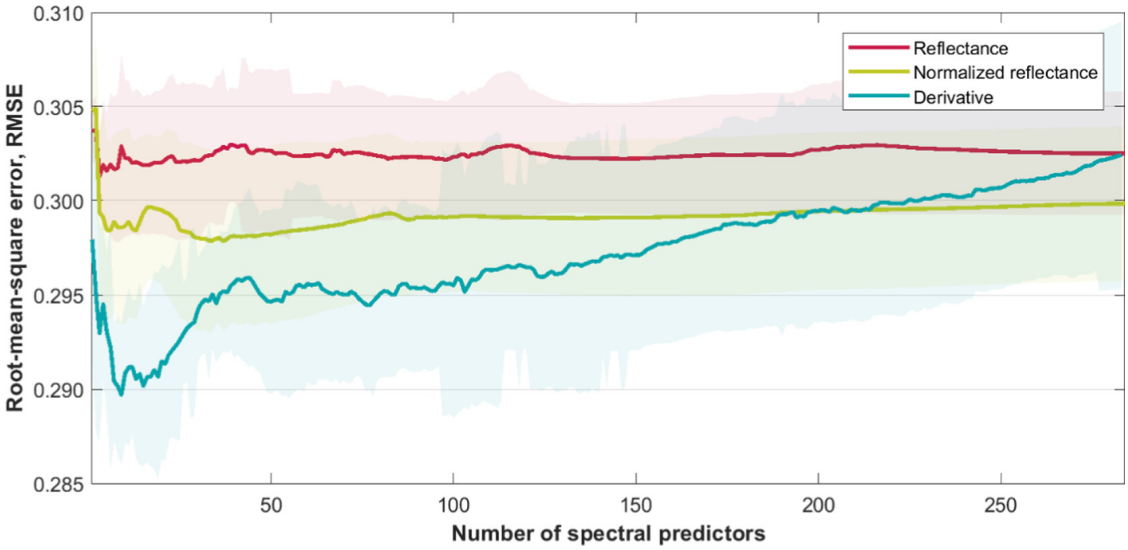

FIGURE 2.4: Root-mean-square error RMSE of genetic structure prediction from PLS models generated from an increasing number of most prominent (defined under 2.3.4) spectral predictors derived based on the VIP score of the models. The red, green, and blue solid lines represent multi-year mean RMSE achieved from analyses made on reflectance (reflectance), z-score of reflectance (normalized reflectance), and a 1^{st} 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

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 (Fig. 2.5).

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.

2.5 Discussion

2.5.1 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.

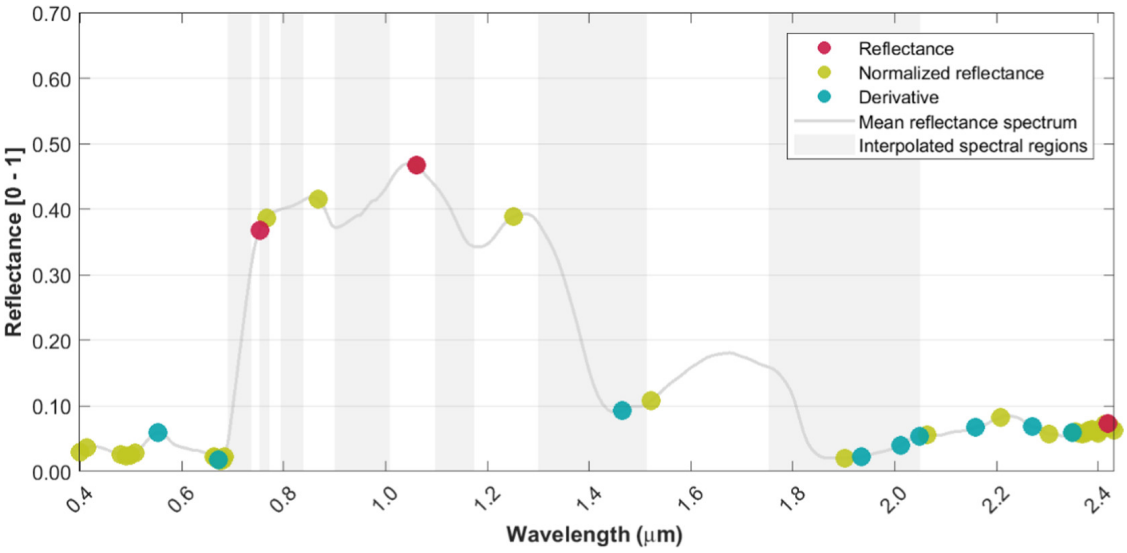

FIGURE 2.5: Spectral predictors selected based on the highest prominence of VIP scores (defined under 2.3.4), from which the genetic structure was predicted with the highest accuracy. The red, green, and blue refers to reflectance (reflectance), z-score of reflectance (normalized reflectance), and a 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

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 and 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 2.3.4) derived over the spectrum (Fig. S.2.1).

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, 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 et al., 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 et al., 1996) and differences in water management (Peuke et al., 2002). Further, phenolic compounds are related to many physiological reactions of plants, including protection against ultraviolet (UV) radiation (Close and McArthur, 2002), microbial (Scalbert, 1991), fungal (Telles et al., 2017), and herbivorous (War et al., 2012) attackers, as well as pollution (Pasqualini et al., 2003) and climatic responses (Stark et al., 2015). The variation in the abundance of phenols is in part under genetic control (J. A. Pereira et al., 2007) and is already used for remote taxonomic identification of trees (Asner et al., 2012), and within-species assessments. Additionally, pigments influence photosynthetic performance (Flexas et al., 2012; Flexas et al., 2008). This performance is also related to the physiological adaptation of the organism and therefore is subject to purifying selection (Arntz and Delph, 2001). Similarly, epicuticular wax content responds to environmental conditions (e.g., Schreiber et al., 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.

2.5.2 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 and Maurer, 2010; Nagendra and Rocchini, 2008; Stickler and 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.

2.5.3 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 and 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 and 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 et al., 2012), or Solar Induced 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 et al., 2018; Schneider et al., 2019) as well as airborne laser scanning (Kaartinen et al., 2012; Schneider et al., 2014; Wang et al., 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.

2.6 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.

Acknowledgments

This study is supported by the University of Zürich Research Priority Program on Global Change and Biodiversity (URPP GCB). The research carried out at the Jet Propulsion Laboratory, California Institute of Technology, was under a contract with the National Aeronautics and Space Administration (80NM0018D0004). Government sponsorship is acknowledged. We thank Eri Yamasaki and Kentaro Shimizu for support. We are also

grateful to Andreas Hueni, Daniel Schläpfer, and Qingyang He for processing parts of our datasets. We thank Gillian Milani and Chengxiu Li for fruitful discussions on early aspects of the manuscript. We acknowledge the comments of the reviewers helping to further improve the manuscript.

Conflict of interest

The authors declare no conflicts of interest.

Supplementary Material

TABLE S.2.1: Microsatellite length statistics for sampled *F. sylvatica* individuals derived from capillary electrophoresis using GeneMapper software.

Length Statistic	FS1-03	FS1-15	FS3-04	FS4-46	FCM5
standard deviation	4.63	8.07	1.73	22.74	12.51
mean	91.55	111.99	200.95	251.75	299.87
variance	21.39	65.09	2.99	517.23	156.45
maximum	108	137	206	328	322
minimum	83	93	194	221	280

TABLE S.2.2: Date, Day Of the Year (DOY) and CGDD for the acquisition of the seven years of Airborne Prism Experiment (APEX) Airborne Imaging Spectrometer (AIS) images.

Date	Day of the year (DOY)	Cumulative Growing Degree Days (CGDD)
17.06.2009	168	429
26.06.2010	177	368
16.06.2012	168	372
12.07.2013	193	461
18.07.2014	199	674
24.06.2015	175	438
07.07.2016	189	461

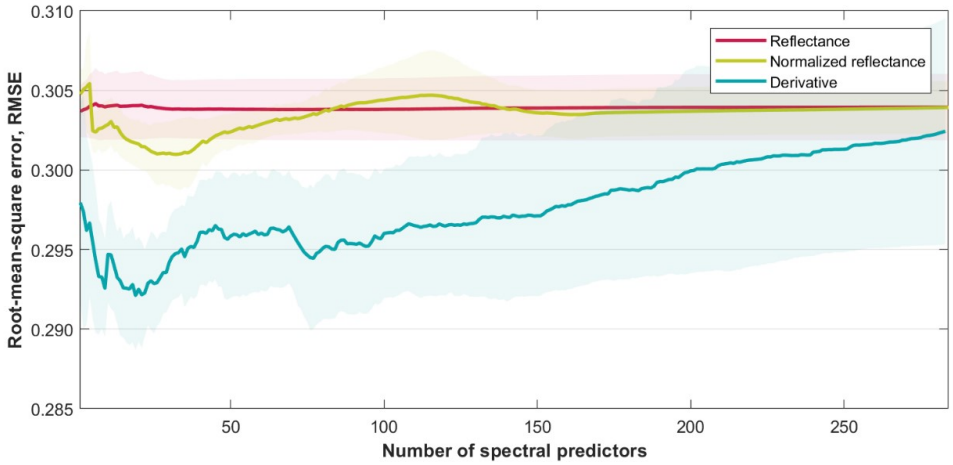

FIGURE S.2.1: Root-mean-square error RMSE of genetic structure prediction from PLS models generated from an increasing number of spectral predictors derived based on the highest VIP score of the models. The red, green and blue solid lines represent multi-year mean RMSE achieved from analyses made on 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.

References

Arntz, M. A., & Delph, L. F. (2001). Pattern and process: Evidence for the evolution of photosynthetic traits in natural populations. *Oecologia*, *127*, 455–467.

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.

Asner, G. P., Martin, R. E., & Suhaili, A. B. (2012). Sources of canopy chemical and spectral diversity in lowland bornean forest. *Ecosystems*, *15*, 504–517.

Bolnick, D. I., Amarasekare, P., Araújo, M. S., Bürger, R., Levine, J. M., Novak, M., Rudolf, V. H., Schreiber, S. J., Urban, M. C., & Vasseur, D. A. (2011). Why intraspecific trait variation matters in community ecology. *Trends in ecology & evolution*, *26*(4), 183–192.

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

Bush, A., Sollmann, R., Wilting, A., Bohmann, K., Cole, B., Balzter, H., Martius, C., Zlinszky, A., Calvignac-Spencer, S., Cobbold, C. A., et al. (2017). Connecting earth observation to high-throughput biodiversity data. *Nature ecology & evolution*, *1*(7), 0176.

Cavender-Bares, J., Meireles, J. E., Couture, J. J., Kaproth, M. A., Kingdon, C. C., Singh, A., Serbin, S. P., Center, A., Zuniga, E., Pilz, G., et al. (2016). Associations of leaf spectra with genetic and phylogenetic variation in oaks: Prospects for remote detection of biodiversity. *Remote Sensing*, *8*(3), 221.

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.

Close, D. C., & McArthur, C. (2002). Rethinking the role of many plant phenolics–protection from photodamage not herbivores? *Oikos*, *99*(1), 166–172.

Davies, N., Meyer, C., Gilbert, J. A., Amaral-Zettler, L., Deck, J., Bicak, M., Rocca-Serra, P., Assunta-Sansone, S., Willis, K., & Field, D. (2012). A call for an international network of genomic observatories (gos). *GigaScience*, *1*(1), 1–5.

Demesure, B., Comps, B., & Petit, R. J. (1996). Chloroplast dna phylogeography of the common beech (fagus sylvatica l.) in europe. *Evolution*, 2515–2520.

Des Roches, S., Post, D. M., Turley, N. E., Bailey, J. K., Hendry, A. P., Kinnison, M. T., Schweitzer, J. A., & Palkovacs, E. P. (2018). The ecological importance of intraspecific variation. *Nature ecology & evolution*, *2*(1), 57–64.

Doyle, J. J. (1990). Isolation of plant dna from faesh tissue. *Focus*, *12*, 13–15.

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.

Etzold, S., Ruehr, N. K., Zweifel, R., Dobbertin, M., Zingg, A., Pluess, P., Häsler, R., Eugster, W., & Buchmann, N. (2011). The carbon balance of two contrasting

mountain forest ecosystems in switzerland: Similar annual trends, but seasonal differences. *Ecosystems*, *14*, 1289–1309.

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.

Flexas, J., Barbour, M. M., Brendel, O., Cabrera, H. M., Carriquı, M., Dıaz-Espejo, A., Douthe, C., Dreyer, E., Ferrio, J. P., Gago, J., et al. (2012). Mesophyll diffusion conductance to co2: An unappreciated central player in photosynthesis. *Plant Science*, *193*, 70–84.

Flexas, J., Ribas-Carbo, M., Diaz-Espejo, A., Galmés, J., & Medrano, H. (2008). Mesophyll conductance to co2: Current knowledge and future prospects. *Plant, cell & environment*, *31*(5), 602–621.

Frankham, R. (2010). Challenges and opportunities of genetic approaches to biological conservation. *Biological conservation*, *143*(9), 1919–1927.

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.

Gausman, H. W. (1985). Plant leaf optical properties in visible and near-infrared light.

Geijzendorffer, I. R., Regan, E. C., Pereira, H. M., Brotons, L., Brummitt, N., Gavish, Y., Haase, P., Martin, C. S., Mihoub, J.-B., Secades, C., et al. (2016). Bridging the gap between biodiversity data and policy reporting needs: An essential biodiversity variables perspective. *Journal of Applied Ecology*, *53*(5), 1341–1350.

Gienapp, P., Teplitsky, C., Alho, J., Mills, J., & Merilä, J. (2008). Climate change and evolution: Disentangling environmental and genetic responses. *Molecular ecology*, *17*(1), 167–178.

Gratani, L. (2014). Plant phenotypic plasticity in response to environmental factors. *Advances in botany*, *2014*.

Guillén-Escribà, C., Schneider, F. D., Schmid, B., Tedder, A., Morsdorf, F., Furrer, R., Hueni, A., Niklaus, P. A., & Schaepman, M. E. (2021). Remotely sensed between-individual functional trait variation in a temperate forest. *Ecology and Evolution*, *11*(16), 10834–10867.

Hoban, S., Arntzen, J. A., Bruford, M. W., Godoy, J. A., Rus Hoelzel, A., Segelbacher, G., Vilà, C., & Bertorelle, G. (2014). Comparative evaluation of potential indicators and temporal sampling protocols for monitoring genetic erosion. *Evolutionary Applications*, *7*(9), 984–998.

Hueni, A., Biesemans, J., Meuleman, K., Dell'Endice, F., Schlapfer, D., Odermatt, D., Kneubuehler, M., Adriaensen, S., Kempenaers, S., Nieke, J., et al. (2008). 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.

Hueni, A., Sterckx, S., Jehle, M., D'Odorico, P., Vreys, K., Bomans, B., Biesemans, J., Meuleman, K., & Schaepman, M. (2012). Operational status of apex and characteristics of the apex open science data set. *2012 IEEE international geoscience and remote sensing symposium*, 5009–5012.

Hueni, A., Woolliams, E., Schlaepfer, D., & Wulf, H. (2017). Apex airborne imaging spectrometer uncertainty budget and vicarious validation method. *AGU Fall Meeting Abstracts*, *2017*, A11A–1860.

Huesca, M., Garcıa, 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.

Hughes, A. R., Inouye, B. D., Johnson, M. T., Underwood, N., & Vellend, M. (2008). Ecological consequences of genetic diversity. *Ecology letters*, *11*(6), 609–623.

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.

Jetz, W., Cavender-Bares, J., Pavlick, R., Schimel, D., Davis, F. W., Asner, G. P., Guralnick, R., Kattge, J., Latimer, A. M., Moorcroft, P., et al. (2016). Monitoring plant functional diversity from space. *Nature plants*, *2*(3), 1–5.

Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., et al. (2012). An international comparison of individual tree detection and extraction using airborne laser scanning. *Remote Sensing*, *4*(4), 950–974.

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*, 591–605.

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.

Kloeti, P., Keller, H., & Guecheva, M. (1989). Effects of forest canopy on throughfall precipitation chemistry. *Atmospheric Deposition. Proceedings of a Symposium held during the Third Scientific Assembly of the International Association of Hydrological Sciences at Baltimore, Maryland May 1989. IAHS Publication*, (179).

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*, S78–S91.

Kruse, F. A., Lefkoff, A., Boardman, J., Heidebrecht, K., Shapiro, A., Barloon, P., & Goetz, A. (1993). The spectral image processing system (sips)—interactive visualization and analysis of imaging spectrometer data. *Remote sensing of environment*, *44*(2-3), 145–163.

Lee, J., Cai, X., Lellmann, J., Dalponte, M., Malhi, Y., Butt, N., Morecroft, M., Schönlieb, C.-B., & 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.

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.

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.

Martin, M., 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.

Merton, R. (1998). Monitoring community hysteresis using spectral shift analysis and the red-edge vegetation stress index. *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.

Nagendra, H., & Rocchini, D. (2008). High resolution satellite imagery for tropical biodiversity studies: The devil is in the detail. *Biodiversity and conservation*, *17*, 3431–3442.

Navarro, L. M., Fernandez, N., Guerra, C., Guralnick, R., Kissling, W. D., Londono, M. C., Muller-Karger, F., Turak, E., Balvanera, P., Costello, M. J., et al. (2017). Monitoring biodiversity change through effective global coordination. *Current opinion in environmental sustainability*, *29*, 158–169.

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.

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.

Pastorelli, R., Smulders, M., Van't Westende, W., Vosman, B., Giannini, R., Vettori, C., & Vendramin, G. (2003). Characterization of microsatellite markers in fagus sylvatica l. and fagus orientalis lipsky. *Molecular Ecology Notes*, *3*(1), 76–78.

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.

Pereira, H. M., Ferrier, S., Walters, M., Geller, G. N., Jongman, R. H., Scholes, R. J., Bruford, M. W., Brummitt, N., Butchart, S. H., Cardoso, A., et al. (2013). Essential biodiversity variables. *Science*, *339*(6117), 277–278.

Pereira, J. A., Oliveira, I., Sousa, A., Valentão, P., Andrade, P. B., Ferreira, I. C., Ferreres, F., Bento, A., Seabra, R., & 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.

Peuke, A., Schraml, C., Hartung, W., & Rennenberg, H. (2002). Identification of drought-sensitive beech ecotypes by physiological parameters. *New phytologist*, *154*(2), 373–387.

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., Schläpfer, D., & Müller, A. (2010). Operational atmospheric correction for imaging spectrometers accounting for the smile effect. *IEEE Transactions on Geoscience and Remote sensing*, *49*(5), 1772–1780.

Rocchini, D., Balkenhol, N., Carter, G. A., Foody, G. M., Gillespie, T. W., He, K. S., Kark, S., Levin, N., Lucas, K., Luoto, M., et al. (2010). Remotely sensed spectral heterogeneity as a proxy of species diversity: Recent advances and open challenges. *Ecological Informatics*, *5*(5), 318–329.

Scalbert, A. (1991). Antimicrobial properties of tannins. *Phytochemistry*, *30*(12), 3875–3883.

Schaepman, M. E., Jehle, M., Hueni, A., D'Odorico, P., Damm, A., Weyermann, J., Schneider, F. D., Laurent, V., Popp, C., Seidel, F. C., et al. (2015). Advanced radiometry measurements and earth science applications with the airborne prism experiment (apex). *Remote Sensing of Environment*, *158*, 207–219.

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.

Schläpfer, D., & Richter, R. (2002). Geo-atmospheric processing of airborne imaging spectrometry data. part 1: Parametric orthorectification. *International Journal of Remote Sensing*, *23*(13), 2609–2630.

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.

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.

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.

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.

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.

Schweiger, A. K., Cavender-Bares, J., Townsend, P. A., Hobbie, S. E., Madritch, M. D., Wang, R., Tilman, D., & 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.

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., Mücher, C. A., O'Connor, B., Paganini, M., Pereira, H. M., et al. (2015). Environmental science: Agree on biodiversity metrics to track from space. *Nature*, *523*(7561), 403–405.

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.

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*, 1993–2005.

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.

Szathmáry, E., Jordán, F., & Pál, C. (2001). Can genes explain biological complexity? *Science*, *292*(5520), 1315–1316.

Telles, A. C., Kupski, L., & Furlong, E. B. (2017). Phenolic compound in beans as protection against mycotoxins. *Food Chemistry*, *214*, 293–299.

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.

Tucker, C. J. (1980). Remote sensing of leaf water content in the near infrared. *Remote sensing of Environment*, *10*(1), 23–32.

Turner, W. (2014). Sensing biodiversity. *Science*, *346*(6207), 301–302.

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.

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*, S67–S77.

Valbuena, R., O'Connor, B., Zellweger, F., Simonson, W., Vihervaara, P., Maltamo, M., Silva, C. A., Almeida, D. R. A. d., Danks, F., Morsdorf, F., et al. (2020). Standardizing ecosystem morphological traits from 3d information sources. *Trends in Ecology & Evolution*, *35*(8), 656–667.

Vihervaara, P., Auvinen, A.-P., Mononen, L., Törmä, M., Ahlroth, P., Anttila, S., Böttcher, K., Forsius, M., Heino, J., Heliölä, J., et al. (2017). How essential biodiversity variables and remote sensing can help national biodiversity monitoring. *Global Ecology and Conservation*, *10*, 43–59.

Wang, Y., Hyyppä, J., Liang, X., Kaartinen, H., Yu, X., Lindberg, E., Holmgren, J., Qin, Y., Mallet, C., Ferraz, A., et al. (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.

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.

Wilson, E. O., et al. (1988). Biodiversity.

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., Ruhe, A., Wold, H., & Dunn, W., Iii. (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.

Wold, S., Sjöström, M., & Eriksson, L. (2001). Pls-regression: A basic tool of chemometrics. *Chemometrics and intelligent laboratory systems*, *58*(2), 109–130.

Yamasaki, E., Altermatt, F., Cavender-Bares, J., Schuman, M. C., Zuppinger-Dingley, D., Garonna, I., Schneider, F. D., Guillén-Escribà, C., van Moorsel, S. J., Hahl, T., et al. (2017). Genomics meets remote sensing in global change studies: Monitoring and predicting phenology, evolution and biodiversity. *Current opinion in environmental sustainability*, *29*, 177–186.

Chapter 3

Genetic constraints on temporal variation of airborne reflectance spectra and their uncertainties over a temperate forest

Ewa A. Czyż, Bernhard Schmid, Andreas Hueni,
Maarten B. Eppinga, Meredith C. Schuman, Fabian D. Schneider,
Carla Guillén Escribà, Michael E. Schaepman

This chapter is based on the peer-reviewed article:
Genetic constraints on temporal variation of airborne reflectance spectra and their uncertainties over a temperate forest.
Remote Sensing of Environment, 2023, 284, 113338.
DOI: 10.1016/j.rse.2022.113338.
It is reprinted as the final submitted manuscript and has been modified to fit the layout of this thesis.

Author contributions. Conceptualization: EAC, BS, MBE, MES. Methodology: EAC, BS, AH, MBE, FDS. Formal analysis: EAC, AH. Investigation: EAC. Resources: MES. Data Curation: EAC, AH, FDS, CGE. Writing - Original Draft: EAC. Writing - Review and Editing: EAC, BS, AH, MBE, MCS, FDS, MES. Visualization: EAC. Supervision: MES. Project administration: MES. Funding acquisition: MES.

Detailed author contributions

EA Czyż conceptualized the study together with ME Schaepman and further contributions of B Schmid, MB Eppinga and MC Schuman. ME Schaepman obtained funding, administrated and supervised the work. C Guillén Escribà derived the georeferenced tree crown map supported by FD Schneider. Airborne imaging spectrometer data were collected with the APEX during yearly campaigns administrated by ME Schaepman and airborne data were processed to reflectance data with their radiometric uncertainties by A Hueni. EA Czyż with B schmid, A Hueni, MB Eppinga and FD Schneider established methodology. EA Czyż derived genetic and spectral distances and incorporated and propagated uncertainties through the analyses, with the latter being validated by MB Eppinga and A Hueni. MB Eppinga and B Schmid validated Partial Mantel correlation, which was the main approach to derive presented spectral-genetic similarities with their uncertainties. The original draft of the manuscript with figures was written by EA Czyż with edits from B Schmid, A Hueni, MC Schuman, MB Eppinga and ME Schaepman. The revision of the manuscript was led by EA Czyż with support from ME Schaepman, MC Schuman, MB Eppinga and B Schmid. Additional contributions not constituting authorship are acknowledged for the support of field sampling, genetic and climatic data processing and generation of APEX data products.

3.1 Abstract

Remote sensing enhances large-scale biodiversity monitoring by overcoming temporal and spatial limitations of ground-based measurements and allows assessment of multiple plant traits simultaneously. The total set of traits and their variation over time is specific for each individual and can reveal information about the genetic composition of forest communities. Measuring trait variation among individuals of one species continuously across space and time is a key component in monitoring genetic diversity but difficult to achieve with ground-based methods. Remote sensing approaches using imaging spectroscopy can provide high spectral, spatial, and temporal coverage to advance the monitoring of genetic diversity, if sufficient relation between spectral and genetic information can be established.

We assessed reflectance spectra from individual *Fagus sylvatica* L. (European beech) trees acquired across eleven years from 69 flights of the Airborne Prism Experiment (APEX) above the same temperate forest in Switzerland. We derived reflectance spectra of 68 canopy trees and correlated differences in these spectra with genetic differences derived from microsatellite markers among the 68 individuals. We calculated these correlations for different points in time, wavelength regions and relative differences between wavelength regions. High correlations indicate high spectral-genetic similarities. We then tested the influence of environmental variables obtained at temporal scales from days to years on spectral-genetic similarities. We performed an uncertainty propagation of radiance measurements to provide a quality indicator for these correlations.

We observed that genetically similar individuals had more similar reflectance spectra, but this varied between wavelength regions and across environmental variables. The shortwave infrared regions of the spectrum, influenced by water absorption, seemed to provide information on the population genetic structure at high temperatures, whereas the visible part of the spectrum, and the near-infrared region affected by scattering properties of tree canopies, showed more consistent patterns with genetic structure across longer time scales. Correlations of genetic similarity with reflectance spectra similarity were easier to detect when investigating relative differences between spectral bands (maximum correlation: 0.40) than reflectance data (maximum correlation: 0.33). Incorporating uncertainties of spectral measurements yielded improvements of spectral-genetic similarities of 36% and 20% for analyses based on single spectral bands, and relative differences between spectral bands, respectively.

This study highlights the potential of dense multi-temporal airborne imaging spectroscopy data to detect the genetic structure of forest communities. We suggest that the observed temporal trajectories of reflectance spectra indicate physiological and possibly genetic constraints on plant responses to environmental change.

3.2 Introduction

Remote sensing has demonstrated its potential for large-scale biodiversity monitoring (e.g. O'Connor et al., 2015; Pettorelli et al., 2014; Skidmore et al., 2021; Skidmore et al., 2015; Turner, 2014). It is especially valuable in the context of current environmental change affecting global biodiversity decline (Blowes et al., 2019; Proença and Pereira, 2017), as it overcomes spatial and temporal limitations of ground-based assessments of biological variation (Gamon et al., 2019; Jetz et al., 2016). Remote assessments of functional diversity (Asner and Martin, 2009; Schimel et al., 2019; Schneider et al., 2017; Singh et al., 2015; Z. Wang et al., 2019; Zheng et al., 2021) may advance our understanding of ecosystem functioning and may be used for ecological modeling. Genetic diversity is an equally important facet of biodiversity, but has received less attention in remote sensing studies.

Genetic diversity is a key component of biodiversity loss (Wilson and Peter, 1988) and has been defined as an Essential Biodiversity Variable (EBV) (Pereira et al., 2013; Skidmore et al., 2021). Greater genetic variation reflects the availability of more alleles, or versions of genes, resulting in a wider range of phenotypic responses and thus increased potential to maintain ecosystem functioning and evolutionary processes under changing environmental conditions (Bolnick et al., 2011; Des Roches et al., 2018; Hughes et al., 2008; Moritz, 2002). Monitoring current genetic diversity and its distribution across populations of interbreeding individuals, and understanding what causes this to change, is essential for preventing further biodiversity loss. Measurement of allele pools is commonly based on field sampling followed by laboratory analyses, which both require substantial time and effort per sample. Thus, the assessment of genetic diversity would benefit from approaches that overcome scale-induced limitations.

Remote sensing is being used to assess phylogenetic diversity in plant communities in direct (Cavender-Bares, 2019; Cavender-Bares et al., 2016; Lin et al., 2021) and indirect ways (Feret and Asner, 2012; Frye et al., 2021; Rocchini et al., 2010; Schweiger

et al., 2018). These assessments are done at species or higher taxonomic levels, while assessments of intraspecific genetic variation are more challenging because trait differences expressed in spectra are likely to be less pronounced at the within- than at the between-species level (Guillén-Escribà et al., 2021; Hulshof and Swenson, 2010). Dense temporal coverage of imaging spectroscopy acquisitions may allow detection of more differences among individuals — thus helping to resolve intraspecific variation — if traits expressed at different times or ontogenetic stages are underpinned by different genes within an individual and if thus a link between temporally variable spectral and temporally constant genetic variation can be established (Czyż et al., 2020; Yamasaki et al., 2017).

Remote sensing allows for systematic monitoring of the land surface on a wide range of temporal scales. Time series of remotely sensed data reveal information about the dynamic behavior of vegetation and have significant advantages over single time-point measurements (Goetz, 2009), which tend to be a limitation of ground-based investigations (Albert et al., 2011). Repeat-visit airborne imaging spectroscopy data can, for example, capture seasonal changes in vegetation spectra. Changes in spectra over time may reflect leaf green-up (Garcia and Ustin, 2001), leaf ontogenetic stages (Chavana-Bryant et al., 2017), mesophyll development (Rivera et al., 2002), seasonal changes in equivalent water thickness (Roberts et al., 1997), lignin and cellulose turnover (Patton and Gieseker, 1942), and start of senescence (Myneni et al., 1997). Reflectance spectra can be used to assess plant water stress (Suárez et al., 2008), and pigment (Ustin et al., 2009) and nitrogen concentration in vegetation (Oldeland et al., 2010; Ollinger et al., 2002). These are all traits, or observable features of individuals, which may be genetically constrained, meaning that they have a genetic basis and thus their variation may in part be explained by genetic variation. Therefore, time-resolved imaging spectroscopy data may help to remotely sense genetic variation even within a species, as it has already been shown to support the identification of vegetation types (Merton, 1998) and plant species (e.g. Asner et al., 2006; Castro-Esau et al., 2006; Somers and Asner, 2013). Dense temporal coverage allows the assessment of plant responses to environmental conditions, which drive adaptation in species (Hut et al., 2012). Some adaptations may be observed in so-called temporal niche partitioning by individuals with varying strategies defined by traits, and relationships between traits, which vary in different ways over time (Angert et al., 2009; Levine and HilleRisLambers, 2009).

Dense multi-temporal imaging spectroscopy data with their rich spectral information, can simultaneously and systematically assess multiple plant traits and relations among

them. Relationships between plant traits are recognized across levels of biological organization (Joswig et al., 2022). On the elementary level, stoichiometry explains the balance among chemical elements as the basic components of living organisms. The specific composition of chemical elements in plants, in the form of molecular building blocks but also as nutrients and cofactors, is related to their morphological and physiological traits. These are in turn related to each other especially in terms of traits describing size or allometry, versus economic investments of nutrients over a plant's lifetime (eg. Wright et al., 2004), and describe a spectrum of form and function across plants (Dıaz et al., 2016). For example, carbon fixed by photosynthesis is allocated differently into shoots and roots, with trade-offs between these, and may be used for energy or energy storage in the form of nonstructural carbohydrates, for structural changes like lignification, or defensive measures such as phenol production (Eichenberg et al., 2015). The set of traits ultimately expressed by individuals builds on specific ecological strategies of plants (Westoby, 1998) and affects ecosystem processes (Marks and Lechowicz, 2006; Niinemets, 2001; Reich et al., 1998; Wright et al., 2003). Natural selection favors strategies which optimize fitness in a given environment, some of which vary within species (Albert et al., 2011; He et al., 2009), while others are characteristic of phylogenetic groups (eg. Kerkhoff et al., 2006; Westoby and Wright, 2006), or of convergent evolution across phylogenies (Gamon et al., 2019; Reich et al., 2003; Ustin et al., 2009). Expressed strategies characteristic for plant phylogenetic groups could be derived by combining information derived from spectral indices such as e.g. Normalized Difference Vegetation Index (NDVI) (Rouse et al., 1974), Normalized Difference Water Index (NDWI) (Gao, 1996) or Photochemical Reflectance Index (PRI) (Gamon et al., 1992), partial least squares regression PLS regression models (e.g. Singh et al., 2015) or potentially also the relationships within the full range of reflectance spectra.

In this study, we present a proof of principle for the use of dense multi-temporal airborne imaging spectroscopy data to investigate intraspecific genetic variation of temperate forest trees. We link multi-temporal aerial imaging spectroscopy with data on the genetic structure of a *Fagus sylvatica* L. (European beech) population derived from five nuclear microsatellites. We hypothesize that genetically similar individuals express similar phenotypes as detected in reflectance spectra at different time points. We investigate (1) which of the time-resolved spectral feature(s) reveal the highest similarity with genetic structure, (2) to what extent relationships between spectral features are genetically constrained, and (3) under which environmental conditions and at what temporal scales

genetic-spectral similarities are largest. Further, we suggest underlying physiological processes that may correlate with genetic structure, and we investigate to what extent the uncertainties associated with the spectral signal propagate to correlations.

We take advantage of detailed in situ and airborne mapping of trees at the study site (Morsdorf et al., 2020). We use genetic markers of intraspecific genetic variation consisting of lengths of five nuclear microsatellite markers, long-term measurements of environmental variables at the study site, and available high-resolution spectral data required for time-series analysis of vegetation spectral measurements. Furthermore, we account for inherent uncertainties of spectral measurements (Schaepman and Dangel, 2000; Schaepman et al., 2015) and propagate them through the analysis. This approach enables us to assess reliability of information derived from spectra (Woolliams et al., 2014), which is of importance in the context of multi-temporal investigations and coordination of upcoming imaging spectroscopy missions and retrievals (Cawse-Nicholson et al., 2021). With that, we hope to contribute to developing scalable approaches to monitor and understand biodiversity change on its basic, intraspecific genetic level.

3.3 Materials and methods

3.3.1 Study area

The study area ('Lägern') covers 12.6 ha of semi-natural temperate mixed forest located on an up to 60°-steep south-facing slope of the Lägern mountain in Switzerland (47°28'N, 8°21'E). Its elevation ranges from 620 to 810 m a.s.l. 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). The vegetation cover is classified as Temperate Deciduous Broadleaf Forest (UNEP-WCMC) with 13 tree species including three conifers and ten angiosperms, with *Fagus sylvatica* L. (European beech) being the most abundant (Guillén-Escribà et al., 2021). Mean height of canopy trees at the site is 30.6 m and the tree age ranges between 53 and 185 years, with a maximum diameter at breast height of 150 cm (Eugster et al., 2007). The canopy is mostly closed and characterized by a complex vertical structure (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 et al., 1989). Individual trees have been reconstructed in 3D using ground and airborne laser scanning, and simulated using 3D radiative transfer models, allowing the validation of airborne data with high accuracy (Kükenbrink et al., 2021; Schneider

et al., 2014; Schneider et al., 2017). The tree trunks were previously georeferenced using a differential Global Positioning System (GPS) system and tied to reference points of the Swiss national grid (LV95), with a polygon traverse and additional fixed points to measure trunk locations in the challenging terrain (Leiterer et al., 2015). Because the slope geometry distorted the nadir projection of crowns to trunks, tree crowns were mapped separately and linked to tree trunks using photogrammetry (Guillén-Escribà et al., 2021; Torabzadeh et al., 2019). The mapped tree crowns of the site allowed us to assign reflectance spectra of multiple pixels and genetic information to 68 analyzed *F. sylvatica* individuals.

3.3.2 Spectral data

The spectral dataset contains 87 acquisitions of the Airborne Prism Experiment (APEX) Airborne Imaging Spectrometer (AIS) (Schaepman et al., 2015) obtained between 2009 and 2019 and covering the range of 23–1160 Cumulative Growing Degree Days [°C] (CGDD) (S.3.2). Preprocessing of the raw APEX data including radiometric calibration and correction for spectral shifts was done using the APEX processing and archiving facility (Hueni et al., 2008; Hueni et al., 2012; Richter et al., 2010; Schaepman et al., 2015). The adjustment for smile effects and atmospheric compensation for surface reflectance accounting for solar zenith and azimuth and viewing angles was done using ATCOR, and the reflectance data were georectified in Parametric Geocoding and Orthorectification for Airborne Optical Scanner Data (PARGE) (Hueni et al., 2017; Schläpfer and Richter, 2002). This resulted, for each acquisition, in a georectified reflectance spectra dataset containing 284 spectral bands with a variable spectral resolution of 7–13.5 nm in the wavelength range from 399 to 2425 nm, with 2 m spatial resolution at canopy level. For the analyses, we selected only acquisitions with NDVI >0.5 and with the area of interest being fully covered. This resulted in a final set of 69 cloud-free acquisitions obtained during 27 airborne missions acquired under leaf-on conditions (Fig. 3.1).

Along with the reflectance spectra data, we obtained data about radiometric uncertainties. The uncertainties were calculated for each pixel and each spectral band based on the radiometric calibration of APEX (Hueni et al., 2017; Jehle et al., 2014; Schläpfer et al., 2000; Schläpfer and Richter, 2002). They account for system noise derived from multiple measurements using an integrating sphere, pressure changes affecting the spectral positions of derived bands, sensitivity to sensor temperature, smear effect, spectral

band binning, electronic anomalies, and estimation errors of radiometric gain and offset (Hueni et al., 2017).

For analysis, we derived reflectance spectra and radiometric uncertainties of pixels that represent the 68 canopy individuals of *F. sylvatica*. For each tree, we selected only those pixels whose centroid lay within the area of the mapped tree crown (see section 3.3.1) and we used average values of reflectance spectra weighted by their inverse uncertainty per tree individual per day (see section 3.3.6).

3.3.3 Genetic data

Relatedness of individuals was assessed using five nuclear microsatellite markers, which are repetitive DNA sequences commonly used for population genetics, from the 68 canopy individuals of *F. sylvatica* as detailed in a previous study (Czyż et al., 2020). Briefly, we sampled leaf disks (diameter 1.15 cm) from each tree in September 2013 and stored them on silica gel. Genomic DNA from the sampled material was extracted using the cetyl trimethylammonium bromide (CTAB) method following the procedure of Doyle, 1990. From the extracted DNA, we amplified five highly variable microsatellite loci (FS1–03, FS1–15, FS3–04, FS4–46, FCM5; Pastorelli et al., 2003). To assess the degree of polymorphism at each microsatellite locus, we performed capillary electrophoresis on an ABI-3720 sequencer (Thermo Fisher Scientific, UK) and we used *GeneMapper* software (Applied Biosystems) to determine the length of the analyzed microsatellites for each sampled tree (S.3.2).

3.3.4 Climatic data

We used climatic data collected on a flux tower 47 m vertically above ground level at the site (Eugster et al., 2007). The data cover the 11-year period 2009–2019 representing measurements at 30-min resolution.

The data used in the present study included air temperature [°C] and Vapor Pressure Deficit (VPD) at different temporal scales as follows. For each airborne acquisition, we calculated the mean temperature (TMP) and mean VPD of the day, aggregated temperature (11TMP) and aggregated Vapor Pressure Deficit (11VPD) over 11 consecutive days before the acquisition, and Cumulative Growing Degree Days (CGDD) and Cumulative Growing Vapor Pressure Deficit Days (CGVPDD). The moderate time scale of 11 days was selected based on a relatively high impact of selected climatic variables on spectral-genetic similarities across all the spectral bands at this time span (Supplementary, 3.7).

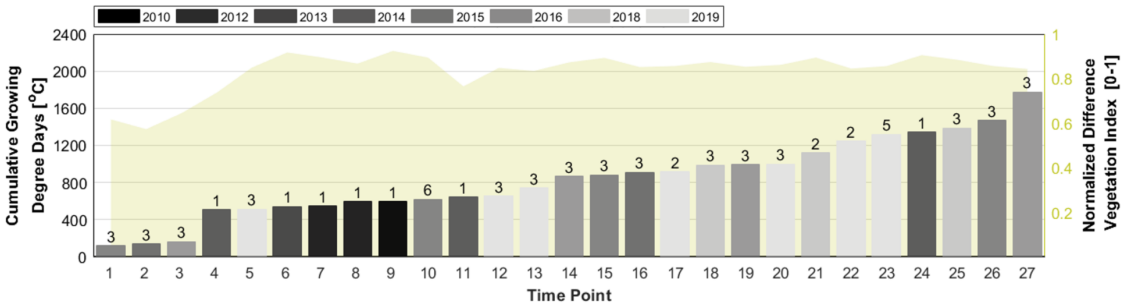

FIGURE 3.1: Airborne imaging spectroscopy missions segregated into time points based on Cumulative Growing Degree Days [°C] (CGDD) and shaded by year of acquisition. Each CGDD bar corresponds to a single day of acquisition. A base temperature of 5 °C was used (see section 3.3.4). The numbers on top of the bars specify the number of acquisitions at that day (multiple acquisitions per day). The background represents the NDVI calculated for each time point.

The CGDD were calculated with a base temperature of 5 °C (Polgar and Primack, 2011). Additionally, we included the Day of the Last Spring Frost (DLSF) [DOY] and Last Year Climatic Water Balance (LYCWB) of March-August calculated as monthly potential evapotranspiration (Hargreaves, 1994) subtracted from the monthly sum of precipitation.

3.3.5 Spectral-genetic similarity estimation

To investigate the correlation between reflectance spectra and genetic structure, we conducted partial Mantel tests between spectral and genetic distances among the 68 individual trees. We refer to the derived correlations as spectral-genetic similarities.

Firstly, we calculated the pairwise distances between trees using (1) the spectral Euclidean distance of each spectral band from i=1 to i=284 for each of the 27 time points (average of multiple pixels and multiple acquisitions per tree per day, see 3.3.2 and 3.3.6), (2) the spectral Euclidean distance of relative differences, d_n between wavelengths calculated for all pairwise combinations of spectral bands for each time point as (3.1):

$$d_n = \frac{|x_i - x_j|}{x_i + x_j} \tag{3.1}$$

where x is the reflectance value of the wavelength at spectral bands i and j respectively, (3) the spatial Euclidean distance based on the geographical position of the stem bases of the individuals, and (4) Nei's genetic distance representing the relationship between individuals based on their genetic divergence from a common ancestor, calculated based on the length of analyzed microsatellites of each *F. sylvatica* individual. We decided to introduce relative differences between wavelengths, because this normalization (1) reduced the impact of brightness, especially important in multi-temporal analysis, which

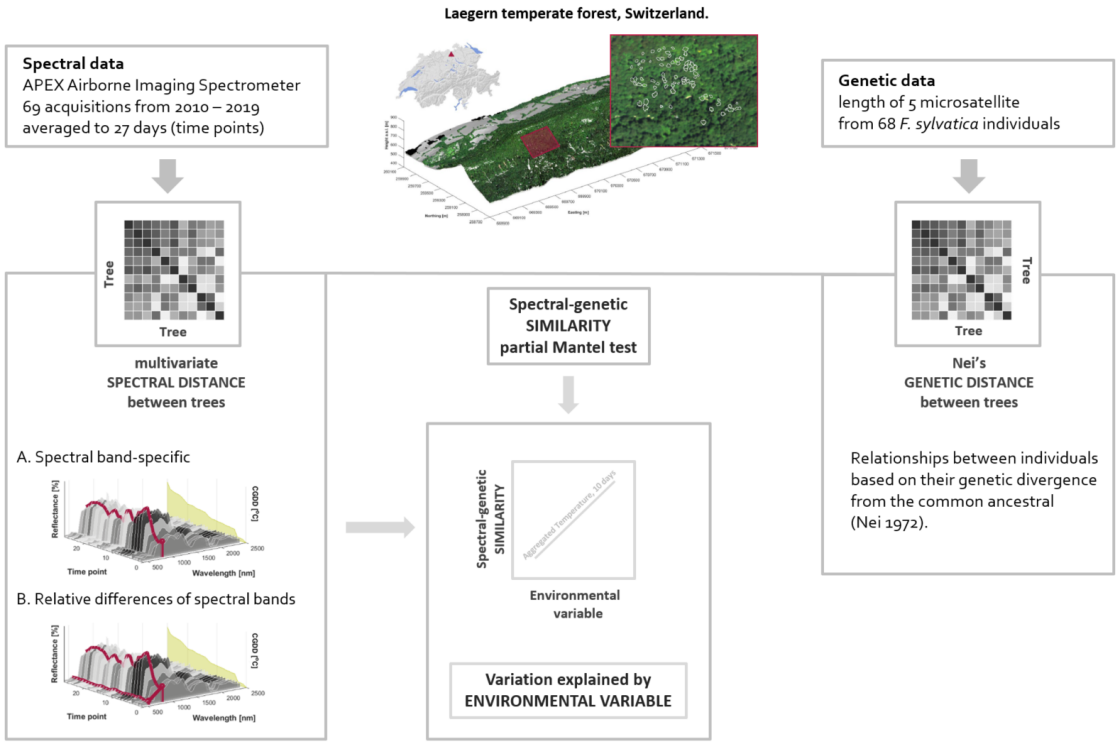

FIGURE 3.2: Schematic representation of the workflow. Spectral-genetic similarities were calculated as similarities between distance matrices derived from multi-temporal spectral and genetic data. The spectral data were considered as a single spectral band and as relative differences between the spectral bands, represented with one and two red lines overlaid on spectral responses averaged per day and shaded by year of acquisition. The Cumulative Growing Degree Days [°C] (CGDD) of each acquisition is represented with green shaded area. The spectral-genetic similarities were finally correlated with environmental variables.

entails variation in sun illumination and geometry of acquisition, and (2) simultaneously accounted for interactions among multiple traits that affect the spectral signatures of analyzed trees and may be indicative about their biological strategies. Nei's genetic distance is used to compare populations where drift and mutation may explain genetic differences (Nei, 1972) and we calculated it using the *nei.distance* function of the R package *poppr* (Kamvar et al., 2014, R v.8.5). This resulted in (1) 284 × 27, (2) (284×(284–1)/2) × 27, (3) one and (4) one symmetric matrices of dimension 68 × 68 representing pairwise distances between the 68 trees.

We then performed partial Mantel tests (Mantel, 1967; Smouse et al., 1986) to derive the spectral-genetic similarities between each spectral distance matrix of each time point and the genetic distance matrix. We assume that trees which are spatially closer together may be exposed to a more similar environment and we thus controlled for spatial distances between trees at the study site by subtracting the correlation of residuals from the linear regression of spectral-spatial correlation and genetic-spatial correlations (Smouse et al., 1986). We calculated partial Mantel correlations in *Matlab* v. R2017b software.

Finally, we calculated to what extent variation in spectral-genetic similarities was explained by variation of different environmental variables at different temporal scales. For that, we calculated the Pearson correlations of time point-specific spectral-genetic similarities and time point-specific environmental variables that represented daily, multi-daily, seasonal and inter-seasonal environmental conditions (Fig. 3.2).

For clarity of representation, we do not visualize the spectral-genetic similarities lower than 0, as such representa negative dependency of spectral signatures on genetic structure (genetically more related individuals being more dissimilar spectrally at these regions), which can not be used to infer genetic constraints. In our analysis, negative spectral-genetic similarities account for 30.7% of results mean of negative correlations 0.05 and lowest value 0.15 with $p = 1$. Furthermore, we discuss only the spectral-genetic similarities and their correlations with environmental variables, which are significant at $p < 0.05$ and of which the relative uncertainties do not exceed 20%.

3.3.6 Uncertainty propagation

Each analyzed tree individual in the spectral data comprises multiple pixels represented as surface reflectance values with their uncertainties of radiance measurements (see 3.3.2). We accounted for the uncertainties as relative reliability of measurements and propagated them to the spectral-genetic similarities and their correlations with environmental variables. More specifically, we incorporated the uncertainties throughout the analysis at two stages. In the first stage, we applied pixel-based uncertainty, u_i (where i stands for pixel within a tree crown) to derive the mean weighted by the inverse uncertainty, $\bar{x}_t$ (3.2) of spectral reflectance with its uncertainty, $u_{\bar{x}_t}$ (3.3) of each tree crown per acquisition.

$$\bar{x}_t = \sum \left(\frac{x_i}{u_i^2} / \sum_i \frac{1}{u_i^2} \right) \tag{3.2}$$

$$u_{\bar{x}_t} = \sqrt{\frac{1}{\sum_i u_i^2}} \tag{3.3}$$

Following the same operations, we calculated the mean weighted by the inverse uncertainty, $\bar{x}_d$ (3.4) of spectral reflectance with its uncertainty, $u_{\bar{x}_d}$ (3.5) of each tree crown per day as the unit of further analyses:

$$\bar{x}_d = \sum \left(\frac{\bar{x}_{t_j}}{u_{\bar{x}_{t_j}}^2} / \sum_t \frac{1}{u_{\bar{x}_{t_j}}^2} \right) \tag{3.4}$$

$$u_{\bar{x}_d} = \sqrt{\frac{1}{\sum_j u_{\bar{x}_{t_j}}^2}} \tag{3.5}$$

where j stands for the acquisition within a day sorted by CGDD.

In the second stage, we propagated the uncertainty of each tree crown per day to derive uncertainties of spectral-genetic similarities and their correlations with environmental variables. For that, we followed the law of propagation of uncertainties (after Joint Committee for Guides in Metrology (JCGM)), assuming the uncertainties to not co-vary, to sequentially calculate absolute uncertainties of pairwise Euclidean distances, z-score of Euclidean distances, correlation coefficients, partial Mantel correlations, relative differences, means of partial Mantel correlations, standard deviations of partial Mantel correlations. The detailed description of the uncertainty propagation as sensitivity analysis of each function to the uncertainty of its driver is described in the supplementary data (Supplementary, 3.7). We represent the derived uncertainties as relative uncertainties to the spectral-genetic similarities and their correlations with environmental variables.

3.4 Results

3.4.1 Spectral-genetic similarities

The similarities between reflectance spectra and genetic structure (partial Mantel statistic r) varied across single spectral bands and time points (Fig. 3.3). The highest spectral-genetic similarity reached 0.33 (relative uncertainty: 5%, $p = 0.01$) for a spectral band centered at 1925 nm at the 19$^{\text{th}}$ time point (993 CGDD, July) and was 36% higher than the highest result (0.21) derived from analyses where the mean reflectance spectra per tree were not weighted by the measurement uncertainty (not shown). Across all time points, the spectral-genetic similarities derived from the spectral regions ranging from 650–680 nm and 1950–2000 nm were relatively stable over time, the rest differed among time points. Across the whole spectrum, the 14$^{\text{th}}$ (866 CGDD, July) time point had the highest (0.10) spectral-genetic similarities. The beginning of the growing season (116–515 CGDD, April until beginning of July) was characterized by relatively low spectral-genetic similarities for spectral bands ranging from 1100–2500 nm. Towards the end of the growing season (1317–1776 CGDD, end of July untill September) spectral-genetic similarities were relatively high for spectral bands ranging from 680–1800 nm. Time points which included of at least five acquisitions — the 10$^{\text{th}}$ and the 23$^{\text{rd}}$ (see

Fig. 3.1) — had higher spectral-genetic similarities derived from spectral bands ranging from 550–650 and 690–1350 nm, respectively. Additional high spectral-genetic similarities occurred at particular time points: at the second time point, for spectral bands ranging from 400–500 nm; at time point 14, for spectral bands around 500 nm; at the time points 10, 12, 20 and 27 for spectral bands around 500–600 nm; at time point 26 for spectral bands around 600–680 nm, and time point 14 for spectral bands around 680–1350 nm, at time point six for spectral bands around 1400–2300 nm and at time points 26 and 27 for spectral bands around 2350–2450 nm. Across the spectrum, the lowest spectral-genetic similarities overtime were found for spectral bands ranging from 400–420, 670–680, 1350–1400 and 1900–2000 nm, which are the spectral bands affected by the scattering and absorption features in the atmosphere and also spectral regions having the largest relative uncertainty. The overall relative uncertainty of spectral-genetic similarities derived from single wavelengths had similar patterns over time and increased at longer wavelengths.

In comparison to similarities between reflectance spectra and genetic structure derived from single spectral bands, spectral-genetic similarities derived from relative differences between spectral bands were on average lower (values of partial Mantel r of 0.002 for relative differences between spectral bands vs 0.03 for single spectral bands), but reached larger value and were moderately more stable (mean standard deviation across time points of 0.06 for relative differences between spectral bands vs 0.07 for single spectral bands) (Fig. S.3.2). We found the highest spectral-genetic similarity (0.40) for relative differences for a spectral band centered at 515 nm with a spectral band centered at 1925 nm (relative uncertainty: 2.9%, $p = 0.001$) during the 19^{th} time point (993 CGDD, July) (Fig. S.3.3). The relative uncertainties of spectral-genetic similarities derived from relative differences are on average higher (8%) than relative uncertainties of spectral-genetic similarities derived from single wavelengths (1.8%).

3.4.2 Spectral-genetic similarities explained by environmental variables

Correlations between spectral-genetic similarities and environmental variables showed different patterns on different temporal scales (Fig. 3.4). Daily mean temperature (TMP) was positively correlated with spectral-genetic similarity derived from spectral bands above 1000 nm and around 1900 nm, which is highly affected by water absorption. The spectral-genetic similarities derived from around the 1900 nm spectral range were also positively correlated with the vapor pressure deficit of the day of the acquisition (VPD). When considering environmental conditions over a longer time period, the correlations

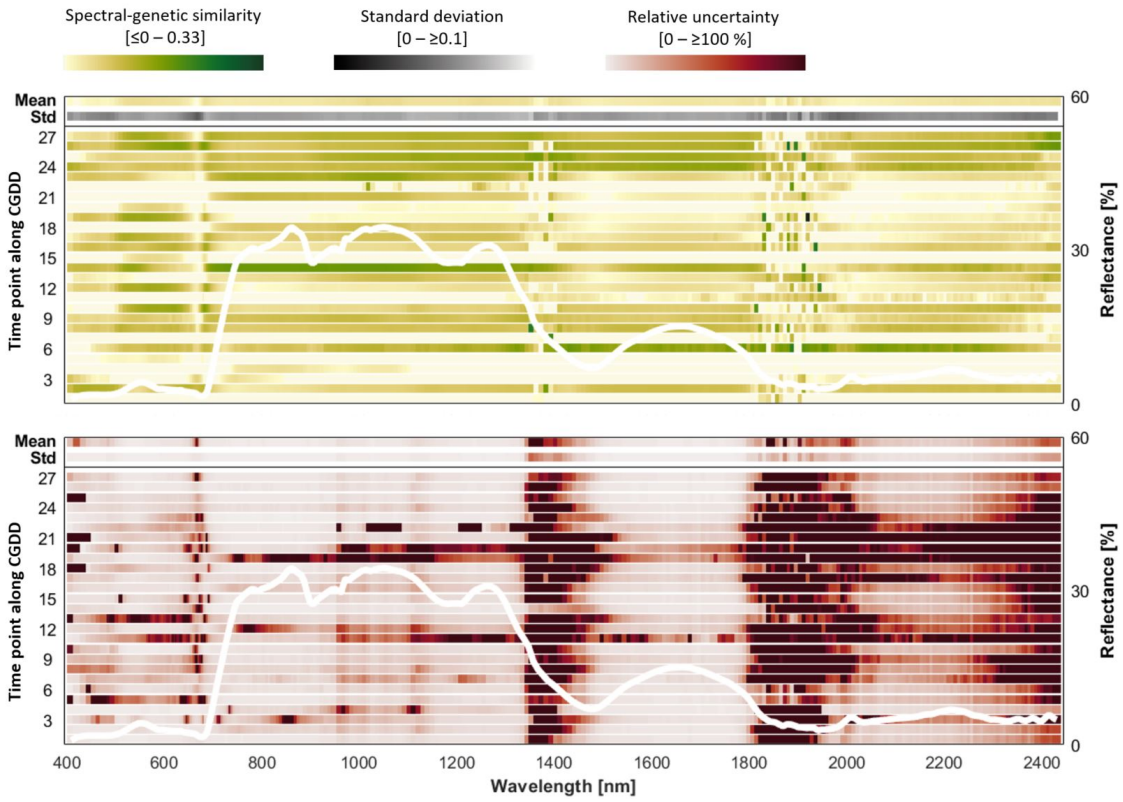

FIGURE 3.3: Spectral-genetic similarities (values of partial Mantel r) for the population of *F. sylvatica* derived from spectral bands for each time point segregated according to CGDD. Additionally, the mean and standard deviation of the similarities from all time points are shown at the top of the panel. The shown values range from 0 (white) to 0.33 (dark green). The relative uncertainties of the spectral-genetic similarities for all the time points across spectral bands are represented on the lower panel. The white lines represent the mean canopy reflectance of all trees.

between environmental variables and spectral-genetic similarities were not significant for Aggregated Temperature over 11 consecutive days before the acquisitions (11TMP) and were partly significant for Aggregated Vapor Pressure Deficit over 11 consecutive days before the acquisitions (11VPD). 11VPD was negatively correlated with spectral-genetic similarities derived from spectral bands around 650–680 nm and 1900–2100 nm. On the seasonal scale, the Cumulative Growing Degree Days (CGDD) and Cumulative Growing Vapor Pressure Deficit Days (CGVPDD) showed similar patterns of positive correlation with spectral-genetic similarities above 450 nm; however, the CGDD showed a stronger correlation on this time scale than the CGVPDD, which in contrast has higher impact on spectral-genetic similarities derived from the visible part of the spectrum (400–700 nm). The Day of the Last Spring Frost (DLSF) was non-significantly ($p > 0.05$) negatively correlated with the spectral-genetic similarity derived from all the spectral bands above 680 nm. Similarly, at the longest investigated temporal scale represented by the Last Year Climatic Water Balance (LYCWB), spectral-genetic similarities were non-significantly (p

> 0.05) correlated with this environmental variable for all spectral bands.

In comparison to the correlation between climatic variables and spectral-genetic similarities derived from single spectral bands, correlations between climatic variables and spectral-genetic similarities derived from relative differences between spectral bands were more pronounce at the shortest (TMP and VPD) and longest time scales (DLSF and LYCWB) (Fig. S.3.4). For both TMP and VPD, we found significant negative correlations, with spectral-genetic similarities derived from relative differences of spectral bands ranging from 400–420 nm, with spectral bands ranging from 1500–1800 nm. At the longest considered time scale, the spring frost (DLSF) and water balance of the preceding year (LYCWB) were, respectively, positively and negatively correlated with spectral-genetic similarities derived from relative differences of spectral bands ranging from 1000–1350 nm, with spectral bands ranging from 1500–1800 nm.

The relative uncertainties of the correlation of environmental variables with spectral-genetic similarities were on average higher for results derived from relative differences of spectral bands than for results derived from single spectral bands, and were higher at longer wavelengths, for spectral bands for which the correlations were weaker, and for spectral bands around 1000 and 1200 nm.

3.5 Discussion

3.5.1 Spectral-genetic similarities suggest underlying biochemical and morphological traits

The observed similarities between information derived from airborne imaging spectrometer data and the genetic structure of a population of *F. sylvatica* vary among the time points and across the spectrum for both single spectral bands, and relative differences between spectral bands. However, some reflectance spectra are similar for genetically similar individuals more systematically across all time points (Fig. 3.5). The spectral regions that express on average high, but not necessarily stable, spectral-genetic similarities lay in the range of 450–680 nm, corresponding to leaf pigment content and composition (eg. Moss and Loomis, 1952), 1000–2300 nm, corresponding to leaf water content that influences the reflectance spectra above 800 nm (Carlson et al., 1971), and 2350–2450 nm, affected by dry matter content (Peterson et al., 1988). Additional spectral-genetic similarities that are indicative of genetic structure and relatively time stable were derived in the 700–1400 nm spectral range. This spectral region is dominated by internal leaf scattering of light,

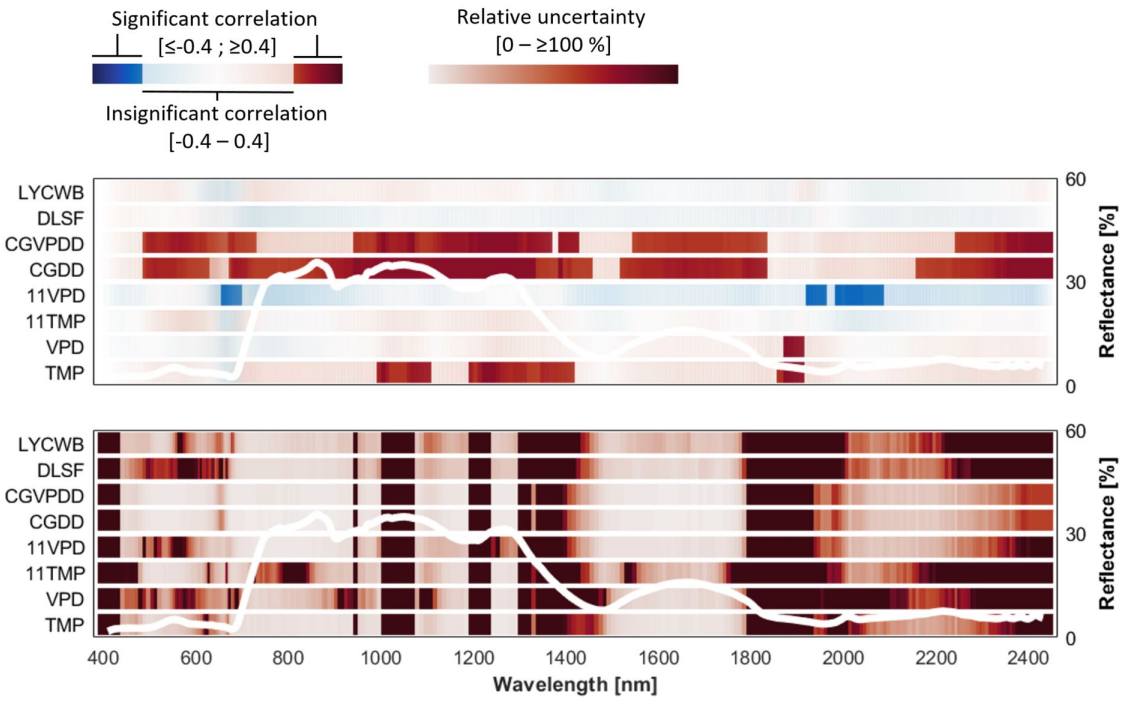

FIGURE 3.4: Correlations between spectral-genetic similarities derived for single spectral bands and: temperature of the day of the acquisition [°C] (TMP), Vapor Pressure Deficit of the day of the acquisition (VPD), Aggregated Temperature over 11 consecutive days [°C] (11TMP), Aggregated Vapor Pressure Deficit over 11 consecutive days (11VPD), Cumulative Growing Degree Days [°C] (CGDD), Cumulative Growing Vapor Pressure Deficit Days (CGVPDD), Day of the Last Spring Frost [DOY] (DLSF), Last Year Climatic Water Balance [DOY] (LYCWB). The shown values range from 0.6 (dark blue) to the 0.6 (dark red). The bottom panel shows the relative uncertainties of correlations between spectral-genetic similarities and environmental variables. The white lines represent the mean canopy reflectance of analyzed trees.

related to the mesophyll structure and thus intercellular air spaces (Gausman et al., 1971), that are accentuated in canopy-level measurements by scattering within the tree canopy (Ollinger, 2011). Beyond similarities derived from these spectral regions, there are similarities between spectral information and genetic structure specific to different time points. At some time points, the spectral region ranging from 420–450 nm related to the leaf carotenoids content (eg. Gitelson et al., 2006) showed similar reflectance values for genetically similar individuals. The absorption feature centered at 680 nm, related to chlorophyll content (eg. Peñuelas et al., 1994), represents low spectral-genetic similarities, which were also stable over time for single spectral band investigations. The relative differences between spectral bands related to carotenoids and leaf waxy layer, influencing the reflectance spectra ranging from 400–420 nm (eg. Long et al., 2003) with spectral bands related to the leaf water content (1400–1900 nm), are on average highly indicative of the genetic structure of the studied population. Strong spectral-genetic similarities at particular time points may occur if a genetically determined trait is only expressed

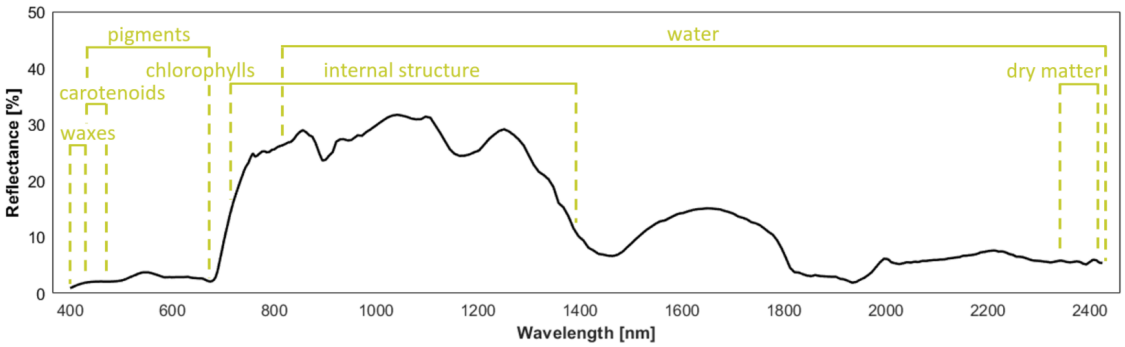

FIGURE 3.5: The revealed spectral features and some underlying biochemical and morphological traits indicative about the genetic structure of the *F. sylvatica* population over time of spectral acquisitions. (The spectral ranges are limited to results of our analysis and do not represent full spectral coverage influenced by particular vegetation traits after: Long et al., 2003; Gitelson et al., 2006; Moss and Loomis, 1952; Peñuelas et al., 1994; Gausman et al., 1971; Carlson et al., 1971; Peterson et al., 1988).

under particular environmental conditions occurring at that time point or if a trait is only expressed at a particular developmental stage of the tree. For example, features of reflectance spectra during pigment breakdown in leaves (Matile et al., 1999), affected by genetic variation between individual trees, can only be assessed before leaf shedding.

3.5.2 Genetically constrained traits and the temporal environments in which they are expressed

Because many genetically determined traits are only expressed under particular environmental conditions (Falconer, 1996), the inclusion of environmental variables over different time scales may help to explain observed links between genetic structure and remotely sensed phenotypes of the *F. sylvatica* population. On a day-to-day time scale, higher daily temperature leads to stronger correlation of reflectance measured at single spectral bands with genetic structure for spectral regions linked with leaf water content (above 800 nm). That may suggest that the water balance in the population is related to genetic structure when the current air temperature is higher. In comparison, the relationships between spectral bands influenced by the waxy layer (400–420 nm) and water content (1500–1800 nm) follow the genetic structure of the population when the vapor pressure deficit is lower and the ambient air is colder. This may indicate that the correlation between genetic structure and photoprotection via the waxy layer (Long et al., 2003) is more apparent (Shepherd and Wynne Griffiths, 2006) when plants are exposed to humid and cold air, differently tolerated by genetically more distant individuals. The leaf waxes provide an important barrier to water loss, and reduce light penetration into the leaf by specular reflection (Petibon et al., 2021; Vanderbilt and Grant, 1985), which may relate

to mechanisms related to the genetic structure. Finally, analyses on the day-to-day time scale suggest that genetically constrained traits are more correlated with daily temperature than with daily vapor pressure deficit. However, more genetically constrained traits seemed to be strongly expressed under different 11-day vapor pressure deficit conditions (11VPD) than different 11-day temperature conditions (11TMP). A lower vapor pressure deficit for 11 days before the acquisition led to observation of higher genetic-spectral similarities of the analyzed population over the whole spectrum, especially the chlorophyll content (680 nm) and water content (1900–2100 nm). This suggests that the balance between water management and light absorption by pigments is more similar for genetically similar individuals under favorable environmental conditions sustained at moderate time scale (11 days). Other studies have shown that water relations of *F. sylvatica* can vary within the species (Gonzalez de Andres et al., 2021) and could be characteristic for genetically different provenances (Peuke et al., 2002). Our observations indicate that water balance may correspond more to genetic structure when the water availability at 11-day time scale is not a limitation — at least within this single population over the observed time points. However, we treat the correlations based on results derived from water and carbon dioxide absorption features (1900–2050 nm) with caution, as the uncertainties presented in our analysis do not account for uncertainties due to atmospheric correction. More specifically, that absorption due to these substances along the path from leaves and canopies, through the atmosphere, to the aerial sensor might not be perfectly accounted in atmospheric compensations.

Across the season, our analyses found that the influences of water availability as measured by CGVPDD and of temperature as measured by CGDD followed a similar pattern and held positive correlations with spectral-genetic similarities derived from single spectral bands for spectral regions related with pigment composition and content (400–680 nm) and water content (above 800 nm). Compared with CGVPDD, CGDD correlated also significantly with spectral-genetic similarities derived from 700–1000 nm, related to canopy scattering properties. This suggests that chlorophyll content (absorption feature at around 680 nm), leaf internal structure (700–1000 nm) and water balance (800–2390 nm) correlate with genetic structure towards the end of a hot and dry growing season. We furthermore speculate that this may be due to more similar foliage development, thus capacity to harvest light, of genetically similar individuals. Finally, the leaf internal structure as inferred from the spectral bands ranging from 1000–1400 nm correlated with water content as inferred from the spectral bands ranging from 1500–1800 nm may be

informative about the genetic structure of the population when the last day of the spring frost was relatively late and the year proceeding the acquisition was relatively dry. This was indicated by the positive correlation with day of last spring frost (DLSF) and negative correlation with Last Year Water Climatic Balance (LYCWB), and the spectral-genetic similarity of this spectral range (Fig. S.3.4). We suggest that there could be a similarity in closely related individuals in their responses to frost and drought, which in turn affects leaf morphology and development (Menzel et al., 2015). It has been demonstrated that juvenile *F. sylvatica* individuals from genetically different provenances vary in sensitivity to frost damage (Kreyling et al., 2012) and the development of leaf buds may be affected by water availability.

Together, our observations show that the spectral information derived from airborne imaging spectrometer data can be indicative for the genetic structure of the population of a dominant tree species and that more genetic structure can be revealed if spectral information is obtained under multiple environmental conditions and over time, which allows plants to express more of their genetically constrained traits than under a single environmental condition or at a single time of development.

3.6 Limitations and outlook

In our study, we focused on observing the temporal dynamics of reflectance spectra, on how those spectra correlate with genetic structure of a tree population, and under which environmental conditions correlations for particular spectral features with genetic structure are particularly strong. The environmental variables we correlated were water availability as measured by either vapor pressure deficit or climatic water balance and temperature on multiple temporal scales. Both of these variables strongly affecting *F. sylvatica*. The investigation of other abiotic and biotic factors, and their temporal consequences on the reflectance spectrum, could give insights to other similarities in the spectral and genetic structures of populations, and the genetically constrained processes that influence these similarities. For example, changes in leaf optical properties may be induced by increases in ambient carbon dioxide and ozone (Carter et al., 1995), aerosols (Farmer, 1993), or wind speed, light availability and thus irradiation that influence the water potential of plant tissue. Additionally, resistance to infections from bacteria, viruses and fungi, or attack by insects may also vary significantly on an intraspecific level and

could be potentially derived from spectral signatures (eg. Carter, 1993). Correlative observations between genetic structure, reflectance spectra and environmental variables are indicators of physiological processes specific for a population adapted to its environment and under given environmental conditions.

While the above suggests that even more environmental conditions will likely affect spectral-genetic similarities, our proof-of-concept study was small and correlative, which at the same time limits confidence regarding the results for multiple spectral features and multiple environmental variables. Ideally, the spectra of a much larger number of replicated tree genotypes (or genetic groups such as seed families) should be analyzed across environmental variables that are not correlated with each other. This could be best done in experimental settings, which we now do with seed families collected from populations of *F. sylvatica* across Europe, or with very large numbers of trees in observational studies by multivariate matched sampling of pixels or trees (Imbens and Rubin, 2015; Sturm et al., 2022).

Further, we approach the analyses in a discrete manner by considering every time point and their environmental characteristic separately. Even though the multi-temporal selected acquisitions were distributed across growing seasons, the ontogenetic stages of the tree crowns were specific for particular years. Thus, segregation of time points by growing degree days does not allow us to follow the trajectories of individuals in time (day-to-day or year-to-year) (eg. Asner et al., 2006; Castro-Esau et al., 2006). Discovering such a trend could be facilitated by a long sequence of acquisitions on a more frequent basis that assess spectral information at regular time intervals (Cawse-Nicholson et al., 2021), although these are limited by instrument lifespan, cloud-free conditions and potentially coarser spatial resolution which would not directly support the assessment of individual trees and thus not help in detecting genetic variation within them. Further, such acquisitions are generally constrained to a particular time of the day as a practical consequence of satellite orbits, which would reduce the variation across acquisitions caused by systematic daily variations (eg. midday depression). From our analyses, we deduce that the mean reflectance spectra of the acquisitions over the day (six acquisitions during the 10^{th} time point and five acquisitions during the 23^{th} time point) may result in relatively low spectral-genetic similarities across the spectrum. Only distinct spectral ranges express relatively higher spectral-genetic similarities (500–650 nm and 700–1800 nm), which could be potentially stable across the day. Unraveling which of the spectral ranges are stable over the day and most informative about the genetic structure would

require more within-day acquisitions. Here we did not analyze the within-day temporal variation because the missions within a day (time point) were not flown at consistently selected times of the day and because this additional possible analysis would have further increased the multiple-testing level for a small sample size of 68 trees.

When investigating multi-temporal datasets, it is especially important to consider differences in the particular geometry of the acquisition and the atmospheric composition. These could cause inherent variations in shading, directional effects (Schaepman-Strub et al., 2006) and atmospheric variation of the signal (Thompson et al., 2019). Combining multiple acquisitions from one day that vary in illumination/observation geometry to represent single time points potentially allows to average the variation due to illumination and observation geometries as well as atmospheric variability. By weighting the spectral information of the pixels by their inverse uncertainties, we account for the full set of spectral information and their relative reliability. Modifying reflectance spectra based on measurement uncertainties according to our results revealed up to 36% more information about biological variation. Still unknown is the impact of inclined surfaces on upright growing trees, their anisotropic behavior (Ma et al., 2020; Vögtli et al., 2020) and the uncertainties of surface reflectance spectra (Carmon et al., 2020; Hueni et al., 2016; Richter and Schläpfer, 2002).

It has to be noted that our proof-of-concept study is species-, site-, and time-specific. We show that prediction of the genetic structure is possible to a certain degree, with degree depending on airborne imaging spectroscopy measurements at specific time points, on combining shorter-term with longer-term observation schemes, and on specific ranges of wavelengths used. Our approach could be transferred to other species, different spatial and temporal scales, and different sensor systems as long as the observed biological units represent distinct genetic identities. In this study, we were able to investigate population-level genetic diversity for European beech trees in natural conditions by incorporating ground-based LiDAR (Morsdorf et al., 2018; Schneider et al., 2019), airborne laser scanning (Kaartinen et al., 2012; Schneider et al., 2014; Y. Wang et al., 2016) and the 2 m-resolution pixel size of the airborne imaging acquisitions which allowed us to delineate *F. sylvatica* individuals (Guillén-Escribà et al., 2021) and to link them with their genetic identities. Working with data with a spatial resolution that is not suited to delineate individual trees or species would require community-based analyses, where the signal is averaged across several individuals and species (Zheng et al., 2021). As genetic differences

between sites over larger areas are likely larger than genetic differences between individuals within a single site, this could still reveal genetic or phylogenetic variation at a cruder level. Finally, a better understanding of the mechanistic links between reflectance spectra, genetic and trait variation could be discovered by genome-wide association studies using e.g. whole-genome sequencing or gene capture datasets.

Our proof-of-concept study is only a first step to establish correlations between remotely sensed spectral information and genetic diversity in a dominant tree species. There is still a long way to go from this to predicting genetic diversity using spectral information in new areas where trees have not been genotyped. Prediction of genetic structure would require relationships calibrated for genetic resolution of interest and likely will need to account for spatially resolved environmental conditions. It would likely be supported by temporally rich acquisitions representing dynamic biological changes, and a well-understood impact of instrument variation on the spectral signal provided with its uncertainties as non-biological variation (Petibon et al., 2021). However, our study indicates wavelength bands which should be considered and under which environmental conditions multiple spectral signatures should be acquired to develop such prediction tools. Further studies will be needed to determine the degree to which calibrations obtained for one species in one area may be generalizable to other species and other areas.

3.7 Conclusions

This study demonstrates the potential of dense multi-temporal imaging spectroscopy data to link spectral information with genetic structure of temperate forest trees. We investigated 69 airborne imaging spectrometer data acquisitions along with genetic information from 68 tree individuals and environmental data to derive spectral-genetic similarities in population structure under different environmental conditions.

We observed that more genetically similar individuals have more similar reflectance in distinct spectral regions correlated with particular environmental conditions such as temperature and water availability on different time scales. From this, we indicate traits and physiological processes that may explain these similarities and may correspond to genetic variation among tree individuals. We could demonstrate that relative differences between spectral bands were more informative about genetic structure than single-spectral band analyses, at least for specific times of the growing season. We were able to assess

the quality of our derivation by incorporating uncertainty propagation at each step of analysis.

With dense multi-temporal observations of reflectance spectra indicative of multiple plant traits, we hope to contribute to the scalable monitoring of genetic diversity via remote sensing. That will allow us to better understand what is driving biodiversity changes, how to stop further biodiversity decline, and how to restore ecosystems where decline has taken place.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study is supported by the University of Zurich Research Priority Program on Global Change and Biodiversity (URPP GCB). The research carried out at the Jet Propulsion Laboratory, California Institute of Technology, was under a contract with the National Aeronautics and Space Administration (80NM0018D0004). Government sponsorship is acknowledged. We express our gratitude to Reinhard Furrer, Maria J. Santos and Joan T. Sturm for their support. We would like to acknowledge Qingyang He, Andrew Tedder, Carmen Meiller, Helena Kühnle, Mike Werfeli, Daniel Schläpfer, and Sabina Keller for processing the datasets and Kentaro Shimizu for providing access to facilities for genetic analyses. Finally, we acknowledge comments of reviewers that greatly contributed to improvement of the manuscript.

Supplementary Material

TABLE S.3.1: List of investigated spectral images acquired by Airborne Prism Experiment (APEX) Airborne Imaging Spectrometer (AIS) above Laegern forest, Switzerland between 2009–2019.

no.	APEX Mission ID	Date	Solar Zenith angle [°]	Solar Azimuth angle [°]
1	MM002.1	17/06/2009	30.4	134.1
2	MM002.2	17/06/2009	29.5	137.4
3	MM017.1	26/06/2010	42.2	109.1
4	MM017.2	26/06/2010	48.1	259.0
5	MM017.3	26/06/2010	49.1	260.3
6	MM017.4	26/06/2010	50.2	261.8
7	MM026.1	29/06/2010	29.3	138.5
8	MM026.2	29/06/2010	28.6	141.1
9	MM026.3	29/06/2010	27.8	144.5
10	MM041.1	16/06/2012	28.6	140.8
11	MM041.2	16/06/2012	27.9	143.0
12	MM054.1	14/05/2013	67.9	273.9
13	MM041.3	16/06/2012	27.1	147.4
14	MM058.1	01/07/2013	24.4	180.5
15	MM058.2	01/07/2013	24.4	183.8
16	MM058.3	01/07/2013	24.6	188.8
17	MM063.1	12/07/2013	54.1	96.1
18	MM084.1	03/09/2013	40.7	166.3
19	MM084.2	03/09/2013	40.4	169.3
20	MM084.3	03/09/2013	40.2	172.0
21	MM086.1	10/04/2014	53.8	122.8
22	MM086.2	10/04/2014	52.3	125.7
23	MM086.3	10/04/2014	51.2	127.8
24	MM092.1	18/07/2014	26.5	179.6
25	MM092.2	18/07/2014	26.5	183.3
26	MM092.3	18/07/2014	26.7	189.0
27	MM102.1	10/04/2015	40.7	161.7

Continues on next page

TABLE S.3.1: (Continued) List of investigated spectral images acquired by Airborne Prism Experiment (APEX) Airborne Imaging Spectrometer (AIS) above Laegern forest, Switzerland between 2009–2019.

no.	APEX Mission ID	Date	Solar Zenith angle [°]	Solar Azimuth angle [°]
28	MM102.2	10/04/2015	40.1	166.5
29	MM102.3	10/04/2015	39	169.2
30	MM107.1	21/04/2015	40.6	216.8
31	MM107.2	21/04/2015	41.5	219.5
32	MM107.3	21/04/2015	42.5	222.4
33	MM110.1	24/06/2015	33.7	125.1
34	MM110.2	24/06/2015	32.5	128.1
35	MM110.3	24/06/2015	31.2	131.7
36	MM119.1	24/06/2015	47.1	257.9
37	MM119.2	24/06/2015	48.6	259.8
38	MM119.3	24/06/2015	50.1	261.7
39	M0136.1	10/07/2015	39.7	244.0
40	M0136.2	10/07/2015	41.1	246.4
41	M0136.3	10/07/2015	42.3	248.5
42	M0142.1	21/08/2015	37.9	154.0
43	M0142.2	21/08/2015	37.3	157.1
44	M0142.3	21/08/2015	36.8	160.2
45	M0150.1	20/04/2016	37.5	201.1
46	M0150.2	20/04/2016	38.2	204.9
47	M0150.3	20/04/2016	38.8	208.3
48	M0159.1	07/07/2016	30.3	137.7
49	M0159.2	07/07/2016	29.4	141.0
50	M0159.3	07/07/2016	28.7	143.9
51	M0166.1	17/07/2016	27.5	200.7
52	M0166.2	17/07/2016	28.1	205.0
53	M0166.3	17/07/2016	28.8	209.2
54	M0181.1	07/09/2016	42.1	167.2
55	M0181.2	07/09/2016	41.9	169.7
56	M0181.3	07/09/2016	41.7	172.7

Continues on next page

TABLE S.3.1: (Continued) List of investigated spectral images acquired by Airborne Prism Experiment (APEX) Airborne Imaging Spectrometer (AIS) above Laegern forest, Switzerland between 2009–2019.

no.	APEX Mission ID	Date	Solar Zenith angle [°]	Solar Azimuth angle [°]
57	M0208.1	30/06/2018	25.9	204.2
58	M0208.2	30/06/2018	26.3	207.2
59	M0208.3	30/06/2018	27.0	211.0
60	M0218.2	09/07/2018	25.7	194.2
61	M0218.3	09/07/2018	26.0	197.3
62	M0227.1	26/07/2018	29	162.7
63	M0227.2	26/07/2018	28.7	165.6
64	M0227.3	26/07/2018	28.4	169.0
65	M0255.1	07/06/2019	34.9	235.9
66	M0255.2	07/06/2019	35.9	238.1
67	M0255.3	07/06/2019	37	240.2
68	M0263.1	19/06/2019	24.3	190.3
69	M0263.2	19/06/2019	24.5	193.6
70	M0263.3	19/06/2019	24.8	197.3
71	M0266.1	25/06/2019	24.1	180.6
72	M0266.2	25/06/2019	24.1	184.5
73	M0266.3	25/06/2019	24.2	187.9
74	M0273.1	04/07/2019	24.8	170.8
75	M0273.2	04/07/2019	24.7	173.5
76	M0278.1	10/07/2019	28.8	214.8
77	M0278.2	10/07/2019	29.4	217.4
78	M0278.3	10/07/2019	29.9	219.5
79	M0287.1	25/07/2019	46.5	249.6
80	M0287.2	25/07/2019	47.6	251.2
81	M0287.3	25/07/2019	48.5	252.4
82	M0292.1	30/07/2019	34.7	138.8
83	M0292.2	30/07/2019	34	141.1
84	M0292.3	30/07/2019	33.5	143.0
85	M0292.4	30/07/2019	32.9	145.3

Continues on next page

TABLE S.3.1: (Continued) List of investigated spectral images acquired by Airborne Prism Experiment (APEX) Airborne Imaging Spectrometer (AIS) above Laegern forest, Switzerland between 2009–2019.

no.	APEX Mission ID	Date	Solar Zenith angle [°]	Solar Azimuth angle [°]
86	M0292.5	30/07/2019	32.2	147.8
87	M0292.6	30/07/2019	31.9	149.8

TABLE S.3.2: The statistics on lengths of five highly variable microsatellite loci (FS1-03, FS1-15, FS3-04, FS4-46, FCM5; Pastorelli et al., 2003) of sampled *F. sylvatica* individuals derived from capillary electrophoresis with *GeneMapper* software (Applied Biosystems).

Length statistic	FS1-03	FS1-15	FS3-04	FS4-46	FCM5
Standard Deviation	4.63	8.07	1.73	22.74	12.51
Mean	91.55	111.99	200.95	251.75	299.87
Variance	21.39	65.09	2.99	517.23	156.45
Maximum	108	137	206	328	322
Minimum	83	93	194	221	280

S.Propagation of uncertainties

In the main text, we explained that uncertainties in radiance measurements were propagated to assess the uncertainty in the key results obtained in our study: the spectral-genetic similarities and their correlations with environmental variables. In this Supplementary section, we explain this procedure in detail. To quantify uncertainty in a variable of interest f, its approximate variance $u^2(f)$ was obtained by applying the general rule of uncertainty propagation (e.g. Ku 1966):

$$u(f) = \sqrt{\sum_{n=1}^{N} \left(\frac{\delta f}{\delta x_n} \right)^2 * u_{x_n}^2} \tag{S.3.1}$$

with $\delta f/(\delta x_n)$ quantifying the sensitivity of variable f to changes in a driver variable x_n, which has its own uncertainty, $u(x_n)$. In the following, we will derive the specific expressions needed to propagate uncertainty for the pairwise Euclidean distance d_{ij} in detail and present the sequential expressions used to derive further key variables used in our study.

Pairwise Euclidean distance, d_{ij}

The pairwise Euclidean distance mentioned in the main text was calculated as:

$$d_{ij} = \sqrt{(x_i - x_j)^2} \tag{S.3.2}$$

Where d_{ij} is the Euclidean distance that quantifies the difference in x_i and x_j, which represent the reflectance within a certain spectral band for trees i and j, for example. These reflectance values are associated with uncertainties u_{x_i} and u_{x_j}, respectively. Hence, the uncertainty in the difference between x_i and x_j can be calculated using (S.3.1) with $f_{diff} = x_i - x_j$, yielding:

$$u_{(x_i - x_j)} = \sqrt{u_{x_i}^2 + u_{x_j}^2} \tag{S.3.3}$$

As the Euclidean distance calculation uses the squared difference between two variables, the uncertainty propagation can be estimated using (S.3.1), with the function $f_{square} = x_{diff}^2$, and $x_{diff} = x_i - x_j$. Including substitution of the right-hand side of (S.3.3), this then yields:

$$\begin{aligned} u_{(x_i - x_j)^2} &= \sqrt{4x_{diff}^2 * u_{x_{diff}}^2} \\ &= \sqrt{4\left(x_i - x_j\right)^2 * \left(u_{(x_i - x_j)}\right)^2} = \sqrt{4\left(x_i - x_j\right)^2 * \left(u_{x_i}^2 + u_{x_j}^2\right)} \end{aligned} \tag{S.3.4}$$

Finally, the Euclidean distance takes the square root of the above distance. The propagation of uncertainty in this step is then again estimated using (S.3.1), this time using $f_{root} = \sqrt{x}_{square}$, with $x_{square} = (x_i - x_j)^2$. Including substitution of the right-hand side of (S.3.4), this then yields:

$$\begin{aligned} u_{\sqrt{(x_i - x_j)^2}} &= \sqrt{\left(\frac{1}{2\sqrt{x_{\text{square}}}}\right)^2 * u_{x_{\text{square}}}^2} \\ &= \sqrt{\frac{1}{4\left(x_i - x_j\right)^2} * 4\left(x_i - x_j\right)^2 * \left(u_{x_i}^2 + u_{x_j}^2\right)} = \sqrt{u_{x_i}^2 + u_{x_j}^2} \end{aligned} \tag{S.3.5}$$

Following the same approach, we calculated Euclidean distances between trees based on relative differences between spectral bands, genetic identities and spatial distribution. We calculated subsequent products and their inherent uncertainties following the same rule (S.3.1) as represented below:

z-score of pairwise Euclidean distance, $\check{d}_{ij}$

In this study, z-scores of Euclidean distances between trees i and j, d_{ij}, were calculated and subsequently used in computations of partial Mantel correlations (see below). These z-scores, $\check{d}_{ij}$ were calculated as:

$$\check{d}_{ij} = \frac{d_{ij} - \bar{d}}{\sqrt{\frac{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{ij}-\bar{d}\right)^2}{N-1}}} \tag{S.3.6}$$

Where N is the number of pairs considered, i.e. $N = n * (n-1)/2$. For which the uncertainty in $\bar{d}_{ij}$ can be propagated as:

$$u_{\left(\sum_{i=1}^{n}\sum_{j=i+1}^{n} d_{ij}\right)} = \sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n} u_{d_{ij}}^2} \tag{S.3.7}$$

And:

$$u_{\bar{d}} = \sqrt{\left(\frac{1}{N}\right)^2 * u^2_{\left(\sum_{i=1}^{n}\sum_{j=i+1}^{n} d_{ij}\right)}} \tag{S.3.8}$$

The uncertainty in the numerator of (S.3.6) can then be calculated as:

$$u_{d_{ij}-\bar{d}} = \sqrt{u_{d_{ij}}^2 + u_{\bar{d}}^2} \tag{S.3.9}$$

While the uncertainty in the denominator of (S.3.6) can be calculated through the following steps:

$$u_{\left(d_{ij}-\bar{d}\right)^2} = \sqrt{\left(2\left(d_{ij}-\bar{d}\right)\right)^2 * u^2_{d_{ij}-\bar{d}}} \tag{S.3.10}$$

$$u_{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{ij}-\bar{d}\right)^2} = \sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n} u^2_{\left(d_{ij}-\bar{d}\right)^2}} \tag{S.3.11}$$

$$u_{\frac{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{ij}-\bar{d}\right)^2}{N-1}} = \sqrt{\left(\frac{1}{N-1}\right)^2 * u^2_{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{ij}-\bar{d}\right)^2}} \tag{S.3.12}$$

$$u_{\sqrt{\frac{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{ij}-\bar{d}\right)^2}{N-1}}} = \sqrt{\left(\frac{1}{\sqrt{\frac{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{ij}-\bar{d}\right)^2}{N-1}}}\right)^2 * u^2_{\frac{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{ij}-\bar{d}\right)^2}{N-1}}} \tag{S.3.13}$$

Finally, the uncertainty in the ratio between numerator and denominator could be calculated as:

$$u_{\check{d}} = \sqrt{\left(\frac{1}{\sqrt{\frac{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{ij}-\bar{d}\right)^2}{N-1}}}\right)^2 * u^2_{\left(d_{ij}-\bar{d}\right)^2} + \left(-\frac{d_{ij}-\bar{d}}{\left(\sqrt{\frac{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{ij}-\bar{d}\right)^2}{N-1}}\right)^2}\right)^2 * u^2_{\sqrt{\frac{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{ij}-\bar{d}\right)^2}{N-1}}}} \tag{S.3.14}$$

Correlation coefficient, r_{AB}

Using the z-scores of pairwise Euclidean distances, $\check{d}_{ij}$, we calculated correlations between spectral, genetic and spatial distances prior to computing partial Mantel correlations and to estimate the variation in spectral-genetic similarities explained by environmental variables. The correlation coefficients, r_{AB} were calculated as:

$$r_{AB} = \frac{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{A_{ij}}-\bar{d}_A\right)\left(d_{B_{ij}}-\bar{d}_B\right)}{\sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{A_{ij}}-\bar{d}_A\right)^2} * \sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{B_{ij}}-\bar{d}_B\right)^2}} \tag{S.3.15}$$

when performing partial Mantel correlations using Euclidean distances between trees i and j regarding variables A and B, which could contain either spectral, genetic or spatial information. The uncertainty in the numerator of (S.3.15) can then be calculated through following steps:

$$u_{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\sigma_{A_{ij}}} = \sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n}u^2_{A_{ij}}} \tag{S.3.16}$$

$$u_{\bar{d}_A} = \sqrt{\left(\frac{1}{n}\right)^2 * u^2_{\sum_{i=1}^{n}\sum_{j=i+1}^{n}u_{A_{ij}}}} \tag{S.3.17}$$

$$u_{d_{A_{ij}}-\bar{d}_A} = \sqrt{\sigma^2_{d_{A_{ij}}} + \sigma^2_{\bar{d}_A}} \tag{S.3.18}$$

$$u_{\left(d_{Ai}-\bar{d}_A\right)\left(d_{Bi}-\bar{d}_B\right)} = \sqrt{\left(d_{A_{ij}}-\bar{d}_A\right)^2 * \sigma^2_{d_{B_{ij}}-\bar{d}_B} + \left(d_{B_{ij}}-\bar{d}_B\right)^2 * \sigma^2_{d_{A_{ij}}-\bar{d}_A}} \tag{S.3.19}$$

$$u_{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{A_{ij}}-\bar{d}_A\right)\left(d_{B_{ij}}-\bar{d}_B\right)} = \sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n} u^2_{\left(d_{A_{ij}}-\bar{d}_A\right)\left(d_{B_{ij}}-\bar{d}_B\right)}} \tag{S.3.20}$$

While the uncertainty in the denominator of (S.3.15) can be calculated through the following steps:

$$u_{\left(d_{A_{ij}}-\bar{d}_A\right)^2} = \sqrt{\left(2\left(d_{A_{ij}}-\bar{d}_A\right)\right)^2 * u^2_{d_{A_{ij}}-\bar{d}_A}} \tag{S.3.21}$$

$$u_{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{A_{ij}}-\bar{d}_A\right)^2} = \sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n} u^2_{\left(d_{A_{ij}}-\bar{d}_A\right)^2}} \tag{S.3.22}$$

$$u_{\sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{A_{ij}}-\bar{d}_A\right)^2}} = \sqrt{\left(\frac{1}{2\sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{A_{ij}}-\bar{d}_A\right)^2}}\right)^2 * u^2_{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{A_{ij}}-\bar{d}_A\right)^2}} \tag{S.3.23}$$

$$u_{\sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{A_{ij}}-\bar{d}_A\right)^2}*\sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{B_{ij}}-\bar{d}_B\right)^2}} = \sqrt{\left(\sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{A_{ij}}-\bar{d}_A\right)^2}\right)^2 * u^2_{\sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{B_{ij}}-\bar{d}_B\right)^2}} + \left(\sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{B_{ij}}-\bar{d}_B\right)^2}\right)^2 * u^2_{\sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{A_{ij}}-\bar{d}_A\right)^2}}} \tag{S.3.24}$$

Finally, the uncertainty in the ratio between numerator and denominator could be calculated as:

$$u_{r_{AB}} = \sqrt{\begin{array}{l} \left(\frac{1}{\sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{A_{ij}}-\bar{d}_A\right)^2} * \sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{B_{ij}}-\bar{d}_B\right)^2}}\right)^2 \\ * u^2_{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{A_{ij}}-\bar{d}_A\right)^2\left(d_{B_{ij}}-\bar{d}_B\right)^2} \\ + \left(-\frac{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{A_{ij}}-\bar{d}_A\right)\left(d_{B_{ij}}-\bar{d}_B\right)}{\left(\sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{A_{ij}}-\bar{d}_A\right)^2} * \sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{B_{ij}}-\bar{d}_B\right)^2}\right)^2}\right)^2 \\ * u^2_{\sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{A_{ij}}-\bar{d}_A\right)^2} * \sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(d_{B_{ij}}-\bar{d}_B\right)^2}} \end{array}} \tag{S.3.25}$$

As noted in the main text, we also estimated the variation in spectral-genetic similarities explained by environmental variables. In these cases, the correlation coefficients, r_{AB} were calculated as:

$$r_{AB} = \frac{\sum_{k=1}^{N}\left(d_{A_k}-\bar{d}_A\right)\left(d_{B_k}-\bar{d}_B\right)}{\sqrt{\sum_{k=1}^{N}\left(d_{A_k}-\bar{d}_A\right)^2} * \sqrt{\sum_{k=1}^{N}\left(d_{B_k}-\bar{d}_B\right)^2}} \tag{S.3.26}$$

Where N is the number of time points, A_k is the spectral-genetic similarity at timepoint k, as derived for a specific spectral band or a relative difference between two spectral bands. Finally, B_k stands for the value of an environmental variable at timepoint k. Note that the calculation of uncertainty propagating through (S.3.26) is analogous to the procedure described by (S.3.16)-(S.3.25).

Partial Mantel correlations, r_p

We calculated Partial Mantel correlations, r_p for standardized Euclidean distances between trees based on genetic and spectral information with extraction of correlation of residuals from the linear regression of spectral-spatial correlation and genetic-spatial correlations to derive spectral-genetic similarities as:

$$r_p = \frac{r_{GS} - r_{GE} * r_{ES}}{\sqrt{1 - r_{GE^2}} * \sqrt{1 - r_{ES^2}}} \tag{S.3.27}$$

Where G corresponds to genetic information, S to spectral information, and E to environmental (spatial) information. In our study, the $u^2_{r_{GE}}$ was equal to zero, thus the uncertainty of partial Mantel correlations, r_p was calculated as:

$$u_{r_p} = \sqrt{\begin{array}{l}\left(\frac{1}{\sqrt{1-r_{GE^2}} * \sqrt{1-r_{ES^2}}}\right)^2 * u^2_{r_{GS}} \\ + \left(\frac{r_{GS} * r_{ES} - r_{GE}}{(1-r_{ES^2})\sqrt{1-r_{GE^2}} * \sqrt{1-r_{ES^2}}}\right)^2 * u^2_{r_{ES}}\end{array}} \tag{S.3.28}$$

Relative differences, D_r

In this study Relative differences, D_r, were calculated between each spectral band for each tree individual and each time point to derive spectral-genetic similarities alongside with spectral-genetic similarities specific for each spectral band. The relative differences were calculated as:

$$D_r = \frac{|x_i - x_j|}{x_i + x_j} \tag{S.3.29}$$

Where x is a value of reflectance spectra and i, j are certain spectral bands. The uncertainties for the nominator and the denominator of (S.3.29) were the same and were calculated as:

$$u_{(|x_i - x_j|)} = u_{(x_i + x_j)} = \sqrt{u^2_{x_i} + u^2_{x_j}} \tag{S.3.30}$$

What yield in a total uncertainty of Relative differences, D_r to be:

$$u_{D_r} = \sqrt{\left(\frac{1}{x_i + x_j}\right)^2 * u^2_{|x_i - x_j|} + \left(-\frac{|x_i - x_j|}{(x_i + x_j)^2}\right)^2 * u^2_{|x_i - x_j|}} \tag{S.3.31}$$

Mean of partial Mantel correlation, $\bar{r}_p$

We calculated the means of derived partial Mantel correlations, $\bar{r}_p$, either derived for single spectral bands or relative differences between spectral bands over time as:

$$\bar{r}_p = \sum_{k=1}^{N} r_{p_k}/N \tag{S.3.32}$$

Where k is a time point. Uncertainties propagating through (S.3.32) can be calculated as:

$$u_{\bar{r}_p} = \sqrt{\left(\frac{1}{N}\right)^2 * u^2_{\sum_{k=1}^{N} u_{\bar{r}_p k}}} \tag{S.3.33}$$

Standard deviation of partial Mantel correlation, Δr_p

In addition to the mean over time of derived spectral-genetic similarities, we also calculated their standard deviations as:

$$\Delta r_p = \sqrt{\frac{\sum_{k=1}^{N} \left(r_{p_k} - \bar{r}_p\right)^2}{N-1}} \tag{S.3.34}$$

Where k is a time point ($N = 27$). Uncertainties propagating through (S.3.34) can be calculated as:

$$u_{\Delta r_p} = \sqrt{\left(\frac{1}{2\sqrt{\frac{\sum_{k=1}^{N}\left(r_{p_k} - \bar{r}_p\right)^2}{N-1}}}\right)^2 * u^2_{\frac{\sum_{k=1}^{N}\left(r_{p_k} - \bar{r}_p\right)^2}{N-1}}} \tag{S.3.35}$$

S. Selection of moderate time scale

We correlated time specific spectral-genetic similarities with time specific climatic variables representing aggregated temperature and vapor pressure deficit for multiple time spans prior to acquisition (Fig. S.3.1a). We selected time spans from one day prior to acquisition up to 43 days prior to acquisition, where 43 days are the minimal time between an acquisition and specific for it day of last spring frost DLSF [DOY]. Subsequently, we summed the absolute correlations across all analyzed wavelengths per each time span for aggregated temperature and aggregated vapor pressure deficit and we took a mean of the correlation of two climatic variables (Fig. S.3.1b). That resulted in highest correlations of spectral-genetic similarities with aggregated temperature and aggregated vapor pressure deficit from 11 days prior to acquisitions.

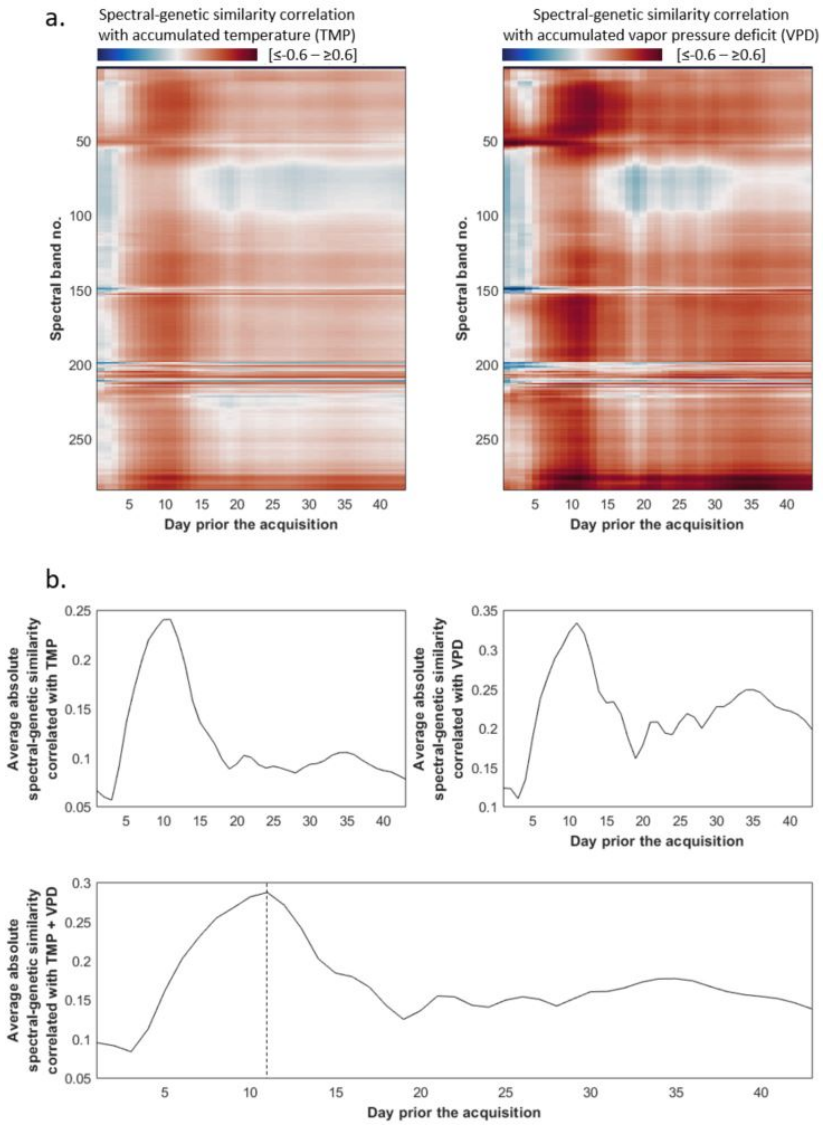

FIGURE S.3.1: Correlations between spectral-genetic similarities derived for single spectral bands and Temperature [°C] (TMP) and Vapor Pressure Deficit [%] (VPD) aggregated for one to 43 days prior to reflectance spectra acquisition (a). Additionally, averaged across all analyzed spectrum absolute correlation per aggregated temperature, aggregated vapor pressure and mean of both climatic variables per each time span is shown (b).

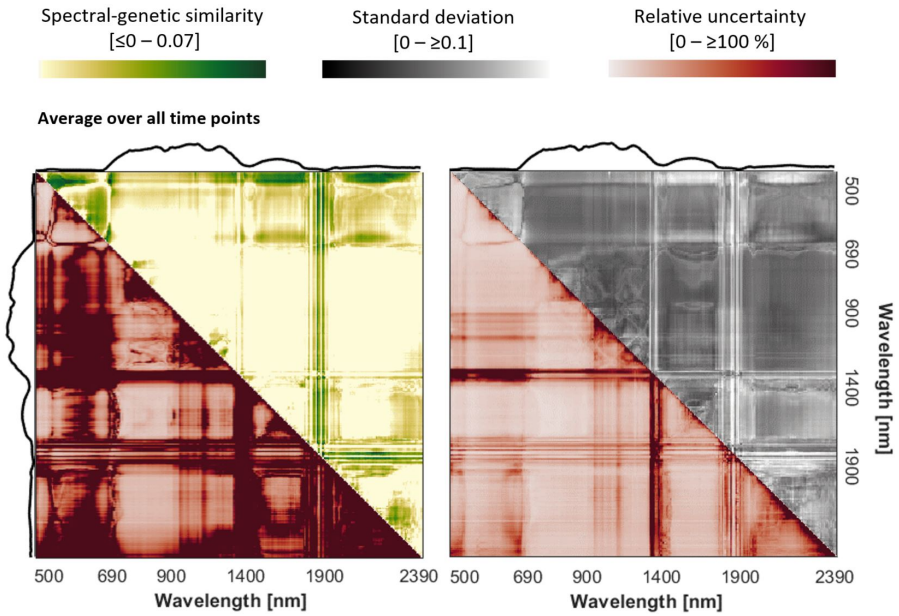

FIGURE S.3.2: Spectral-genetic similarities (values of partial Mantel r) for the population of *F. sylvatica* derived from normalized differences between spectral bands for mean (left panel) and standard deviation (right panel) from all time points. The shown values range from 0 to 0.07 (dark green). The relative uncertainties of spectral-genetic similarities for mean and standard deviation from all time points are represented on the lower left of panels. The shape of the mean canopy reflectance spectrum of all trees is shown at the top and left of the panels.

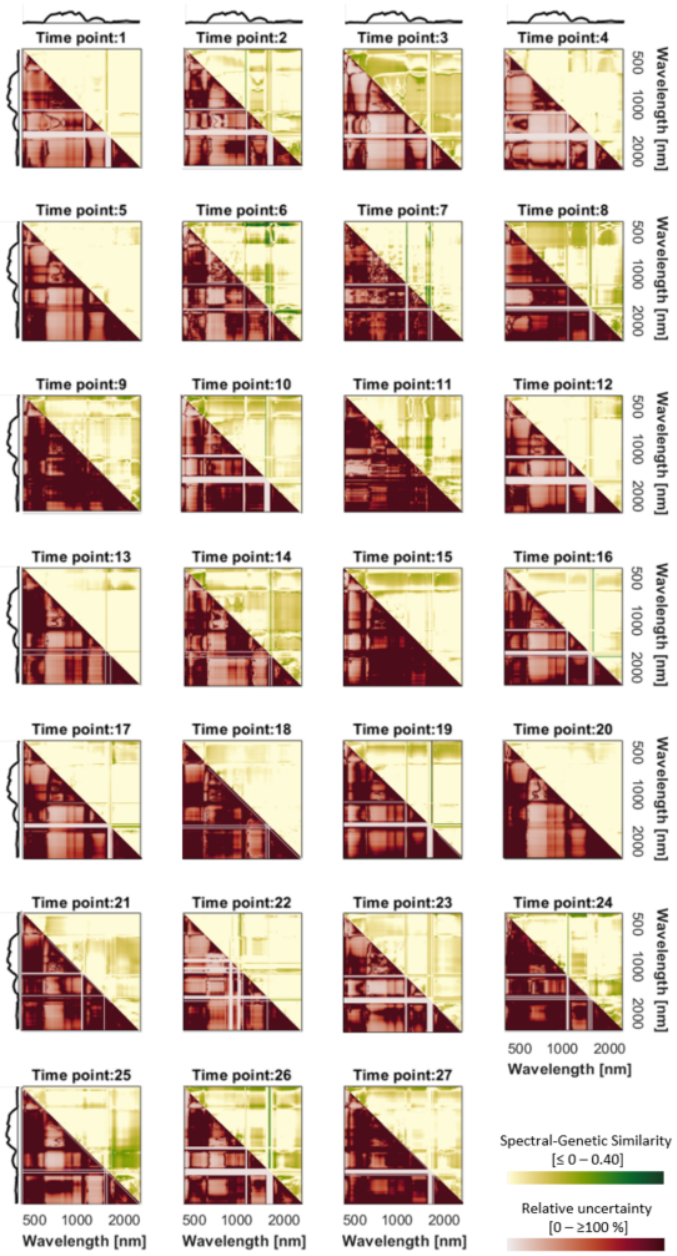

FIGURE S.3.3: Spectral-genetic similarities (values of partial Mantel r) for the population of *F. sylvatica* derived from relative differences between spectral bands for all time points. The shown values range from 0 to 0.4005 (dark green). The relative uncertainties of spectral-genetic similarities are represented on the lower left of panels. The shape of the mean canopy reflectance spectrum of all trees is shown at the top and left of the panels.

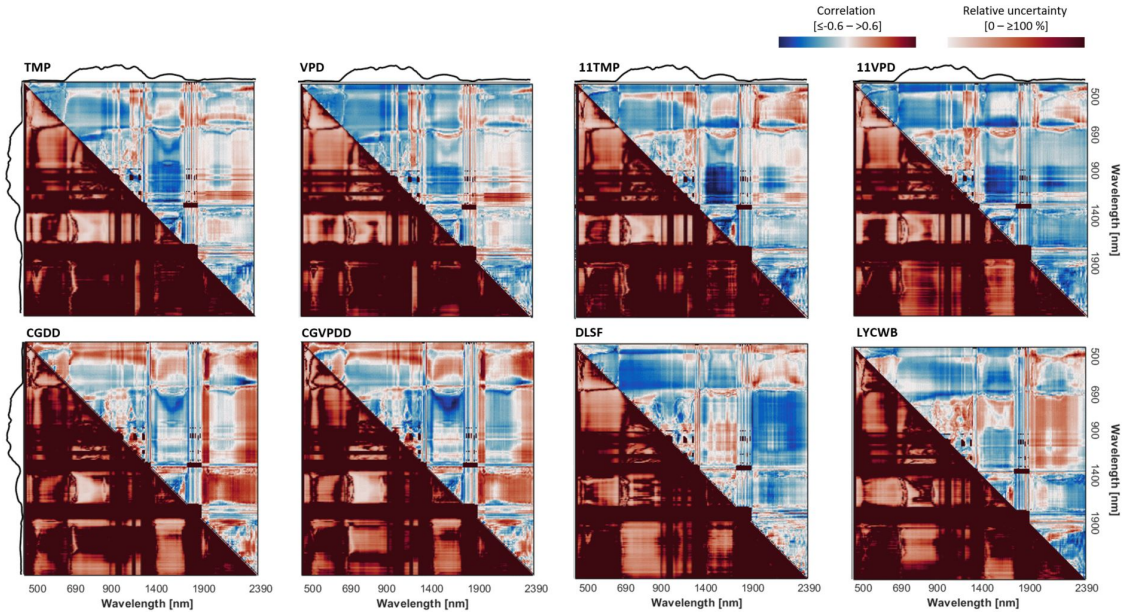

FIGURE S.3.4: Correlations between spectral-genetic similarities derived from relative differences between spectral bands and from left to right: temperature of the day of the acquisition [°C] (TMP) and vapor pressure deficit of the day of the acquisition (VPD), Aggregated Temperature over 110 consecutive days [°C] (11TMP) and Aggregated Vapor Pressure Deficit over 11 consecutive days (11VPD), Cumulative Growing Degree Days [°C] (CGDD) and Cumulative Growing Vapor Pressure Deficit Days (CGVPDD), and Day of the Last Spring Frost [DOY] (DLSF) and Last Year Climatic Water Balance [DOY] (LYCWB). The shown values range from -0.6 (dark blue) to the 0.6 (dark red). The bottom panels show the relative uncertainties of correlations between spectral-genetic similarities and environmental variables. The shape of the mean canopy reflectance spectrum of all trees is shown at the top and left of the panels.

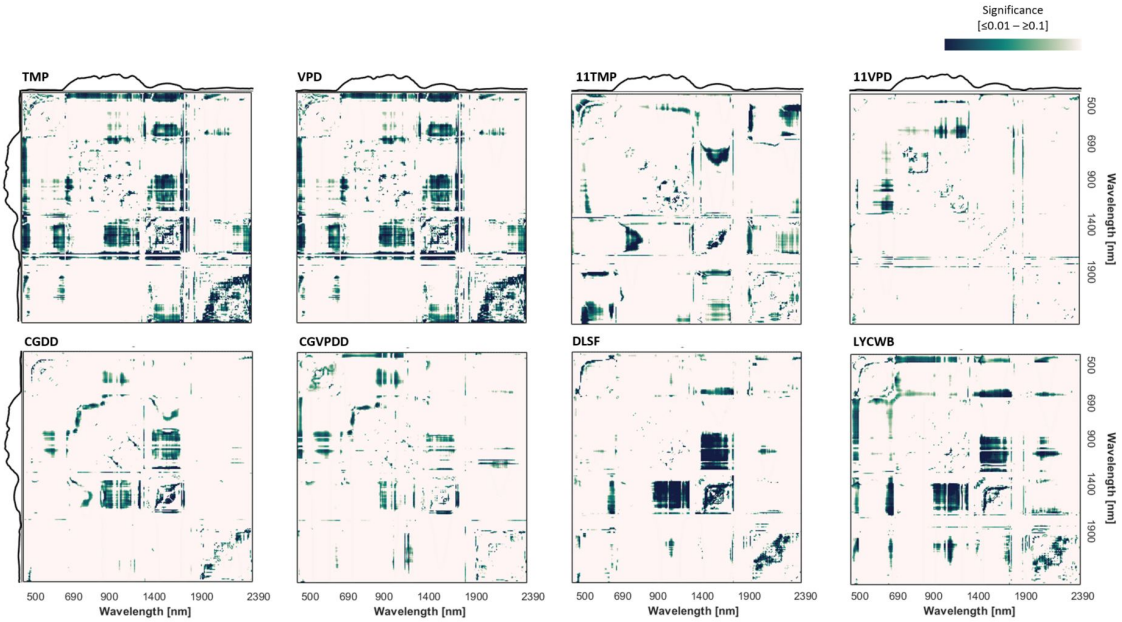

FIGURE S.3.5: Significance as p-values of correlations between spectral-genetic similarities derived from relative differences between spectral bands and from: temperature of the day of the acquisition [°C] (TMP) and vapor pressure deficit of the day of the acquisition (VPD), Aggregated Temperature over 11 consecutive days [°C] (11TMP) and Aggregated Vapor Pressure Deficit over 11 consecutive days (11VPD), Cumulative Growing Degree Days [°C] (CGDD) and Cumulative Growing Vapor Pressure Deficit Days (CGVPDD), and Day of the Last Spring Frost [DOY] (DLSF) and Last Year Climatic Water Balance [DOY] (LYCWB). The shown values range from $p<0.01$ (dark blue) to $p>0.1$ (bright blue).

References

Albert, C. H., Grassein, F., Schurr, F. M., Vieilledent, G., & Violle, C. (2011). When and how should intraspecific variability be considered in trait-based plant ecology? *Perspectives in Plant Ecology, Evolution and Systematics*, *13*(3), 217–225.

Angert, A. L., Huxman, T. E., Chesson, P., & Venable, D. L. (2009). Functional tradeoffs determine species coexistence via the storage effect. *Proceedings of the National Academy of Sciences*, *106*(28), 11641–11645.

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.

Asner, G. P., Martin, R. E., Carlson, K. M., Rascher, U., & Vitousek, P. M. (2006). Vegetation–climate interactions among native and invasive species in hawaiian rainforest. *Ecosystems*, *9*, 1106–1117.

Blowes, S. A., Supp, S. R., Antão, L. H., Bates, A., Bruelheide, H., Chase, J. M., Moyes, F., Magurran, A., McGill, B., Myers-Smith, I. H., et al. (2019). The geography of biodiversity change in marine and terrestrial assemblages. *Science*, *366*(6463), 339–345.

Bolnick, D. I., Amarasekare, P., Araújo, M. S., Bürger, R., Levine, J. M., Novak, M., Rudolf, V. H., Schreiber, S. J., Urban, M. C., & Vasseur, D. A. (2011). Why intraspecific trait variation matters in community ecology. *Trends in ecology & evolution*, *26*(4), 183–192.

Carlson, R. E., Yarger, D. N., & Shaw, R. H. (1971). Factors affecting the spectral properties of leaves with special emphasis on leaf water status 1. *Agronomy Journal*, *63*(3), 486–489.

Carmon, N., Thompson, D. R., Bohn, N., Susiluoto, J., Turmon, M., Brodrick, P. G., Connelly, D. S., Braverman, A., Cawse-Nicholson, K., Green, R. O., et al. (2020). Uncertainty quantification for a global imaging spectroscopy surface composition investigation. *Remote Sensing of Environment*, *251*, 112038.

Carter, G. A. (1993). Responses of leaf spectral reflectance to plant stress. *American journal of botany*, *80*(3), 239–243.

Carter, G. A., Rebbeck, J., & Percy, K. E. (1995). Leaf optical properties in liriodendron tulipifera and pinus strobus as influenced by increased atmospheric ozone and carbon dioxide. *Canadian Journal of Forest Research*, *25*(3), 407–412.

Castro-Esau, K. L., Sánchez-Azofeifa, G. A., Rivard, B., Wright, S. J., & Quesada, M. (2006). Variability in leaf optical properties of mesoamerican trees and the potential for species classification. *American Journal of Botany*, *93*(4), 517–530.

Cavender-Bares, J. (2019). Diversification, adaptation, and community assembly of the american oaks (quercus), a model clade for integrating ecology and evolution. *New Phytologist*, *221*(2), 669–692.

Cavender-Bares, J., Meireles, J. E., Couture, J. J., Kaproth, M. A., Kingdon, C. C., Singh, A., Serbin, S. P., Center, A., Zuniga, E., Pilz, G., et al. (2016). Associations of leaf spectra with genetic and phylogenetic variation in oaks: Prospects for remote detection of biodiversity. *Remote Sensing*, *8*(3), 221.

Cawse-Nicholson, K., Townsend, P. A., Schimel, D., Assiri, A. M., Blake, P. L., Buongiorno, M. F., Campbell, P., Carmon, N., Casey, K. A., Correa-Pabón, R. E., et al. (2021). Nasa's surface biology and geology designated observable: A perspective on surface imaging algorithms. *Remote Sensing of Environment*, *257*, 112349.

Chavana-Bryant, C., Malhi, Y., Wu, J., Asner, G. P., Anastasiou, A., Enquist, B. J., Cosio Caravasi, E. G., Doughty, C. E., Saleska, S. R., Martin, R. E., et al. (2017). Leaf aging of amazonian canopy trees as revealed by spectral and physiochemical measurements. *New Phytologist*, *214*(3), 1049–1063.

Czyż, E. A., Guillén Escribà, C., Wulf, H., Tedder, A., Schuman, M. C., Schneider, F. D., & Schaepman, M. 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.

Des Roches, S., Post, D. M., Turley, N. E., Bailey, J. K., Hendry, A. P., Kinnison, M. T., Schweitzer, J. A., & Palkovacs, E. P. (2018). The ecological importance of intraspecific variation. *Nature ecology & evolution*, *2*(1), 57–64.

Dıaz, S., Kattge, J., Cornelissen, J. H., Wright, I. J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C., Colin Prentice, I., et al. (2016). The global spectrum of plant form and function. *Nature*, *529*(7585), 167–171.

Doyle, J. J. (1990). Isolation of plant dna from faesh tissue. *Focus*, *12*, 13–15.

Eichenberg, D., Purschke, O., Ristok, C., Wessjohann, L., & Bruelheide, H. (2015). Trade-offs between physical and chemical carbon-based leaf defence: Of intraspecific variation and trait evolution. *Journal of Ecology*, *103*(6), 1667–1679.

Etzold, S., Ruehr, N. K., Zweifel, R., Dobbertin, M., Zingg, A., Pluess, P., Häsler, R., Eugster, W., & Buchmann, N. (2011). The carbon balance of two contrasting mountain forest ecosystems in switzerland: Similar annual trends, but seasonal differences. *Ecosystems*, *14*, 1289–1309.

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.

Falconer, D. S. (1996). *Introduction to quantitative genetics*. Pearson Education India.

Farmer, A. M. (1993). The effects of dust on vegetation—a review. *Environmental pollution*, *79*(1), 63–75.

Feret, J.-B., & Asner, G. P. (2012). Tree species discrimination in tropical forests using airborne imaging spectroscopy. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(1), 73–84.

Frye, H. A., Aiello-Lammens, M. E., Euston-Brown, D., Jones, C. S., Kilroy Mollmann, H., Merow, C., Slingsby, J. A., van Der Merwe, H., Wilson, A. M., & Silander Jr, J. A. (2021). Plant spectral diversity as a surrogate for species, functional and phylogenetic diversity across a hyper-diverse biogeographic region. *Global Ecology and Biogeography*, *30*(7), 1403–1417.

Gamon, J., Penuelas, J., & Field, C. (1992). A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing of environment*, *41*(1), 35–44.

Gamon, J., Somers, B., Malenovskỳ, Z., Middleton, E., Rascher, U., & Schaepman, M. E. (2019). Assessing vegetation function with imaging spectroscopy. *Surveys in Geophysics*, *40*, 489–513.

Gao, B.-C. (1996). Ndwi—a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote sensing of environment*, *58*(3), 257–266.

Garcia, M., & Ustin, S. L. (2001). Detection of interannual vegetation responses to climatic variability using aviris data in a coastal savanna in california. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(7), 1480–1490.

Gausman, H., Allen, W., Escobar, D., Rodriguez, R., & Cardenas, R. (1971). Age effects of cotton leaves on light reflectance, transmittance, and absorptance and on water content and thickness 1. *Agronomy Journal*, *63*(3), 465–469.

Gitelson, A. A., Keydan, G. P., & Merzlyak, M. N. (2006). Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves. *Geophysical research letters*, *33*(11).

Goetz, A. F. (2009). Three decades of hyperspectral remote sensing of the earth: A personal view. *Remote sensing of environment*, *113*, S5–S16.

Gonzalez de Andres, E., Rosas, T., Camarero, J. J., & Martınez-Vilalta, J. (2021). The intraspecific variation of functional traits modulates drought resilience of european beech and pubescent oak. *Journal of Ecology*, *109*(10), 3652–3669.

Guillén-Escribà, C., Schneider, F. D., Schmid, B., Tedder, A., Morsdorf, F., Furrer, R., Hueni, A., Niklaus, P. A., & Schaepman, M. E. (2021). Remotely sensed between-individual functional trait variation in a temperate forest. *Ecology and Evolution*, *11*(16), 10834–10867.

Hargreaves, G. H. (1994). Defining and using reference evapotranspiration. *Journal of irrigation and drainage engineering*, *120*(6), 1132–1139.

He, J.-S., Wang, X., Flynn, D. F., Wang, L., Schmid, B., & Fang, J. (2009). Taxonomic, phylogenetic, and environmental trade-offs between leaf productivity and persistence. *Ecology*, *90*(10), 2779–2791.

Hueni, A., Biesemans, J., Meuleman, K., Dell'Endice, F., Schlapfer, D., Odermatt, D., Kneubuehler, M., Adriaensen, S., Kempenaers, S., Nieke, J., et al. (2008). 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.

Hueni, A., Damm, A., Kneubuehler, M., Schläpfer, D., & Schaepman, M. E. (2016). Field and airborne spectroscopy cross validation—some considerations. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(3), 1117–1135.

Hueni, A., Sterckx, S., Jehle, M., D'Odorico, P., Vreys, K., Bomans, B., Biesemans, J., Meuleman, K., & Schaepman, M. (2012). Operational status of apex and characteristics of the apex open science data set. *2012 IEEE international geoscience and remote sensing symposium*, 5009–5012.

Hueni, A., Woolliams, E., Schlaepfer, D., & Wulf, H. (2017). Apex airborne imaging spectrometer uncertainty budget and vicarious validation method. *AGU Fall Meeting Abstracts*, *2017*, A11A–1860.

Hughes, A. R., Inouye, B. D., Johnson, M. T., Underwood, N., & Vellend, M. (2008). Ecological consequences of genetic diversity. *Ecology letters*, *11*(6), 609–623.

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.

Hut, R. A., Kronfeld-Schor, N., van der Vinne, V., & De la Iglesia, H. (2012). In search of a temporal niche: Environmental factors. *Progress in brain research*, *199*, 281–304.

Imbens, G. W., & Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Jehle, M., Hueni, A., Lenhard, K., Baumgartner, A., & Schaepman, M. E. (2014). Detection and correction of radiance variations during spectral calibration in apex. *IEEE Geoscience and Remote Sensing Letters*, *12*(5), 1023–1027.

Jetz, W., Cavender-Bares, J., Pavlick, R., Schimel, D., Davis, F. W., Asner, G. P., Guralnick, R., Kattge, J., Latimer, A. M., Moorcroft, P., et al. (2016). Monitoring plant functional diversity from space. *Nature plants*, *2*(3), 1–5.

Joswig, J. S., Wirth, C., Schuman, M. C., Kattge, J., Reu, B., Wright, I. J., Sippel, S. D., Rüger, N., Richter, R., Schaepman, M. E., et al. (2022). Climatic and soil factors explain the two-dimensional spectrum of global plant trait variation. *Nature ecology & evolution*, *6*(1), 36–50.

Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., et al. (2012). An international comparison of individual tree detection and extraction using airborne laser scanning. *Remote Sensing*, *4*(4), 950–974.

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An r package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281.

Kerkhoff, A. J., Fagan, W. F., Elser, J. J., & Enquist, B. J. (2006). Phylogenetic and growth form variation in the scaling of nitrogen and phosphorus in the seed plants. *The American Naturalist*, *168*(4), E103–E122.

Kloeti, P., Keller, H., & Guecheva, M. (1989). Effects of forest canopy on throughfall precipitation chemistry. *Atmospheric Deposition. Proceedings of a Symposium held during the Third Scientific Assembly of the International Association of Hydrological Sciences at Baltimore, Maryland May 1989. IAHS Publication*, (179).

Kreyling, J., Thiel, D., Nagy, L., Jentsch, A., Huber, G., Konnert, M., & Beierkuhnlein, C. (2012). Late frost sensitivity of juvenile fagus sylvatica l. differs between southern germany and bulgaria and depends on preceding air temperature. *European Journal of Forest Research*, *131*, 717–725.

Kükenbrink, D., Schneider, F. D., Schmid, B., Gastellu-Etchegorry, J.-P., Schaepman, M. E., & Morsdorf, F. (2021). Modelling of three-dimensional, diurnal light extinction in two contrasting forests. *Agricultural and Forest Meteorology*, *296*, 108230.

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.

Levine, J. M., & HilleRisLambers, J. (2009). The importance of niches for the maintenance of species diversity. *Nature*, *461*(7261), 254–257.

Lin, M., Simons, A. L., Harrigan, R. J., Curd, E. E., Schneider, F. D., Ruiz-Ramos, D. V., Gold, Z., Osborne, M. G., Shirazi, S., Schweizer, T. M., et al. (2021). Landscape analyses using edna metabarcoding and earth observation predict community biodiversity in california. *Ecological Applications*, *31*(6), e02379.

Long, L. M., Patel, H. P., Cory, W. C., & Stapleton, A. E. (2003). The maize epicuticular wax layer provides uv protection. *Functional plant biology*, *30*(1), 75–81.

Ma, Z., Jia, G., Schaepman, M. E., & Zhao, H. (2020). Uncertainty analysis for topographic correction of hyperspectral remote sensing images. *Remote Sensing*, *12*(4), 705.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer research*, *27*(2_Part_1), 209–220.

Marks, C. O., & Lechowicz, M. J. (2006). Alternative designs and the evolution of functional diversity. *The American Naturalist*, *167*(1), 55–66.

Matile, P., Hörtensteiner, S., & Thomas, H. (1999). Chlorophyll degradation. *Annual review of plant biology*, *50*(1), 67–95.

Menzel, A., Helm, R., & Zang, C. (2015). Patterns of late spring frost leaf damage and recovery in a european beech (fagus sylvatica l.) stand in south-eastern germany based on repeated digital photographs. *Frontiers in Plant Science*, *6*, 110.

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

Moritz, C. (2002). Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic biology*, *51*(2), 238–254.

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.

Morsdorf, F., Schneider, F. D., Gullien, C., Kükenbrink, D., Leiterer, R., & Schaepman, M. E. (2020). The laegeren site: An augmented forest laboratory: Combining 3-d reconstruction and radiative transfer models for trait-based assessment of functional diversity. *Remote sensing of plant biodiversity*, 83–104.

Moss, R., & Loomis, W. (1952). Absorption spectra of leaves. i. the visible spectrum. *Plant Physiology*, *27*(2), 370.

Myneni, R. B., Keeling, C., Tucker, C. J., Asrar, G., & Nemani, R. R. (1997). Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature*, *386*(6626), 698–702.

Nei, M. (1972). Genetic distance between populations. *The American Naturalist*, *106*(949), 283–292.

Niinemets, Ü. (2001). Global-scale climatic controls of leaf dry mass per area, density, and thickness in trees and shrubs. *Ecology*, *82*(2), 453–469.

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.

Oldeland, J., Dorigo, W., Wesuls, D., & Jürgens, N. (2010). Mapping bush encroaching species by seasonal differences in hyperspectral imagery. *Remote Sensing*, *2*(6), 1416–1438.

Ollinger, S. V. (2011). Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytologist*, *189*(2), 375–394.

Ollinger, S. V., Smith, M.-L., Martin, M. E., Hallett, R. A., Goodale, C. L., & Aber, J. D. (2002). Regional variation in foliar chemistry and n cycling among forests of diverse history and composition. *Ecology*, *83*(2), 339–355.

Pastorelli, R., Smulders, M., Van't Westende, W., Vosman, B., Giannini, R., Vettori, C., & Vendramin, G. (2003). Characterization of microsatellite markers in fagus sylvatica l. and fagus orientalis lipsky. *Molecular Ecology Notes*, *3*(1), 76–78.

Patton, A., & Gieseker, L. (1942). Seasonal changes in the lignin and cellulose content of some montana grasses. *Journal of Animal Science*, *1*(1), 22–26.

Peñuelas, J., Gamon, J., Fredeen, A., Merino, J., & Field, C. (1994). Reflectance indices associated with physiological changes in nitrogen-and water-limited sunflower leaves. *Remote sensing of Environment*, *48*(2), 135–146.

Pereira, H. M., Ferrier, S., Walters, M., Geller, G. N., Jongman, R. H., Scholes, R. J., Bruford, M. W., Brummitt, N., Butchart, S. H., Cardoso, A., et al. (2013). Essential biodiversity variables. *Science*, *339*(6117), 277–278.

Peterson, D. L., Aber, J. D., Matson, P. A., Card, D. H., Swanberg, N., Wessman, C., & Spanner, M. (1988). Remote sensing of forest canopy and leaf biochemical contents. *Remote Sensing of Environment*, *24*(1), 85–108.

Petibon, F., Czyż, E. A., Ghielmetti, G., Hueni, A., Kneubühler, M., Schaepman, M. E., & Schuman, M. C. (2021). Uncertainties in measurements of leaf optical properties are small compared to the biological variation within and between individuals of european beech. *Remote Sensing of Environment*, *264*, 112601.

Pettorelli, N., Safi, K., & Turner, W. (2014). Satellite remote sensing, biodiversity research and conservation of the future.

Peuke, A., Schraml, C., Hartung, W., & Rennenberg, H. (2002). Identification of drought-sensitive beech ecotypes by physiological parameters. *New phytologist*, *154*(2), 373–387.

Polgar, C. A., & Primack, R. B. (2011). Leaf-out phenology of temperate woody plants: From trees to ecosystems. *New phytologist*, *191*(4), 926–941.

Proença, V., & Pereira, H. M. (2017). Comparing extinction rates: Past, present, and future.

Reich, P. B., Walters, M. B., Ellsworth, D. S., Vose, J. M., Volin, J. C., Gresham, C., & Bowman, W. D. (1998). Relationships of leaf dark respiration to leaf nitrogen, specific leaf area and leaf life-span: A test across biomes and functional groups. *Oecologia*, *114*, 471–482.

Reich, P. B., Wright, I. J., Cavender-Bares, J., Craine, J., Oleksyn, J., Westoby, M., & Walters, M. (2003). The evolution of plant functional variation: Traits, spectra, and strategies. *International Journal of Plant Sciences*, *164*(S3), S143–S164.

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., Schläpfer, D., & Müller, A. (2010). Operational atmospheric correction for imaging spectrometers accounting for the smile effect. *IEEE Transactions on Geoscience and Remote sensing*, *49*(5), 1772–1780.

Rivera, G., Elliott, S., Caldas, L. S., Nicolossi, G., Coradin, V. T., & Borchert, R. (2002). Increasing day-length induces spring flushing of tropical dry forest trees in the absence of rain. *Trees*, *16*(7), 445.

Roberts, D., Green, R., & Adams, J. (1997). Temporal and spatial patterns in vegetation and atmospheric properties from aviris. *Remote sensing of environment*, *62*(3), 223–240.

Rocchini, D., Balkenhol, N., Carter, G. A., Foody, G. M., Gillespie, T. W., He, K. S., Kark, S., Levin, N., Lucas, K., Luoto, M., et al. (2010). Remotely sensed spectral heterogeneity as a proxy of species diversity: Recent advances and open challenges. *Ecological Informatics*, *5*(5), 318–329.

Schaepman, M. E., & Dangel, S. (2000). Solid laboratory calibration of a nonimaging spectroradiometer. *Applied Optics*, *39*(21), 3754–3764.

Schaepman, M. E., Jehle, M., Hueni, A., D'Odorico, P., Damm, A., Weyermann, J., Schneider, F. D., Laurent, V., Popp, C., Seidel, F. C., et al. (2015). Advanced radiometry measurements and earth science applications with the airborne prism experiment (apex). *Remote Sensing of Environment*, *158*, 207–219.

Schaepman-Strub, G., Schaepman, M. E., Painter, T. H., Dangel, S., & Martonchik, J. V. (2006). Reflectance quantities in optical remote sensing—definitions and case studies. *Remote sensing of environment*, *103*(1), 27–42.

Schimel, D., Schneider, F. D., Carbon, J., & Participants, E. (2019). Flux towers in the sky: Global ecology from space. *New Phytologist*, *224*(2), 570–584.

Schläpfcr, D., Schaepman, M., Bojinski, S., & Börner, A. (2000). Calibration and validation concept for the airborne prism experiment (apex). *Canadian journal of remote sensing*, *26*(5), 455–465.

Schläpfer, D., & Richter, R. (2002). Geo-atmospheric processing of airborne imaging spectrometry data. part 1: Parametric orthorectification. *International Journal of Remote Sensing*, *23*(13), 2609–2630.

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.

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.

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.

Schweiger, A. K., Cavender-Bares, J., Townsend, P. A., Hobbie, S. E., Madritch, M. D., Wang, R., Tilman, D., & 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.

Shepherd, T., & Wynne Griffiths, D. (2006). The effects of stress on plant cuticular waxes. *New Phytologist*, *171*(3), 469–499.

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., Coops, N. C., Neinavaz, E., Ali, A., Schaepman, M. E., Paganini, M., Kissling, W. D., Vihervaara, P., Darvishzadeh, R., Feilhauer, H., et al. (2021). Priority list of biodiversity metrics to observe from space. *Nature ecology & evolution*, *5*(7), 896–906.

Skidmore, A. K., Pettorelli, N., Coops, N. C., Geller, G. N., Hansen, M., Lucas, R., Mücher, C. A., O'Connor, B., Paganini, M., Pereira, H. M., et al. (2015). Environmental science: Agree on biodiversity metrics to track from space. *Nature*, *523*(7561), 403–405.

Smouse, P. E., Long, J. C., & Sokal, R. R. (1986). Multiple regression and correlation extensions of the mantel test of matrix correspondence. *Systematic zoology*, *35*(4), 627–632.

Somers, B., & Asner, G. P. (2013). Multi-temporal hyperspectral mixture analysis and feature selection for invasive species mapping in rainforests. *Remote Sensing of Environment*, *136*, 14–27.

Sturm, J., Santos, M. J., Schmid, B., & Damm, A. (2022). Satellite data reveal differential responses of swiss forests to unprecedented 2018 drought. *Global Change Biology*, *28*(9), 2956–2978.

Suárez, L., Zarco-Tejada, P. J., Sepulcre-Cantó, G., Pérez-Priego, O., Miller, J., Jiménez-Muñoz, J., & Sobrino, J. (2008). Assessing canopy pri for water stress detection with diurnal airborne imagery. *Remote Sensing of Environment*, *112*(2), 560–575.

Thompson, D. R., Babu, K., Braverman, A. J., Eastwood, M. L., Green, R. O., Hobbs, J. M., Jewell, J. B., Kindel, B., Massie, S., Mishra, M., et al. (2019). Optimal estimation of spectral surface reflectance in challenging atmospheres. *Remote Sensing of Environment*, *232*, 111258.

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.

Turner, W. (2014). Sensing biodiversity. *Science*, *346*(6207), 301–302.

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*, S67–S77.

Vanderbilt, V. C., & Grant, L. (1985). Plant canopy specular reflectance model. *IEEE Transactions on Geoscience and Remote Sensing*, (5), 722–730.

Vögtli, M., Schläpfer, D., Richter, R., Hueni, A., Schaepman, M. E., & Kneubühler, M. (2020). About the transferability of topographic correction methods from spaceborne to airborne optical data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 1348–1362.

Wang, Y., Hyyppä, J., Liang, X., Kaartinen, H., Yu, X., Lindberg, E., Holmgren, J., Qin, Y., Mallet, C., Ferraz, A., et al. (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.

Wang, Z., Townsend, P. A., Schweiger, A. K., Couture, J. J., Singh, A., Hobbie, S. E., & Cavender-Bares, J. (2019). Mapping foliar functional traits and their uncertainties across three years in a grassland experiment. *Remote Sensing of Environment*, *221*, 405–416.

Westoby, M. (1998). A leaf-height-seed (lhs) plant ecology strategy scheme. *Plant and soil*, *199*, 213–227.

Westoby, M., & Wright, I. J. (2006). Land-plant ecology on the basis of functional traits. *Trends in ecology & evolution*, *21*(5), 261–268.

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

Woolliams, E., Hueni, A., & Gorroño, J. (2014). Intermediate uncertainty analysis for earth observation (instrument calibration module): Training course textbook.

Wright, I. J., Reich, P. B., & Westoby, M. (2003). Least-cost input mixtures of water and nitrogen for photosynthesis. *The American Naturalist*, *161*(1), 98–111.

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J. H., Diemer, M., et al. (2004). The worldwide leaf economics spectrum. *Nature*, *428*(6985), 821–827.

Yamasaki, E., Altermatt, F., Cavender-Bares, J., Schuman, M. C., Zuppinger-Dingley, D., Garonna, I., Schneider, F. D., Guillén-Escribà, C., van Moorsel, S. J., Hahl, T., et al. (2017). Genomics meets remote sensing in global change studies: Monitoring and predicting phenology, evolution and biodiversity. *Current opinion in environmental sustainability*, *29*, 177–186.

Zheng, Z., Zeng, Y., Schneider, F. D., Zhao, Y., Zhao, D., Schmid, B., Schaepman, M. E., & Morsdorf, F. (2021). Mapping functional diversity using individual tree-based morphological and physiological traits in a subtropical forest. *Remote Sensing of Environment*, *252*, 112170.

Chapter 4

Remote observation of genetic variation in a forest tree species (*Fagus sylvatica*) across its natural range

Ewa A. Czyż, Bernhard Schmid, Maarten B. Eppinga, Marylaure de la Harpe, Aboubakr Moradi, Cheng Li, Michael E. Schaepman, Meredith C. Schuman

To be submitted as:
Remote observation of genetic variation in a forest tree species (Fagus sylvatica) across its natural range.
It is reprinted as the near final version of the manuscript and has been modified to fit the layout of this thesis.

Author contributions. Conceptualization: EAC, BS, MBE, MH, MCS, MES. Methodology: EAC, BS, MBE, AM, MH, MCS, MES. Software: EAC, BS, CL. Validation: EAC, BS, MBE. Formal analysis: EAC, BS, MBE. Investigation: EAC, BS, MBE, MH, CL, MCS, MES. Resources: MCS, MES. Data Curation: EAC. Writing - Original Draft: EAC. Writing - Review and Editing: EAC, BS, MH, MCS, CL, MES. Visualization: EAC. Supervision: BS, MCS, MES. Project administration: MCS, MES. Funding acquisition: MES.

Detailed author contributions

EA Czyż conceptualized the study together with ME Schaepman and further contributions of B Schmid, MB Eppinga, MC Schuman and M de La Harpe. The project was administrated by ME Schaepman and data collection was administrated by MC Schuman. ME Schaepman obtained funding. ME Schaepman, MC Schuman and B Schmid supervised the work. EA Czyż estabilished contact at Bracciano Natural Reserve (IT), Mt. Etna (IT), Sonian Forest (BE), Bokskog Forest (SE), Maskulińskie Forestry (PL), Mt. Bukowa (PL), Bieszczadski NP (PL), Paklenica NP (HR) and Krokar Forest (SI). M de La Harpe estabilished contact at Mt. Vulture (IT), Mt. Pollino NP (IT), Mt. Pierfaone (IT), Madonie NP (IT), Vizzavona (FR), Plateau du Coscione (FR), Sainte-Baume (FR), Mt. Moncayo (ES) and Mt. Peña (ES). MC Schuman estabilished contact at Fraska Gora (RS), Mt. Boranja (RS) and Massane Natural Reserve (FR). EA Czyż, MC Schuman and M de La Harpe estabilished scheme of sampling genetic and spectral data. EA Czyż and M de La Harpe collected the data with support from C Li. MC Schuman and further contributions of M de La Harpe and A Moradi estabilished the methodology on storage, extraction, sequencing and analysis of genetic data. EA Czyż extracted DNA. EA Czyż supported by C Li developed bioinformatic pipeline to derive Single-Nucleotide Polymorphisms (SNPs) among genomes of sampled individuals. MC Schuman provided access to the computing resources. EA Czyż curated spectral data. EA Czyż, B Schmid, MB Eppinga, MC Schuman and ME Schaepman established analysis on linking spectral variation with genetic and environmental variation. EA Czyż conducted the analysis that were validated by MB Eppinga and B Schmid. The original draft of the manuscript with figures was written by EA Czyż with edits from MC Schuman, M de La Harpe, MB Eppinga, B Schmid, ME Schaepman and C Li. Additional contributions not constituting authorship are acknowledged for the permissions to access the sites and support in sampling.

4.1 Abstract

Remote observations of spectral phenotypes might facilitate global monitoring of within-species genetic variation, a key factor in biodiversity loss. We investigated genetic variation in the forest tree species European beech (*Fagus sylvatica* L.) using 1,380,310 SNPs derived from whole-genome sequences sampled from more than 200 individuals at 22 sites across the species' natural range. From the same top-of-canopy leaves used for nuclear DNA extraction, we measured leaf level reflectance spectra using a contact probe and a field spectroradiometer. We found that information from leaf spectral phenotypes improved predictions of genetic structure made using environmental variables across the sites in our dataset. Combining spectra with environmental variables explained 77% of variation in sample identity along the first two genetic principle coordinates. We propose that imaging spectrometrometers repeatedly acquiring tree canopy spectra across species ranges could contribute to efficient monitoring of genetic diversity in trees.

4.2 Introduction

Biodiversity loss presents a large and growing risk for the global environment and its ability to support life. Plant diversity loss can negatively affect ecosystem functioning (Cardinale et al., 2012; Isbell et al., 2011; Hector and Bagchi, 2007; Reich et al., 2012; Purvis and Hector, 2000), reduce ecosystem stability, e.g. in the face of disturbance (Tilman and Downing, 1994; Hector et al., 2010; Isbell et al., 2015; van Moorsel et al., 2021), and reduce the potential of populations and communities to adapt to environmental change (McNeely et al., 1990; Des Roches et al., 2021). Biodiversity, and the related composition, structure and functioning of ecosystems (Franklin, 1981) are under pressure from human activity (Johnson et al., 2017). To mitigate human impacts on biodiversity decline, many steps need to be taken (Jenkins and Joppa, 2009, Mora et al., 2006); and targets have been set several times aiming to mitigate biodiversity loss worldwide (IPBES, 2019; Dıaz et al., 2020). One important approach towards protecting and restoring biodiversity is its continuous and standardized monitoring (Scholes et al., 2008; Scholes et al., 2008; Walpole et al., 2009; Turner, 2014). Monitoring indicators of biodiversity would help to set conservation priorities (Scott et al., 1989, Reid, 1998), evaluate actions taken, and develop adaptive management (Sutherland et al., 2004) and policy adjustments (Perrings et al., 2011; Sarkar, 2002) to preserve and restore biodiversity. Monitoring schemes should span levels of biological organization from ecosystems

to species and genes (Norse et al., 1986). Genetic diversity is especially challenging to monitor, but can provide valuable information about a species' contribution to ecosystem functioning (Fridley and Grime, 2010; Tang et al., 2022) and its evolutionary potential, and thus is a key component when accounting for biodiversity loss (Wilson and Peter, 1988; Exposito-Alonso et al., 2022).

The assessment of genetic diversity is generally based on field sample collection, with limited temporal and spatial coverage. Furthermore, even with recent advances in field-mobile genomics, sampled material usually requires laboratory extraction and analyses, and is still commonly limited to a portion of whole genomes to save costs, facilitate extraction of data from difficult samples, and allow comparison to existing sequence data (Lou et al., 2021; Pomerantz et al., 2018). These scale-induced limitations result in sparse and biased representations of genetic diversity, which lead to inaccurate representation of its current status, omission of unique but difficult-to-sample sites, and partially misses components of the variation that reflects total adaptation potential of species to environmental change (Hughes et al., 2008; Des Roches et al., 2018). Whole-genome sequencing provides a means to assess intraspecific genetic diversity without the bias of more restricted sets of molecular markers, and the required per-sample investment is approaching the efforts of marker-based approaches (Lou et al., 2021). However, the spatiotemporal limitations still apply. Phenotypic variation acquired over space and time may offer a useful indicator of a species' genetic diversity if a link between the two can be established across different test sites and environmental conditions.

To collect these phenotypic data, imaging spectrometers can provide repeated and spatially contiguous acquisitions of spectral phenotypes comprising integrated signatures of plant traits (Schaepman et al., 2009; Schaepman et al., 2015). Plants reflect solar radiation as a function of absorption and transmission by organic (e.g. pigments, proteins, phenolics) and inorganic molecules (water) and scattering by plant structures (Jacquemoud and Ustin, 2019; Allen et al., 1970; Guyot et al., 1992). Biochemical and structural traits result from the interaction of plant genes with the environment, and might affect the spectral phenotypes of individuals from leaf to canopy levels (Meireles et al., 2020). Leaf optical properties are commonly estimated from contact probe measurements using field spectroradiometers, while the optical properties of entire canopies are derived from remote acquisitions of reflected radiation by imaging spectrometers (Kupiec and Curran, 1995). Airborne and spaceborne imaging spectrometers overcome spatiotemporal limitations of ground measurements and provide a more representative sampling of tree

canopies, acquiring spatially explicit, landscape-scale information about the distribution of plants traits (Peterson et al., 1988; Wessman, 1987; Schneider et al., 2017; Singh et al., 2015). Spectral phenotypes can indicate genetic groups and variants, and thus spectral acquisitions might facilitate monitoring of genetic diversity under natural conditions (Madritch et al., 2014; Czyż et al., 2020; Czyż et al., 2023), if a link between spectral and genetic variation can be established (Cavender-Bares et al., 2016; Yamasaki et al., 2017). Until now, however, this potential remains uncertain as previous studies associating genetic and spectroscopy data collected under natural conditions have focused on relatively small numbers of sites covering a small proportion of the studied species' range, and did not integrate whole-genome information.

Here, we demonstrate the potential of spectral information to contribute to assessing genetic diversity of a plant species across its natural range. We use *Fagus sylvatica* L. (European beech) as our test species because it is a dominant, ecologically and economically important tree species in Europe which has gained interest as a model for ecological genomics of tree species under global change (e.g. Capblancq et al., 2020; Pfenninger et al., 2020; Mishra et al., 2022). Patterns of *F. sylvatica* genetic diversity and structure have been studied with marker-based approaches including allozymes (Gallois et al., 1998), AFLP (Guevara et al., 2022), microsatellites and, more recently, SNPs in candidate genes (Meger et al., 2021, reviewed in Stefanini et al., 2023). However, no study so far used whole-genome sequences from across the range of *F. sylvatica* range. The *F. sylvatica* genome size is estimated to be 541 Mbp, and a chromosome-level assembly is available (Mishra et al., 2022).

In the present study we used nuclear whole-genome sequences of trees sampled across a large portion of the species range to derive genetic structure, which is related to the phylogeographic history of the species (Capblancq et al., 2020) and its adaptation to site environmental conditions (Pluess and Weber, 2012). We incorporate spectral phenotypes derived from handheld measurements and use these in combination with environmental variables to predict the genetic structure. We thereby (1) contribute to current knowledge about genomic variation and structure in *F. sylvatica*; and (2) indicate how spectral phenotypes acquired in natural environments may help to predict it. The results suggest that aerial and satellite-based imaging spectroscopy could be leveraged for genetic diversity monitoring and conservation over larger areas and at higher temporal resolution than typically attainable when solely relying on field sampling.

4.3 Materials and Methods

4.3.1 Study areas and sample collection

We selected 22 sites dominated by *Fagus sylvatica* L. (common beech) across its entire natural range in Europe (Fig. 4.1a). We abbreviate the site names to the first two letters referring to the country that the site is located in, and the third letter referring to the particular location in the country. Most of the sites are covered by protected, old-growth beech forests. The variation among the sites accounts for climatic (mean 50% of the range and 95% of the distribution of environmental values across the range; Fig. S.4.1), edaphic and phylogeographic variation (Caudullo et al., 2017). Data were collected *in situ* during August–September 2021. At each site we sampled 10 *F. sylvatica* trees that occupied canopy space along two ca. 10 x 200 m transects separated by 50 m. Each transect had 5 trees with at least 30 m between each other. At some sites we sampled fewer than 10 trees due to a lack of *F. sylvatica* trees in the canopy within the transect (4 trees at RSB, 8 trees at ESP) and in some cases the sampled material was of a poor quality for the DNA extraction (1 tree at PLM, 1 tree at ITP). When possible, we selected transects on south-west to south-east-facing slopes to normalize for local adaptation to exposition to solar radiation. From each tree, we collected a top-of-canopy branch of 20 to 100 cm length with the help of tree climbers or, when tree height did not exceed 15 m, with a telescoping carbon fiber pruner (Takeni Trading Co., Osaka, Japan). From each branch, we randomly selected three undamaged leaves separated by at least 5 cm, from which we acquired spectral data after harvesting (see 4.3.3) and which we subsequently stored on silica gel for later genetic analysis (see 4.3.2).

4.3.2 Genetic data

The collected leaves were taken back to the laboratory where we took 50 mg of undamaged leaf lamina and then stored these samples in tubes at –20 °C until grinding and DNA extraction. The DNA was purified using a Norgen Plant and Fungi genomic DNA extraction kit (Norgen Biotek, Thorold, ON, Canada), modified to reduce RNA contamination and to improve lysis, yield, and purity. Specifically, samples were flash-frozen in liquid nitrogen immediately prior to grinding and ground using 3 mm glass beads at 30 Hz for 3 min in a TissueLyser II (Qiagen) into a fine powder. The powder was briefly centrifuged before adding lysis buffer and vortexed to mix with the buffer. One μl of RNAse A (DNAse-free, 100,000 units/ml in 50% glycerol, 10 mM Tris-HCl, pH 8.0) and

3 μl Proteinase K (>600 μl 20 mg/ml in 10 mM Tris-HCl pH 7.5, containing calcium acetate and 50% v/v glycerol) were added to the lysis buffer prior to 10 min incubation at 56°C with 1000 rpm agitation. An additional 1 μl of RNAse A was subsequently added to the reaction, followed by a 10 min incubation at 65°C with 2000 rpm agitation. The samples were then processed according to the manufacturer's protocol with the modification that 70% ethanol was mixed with lysate by pipetting ca. 10x/sample instead of vortexing. The DNA from clarified lysate was captured on spin columns (centrifugation until all lysate passed) and columns were washed three times with Solution A prior to elution with Elution Buffer B (100 μl, 5 min RT incubation, spinning until all elution buffer was recovered).

The purified DNA was sequenced by Illumina NovaSeq PE150 paired-end sequencing (Novogene, Cambridge, UK), with read length 150 bp and PCR-based library preparation, to produce ca. 11 Gbp/sample, resulting in a mean ca. 20x coverage of the 541 Mbp nuclear genome (Mishra et al., 2022). The Illumina adapters were trimmed from the raw reads, and reads with quality below 3 and windows of 4 bases with average quality below 15 were also removed using *trimmomatic* v. 0.40. We aligned reads against the chromosome-level reference genome (Mishra et al., 2022) in *bowtie2* v. 2.4.5 and removed alignment duplicates and unaligned reads in *picard* v. 2.0.1. Sequentially, we derived sets of alleles per each individual tree by calling variants against the chromosome-level reference genome (Mishra et al., 2022) using the function *HaplotypeCaller* in *GATK4*. We consolidated the derived alleles and obtained the joint SNPs among all individuals using the function *GenotypeGVCFs* in *GATK4*. From the derived SNPs, we selected those with phred scores above 30 (i.e., accuracy higher than 99.9%), minor allele frequencies of more than 0.1, and minimum and maximum allele depths of 10 and 50, respectively. We pruned the derived SNPs for linkage disequilibrium by removing SNPs that had a correlation coefficient of greater than 0.1 with each other in *plink* v. 1.9 software. This resulted in 1,380,310 SNPs across all sampled individuals used in further analyses.

4.3.3 Spectral data

Spectral data comprised leaf reflectance spectra from 350–2500 nm calculated from measurements taken using an ASD FieldSpec 4 field spectroradiometer (serial n° 18739, ASD Inc., Boulder, CO, USA), coupled to a plant probe and leaf clip attachment equipped with a standardized halogen light source and a set of standard black and white Spectralon reference backgrounds (serial n°445, ASD Inc., Boulder, CO, USA). Measurements of three

leaf samples per tree were acquired within minutes and up to 3 h (for ITE, ITM, 5 trees at ITV, and 5 trees at PLB, which were restricted by harvesting time) from the time of branch harvest using the contact probe positioned in the middle of the leaf and to the right of the midrib as viewed from the adaxial side (the side facing the light source and measurement optics). Measurements comprised five successive readings per leaf for four types of acquisitions: the white reference (R_w), leaf with white reference as background (T_w), the black reference (R_b), and the leaf with black reference as background (T_b) (cf., 4.1). For each reading, five scans, each with an integration time of 8.5 ms, were recorded and averaged per leaf. The averaged reflectance spectra per tree was calculated from the mean of leaf reflectance (R) using a formula for background-differentiated measurements (Miller et al., 1992):

$$R = \frac{T_b * R_w - T_w * R_b}{R_w - R_b} * 100[\%] \tag{4.1}$$

The ASD instrument interpolates the measured data to provide a read-out that is sampled to 1 nm spectral sampling interval steps, although the spectral resolution of the instrument ranges from 3–10 nm. To reduce the collinearity of neighboring spectral bands, we convolved the spectral signal from the original 2151 spectral bands to 216 spectral bands by linearly interpolating the original signal using *interp1* function in *Matlab* (v. R2017b) followed by multiplication with a Spectral Response Function (SRF) of Gaussian shape with Full Width at Half Maximum (FWHM) of 10 nm.

4.3.4 Climatic data

Climatic data comprised 19 bioclimatic variables derived from WorldClim Version2 (Fick et al. 2017) at a resolution of 1 km with global coverage. We extracted the value of each environmental variable per site based on the geographic locations of sampled sites and applied min-max normalization of extracted values for each variable.

4.3.5 Statistical analysis

First, we used the filtered SNPs (see 4.3.2) to estimate genetic structure and relatedness of individuals. Genetic structure was calculated in *ADMIXTURE* v. 1.3 (Alexander et al., 2015) and the number of genetic clusters was set to 4 based on the smallest cross-validation error. To derive the relationships among trees, we calculated the SNPs-based

pairwise distance between individuals using *plink* v. 1.9, and applied the *cmdscale* function of *Matlab* (v. R2017b) to perform Principal coordinate analysis (PCoA) and to derive the genetic Principal Coordinate (PCO). Based on pairwise distances between individuals, we additionally calculated within-site standard deviations of genetic relatedness. Furthermore, we estimated independent and common contributions of climatic and spectral information to model genetic structure among sampled *F. sylvatica* sites. We used redundancy analysis (RDA) as a multivariate ordination method to simultaneously account for multiple response variables (Capblancq et al. 2021). We used Akaike Information Criterion (AIC) values for forward variable selection and a threshold of 2.5 for Variance Inflation Factor (VIF) scores to reduce the impact of collinearity of variables (Johnston et al., 2018). We limited the response variables to *F. sylvatica* distribution along the first and the second genetic PCO, which explained together 7.7% of total variation in derived genetic structure (Fig. S.4.3). Further genetic PCOs (3-6) did not correlate strongly with environmental and spectral variation (Fig. S.4.4). We constructed RDA models from environmental and spectral information separately, which resulted in the construction of environment-based (including 4 out of 19 explanatory environmental variables) and spectral-based (including 5 out of 216 explanatory spectral variables) models. For the full model, we considered all 9 variables selected for the environment- and spectral-based models to construct RDA models including both environmental and spectral information as explanatory variables. Following the variable selection procedures described above yielded a full model that include three environmental and five spectral explanatory variables. We derived the independent contributions of each environmental and spectral variable to the full model and we partitioned the genetic variance (as variance along genetic PCO1 and PCO2 to within- and between-site variation using analysis of variance (ANOVA) with explanatory terms for environment and spectral information and for site identity. We used the *rda* function of the *vegan* package (v. 2.6-4) in *R* (v. 4.1.3) to construct the models and to find the model fits.

To determine the contribution of spectral information to genetic structure predictions, we conducted leave-one-out validation. Accordingly, we created a model of 21 populations to predict the distribution of individuals of the 22nd population along genetic PCOs from (1) environmental information only, and (2) environmental information merged with spectral information. We present the contribution of spectral information to genetic structure prediction as the percentage change in predicted versus actual distribution along genetic PCO1 and PCO2 from the model based on (1) environmental and (2) environmental with

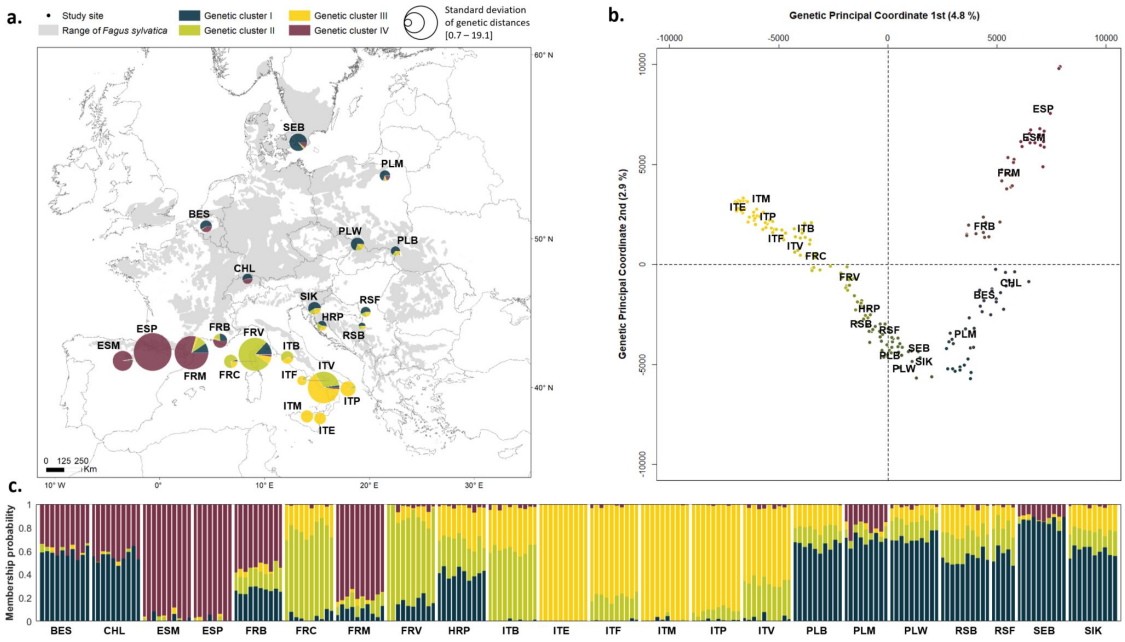

FIGURE 4.1: Genetic structure among 22 *F. sylvatica* sites (a) and 209 *F. sylvatica* individuals (c), showing the distribution of K=4 genetic clusters with the smallest cross-validation error in *ADMIXTURE*. The size of circles at each site in a) indicates the standard deviation of genetic variation among the individuals from that site (N= 10 except for RSB, N= 4, ESP, N= 8, PLM and ITP, N=9). In b) the relationship between individuals is plotted along the first and second genetic PCO. Color indicates kinship probability to belong to one of the genetic clusters. All results are inferred from 1,380,310 SNPs across the nuclear genome (see 4.3.2).

spectral information, respectively.

4.4 Results and discussion

Derived genetic structure and species phylogeographic history

We present the first assessment of genetic structure of a broad-leaf tree species, *Fagus sylvatica* L., across its natural distribution derived from whole-genome sequences of nuclear DNA.

Using *ADMIXTURE* cross-validation, we identified four genetic clusters representing the composition of current populations (Fig. 4.1a). The high abundance of one genetic cluster in southern Italy (ITE, ITM, ITP, ITV; cluster indicated in yellow) and the southern Pyrenees (ESP, ESM, FRM; cluster indicated in maroon) may correspond to previously reported glacial refugia (Magri et al., 2006; Magri, 2008; López-Merino et al., 2008). The analysis identified relatively high genetic dissimilarities between individuals at the Corsican site FRV (including a third cluster indicated in green), indicating that further genetic and phylogeographic studies would be valuable in Corsica to determine whether this may represent intermixing between refugia or an additional refugium (Petit et al.,

2003). The origin of the most widespread (blue) genetic cluster in central and northern Europe remains uncertain and might be assigned to the eastern range of the Alps (Gomory et al., 2003) or southern Moravia (Magri et al., 2006). Thus, these areas will be of interest for future genomic analyses, in addition to available microsatellite-based genetic information (Stefanini et al., 2023). It has to be noted that we sampled 10 individuals per site (sometimes fewer) from a ca. 50 x 200 m patch within a population, and thus only relative comparisons between sites genetic distances are possible (Fig. S.4.2), and the dominance of single clusters at a site may be due to sampling few individuals. However, similarities in cluster composition at neighboring sites (for example, ITE and ITM, ESM and ESP, PLB and PLW), and clusters transitions between the sites, may support representation of sites with derived genetic clusters.

Unraveled genetic composition across the species range could further help to understand the species phylogeographic history. Shared nDNA kinship as indicated by the cluster shared between Corsica and nearby sites at the French Mediterranean basin (FRB, FRM) and Appenninan Peninsula (ITB, ITV, ITF, ITP), and reaching into the Balkans (HRP, RSB, RSF, SIK), Carpathians (PLB, PLW), may be explained by terrestrial connections between the Adriatic and Mediterranean coasts during the last glaciation (Hazler et al., 1997) and gradual further expansion of *F. sylvatica* genetic pools across mountain ranges at its southern and north-eastern range (Magri et al., 2006). The mixture of clusters in the central and northern range of *F. sylvatica* is thought to result from post-glacial expansion from the south across mountain ranges (yellow cluster), and from the west across the central European plains (maroon cluster). The northernmost sampled site shows evidence of bottlenecks from recent expansion out of central European plains (Demesure et al., 1996). In contrast to lower mountain ranges, the Alps act as a barrier for gene flow between southern and northern populations (Paule, 1995).

Per-individual kinship probabilities to different genetic clusters (Fig. 4.1c), and pairwise genetic dissimilarities based on allele frequencies between individuals, demonstrate relatively higher between-population variation of southern than northern samples (Fig. 4.1a and b, Supplementary Fig. S.4.2). The continuous distribution of the species in the northern part of its range likely allows greater gene flow across the landscape, while the patchier distribution in the south limits gene flow and maintains genetic distance between populations. This may be due to post-glacial climatic changes causing migration to higher altitudes in the southern part of the species range. Moreover, environmental pressure may drive the diversification of *F. sylvatica* populations at their southern limit (Teau et al.,

1982, Capblancq et al., 2020), which could promote diversification among sites.

Observed genetic structure explained by environmental and spectral information

Together, environmental and spectral information explained 77.7% of the variance in the distribution of sampled *F. sylvatica* along the first and second genetic PCOs (Fig. 4.1). In contrast, 39.8% of the variation was explained by environmental variables independently of spectral information. The environmental variables contributing most were Mean Temperature of Driest Quarter (BIO 09), Isothermality (BIO 03), Precipitation Seasonality (BIO 15) and Precipitation of Driest Quarter (BIO 17). These variables define environmental niches which may determine adaptive patterns in European beech (Müller-Starck, 1985; Fang and Lechowicz, 2006; Bolte et al., 2007; Meger et al., 2021; Csilléry et al., 2014; Jump et al., 2006; Pluess and Weber, 2012; Cuervo-Alarcon et al., 2021; Pfenninger et al., 2020). In contrast, variables from leaf spectra explained 12.2% variation in genetic structure among trees (genetic PCO1 and PCO2) independently of environmental information. The spectral variation explained 12.1% in genetic variation between sites and only 0.1% in genetic variation within sites (Fig. S.4.6, see also S.4.1). This suggests that spectral information mostly explains genetic variation between sites, including an additional proportion of variation that could not be explained by available environmental information on a coarser spatial resolution capturing only the between-site variation. Portions of spectra which were most informative about between-site genetic structure correspond to carotenoid absorption, related to electronic transitions induced by light at 410 nm and 530 nm (Gitelson et al., 2006), leaf surface morphological properties affecting the reflection of UV light at 350 nm, leaf water content related to vibrational absorption of photons at 1500 nm and 790 nm (Kaye, 1954), and leaf inner structure related to scattering of light at 790 nm (Gausman et al., 1971). Calculated reflectance centered at 350 nm was positively correlated with Precipitation of Driest Quarter at the sites, thus based on VIF scores this environmental variable was excluded from the full model built from environmental and spectral information together. Environmental and spectral variation combined contributed an additional 25.6% to predicting genetic variation along genetic PCO1 and PCO2.

Spectral information contributed most to prediction of the first genetic PCO, especially for populations from the southernmost and northernmost parts of the sampling area (Fig. 4.3a). The introduction of predictors derived from leaf reflectance improved

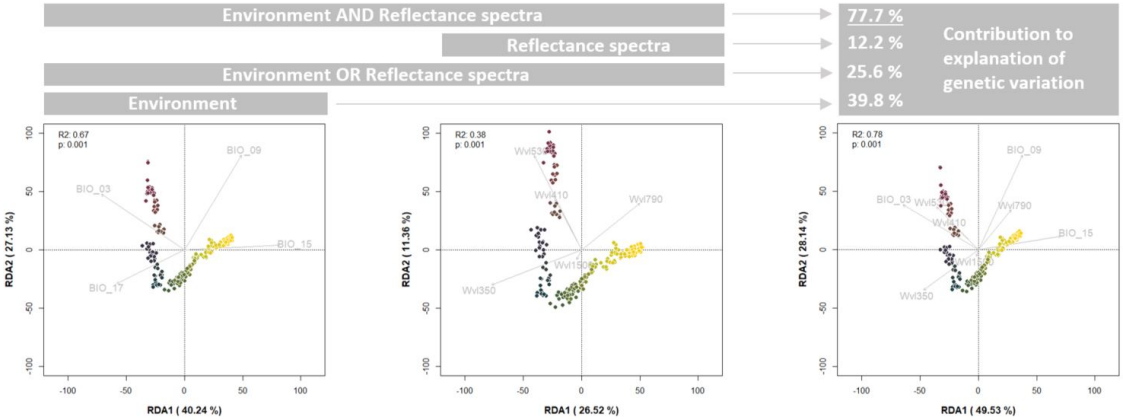

FIGURE 4.2: RDA models for *F. sylvatica* genetic variation as captured by the first (PCO1) and second (PCO2) genetic principal coordinate across the species range derived from environmental information, spectral information, and environmental with spectral information combined. The variables were chosen based on AIC and VIF <2.5. Out of 19 environmental variables, four were identified to be indicative of the genetic structure between sites, and out of 216 spectral bands, five were indicative of the genetic structure between and within sites. The full model combined three out of four selected environmental variables and five out of the five spectral variables selected for the spectral-based model. The incorporation of spectral information added 12.2% explanatory power to the model created based on environmental (with contribution of 39.8%) and spectral information, which together explained 77.7% of the variation in genetic structure among sampled *F. sylvatica* individuals (mean of R^2 for PCO1 and PCO2). The *F. sylvatica* individuals in RDA space are colored by combined kindship probabilities to four genetic clusters.

the genetic structure prediction by 59.7% along genetic PCO1 in comparison to the prediction based on environmental variables only (R^2 of 0.79 versus 0.63) (Fig. S.4.5). The spectral indicators contributing most to prediction of the distribution of individuals along this genetic PCO1 relate to leaf interaction with spectral bands centered at 530 nm and 410 nm in the visible region, influenced by plant pigments. The contribution of spectral information to prediction of the second genetic PCO also varied among sites, but produced a mean 7.9% improvement in comparison to the prediction based on environmental variables only (R^2 of 0.76 versus 0.74) (Fig. S.4.5). The smaller contribution of spectral information to prediction of the second genetic PCO might correspond to the shared explanatory power of wavelengths related to leaf structural properties (i.e spectral bands centered at 350 nm and 790 nm), which contribute most to the second genetic PCO axis of the full RDA model.

Perspective for imaging spectroscopy supporting the monitoring of genetic variation

We present contributions of spectral phenotype data in combination with environmental data to predict intraspecific genetic variation derived from whole-genome sequences of a

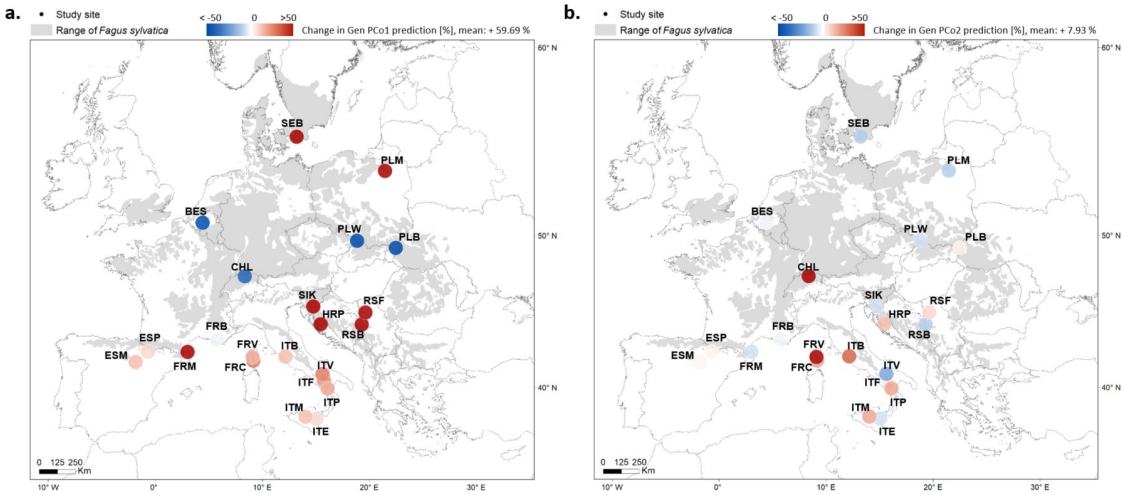

FIGURE 4.3: Change in *F. sylvatica* genetic identity prediction [%] for the first (a) and second (b) genetic principal coordinates after incorporation of spectral information to the prediction model based purely on environmental information. The spectral information contributes to improvement of prediction by 59.7% along the first Principal Coordinate (genetic PCO1) and 7.9% along the second Principal Coordinate (genetic PCO2). The within-site variation in prediction accuracy is presented in Fig. S.4.7. Red indicates improvement and blue indicates deterioration of predictions.

forest tree species across its natural range. While spectra carry information about current genetic composition, further incorporation of information about the phylogeographic history of the species would constrain model predictions to the available gene pool of particular areas. Merging species phylogeographic history, current genetic structure and past and current environmental conditions with expressed phenotypic variation will allow more accurate prediction of genetic variation within tree species across their range. Moreover, here we have shown the contribution of spectral phenotypes to predict genetic variation without discriminating the adaptive component. Currently ongoing work on establishing a common garden experiment from maternal seed families collected at the sites would support further associations of genome-wide sequences with spectral and environmental variations and may facilitate gene associations by observed phenotypes and spectra associations by genes of known function. Such associations would advance the understanding of selective processes that underlay genetic adaptations across wide spatial and temporal scales.

Our measurements were spatially and temporally limited to single *in situ* sampling campaigns per site, whereas spectral phenotypes at canopy level acquired from remote observations of airborne and spaceborne imaging spectrometers will allow for repeated acquisitions that capture dynamic phenotypic responses on the landscape scale. Remote observations of solar radiation reflected from tree canopies come with greater uncertainty than contact measurements (Czyż et al., 2023, Petibon et al., 2021), but can integrate leaf

absorption signals across singular crowns – given the spatial resolution of the instruments allow for an unambiguous identification and delineation of individual crown shapes – and reveal additional canopy-level traits related, for example, to canopy architecture (Zheng et al., 2021) or function (Gamon et al., 2019), that may also be informative about genetic identity (Ollinger, 2011). Current data of spaceborne imaging spectrometers (e.g., DESIS, HISUI, EMIT, EnMAP, PRISMA, etc.), and the wealth of planned and upcoming imaging spectrometers missions (SBG, Cawse-Nicholson et al., 2021; CHIME, Celesti et al., 2022, etc.), will bring spatially contiguous and temporally dense information about spectral properties of the Earth surface resolved at spectral intervals which are sufficiently narrow to be interpreted in terms of molecular absorption features. Spectrally resolved information acquired repeatedly over landscapes may further help to calibrate predictive models of intraspecific genetic variation across species ranges and might facilitate spatial mapping of gene frequencies. Such spectral-based modelling might be further applied across different phylogenetic groups (Meireles et al., 2020) and might be used to calibrate models predicting genetic distribution under global change.

Predictions of genetic diversity and gene frequency distributions across space and time will help to target environmentally, phenotypically and genetically underrepresented sites or hotspots for sampling of genetic resources, whose sampling will further contribute to calibration of predictive models. The improved predictions of genetic variation with repeat spectral monitoring would indicate changes in genetic diversity and adaptation over time, and could deliver information for the adjustment of current efforts to safeguard biodiversity at the global scale, at unprecedented speed and spatial completeness.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was conceptualized by EA Czyż and ME Schaepman with further contributions of B Schmid, MB Eppinga, MC Schuman and M de La Harpe. The project was administrated by ME Schaepman and data collection was administrated by MC Schuman. MC Schuman lead the methodology on storage, extraction, sequencing and analysis of genetic data. This study is supported by the University of Zurich Research Priority

Program on Global Change and Biodiversity (URPP GCB) and found by NOMIS foundation, grant to ME Schaepman. We express our gratitude to Emmanuelle Argenti, Ventura Gennaro, Hugot Laetitia, Elodie Magnanou, Joseph Garrigue, Kris Vandekerkhove, Arne Verstraeten, Nilsson Jorgen, Piotr Małz, Katarzyna Gierad, Dejan Stojanovic, Gordan Lukac, Katja Konecnik, Janez Subic, Eustaquio Gil Pelegrın, Juan Pedro Ferrio, Stanislaw Kucharzyk, Piotr Poloczek, Zbigniew Stebel, Nicu Ghircoias and Roman George for allowing the access to the sampling sites and support on investigation and to Romeo Galiano, Joseph Garrigue, Krzystof Poloczek, Helgessons TrädTjänst AB company, Stefan Nica, Michal Szydlowski for assistance in sample collection. We express our gratitude to Reinhard Furrer, Anna K. Schweiger, Benjamin Dauphin and Ingmar Staude for helpful discussions.

Supplementary Material

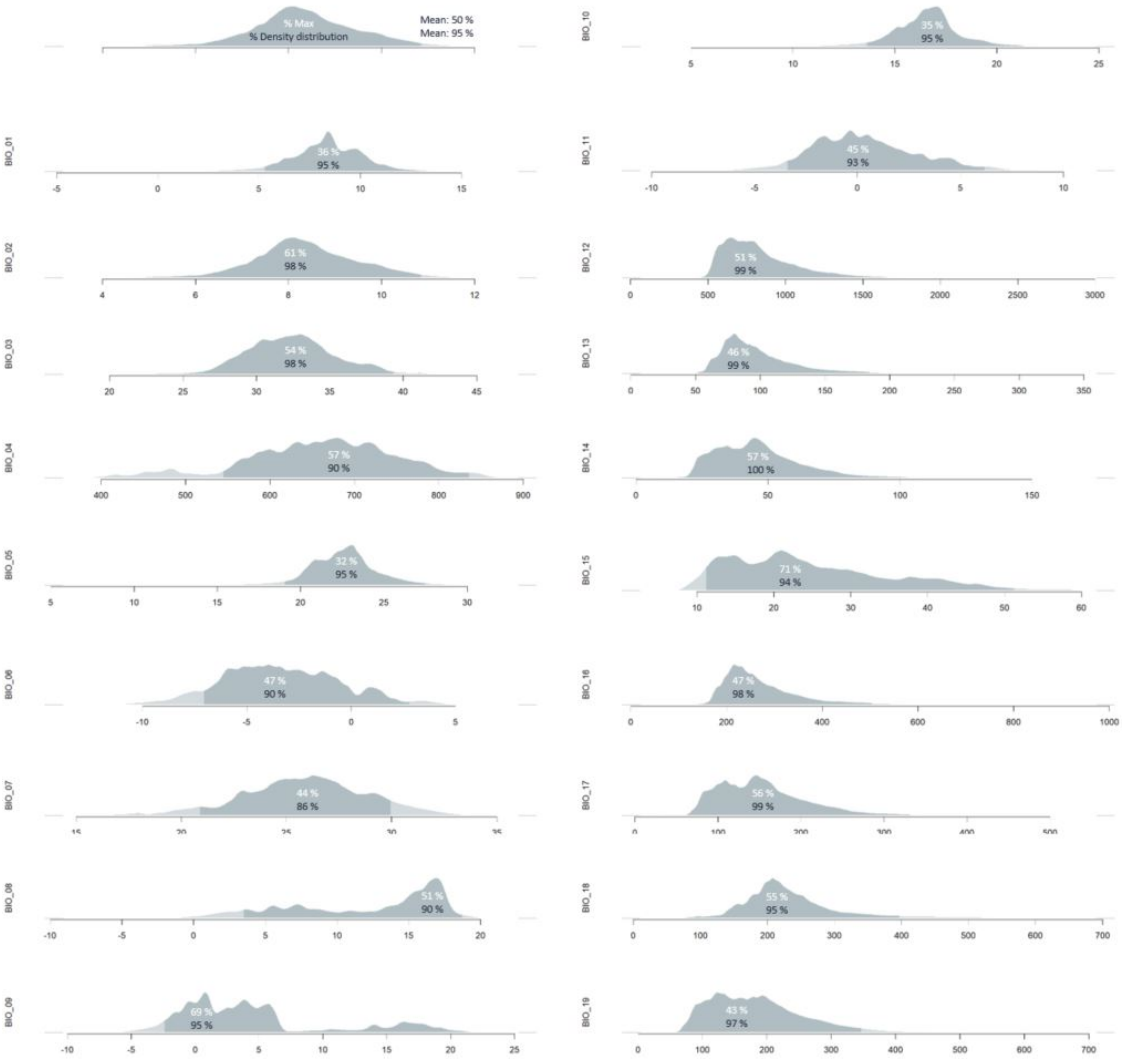

FIGURE S.4.1: Representation of the *F. sylvatica* environmental envelope for bioclimatic variables across the species range (represented with brighter blue color). The darker color indicates the envelopes limited by the maximum and minimum value for bioclimatic variables of the sampled sites. The white overlaid numbers indicate the percentage covered by the sampling variable range in relation to the total variable range (mean of 50%). The blue overlaid numbers indicate the percentage covered by the sampling variable distribution in relation to the total variable distribution (mean of 95%).

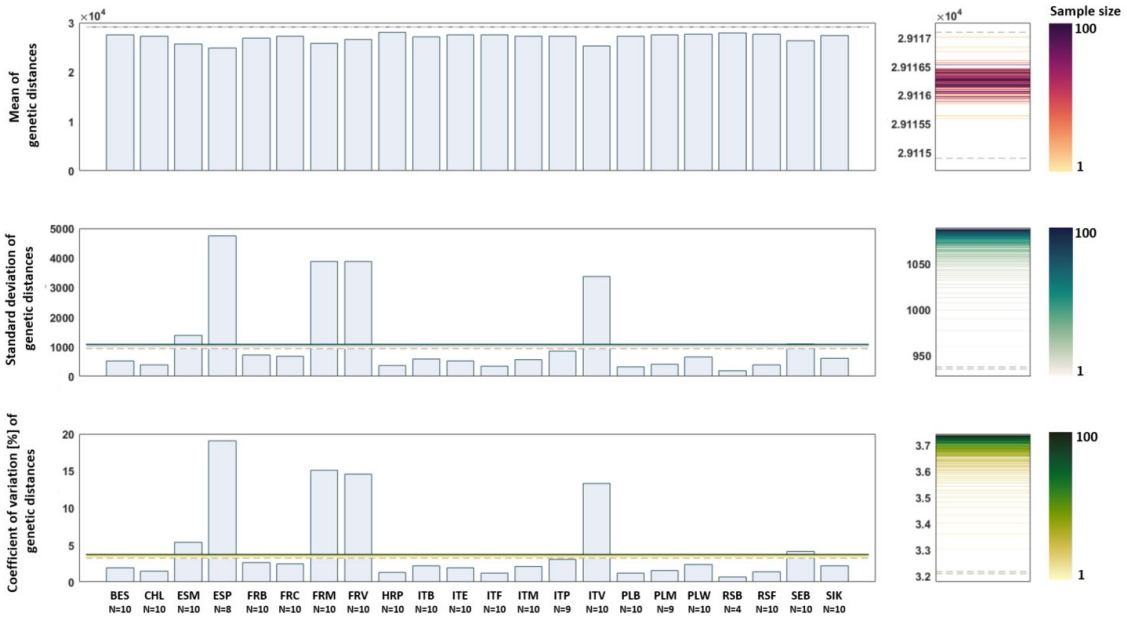

FIGURE S.4.2: Left, Mean, standard deviation and coefficient of variation [%] of within-site genetic distances between sampled *F. sylvatica* individuals derived from pairwise allele counts. Right, the impact of sampling size as indicated by mean of 10 000 bootstraps for 4 to 100 samples selected from all the pairwise distances between *F. sylvatica* individuals is indicated with solid lines. The 95% confidence interval in bootstrapping with a sampling of size 4 (the lowest sampled size, for the RSB site) is indicated with dashed lines.

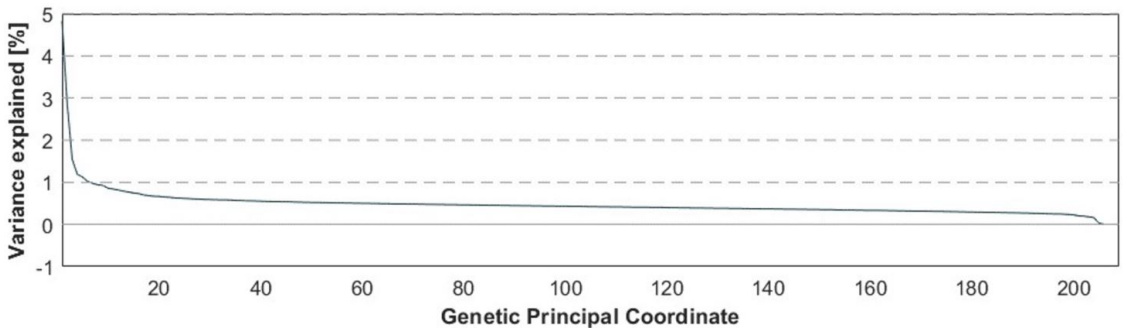

FIGURE S.4.3: Contribution of genetic PCOs derived from PCoA performed on genetic distances between sampled *F. sylvatica* individuals derived from pairwise allele counts. The first six PCOs explain more than 1% of variation as genetic distances between individuals.

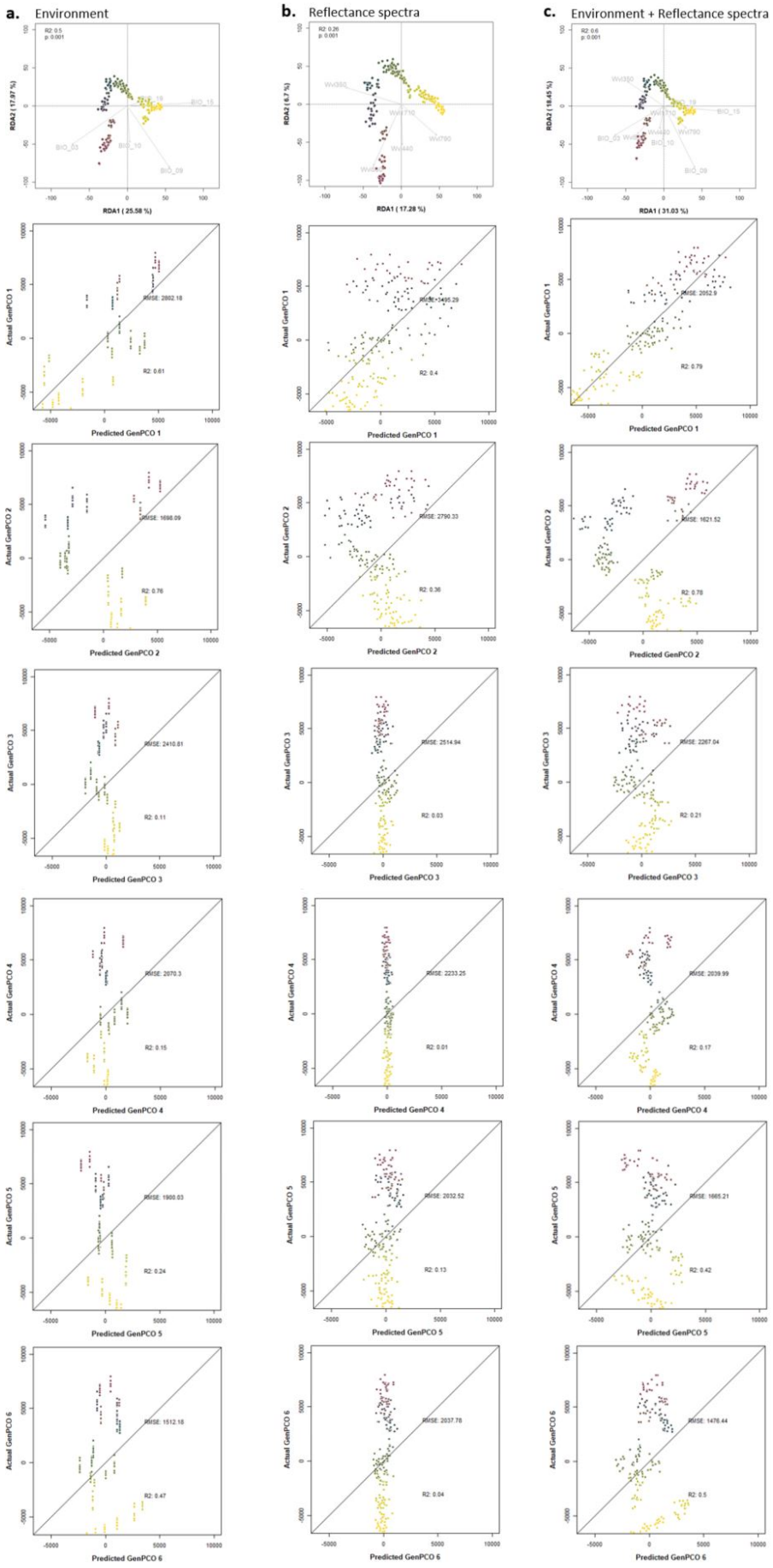

FIGURE S.4.4: Prediction of genetic structure along the first six genetic PCOs derived from PCoA performed on genetic distances between sampled *F. sylvatica* individuals derived from pairwise allele counts using environmental information (a), spectral information (b) and combined environmental with spectral information (c).

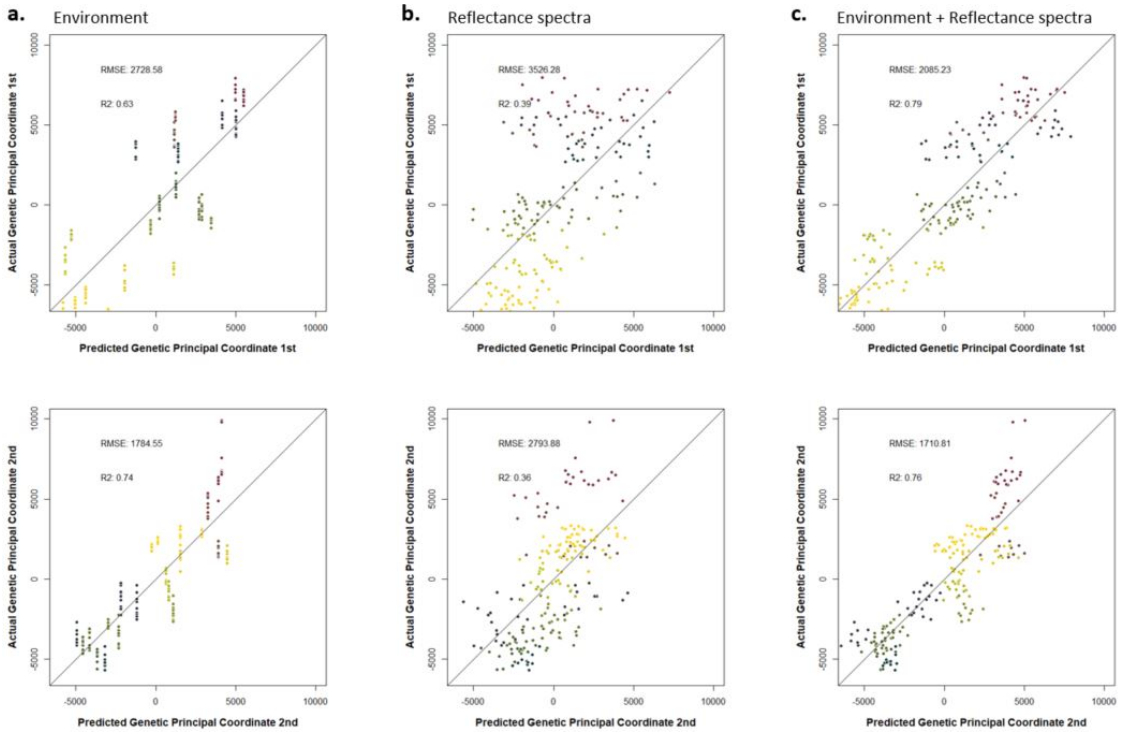

FIGURE S.4.5: Prediction of genetic structure along the first and second genetic PCOs derived from PCoA using environmental information (a), spectral information (b) and combined environmental with spectral information (c) to predict genetic distances between sampled *F. sylvatica* individuals derived from pairwise allele counts.

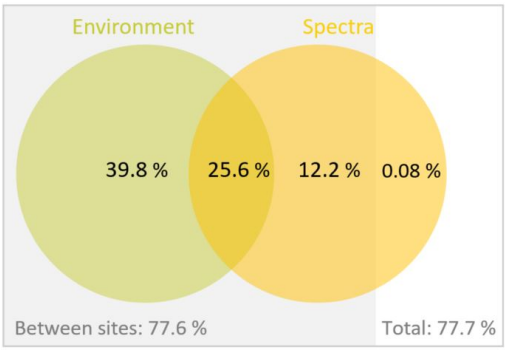

FIGURE S.4.6: Contribution to explanation of variation in derived genetic structure by independent variation in environmental and spectral variables and their interaction.

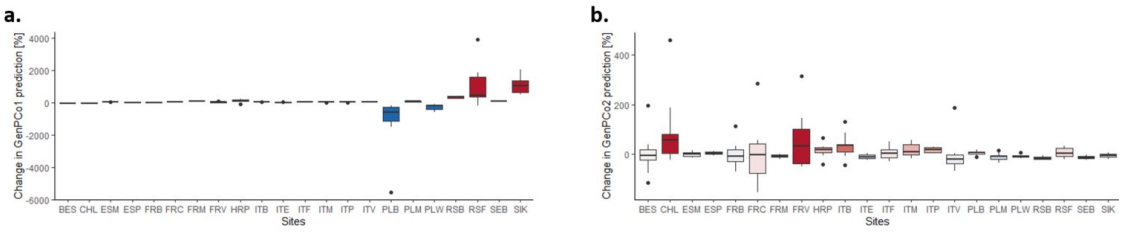

FIGURE S.4.7: Per-site standard deviation in changes in *F. sylvatica* genetic identity prediction[%] for the first and second genetic PCO after incorporation of spectral information to the prediction model based purely on environmental information.

TABLE S.4.1: Percentage of generic variation along genetic PCO1 and genetic PCO2 explained by spectral variables corresponding to the five first spectral principal components (PCs). The principal components were derived from variation across the whole analyzed spectral range and all the sampled F. sylvatica trees. The most prominent, second principal component (PC2) explain 5.21 % of the genetic variation along the 2nd genetic principal coordinate (PCO2), which is half of the variation explained by 5 spectral predictors selected from the whole analyzed spectral range (12.2 %) as inferred from the RDA models.

Genetic	**Spectral**				
	PC1	**PC2**	**PC3**	**PC4**	**PC5**
PCO1	0.10	0.33	1.64'	0.15	0.00
PCO2	2.07*	5.21***	0.46	0.94	0.42

References

Alexander, D. H., Shringarpure, S. S., Novembre, J., & Lange, K. (2015). Admixture 1.3 software manual. *Los Angeles: UCLA Human Genetics Software Distribution.*

Allen, W. A., Gausman, H., Richardson, A., & Wiegand, C. (1970). Mean effective optical constants of thirteen kinds of plant leaves. *Applied Optics*, *9*(11), 2573–2577.

Bolte, A., Czajkowski, T., & Kompa, T. (2007). The north-eastern distribution range of european beech—a review. *Forestry*, *80*(4), 413–429.

Capblancq, T., Morin, X., Gueguen, M., Renaud, J., Lobreaux, S., & Bazin, E. (2020). Climate-associated genetic variation in fagus sylvatica and potential responses to climate change in the french alps. *Journal of Evolutionary Biology*, *33*(6), 783–796.

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., Narwani, A., Mace, G. M., Tilman, D., Wardle, D. A., et al. (2012). Biodiversity loss and its impact on humanity. *Nature*, *486*(7401), 59–67.

Caudullo, G., Welk, E., & San-Miguel-Ayanz, J. (2017). Chorological maps for the main european woody species. *Data in brief*, *12*, 662–666.

Cavender-Bares, J., Meireles, J. E., Couture, J. J., Kaproth, M. A., Kingdon, C. C., Singh, A., Serbin, S. P., Center, A., Zuniga, E., Pilz, G., et al. (2016). Associations of leaf spectra with genetic and phylogenetic variation in oaks: Prospects for remote detection of biodiversity. *Remote Sensing*, *8*(3), 221.

Cawse-Nicholson, K., Townsend, P. A., Schimel, D., Assiri, A. M., Blake, P. L., Buongiorno, M. F., Campbell, P., Carmon, N., Casey, K. A., Correa-Pabón, R. E., et al. (2021). Nasa's surface biology and geology designated observable: A perspective on surface imaging algorithms. *Remote Sensing of Environment*, *257*, 112349.

Celesti, M., Rast, M., Adams, J., Boccia, V., Gascon, F., Isola, C., & Nieke, J. (2022). The copernicus hyperspectral imaging mission for the environment (chime): Status and planning. *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium*, 5011–5014.

Csilléry, K., Lalagüe, H., Vendramin, G. G., González-Martınez, S. C., Fady, B., & Oddou-Muratorio, S. (2014). Detecting short spatial scale local adaptation and epistatic selection in climate-related candidate genes in e uropean beech (f agus sylvatica) populations. *Molecular Ecology*, *23*(19), 4696–4708.

Cuervo-Alarcon, L., Arend, M., Müller, M., Sperisen, C., Finkeldey, R., & Krutovsky, K. V. (2021). A candidate gene association analysis identifies snps potentially involved in drought tolerance in european beech (fagus sylvatica l.) *Scientific Reports*, *11*(1), 1–15.

Czyż, E. A., Guillén Escribà, C., Wulf, H., Tedder, A., Schuman, M. C., Schneider, F. D., & Schaepman, M. 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.

Czyż, E. A., Schmid, B., Hueni, A., Eppinga, M. B., Schuman, M. C., Schneider, F. D., Guillén-Escribà, C., & Schaepman, M. E. (2023). Genetic constraints on temporal variation of airborne reflectance spectra and their uncertainties over a temperate forest. *Remote Sensing of Environment*, *284*, 113338.

Demesure, B., Comps, B., & Petit, R. J. (1996). Chloroplast dna phylogeography of the common beech (fagus sylvatica l.) in europe. *Evolution*, 2515–2520.

Des Roches, S., Pendleton, L. H., Shapiro, B., & Palkovacs, E. P. (2021). Conserving intraspecific variation for nature's contributions to people. *Nature Ecology & Evolution*, *5*(5), 574–582.

Des Roches, S., Post, D. M., Turley, N. E., Bailey, J. K., Hendry, A. P., Kinnison, M. T., Schweitzer, J. A., & Palkovacs, E. P. (2018). The ecological importance of intraspecific variation. *Nature ecology & evolution*, *2*(1), 57–64.

Dıaz, S., Zafra-Calvo, N., Purvis, A., Verburg, P. H., Obura, D., Leadley, P., Chaplin-Kramer, R., De Meester, L., Dulloo, E., Martın-López, B., et al. (2020). Set ambitious goals for biodiversity and sustainability. *Science*, *370*(6515), 411–413.

Exposito-Alonso, M., Booker, T. R., Czech, L., Gillespie, L., Hateley, S., Kyriazis, C. C., Lang, P. L., Leventhal, L., Nogues-Bravo, D., Pagowski, V., et al. (2022). Genetic diversity loss in the anthropocene. *Science*, *377*(6613), 1431–1435.

Fang, J., & Lechowicz, M. J. (2006). Climatic limits for the present distribution of beech (fagus l.) species in the world. *Journal of Biogeography*, *33*(10), 1804–1819.

Franklin, J. F. (1981). *Ecological characteristics of old-growth douglas-fir forests* (Vol. 118). US Department of Agriculture, Forest Service, Pacific Northwest Forest and ...

Fridley, J. D., & Grime, J. P. (2010). Community and ecosystem effects of intraspecific genetic diversity in grassland microcosms of varying species diversity. *Ecology*, *91*(8), 2272–2283.

Gallois, A., Audran, J., & Burrus, M. (1998). Assessment of genetic relationships and population discrimination among fagus sylvatica l. by rapd. *Theoretical and applied genetics*, *97*, 211–219.

Gamon, J., Somers, B., Malenovskỳ, Z., Middleton, E., Rascher, U., & Schaepman, M. E. (2019). Assessing vegetation function with imaging spectroscopy. *Surveys in Geophysics*, *40*, 489–513.

Gausman, H., Allen, W., Escobar, D., Rodriguez, R., & Cardenas, R. (1971). Age effects of cotton leaves on light reflectance, transmittance, and absorptance and on water content and thickness 1. *Agronomy Journal*, *63*(3), 465–469.

Gitelson, A. A., Keydan, G. P., & Merzlyak, M. N. (2006). Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves. *Geophysical research letters*, *33*(11).

Gomory, D., Paule, L., Shvadchak, I., Popescu, F., Sulkowska, M., Hynek, V., & Longauer, R. (2003). Spatial patterns of the genetic differentiation in european beech (fagus sylvatica l.) at allozyme loci in the carpathians and the adjacent regions. *Silvae Genetica*, *52*(2), 78–83.

Guevara, M. Á., Sánchez-Gómez, D., Vélez, M. D., de Marıa, N., Dıaz, L. M., Ramırez-Valiente, J. A., Mancha, J. A., Aranda, I., & Cervera, M. T. (2022). Epigenetic and genetic variability in contrasting latitudinal fagus sylvatica l. provenances. *Forests*, *13*(12), 1971.

Guyot, G., Baret, F., & Jacquemoud, S. (1992). *Imaging spectroscopy for vegetation studies* (Vol. 2). Kluwer Academic Publishers: Norwell, MA, USA.

Hazler, K., Comps, B., Sugar, I., Melovski, L., Tashev, A., & Gracan, J. (1997). Genetic structure of fagus sylvatica l. populations in southeastern europe. *Silvae Genetica*, *46*(4), 229–235.

Hector, A., Hautier, Y., Saner, P., Wacker, L., Bagchi, R., Joshi, J., Scherer-Lorenzen, M., Spehn, E. M., Bazeley-White, E., Weilenmann, M., et al. (2010). General stabilizing effects of plant diversity on grassland productivity through population asynchrony and overyielding. *Ecology*, *91*(8), 2213–2220.

Hector, A., & Bagchi, R. (2007). Biodiversity and ecosystem multifunctionality. *Nature*, *448*(7150), 188–190.

Hughes, A. R., Inouye, B. D., Johnson, M. T., Underwood, N., & Vellend, M. (2008). Ecological consequences of genetic diversity. *Ecology letters*, *11*(6), 609–623.

IPBES. (2019). *Global assessment report of the intergovernmental science-policy platform on biodiversity and ecosystem services* (E. Brondizio, J. Settele, S. Díaz, & H. T. Ngo, Eds.). https://doi.org/10.5281/zenodo.3831673

Isbell, F., Calcagno, V., Hector, A., Connolly, J., Harpole, W. S., Reich, P. B., Scherer-Lorenzen, M., Schmid, B., Tilman, D., Van Ruijven, J., et al. (2011). High plant diversity is needed to maintain ecosystem services. *Nature*, *477*(7363), 199–202.

Isbell, F., Craven, D., Connolly, J., Loreau, M., Schmid, B., Beierkuhnlein, C., Bezemer, T. M., Bonin, C., Bruelheide, H., De Luca, E., et al. (2015). Biodiversity increases the resistance of ecosystem productivity to climate extremes. *Nature*, *526*(7574), 574–577.

Jacquemoud, S., & Ustin, S. (2019). *Leaf optical properties*. Cambridge University Press.

Jenkins, C. N., & Joppa, L. (2009). Expansion of the global terrestrial protected area system. *Biological conservation*, *142*(10), 2166–2174.

Johnson, C. N., Balmford, A., Brook, B. W., Buettel, J. C., Galetti, M., Guangchun, L., & Wilmshurst, J. M. (2017). Biodiversity losses and conservation responses in the anthropocene. *Science*, *356*(6335), 270–275.

Johnston, R., Jones, K., & Manley, D. (2018). Confounding and collinearity in regression analysis: A cautionary tale and an alternative procedure, illustrated by studies of british voting behaviour. *Quality & quantity*, *52*, 1957–1976.

Jump, A. S., Hunt, J. M., MARTÍNEZ-IZQUIERDO, J. A., & Peñuelas, J. (2006). Natural selection and climate change: Temperature-linked spatial and temporal trends in gene frequency in fagus sylvatica. *Molecular Ecology*, *15*(11), 3469–3480.

Kaye, W. (1954). Near-infrared spectroscopy: I. spectral identification and analytical applications. *Spectrochimica acta*, *6*(4), 257–E2.

Kupiec, J., & Curran, P. (1995). Decoupling effects of the canopy and foliar biochemicals in aviris spectra. *International Journal of Remote Sensing*, *16*(9), 1731–1739.

López-Merino, L., López-Sáez, J. A., Zapata, M. R., & Garcıa, M. G. (2008). Reconstructing the history of beech (fagus sylvatica l.) in the north-western iberian range (spain): From late-glacial refugia to the holocene anthropic-induced forests. *Review of Palaeobotany and Palynology*, *152*(1-2), 58–65.

Lou, R. N., Jacobs, A., Wilder, A. P., & Therkildsen, N. O. (2021). A beginner's guide to low-coverage whole genome sequencing for population genomics. *Molecular Ecology*, *30*(23), 5966–5993.

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.

Magri, D. (2008). Patterns of post-glacial spread and the extent of glacial refugia of european beech (fagus sylvatica). *Journal of Biogeography*, *35*(3), 450–463.

Magri, D., Vendramin, G. G., Comps, B., Dupanloup, I., Geburek, T., Gömöry, D., Latałowa, M., Litt, T., Paule, L., Roure, J. M., et al. (2006). A new scenario for the quaternary history of european beech populations: Palaeobotanical evidence and genetic consequences. *New phytologist*, *171*(1), 199–221.

McNeely, J. A., Miller, K. R., Reid, W. V., Mittermeier, R. A., & Werner, T. B. (1990). Strategies for conserving biodiversity. *Environment: Science and Policy for Sustainable Development*, *32*(3), 16–40.

Meger, J., Ulaszewski, B., & Burczyk, J. (2021). Genomic signatures of natural selection at phenology-related genes in a widely distributed tree species fagus sylvatica l. *BMC genomics*, *22*(1), 1–20.

Meireles, J. E., Cavender-Bares, J., Townsend, P. A., Ustin, S., Gamon, J. A., Schweiger, A. K., Schaepman, M. E., Asner, G. P., Martin, R. E., Singh, A., et al. (2020). Leaf reflectance spectra capture the evolutionary history of seed plants. *New Phytologist*, *228*(2), 485–493.

Miller, J., Steven, M., & Demetriades-Shah, T. (1992). Reflection of layered bean leaves over different soil backgrounds: Measured and simulated spectra. *International Journal of Remote Sensing*, *13*(17), 3273–3286.

Mishra, B., Ulaszewski, B., Meger, J., Aury, J.-M., Bodénès, C., Lesur-Kupin, I., Pfenninger, M., Da Silva, C., Gupta, D. K., Guichoux, E., et al. (2022). A chromosome-level genome assembly of the european beech (fagus sylvatica) reveals anomalies for organelle dna integration, repeat content and distribution of snps. *Frontiers in Genetics*, *12*, 2748.

Mora, C., Andréfouët, S., Costello, M. J., Kranenburg, C., Rollo, A., Veron, J., Gaston, K. J., & Myers, R. A. (2006). Coral reefs and the global network of marine protected areas.

Müller-Starck, G. (1985). Genetic differences between" tolerant" and" sensitive" beeches(fagus sylvatica l.) in an environmentally stressed adult forest stand. *Silvae Genetica*, *34*(6), 241–246.

Norse, E. A., Rosenbaum, K. L., Wilcove, D. S., & Wilcox, B. A. (1986). *Conserving biological diversity in our national forests* (tech. rep.).

Ollinger, S. V. (2011). Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytologist*, *189*(2), 375–394.

Paule, L. (1995). Gene conservation in european beech (fagus sylvatica l.) *Forest Genetics*, *2*(3), 161–170.

Perrings, C., Duraiappah, A., Larigauderie, A., & Mooney, H. (2011). The biodiversity and ecosystem services science-policy interface. *Science*, *331*(6021), 1139–1140.

Peterson, D. L., Aber, J. D., Matson, P. A., Card, D. H., Swanberg, N., Wessman, C., & Spanner, M. (1988). Remote sensing of forest canopy and leaf biochemical contents. *Remote Sensing of Environment*, *24*(1), 85–108.

Petibon, F., Czyż, E. A., Ghielmetti, G., Hueni, A., Kneubühler, M., Schaepman, M. E., & Schuman, M. C. (2021). Uncertainties in measurements of leaf optical properties are small compared to the biological variation within and between individuals of european beech. *Remote Sensing of Environment*, *264*, 112601.

Petit, R. J., Aguinagalde, I., de Beaulieu, J.-L., Bittkau, C., Brewer, S., Cheddadi, R., Ennos, R., Fineschi, S., Grivet, D., Lascoux, M., et al. (2003). Glacial refugia: Hotspots but not melting pots of genetic diversity. *science*, *300*(5625), 1563–1565.

Pfenninger, M., Reuss, F., Kiebler, A., Schönnenbeck, P., Caliendo, C., Gerber, S., Cocchiararo, B., Reuter, S., Blüthgen, N., Mody, K., et al. (2020). Genomic basis of drought resistance in fagus sylvatica. *bioRxiv*, 2020–12.

Pluess, A. R., & Weber, P. (2012). Drought-adaptation potential in fagus sylvatica: Linking moisture availability with genetic diversity and dendrochronology. *PloS one*, *7*(3), e33636.

Pomerantz, A., Peñafiel, N., Arteaga, A., Bustamante, L., Pichardo, F., Coloma, L. A., Barrio-Amorós, C. L., Salazar-Valenzuela, D., & Prost, S. (2018). Real-time dna barcoding in a rainforest using nanopore sequencing: Opportunities for rapid biodiversity assessments and local capacity building. *GigaScience*, *7*(4), giy033.

Purvis, A., & Hector, A. (2000). Getting the measure of biodiversity. *Nature*, *405*(6783), 212–219.

Reich, P. B., Tilman, D., Isbell, F., Mueller, K., Hobbie, S. E., Flynn, D. F., & Eisenhauer, N. (2012). Impacts of biodiversity loss escalate through time as redundancy fades. *Science*, *336*(6081), 589–592.

Reid, W. V. (1998). Biodiversity hotspots. *Trends in Ecology & Evolution*, *13*(7), 275–280.

Sarkar, S. (2002). Defining "biodiversity"; assessing biodiversity. *The Monist*, *85*(1), 131–155.

Schaepman, M. E., Jehle, M., Hueni, A., D'Odorico, P., Damm, A., Weyermann, J., Schneider, F. D., Laurent, V., Popp, C., Seidel, F. C., et al. (2015). Advanced radiometry measurements and earth science applications with the airborne prism experiment (apex). *Remote Sensing of Environment*, *158*, 207–219.

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.

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.

Scholes, R. J., Mace, G. M., Turner, W., Geller, G. N., Jürgens, N., Larigauderie, A., Muchoney, D., Walther, B. A., & Mooney, H. (2008). Toward a global biodiversity observing system.

Scott, J. M., Csuti, B., Estes, J., & Anderson, H. (1989). Status assessment of biodiversity protection. *Conservation Biology*, *3*(1), 85–87.

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.

Stefanini, C., Csilléry, K., Ulaszewski, B., Burczyk, J., Schaepman, M. E., & Schuman, M. C. (2023). A novel synthesis of two decades of microsatellite studies on european beech reveals decreasing genetic diversity from glacial refugia. *Tree Genetics & Genomes*, *19*(1), 3.

Sutherland, W. J., Pullin, A. S., Dolman, P. M., & Knight, T. M. (2004). The need for evidence-based conservation. *Trends in ecology & evolution*, *19*(6), 305–308.

Tang, T., Zhang, N., Bongers, F. J., Staab, M., Schuldt, A., Fornoff, F., Lin, H., Cavender-Bares, J., Hipp, A. L., Li, S., et al. (2022). Tree species and genetic diversity increase productivity via functional diversity and trophic feedbacks. *Elife*, *11*, e78703.

Teau, B., LUARE, R., & Venstr, P. (1982). The bud enzymes of beech (fagus sylvatica l.) genetic distinction and analysis of polymorphism in several french populations. *Silvae genetica*, *31*, 2–3.

Tilman, D., & Downing, J. A. (1994). Biodiversity and stability in grasslands. *Nature*, *367*(6461), 363–365.

Turner, W. (2014). Sensing biodiversity. *Science*, *346*(6207), 301–302.

van Moorsel, S. J., Hahl, T., Petchey, O. L., Ebeling, A., Eisenhauer, N., Schmid, B., & Wagg, C. (2021). Co-occurrence history increases ecosystem stability and resilience in experimental plant communities. *Ecology*, *102*(1), e03205.

Walpole, M., Almond, R. E., Besançon, C., Butchart, S. H., Campbell-Lendrum, D., Carr, G. M., Collen, B., Collette, L., Davidson, N. C., Dulloo, E., et al. (2009). Tracking progress toward the 2010 biodiversity target and beyond. *Science*, *325*(5947), 1503–1504.

Wessman, C. A. (1987). *Estimating key forest ecosystem parameters through remote sensing (wisconsin)*. The University of Wisconsin-Madison.

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

Yamasaki, E., Altermatt, F., Cavender-Bares, J., Schuman, M. C., Zuppinger-Dingley, D., Garonna, I., Schneider, F. D., Guillén-Escribà, C., van Moorsel, S. J., Hahl, T., et al. (2017). Genomics meets remote sensing in global change studies: Monitoring and predicting phenology, evolution and biodiversity. *Current opinion in environmental sustainability*, *29*, 177–186.

Zheng, Z., Zeng, Y., Schneider, F. D., Zhao, Y., Zhao, D., Schmid, B., Schaepman, M. E., & Morsdorf, F. (2021). Mapping functional diversity using individual tree-based morphological and physiological traits in a subtropical forest. *Remote Sensing of Environment*, *252*, 112170.

Chapter 5

Synthesis

5.1 Contributions

Each of the presented studies explores the correlation between spectral information and plant intraspecific genetic variation. While such linkages have been shown for laboratory conditions (Matsuda et al., 2012), experimental designs (Cavender-Bares et al., 2016), or clonal species (Madritch et al., 2014), the studies of this dissertation target variation in a non-clonal tree species observed under natural conditions. With that, the studies take the opportunity to explore spectral, temporal, spatial and genetic ranges and resolutions. In these contexts, the contributions are presented in the following chapters.

5.1.1 Spectral interpretation of biological variation

In each presented study, the spectral variables are treated as an independent set of spectral phenotypes that relate to a radiation-plant interaction across the whole measured solar spectral range. Further, the studies partition the explained variance in intraspecific plant reflectance spectra into genetic, environmental and technical contributions. This allows for associations of spectral information with biological variation restricted to the genetic scale at the intraspecific level.

The studies treat each spectral variable as a unique plant trait and let the genetic variation to indicate spectral regions that represent biological variation between individuals and correlate with the intraspecific genetic structure (Czyż et al., 2020; Czyż et al., 2023, Czyż et al. in preparation). Based on the post-hoc interpretation of recognized spectral features as particular plant traits, the studies further propose ecophysiological processes that likely underpin the spectral variation (Czyż et al., 2023). Finally, the studies address the intraspecific variation of spectral properties of a species across its natural range and, by statistically accounting for effects of environmental variation over the species range,

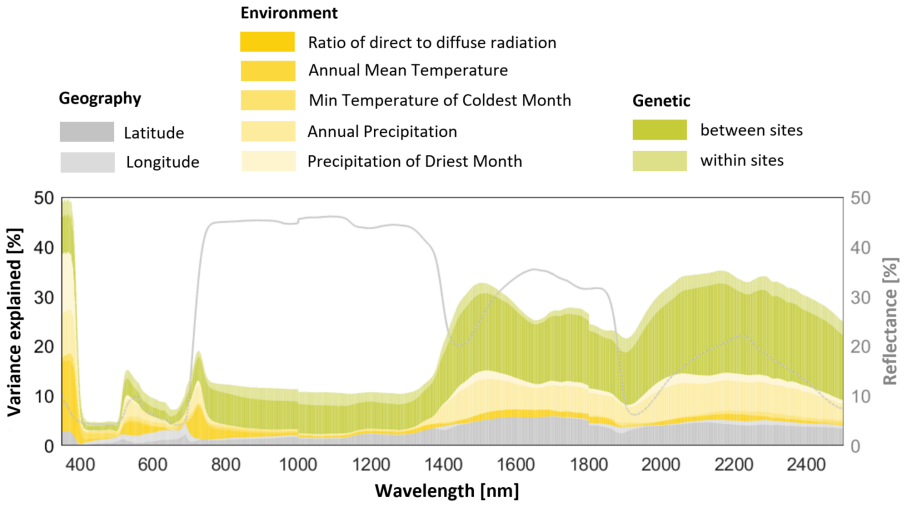

FIGURE 5.1: Hierarchical partitioning of variation in reflectance spectra explained by variation in selected geographical and environmental variation, and genetic variation between and within sites. The analysis was conducted on leaf-level reflectance spectra of 209 *F. sylvatica* individuals sampled across the species range for the study presented in the fourth chapter of this dissertation.

indicate spectral regions whose variation may be most indicative of intraspecific genetic variation (Fig. 5.1).

In addition to evaluating the spectral variation as a function of intraspecific phenotypic and genetic variation, the studies evaluate the impact of technical variation on spectral signals as presented in chapter three (Czyż et al., 2023) and in a co-authored study (Petibon et al., 2021). Further, the studies show that the transformation of the reflectance spectra by per-band normalization, first derivatives, or the relative differences between spectral features, both reducing the impact of observation geometry, enhance the derived correlation between spectral and genetic variation (Czyż et al., 2020; Czyż et al., 2023). The spectral information was further modified to the range of values defined by the measurement uncertainties (Czyż et al., 2023), and the analytical approach to propagate the uncertainties was introduced. With that, the study demonstrates that constraining a spectral information by its inherent uncertainties enhances the biological information derived from spectral data and indicates the fidelity of the products derived from spectra.

5.1.2 Temporal dynamics constrained by genetic composition

Each spectral acquisition captures phenotypes that result from gene expression under the environmental conditions at the particular point in time. The presented studies expand the spectral data to multiple acquisitions over a decade-long period to account for the temporal dynamics of plant traits, underpinned by genes and their expression. For that, the studies use dense-multitemporal airborne spectral acquisitions. The thesis thereby

contributes to new approaches towards upcoming systematic acquisitions of spaceborne spectral data, which will potentially provide repeated acquisitions of the same sites and trees representing the temporal dynamic of phenotypic responses.

The dissertation accounts for the temporal dynamics of phenotypic responses twofold: as consistent responses constrained by genetic composition (Czyż et al., 2020) and as dynamic responses driven by changing environment (Czyż et al., 2023). In the first approach, the study assumes that the reflectance spectra of individuals acquired during the peak of growing season are consistent over years because genotypes of individuals do not change over their lifetimes. In the second approach, the study identifies spectral features resulting from gene expression over the growing season and determines the environmental context under which the similarities in spectral features correlate most with genetic similarities between individuals. The studies respectively distinguish phenotypic responses specific for genetic groups that are consistent over years from others that are only expressed under environmental conditions at a particular point in time. By introducing the dynamic environmental context, the studies also present opportunities for improved understanding of the ecophysiological basis of intraspecific spectral variation at time scales from days to years.

For both studies, decade-long spectral acquisitions over the same population of *F. sylvatica* trees in the Laegern forest in Switzerland, including 69 airborne imaging spectrometer images, were used. Through analysis of these temporally extensive spectral data, the thesis contributes to addressing problems associated with multi-temporal approaches such as imaging spectrometer stability, georeferencing mismatches, and anisotropic effects. To reduce the effects of these problems, the studies propose and respectively apply approaches such as accounting for radiometric uncertainties, spatial calibration, brightness normalization, and multi-directional averaging of the spectral information.

5.1.3 Spatial representation of genetic structure

Wide spatial coverage of spectral acquisitions allows for a collection of spectral identities of a large number of individuals from single or multiple populations of a single species across its range. By observing one species across its natural range under various environmental settings, it is possible to describe intraspecific spectral and presumed genetic variation. The thesis thus presents the potential of imaging spectroscopy to contribute to genetic diversity monitoring across wide spatial extents and it contributes to the description of

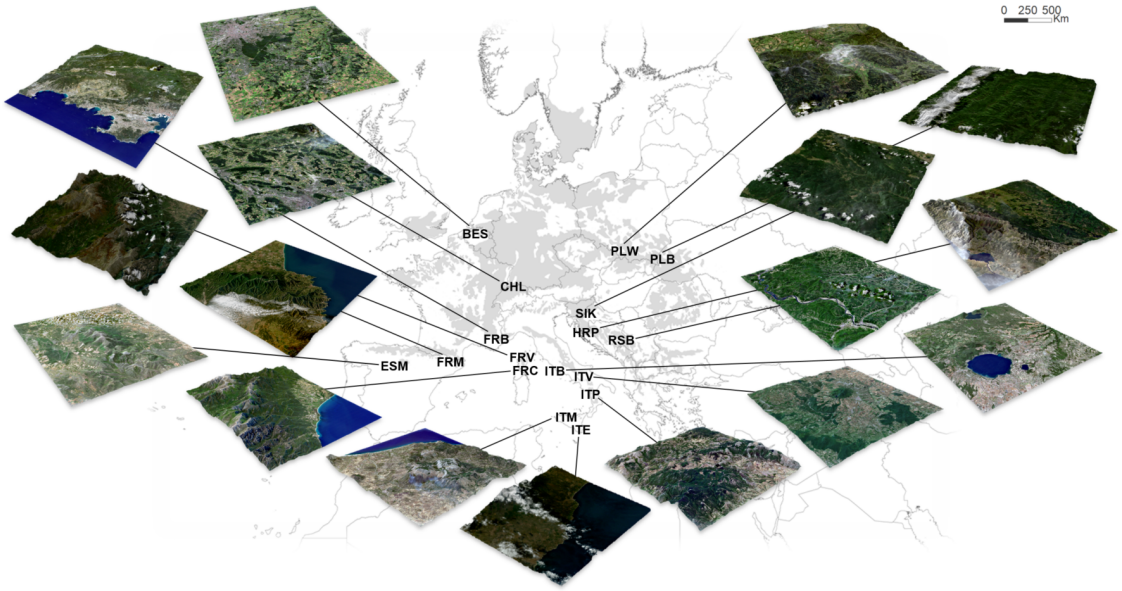

FIGURE 5.2: Representation of spaceborne acquisitions of the DESIS imaging spectrometer requested above selected *F. sylvatica* sites across the species range. The first two letters of abbreviations refer to the country that the site is located in, and the third letter refers to the particular location in the country. The images were acquired above sites investigated in chapter four of this thesis and hold potential for further associations of *in situ* collected genetic information with spectral information, and sensitivity analyses of spaceborne imaging spectrometers for the detection of genetic variation.

the sensitivity of imaging spectrometers for within-population and across-population genetic variation.

The spatial resolution of spectral images acquired from an airborne imaging spectrometer allowed the assigning of spectral identities to tree individuals of the population in the Laegern forest. By correlating the acquired spectral data with the genetic data derived from *in situ* collected leaf samples, the studies show that information derived from airborne imaging spectrometer data holds information on the genetic structure of a single population of genetically relatively similar individuals. While two of the studies target this single population, the third study expands the spatial coverage to the species range. Evaluating the spectral variation of one species across its natural range contributes not only to the description of intraspecific variation among populations but also facilitates the analysis of broad-scale genotype-environment interactions. Partitioning spectral variation into environmental and genetic contributions is essential for establishing correlation between genetic and spectral variation only. To better analyze the contribution of genetic variation to spectral responses, the work contributed to establishing a common garden experiment in which seed families from different natural populations are being grown in a common environment. The collection of the seeds was conducted along with sampling

sites represented in chapter four of this dissertation and is the topic of ongoing work. Correlating the assessed spectral, genetic and environmental variations, the studies provide a proof of concept and potential sensitivity for the prediction of genetic structure across the species range from remotely sensed spectral information. This demonstrates that the incorporation of spectral information can improve the assessment of a species' genetic variation and structure and thus may contribute to the future development of global-scale monitoring and protection of plant biodiversity (Czyż et al., in preparation). In addition to expanding the spatial range of analyses, the studies also investigated multiple spatial resolutions provided by sensor characteristics. The conducted work contributed to the airborne Airborne Visible/Infrared Imaging Spectrometer-Next Generation (AVIRIS-NG) data collection at 1-m spatial resolution and spaceborne DESIS, PRISMA and HISUI data of 30-m spatial resolution above the sampled sites of the species were requested (Fig. 5.2). Integrative approaches to analyze *in situ* collected genetic and spectral data together with imaging spectrometer acquisitions on various spatial resolutions open a path to better match data acquired by different airborne and spaceborne imaging spectrometers and understand their sensitivity to intraspecific genetic variation (Fig. 5.3). Furthermore, such data may contribute to scoping future spaceborne imaging spectrometer missions in the context of genetic diversity assessments.

5.1.4 Description of intraspecific genetic variation

Each study presented in this thesis took advantage of the *in situ* collected plant material from which the genetic information of sampled individuals was derived. The genotype of an individual drives each living process and thus underpins each trait that may affect a spectral phenotype of that individual. Along with establishing a linkage of spectral with genetic variations, the studies describe still underrepresented intraspecific genetic variation and relatedness between individuals within one species.

This thesis presents multiple analytical approaches to linking genetic information with spectral information collected under natural conditions. It uses the molecular genetic information of individuals to assess the genetic structure of *F. sylvatica* at Laegern forest. In one approach, the individuals are assigned to genetic clusters based on kinship probability (Czyż et al., 2020) and in another approach, Nei's genetic distances between individuals are used (Czyż et al., 2023). In the case of both descriptions, the genetic structure is derived from microsatellite data representing selected fragments of DNA. The following study then expands both the coverage of sampled DNA to the whole-genome sequences

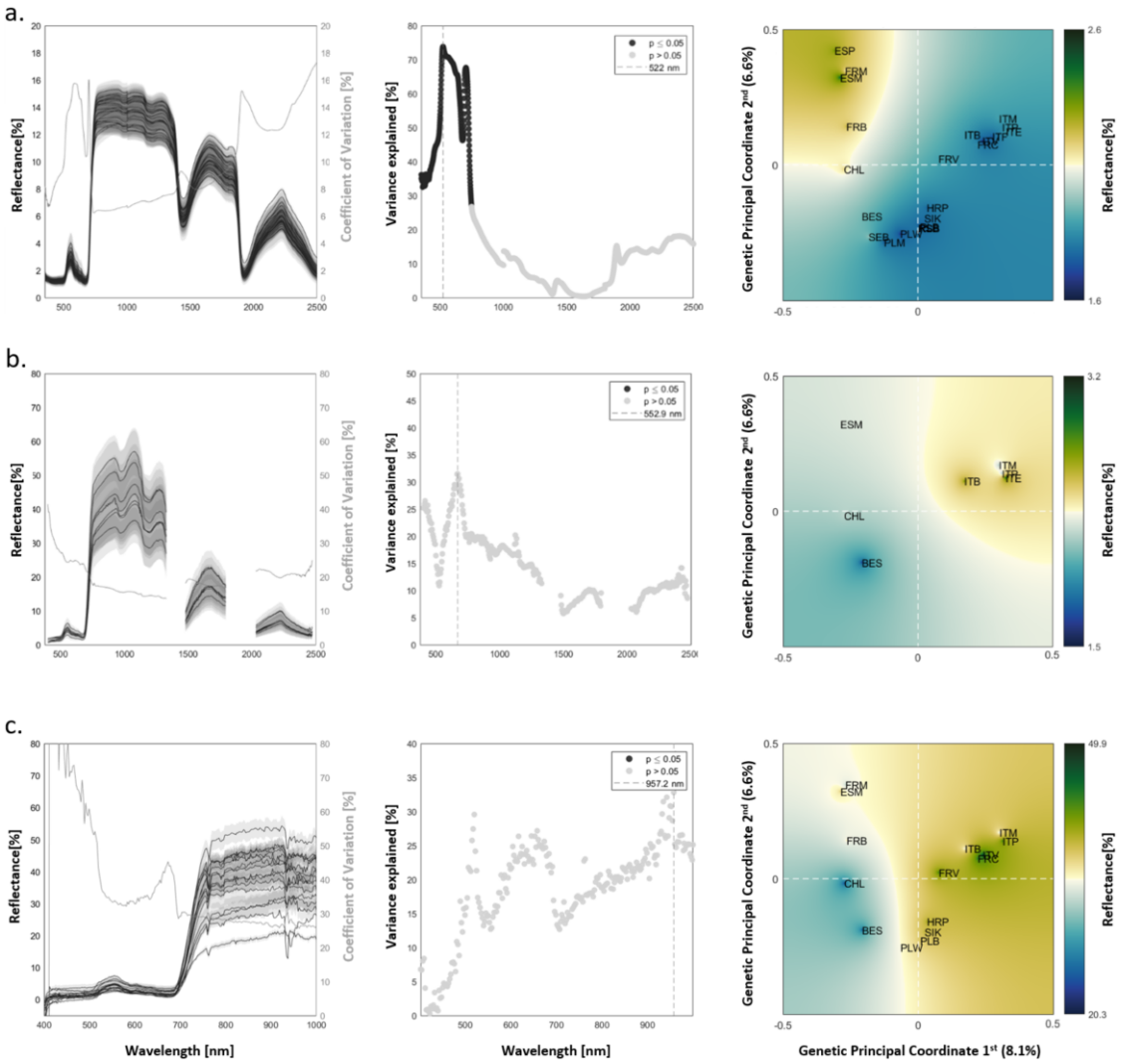

FIGURE 5.3: Explanation of genetic variation among *F. sylvatica* individuals from across the species range by reflectance spectra derived from spectroradiometer attached to a leaf clip (a), airborne imaging spectrometer; AVIRIS-NG (b), and spaceborne imaging spectrometer; DESIS (c). The left panels present mean reflectance spectra across sampled sites; 22 for leaf-level measurements, 7 for airborne acquisitions, and 15 for spaceborne acquisitions. The middle panels present the genetic variance explained by variance at each spectral band. The right panels present the reflectance at the spectral band with the highest explanatory power interpolated between sites located in genetic principal component space. The first two letters of abbreviations refer to the country that the site is located in, and the third letter refers to the particular location in the country. The analyses are preliminary and use leaf-level reflectance spectra and genetic material of 209 *F. sylvatica* individuals sampled across the species range for the study presented in the fourth chapter of this dissertation.

and the range of sampled populations to the species' natural range, while the distribution of sampling sites covers a wider range of genetic variation and environmental conditions to which the species is potentially adapted, and studies the species' phylogeographic history. For that, adjustments of the DNA extraction protocol were introduced to reduce RNA contamination and improve lysis, yield, and DNA purity. Additionally, a bioinformatics pipeline to derive SNPs from whole genomes across sampled individuals followed by their description in genetic principal coordinate space was established. With that, the studies build on current knowledge of genetic structure of a tree species — which until now was commonly derived from short molecular markers — to the genetic structure derived from whole-genome sequences of nuclear DNA from individuals across a high proportion of the species' natural range. This has contributed to current understanding of the spatial pattern of genetic variation of *F. sylvatica*, its potential dispersal from glacial refuges (Magri, 2008; Magri et al., 2006; Stefanini et al., 2023), and often-neglected (Des Roches et al., 2021; Newton et al., 1993) consequences of intraspecific genetic composition for biodiversity protection and restoration.

5.2 Limitations and outlook

This thesis consequently expanded (1) the understanding of spectral signatures, (2) temporal representations of acquired spectral data, (3) the spatial range of assessing intraspecific genetic variation, and (4) the coverage of genetic information used along the whole genome. While all the presented studies analyze spectral data resolved at narrow intervals along the solar spectral range, each study represents results of trade-offs between resources invested into spatial, temporal and genetic ranges and resolution. In the following, some limitations of the studies regarding these four aspects and further potential for contributions emerging from observations linking spectral and genetic data are discussed.

5.2.1 Dimensions of spectral signal

In each of the studies, the spectral information was treated as a set of independent variables representing phenotypic expression of genes. The studies were thereby limited to a purely quantitative representation of the reflected energy without adding the biological context of absorptive spectral features or the integrated reflected energy across the spectral range. Spectral features result from energy absorption by molecules, which affects radiation reflection at multiple, neighboring spectral bands. The description of location,

magnitude, and shape of the absorptive spectral features (Clark and Roush, 1984; Kokaly et al., 2007) may provide more information than single spectral band-based associations with plant genotypes and their interactions with the environment that modifies gene expression and absorption features (Kokaly, 2001). Furthermore, the absorptive properties of different molecules may cause overlapping spectral features of which decoupling by using, for example, a wavelet transform (Ferwerda and Jones, 2006) may unravel further independent absorptions by particular molecules. Finally, the studies did not consider wavelength-independent total reflected energy by the plant leaf or tree canopy relating to their scattering properties and neither the reflected energy was partitioned into direct and diffuse components. The second could be done by decoupling single and multiple scatterings via multi-angular (Bousquet et al., 2005; Grant, 1987) or polarization measurements (Grant et al., 1993). Such scattering characteristics of plants can be indicative of their structural properties (Forsström et al., 2021; Lewis and Disney, 2007) and thus might be temporally more stable and informative of genetic variation than commonly more dynamic physiological traits related to molecular absorptions.

Further, imaging spectrometers with the capabilities of simultaneous measurement of multiple plant traits that affect various spectral ranges distributed across the spectrum hold unique opportunities for plant descriptions. While the studies accounted for normalized differences between single spectral bands across the whole measured spectrum (Czyż et al., 2023), further interrelation between absorptive spectral feature descriptors may unravel additional plant traits and plant-specific traits correlation which could relate to plant strategies (Westoby, 1998) or drive plants' short-term acclimatization and ultimately evolutionary adaptations (He et al., 2009). Such approaches are still based on spectral phenotypes and post-hoc interpretation of unraveled spectral features to particular plant traits. Alternatively, using spectral indexes or multivariate empirical models (Singh et al., 2015) to derive particular plant traits from spectral signal may help to find which spectrally derived plant traits are associated with genetic variation. Such an approach would be limited by the specificity of developed spectral indexes or models, however, they would potentially reduce the impact of technical variation on the associations found.

The ratio of biological to technical variation may also be enhanced by further correction for the fidelity of the spectral signal. One of the presented studies accounted for radiometric uncertainties to enhance associations between the genetic and spectral information (Czyż et al., 2023) but it did not account for all uncertainties possibly affecting the whole data

processing chain (Braverman et al., 2021; Carmon et al., 2020; Hueni et al., 2016; Richter and Schläpfer, 2002). Incorporating all uncertainties of processed products would, for example, also have to account for the accuracy of spatial georegistration and atmospheric correction. Correcting for these would potentially further reduce technical variation and thus enhance the observed associations between spectral and genetic information.

5.2.2 Genetic drivers of temporal trajectories

Through the analysis, the studies take advantage of the extensive airborne datasets collected over the same forest site at multiple points in time. Even though the acquisitions cover a whole decade, the multiple points in time were not regularly spaced. The studies thus only had limited potential to compare temporal trajectories of individual's reflectance spectra over years, seasons, and days representative of ontogenetic shifts and seasonal foliage development (Chavana-Bryant et al., 2017; Garcia and Ustin, 2001; Rivera et al., 2002), both of which may vary between genetically different individuals (Lechowicz, 1984; Menzel et al., 2015; Nienstaedt, 1974).

Temporal trajectories of reflectance spectra are affected by various ecophysiological processes acting on particular temporal scales, thus interpretation of intraspecific genetic variation and underlying processes over a longer temporal scale should be corrected for variation on a shorter temporal scale. The presented study normalized yearly analyses to the seasonal scale by selecting acquisitions representing the peak of the growing season (Czyż et al., 2020), but as for the study on seasonal analysis (Czyż et al., 2023), the study could not normalize completely to the daily scale. This may have weakened the averaged spectral-genetic linkage because likely different short-term ecophysiological processes driving spectral phenotypes may be captured. Further investigation of dense acquisitions at multiple times of the day would help to understand biological variation resulting from genetically constrained daily rhythms (Miller et al., 2015). Additionally, evaluating a magnitude of change in temporal trajectories of reflectance spectra continuously acquired at different temporal scales may further help to interpret unique for genotype changes from one point in time to the other. Alternatively, Fourier analysis of temporal trajectories of reflectance spectra could be used to derive frequencies representing potentially genetically underpinned temporal patterns of plant traits. The deviations from the daily, seasonal, or yearly patterns of reflectance spectra trajectories would depend on the impact of external factors on individuals and on an individual's acclimatization or adaptation thereto.

Exposition of genetically different individuals to the same environmental pressure may

have different influences on their reflectance spectra depending on their tolerance to the stressor and thus may help to unravel a genetic structure between them. The presented studies accounted for the impact of environmental variation on plant traits, but time during or after a higher environmental pressure was not deliberately selected. Alternatively, exposition of different genotypes of the same species to the same environmental stressor under experimental conditions could be used to better assess heritable plant properties and genetic structure among genotypes. Furthermore, in such experimental settings individuals could also be matched for age and size, which could reduce phenotypic variation due to ontogenetic rather than genetic causes, something that was not considered in the present studies, while such knowledge may help to interpret the spectral-genetic link under natural conditions.

The lack of multitemporal spectral signatures and information on the age, size and development history of the over 200 individuals of the tree species investigated across its range prevented a complete distinction between variation in reflectance spectra that resulted from phenological changes or age and size differences, and variation that resulted from environmental conditions and genetic variation. To enhance the confidence of the derived associations of spectral with genetic variation under natural conditions, it may be necessary to expand the temporal coverage of the acquisitions to capture the dynamics of spectral properties or to expand the spatial coverage for potential representation of genetically similar individuals under different environmental conditions. Both of these may be facilitated by spaceborne imaging spectrometer data that would potentially allow for repeated acquisitions over wide spatial scales (Fig. 5.4).

5.2.3 Spatial representation of biological organization

The spatial resolution of acquired images determines the level of biological organization to which the spectral information can be assigned. Among the presented analyses, it was possible to derive spectral signatures of each tree individual. However, this required point measurements at the leaf level and airborne images at 2-m spatial resolution at the canopy level along with Light Detection and Ranging (LiDAR) data (Kaartinen et al., 2012; Schneider et al., 2014; Wang et al., 2001) to delineate single tree crowns (Guillén-Escribà et al., 2021). Extrapolation of the analyses to sites that lack the LiDAR point cloud would require using manual or automatic tree crown segmentation on high-resolution datasets (Warner et al., 2006). Using spectral images with spatial resolution

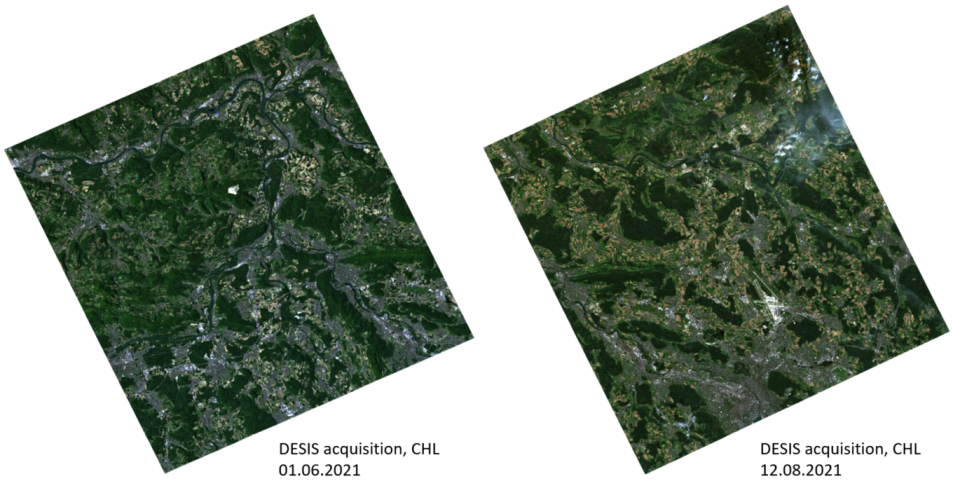

FIGURE 5.4: The DESIS spaceborne imaging spectrometer acquisitions over Laegern, Switzerland, at two points in time: 1st June and 12th August 2021. The spaceborne acquisitions hold potential for repeated data collection and were requested as part of the European-wide CHIME-ESA campaign together with NASA-Jet Propulsion Laboratory (JPL) and University of Zürich (UZH) during which the airborne imaging spectrometer data were collected above multiple *F. sylvatica* forest sites across the species range.

coarser than tree crown diameter – such as images acquired from current spaceborne imaging spectrometers of 30-m spatial resolution – may require spectral unmixing (Roberts et al., 1998), or would have to account for areas representing monospecific patches of multiple neighboring trees. In the latter case, patch-based analyses would reduce the spectral and genetic variation between units of analyses due to averaging among different tree individuals within patches. This may be compensated by expansion of spatial coverage to establish a sufficient link between potentially enhanced spectral and genetic variations.

Furthermore, the presented studies do not account fully for spectral heterogeneity within individual tree crowns (Petibon et al., 2021). By investigating only the reflectance spectra from the top of the canopy, the spectral sampling was normalized only vertically but not to different parts of tree crowns horizontally. Variation within tree crowns in airborne acquisitions also occurs between sun-exposed and shaded parts, which were not distinguished but may allow for the partitioning of the spectral information into direct and diffuse components of reflected radiation (Richter and Müller, 2005) and the indication of tree anisotropic traits (Forsström et al., 2021). Such information may reveal additional genetic constrains on radiation-leaf interactions. Finally, the leaf spectral measurements selected from one tree branch from the top of the canopy represent gene expression characteristic for the branch and its exposition (Lichtenthaler et al., 2007; Watson and Casper, 1984). Thus either the leaf reflectance spectra may be averaged from branches of various expositions to reduce within-individual variation, or branches of similar expositions

should be investigated (Gausman, 1984). Ultimately, the reflectance spectra acquired at the canopy level from airborne or spaceborne platforms would average the spectral signal across branches and may enhance the detection of variation between tree individuals of different genotypes via enhancing proportions of absorptive-to-scattering plant canopy properties (Donaldson et al., 2004).

The studies are also limited by the lack of geospatial information representing environmental conditions at the level of single individuals. An individual's traits exposed in their reflectance spectra are the result of not only their genotypes, but also of the local environmental conditions and their interactions with the genotypes. Thus, controlling for local environmental variation would help to derive the spectral variation that results mainly from intraspecific genetic variation. In the case of the analyzed population from the Laegern forest, there was the assumption that spatial distance partly correlates with environmental variation. In the case of the species range-wide investigations, the restriction of the analysis to between-population levels was applied. This allowed derivation of the percentage of variation in reflectance spectra explained by between-site environmental and genetic variation. Extension of these studies to the within-site level may be achieved in the future by obtaining tree-level environmental data with the consent of site authorities, or derivation from, for example, the available topographic models' environmental layers at the spatial scale of minimal distances between sampled tree individuals.

5.2.4 Genetic associations with the spectral signal

Genetic variation consists of differences in genomes among individuals. Such differences are distributed across the whole genome and may be located in its coding or non-coding regions. The coding regions of the genome correspond to genes underpinning plant traits, whereas non-coding regions can be indicative of the demographic relationship of individuals. Furthermore, together with relatively neutral variation at coding regions, variation at non-coding regions correlate with the potential to adapt to environmental changes (Bolnick et al., 2011; Des Roches et al., 2021; Frankham, 2010).

The presented analyses were limited to the correlation between spectral (phenotypic) variation with genetic relatedness derived from microsatellites or the SNPs derived from short-read whole-genome sequences (Holderegger et al., 2006; Pearman, 2001; Reed and Frankham, 2001). To derive genetic structure among individuals, in the microsatellite-based approach, it was assumed that genetic variation at the chosen genetic markers was mainly neutral, and the study was also limited to the use of the microsatellites proposed

in another study (Pastorelli et al., 2003). Variation in the quality and quantity of selected microsatellite markers may affect the derivation of genetic structure (Kashi and King, 2006; Vinces et al., 2009). To overcome this limitation, either the sensitivity of the available microsatellites should be assessed or further expansion of the coverage to other non-coding regions of the genome should be done to reduce the potential DNA sampling bias. Similarly, for the description of genetic structure based on the whole-genome analyses, it was again assumed that variation at the intraspecific level was mainly neutral. Even though the proportion of neutral variation to adaptive variation at the whole genome scale is normally high, an obvious limitation of this approach was that it did not partition neutral and adaptive variation (Akey et al., 2002; Coop et al., 2010). The derived genetic structure may thus have reflected both the species' phylogeographic history as well as adaptive events.

The span of the studies was limited to the correlation of genetic relatedness between individuals with spectral phenotypes. The derived whole-genome sequences hold potential for direct associations of genetic information with spectral information as phenotypic expression of genes. The studies could be expanded to association studies between genetic variation and phenotypic (spectral) variation to detect functional genes by assigning regions of the genome with spectral features of known molecular absorption. On the other hand, the spectral features of unknown origin may be assigned by annotated genes with which (or in whose surroundings) the derived spectral-genetic correlation is strong. Further, the spectral-genetic correlation discovered at various temporal and spatial scales may be indicative of adaptive genetic variation as differences in allele frequencies. In order to improve the robustness of spectral-adaptive association studies, it would be important to (1) normalize the environment of multiple genetic provenances, as could be done in a common garden experiment, or across transplant experiments, or (2) maximize the range of the studies to increase the representation of genetic and environmental variation to constrain the spectral-genetic correlation. Furthermore, RNA sequences as a direct product of the DNA transcription could be analyzed under different environmental contexts to better understand the control of gene expression of putative functional genes affecting observable spectral phenotypic responses. Such an approach is, however, still limited by efficiency and affordability of current technological advances. In contrast, the incorporation of long-read genome sequences is becoming more accurate and accessible, and their analyses may reduce the uncertainties associated with short-read sequences and expose structural variations across genomes that potentially hold additional information

about genetic variants for which we cannot reliably account based on short-read DNA sequences. The associations between spectral phenotypes and non-neutral genes may further allow for mapping gene variant distribution across space and time, which will serve to better understand and predict ecosystems functioning under current conditions and adaptive potential in changing environments.

Associations of genetic information with the spectral information will further benefit from expanding genetic variation beyond the level of a single species. Additionally, evaluating matches of recognized associations between different clades across the phylogenetic tree may contribute to the understanding of evolutionary processes (Meireles et al., 2020) and the future of biological variation under global change (Cavender-Bares et al., 2022; Exposito-Alonso et al., 2022).

5.3 Conclusions

This thesis integrates spectral information reflecting radiation-plant interactions with genetic information underpinning living processes and the adaptive potential of plants to environmental change. Unraveled associations between the genetic structure of a population or species and its reflectance features along the solar spectral range expose the potential of remote observations to assessing and monitoring biological variation effectively over large areas and over time. By expanding datasets towards denser temporal coverage, more reliable measurements, wider spatial ranges, higher genetic resolution, and by incorporating the environmental context, the studies of this dissertation present a proof of concept on how the spectral signals obtained from tree individuals may become useable for this task, i.e. to assess and monitor genetic variation within species. This thesis also shows multiple limitations of current imaging spectrometers for assessing intraspecific genetic variation. To overcome trade-offs between spectral, temporal, spatial and genetic limitations, it may be necessary to integrate different sensors of the growing fleet of spaceborne instruments for Earth observations and expand the *in situ* sample collection. Spaceborne imaging spectrometers repeatedly acquiring spectral information over wide spatial ranges and the integration of *in situ* sampled genotypic information from extended spatial coverage may further improve models predicting genetic diversity and ultimately facilitate gene variant mapping over large contiguous areas. Continuous spectral monitoring of genetic composition and adaptation over space and time would enhance our understanding of ecological processes, contribute to defining priority areas for gene pool maintenance and restoration, and could help improving strategies towards protecting biodiversity and its services on local to global scales.

References

Akey, J. M., Zhang, G., Zhang, K., Jin, L., & Shriver, M. D. (2002). Interrogating a high-density snp map for signatures of natural selection. *Genome research*, *12*(12), 1805–1814.

Bolnick, D. I., Amarasekare, P., Araújo, M. S., Bürger, R., Levine, J. M., Novak, M., Rudolf, V. H., Schreiber, S. J., Urban, M. C., & Vasseur, D. A. (2011). Why intraspecific trait variation matters in community ecology. *Trends in ecology & evolution*, *26*(4), 183–192.

Bousquet, L., Lachérade, S., Jacquemoud, S., & Moya, I. (2005). Leaf brdf measurements and model for specular and diffuse components differentiation. *Remote Sensing of Environment*, *98*(2-3), 201–211.

Braverman, A., Hobbs, J., Teixeira, J., & Gunson, M. (2021). Post hoc uncertainty quantification for remote sensing observing systems. *SIAM/ASA Journal on Uncertainty Quantification*, *9*(3), 1064–1093.

Carmon, N., Thompson, D. R., Bohn, N., Susiluoto, J., Turmon, M., Brodrick, P. G., Connelly, D. S., Braverman, A., Cawse-Nicholson, K., Green, R. O., et al. (2020). Uncertainty quantification for a global imaging spectroscopy surface composition investigation. *Remote Sensing of Environment*, *251*, 112038.

Cavender-Bares, J., Meireles, J. E., Couture, J. J., Kaproth, M. A., Kingdon, C. C., Singh, A., Serbin, S. P., Center, A., Zuniga, E., Pilz, G., et al. (2016). Associations of leaf spectra with genetic and phylogenetic variation in oaks: Prospects for remote detection of biodiversity. *Remote Sensing*, *8*(3), 221.

Cavender-Bares, J., Schneider, F. D., Santos, M. J., Armstrong, A., Carnaval, A., Dahlin, K. M., Fatoyinbo, L., Hurtt, G. C., Schimel, D., Townsend, P. A., et al. (2022). Integrating remote sensing with ecology and evolution to advance biodiversity conservation. *Nature Ecology & Evolution*, *6*(5), 506–519.

Chavana-Bryant, C., Malhi, Y., Wu, J., Asner, G. P., Anastasiou, A., Enquist, B. J., Cosio Caravasi, E. G., Doughty, C. E., Saleska, S. R., Martin, R. E., et al. (2017). Leaf aging of amazonian canopy trees as revealed by spectral and physiochemical measurements. *New Phytologist*, *214*(3), 1049–1063.

Clark, R. N., & Roush, T. L. (1984). Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *Journal of Geophysical Research: Solid Earth*, *89*(B7), 6329–6340.

Coop, G., Witonsky, D., Di Rienzo, A., & Pritchard, J. K. (2010). Using environmental correlations to identify loci underlying local adaptation. *Genetics*, *185*(4), 1411–1423.

Czyż, E. A., Guillén Escribà, C., Wulf, H., Tedder, A., Schuman, M. C., Schneider, F. D., & Schaepman, M. 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.

Czyż, E. A., Schmid, B., Hueni, A., Eppinga, M. B., Schuman, M. C., Schneider, F. D., Guillén-Escribà, C., & Schaepman, M. E. (2023). Genetic constraints on temporal variation of airborne reflectance spectra and their uncertainties over a temperate forest. *Remote Sensing of Environment*, *284*, 113338.

Des Roches, S., Pendleton, L. H., Shapiro, B., & Palkovacs, E. P. (2021). Conserving intraspecific variation for nature's contributions to people. *Nature Ecology & Evolution*, *5*(5), 574–582.

Donaldson, L. A., Grace, J., & Downes, G. M. (2004). Within-tree variation in anatomical properties of compression wood in radiata pine. *IAWA journal*, *25*(3), 253–271.

Exposito-Alonso, M., Booker, T. R., Czech, L., Gillespie, L., Hateley, S., Kyriazis, C. C., Lang, P. L., Leventhal, L., Nogues-Bravo, D., Pagowski, V., et al. (2022). Genetic diversity loss in the anthropocene. *Science*, *377*(6613), 1431–1435.

Ferwerda, J. G., & Jones, S. D. (2006). Continuous wavelet transformations for hyperspectral feature detection. *Progress in Spatial Data Handling: 12th International Symposium on Spatial Data Handling*, 167–178.

Forsström, P. R., Hovi, A., Ghielmetti, G., Schaepman, M. E., & Rautiainen, M. (2021). Multi-angular reflectance spectra of small single trees. *Remote Sensing of Environment*, *255*, 112302.

Frankham, R. (2010). Challenges and opportunities of genetic approaches to biological conservation. *Biological conservation*, *143*(9), 1919–1927.

Garcia, M., & Ustin, S. L. (2001). Detection of interannual vegetation responses to climatic variability using aviris data in a coastal savanna in california. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(7), 1480–1490.

Gausman, H. W. (1984). Evaluation of factors causing reflectance differences between sun and shade leaves. *Remote Sensing of Environment*, *15*(2), 177–181.

Grant, L. (1987). Diffuse and specular characteristics of leaf reflectance. *Remote Sensing of Environment*, *22*(2), 309–322.

Grant, L., Daughtry, C., & Vanderbilt, V. (1993). Polarized and specular reflectance variation with leaf surface features. *Physiologia Plantarum*, *88*(1), 1–9.

Guillén-Escribà, C., Schneider, F. D., Schmid, B., Tedder, A., Morsdorf, F., Furrer, R., Hueni, A., Niklaus, P. A., & Schaepman, M. E. (2021). Remotely sensed between-individual functional trait variation in a temperate forest. *Ecology and Evolution*, *11*(16), 10834–10867.

He, J.-S., Wang, X., Flynn, D. F., Wang, L., Schmid, B., & Fang, J. (2009). Taxonomic, phylogenetic, and environmental trade-offs between leaf productivity and persistence. *Ecology*, *90*(10), 2779–2791.

Holderegger, R., Kamm, U., & Gugerli, F. (2006). Adaptive vs. neutral genetic diversity: Implications for landscape genetics. *Landscape Ecology*, *21*, 797–807.

Hueni, A., Damm, A., Kneubuehler, M., Schläpfer, D., & Schaepman, M. E. (2016). Field and airborne spectroscopy cross validation—some considerations. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(3), 1117–1135.

Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., et al. (2012). An international comparison of individual tree detection and extraction using airborne laser scanning. *Remote Sensing*, *4*(4), 950–974.

Kashi, Y., & King, D. G. (2006). Simple sequence repeats as advantageous mutators in evolution. *TRENDS in Genetics*, *22*(5), 253–259.

Kokaly, R. F. (2001). Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. *Remote Sensing of Environment*, *75*(2), 153–161.

Kokaly, R. F., Despain, D. G., Clark, R. N., & Livo, K. E. (2007). Spectral analysis of absorption features for mapping vegetation cover and microbial communities in yellowstone national park using aviris data.

Lechowicz, M. J. (1984). Why do temperate deciduous trees leaf out at different times? adaptation and ecology of forest communities. *The American Naturalist*, *124*(6), 821–842.

Lewis, P., & Disney, M. (2007). Spectral invariants and scattering across multiple scales from within-leaf to canopy. *Remote Sensing of Environment*, *109*(2), 196–206.

Lichtenthaler, H. K., Ač, A., Marek, M. V., Kalina, J., & Urban, O. (2007). Differences in pigment composition, photosynthetic rates and chlorophyll fluorescence images of sun and shade leaves of four tree species. *Plant Physiology and Biochemistry*, *45*(8), 577–588.

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.

Magri, D. (2008). Patterns of post-glacial spread and the extent of glacial refugia of european beech (fagus sylvatica). *Journal of Biogeography*, *35*(3), 450–463.

Magri, D., Vendramin, G. G., Comps, B., Dupanloup, I., Geburek, T., Gömöry, D., Latałowa, M., Litt, T., Paule, L., Roure, J. M., et al. (2006). A new scenario for the quaternary history of european beech populations: Palaeobotanical evidence and genetic consequences. *New phytologist*, *171*(1), 199–221.

Matsuda, O., Tanaka, A., Fujita, T., & Iba, K. (2012). Hyperspectral imaging techniques for rapid identification of arabidopsis mutants with altered leaf pigment status. *Plant and Cell Physiology*, *53*(6), 1154–1170.

Meireles, J. E., Cavender-Bares, J., Townsend, P. A., Ustin, S., Gamon, J. A., Schweiger, A. K., Schaepman, M. E., Asner, G. P., Martin, R. E., Singh, A., et al. (2020). Leaf reflectance spectra capture the evolutionary history of seed plants. *New Phytologist*, *228*(2), 485–493.

Menzel, A., Helm, R., & Zang, C. (2015). Patterns of late spring frost leaf damage and recovery in a european beech (fagus sylvatica l.) stand in south-eastern germany based on repeated digital photographs. *Frontiers in Plant Science*, *6*, 110.

Miller, M., Song, Q., Shi, X., Juenger, T. E., & Chen, Z. J. (2015). Natural variation in timing of stress-responsive gene expression predicts heterosis in intraspecific hybrids of arabidopsis. *Nature communications*, *6*(1), 7453.

Newton, A. C., Leakey, R. R., & Mesén, J. F. (1993). Genetic variation in mahoganies: Its importance, capture and utilization. *Biodiversity & Conservation*, *2*, 114–126.

Nienstaedt, H. (1974). Genetic variation in some phenological characteristics of forest trees. *Phenology and seasonality modeling*, 389–400.

Pastorelli, R., Smulders, M., Van't Westende, W., Vosman, B., Giannini, R., Vettori, C., & Vendramin, G. (2003). Characterization of microsatellite markers in fagus sylvatica l. and fagus orientalis lipsky. *Molecular Ecology Notes*, *3*(1), 76–78.

Pearman, P. B. (2001). Conservation value of independently evolving units: Sacred cow or testable hypothesis? *Conservation Biology*, *15*(3), 780–783.

Petibon, F., Czyż, E. A., Ghielmetti, G., Hueni, A., Kneubühler, M., Schaepman, M. E., & Schuman, M. C. (2021). Uncertainties in measurements of leaf optical properties are small compared to the biological variation within and between individuals of european beech. *Remote Sensing of Environment*, *264*, 112601.

Reed, D. H., & Frankham, R. (2001). How closely correlated are molecular and quantitative measures of genetic variation? a meta-analysis. *Evolution*, *55*(6), 1095–1103.

Richter, R., & Müller, A. (2005). De-shadowing of satellite/airborne imagery. *International Journal of Remote Sensing*, *26*(15), 3137–3148.

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.

Rivera, G., Elliott, S., Caldas, L. S., Nicolossi, G., Coradin, V. T., & Borchert, R. (2002). Increasing day-length induces spring flushing of tropical dry forest trees in the absence of rain. *Trees*, *16*(7), 445.

Roberts, D. A., Gardner, M., Church, R., Ustin, S., Scheer, G., & Green, R. (1998). Mapping chaparral in the santa monica mountains using multiple endmember spectral mixture models. *Remote sensing of environment*, *65*(3), 267–279.

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.

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.

Stefanini, C., Csilléry, K., Ulaszewski, B., Burczyk, J., Schaepman, M. E., & Schuman, M. C. (2023). A novel synthesis of two decades of microsatellite studies on european beech reveals decreasing genetic diversity from glacial refugia. *Tree Genetics & Genomes*, *19*(1), 3.

Vinces, M. D., Legendre, M., Caldara, M., Hagihara, M., & Verstrepen, K. J. (2009). Unstable tandem repeats in promoters confer transcriptional evolvability. *Science*, *324*(5931), 1213–1216.

Wang, G., Xia, X.-G., & Chen, V. C. (2001). Three-dimensional ISAR imaging of maneuvering targets using three receivers. *IEEE Transactions on Image Processing*, *10*(3), 436–447.

Warner, T. A., McGraw, J. B., & Landenberger, R. (2006). Segmentation and classification of high resolution imagery for mapping individual species in a closed canopy, deciduous forest. *Science in China Series E: Technological Sciences*, *49*, 128–139.

Watson, M. A., & Casper, B. B. (1984). Morphogenetic constraints on patterns of carbon distribution in plants. *Annual Review of Ecology and Systematics*, *15*(1), 233–258.

Westoby, M. (1998). A leaf-height-seed (lhs) plant ecology strategy scheme. *Plant and soil*, *199*, 213–227.

ISBN Nr. 978-3-906894-29-4