Assessing Plant and Lichen Diversity Using Reflectance Spectra

Lance Stasinski

University of Maine, lance.stasinski@maine.edu

Follow this and additional works at: https://digitalcommons.library.umaine.edu/etd

Part of the Botany Commons

Recommended Citation


https://digitalcommons.library.umaine.edu/etd/3539

This Open-Access Thesis is brought to you for free and open access by DigitalCommons@UMaine. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of DigitalCommons@UMaine. For more information, please contact um.library.technical.services@maine.edu.
ASSESSING PLANT AND LICHEN DIVERSITY USING REFLECTANCE SPECTRA

By

Lance Stasinski

B.S. Chadron State College, 2019

A thesis

Submitted in partial fulfillment of the
Requirements for the Degree of
Master of Science
(in Botany and Plant Pathology)

The Graduate School
The University of Maine
December 2021

Advisory Committee:

José Eduardo Meireles, Assistant Professor of Plant Evolution and Systematics

Peter R. Nelson, Forest Ecology Director at the Schoodic Institute

Daniel Stanton, Assistant Professor of Ecology, Evolution and Behavior (UMN)

Jacquelyn Gill, Associate Professor of Paleoecology and Plant Ecology

Brian McGill, Professor of Biological Science
Biodiversity is changing and it is imperative that we continually assess it in order to preserve ecological services that we rely on. Spectral platforms are increasingly being used to assess biodiversity due to the fact that light reflected from an organism’s surface carries much of information about that organism. Despite the promise spectroscopy shows, two gaps in our knowledge remain. First, we do not know how well reflectance spectra can be used to estimate fine-scale diversity – intraspecific genetic and phenotypic diversity – that is fundamental to ecological and evolutionary processes. Second, spectral libraries, used to construct models to estimate diversity, have largely been built from plant spectra and have neglected other ecologically important organisms such as lichens.

To investigate the first problem, my colleagues and I tested the utility of reflectance spectra for distinguishing genomically defined populations. We collected spectra (400–2400 nm) and samples from co-occurring *Dryas alaskensis*, *Dryas ajanensis*, and hybrid individuals from six different mountaintops in the interior of Alaska, United States. We used partial least squares discriminant analysis (PLS-DA) to classify leaf reflectance spectra into six populations defined by STRUCTURE and PCA analyses using genomic data. We also estimated the phylogenetic signal carried by the spectra, and we used PLS beta regression to estimate the proportion of ancestry for each individual from the reflectance spectra. We found that the two
species and their six populations could be distinguished with 99.7% and 98.9% overall accuracy, respectively. A significant phylogenetic signal was found for all regions of the spectrum, and the model for estimating the proportion of ancestry explained 91% of the variation with an RMSE of 0.13. Hybrids were classified with 80% accuracy, and this is thought to be due to a lack of strong trait correlations. These findings suggest that fine-scale diversity can be retrieved from reflectance spectra and this should be considered in future spectrally-based biodiversity assessments.

To address the second problem, I investigated whether herbarium specimens would be valuable for building a spectral library for lichens. Specifically, I investigated whether lichen specimens were altered by the long-term desiccation inherent with herbarium storage and if that influenced the classification of herbarium specimens. I used a spectral dataset of 30 lichen species that covered an age range of 126 years, and used linear mixed-effects models and PLS-DA to determine 1) how reflectance changed with age, and 2) the influence of age on classification accuracy. I found that the reflectance for wavelengths between 700 and 1900 nm decreased by less than 0.2% reflectance per year, but wavelengths outside this range did not clearly respond to aging. This implies a gradual change in thallus structure over time in herbarium storage. Species, families, orders, and classes were classified with 77.0 to 94.5% accuracy, and these models were only marginally influenced by specimen age. These results indicate that lichen specimens do change over time, but these changes do not negate their utility for building spectral libraries for lichens.
ACKNOWLEDGEMENTS

First and foremost, I would like to thank Dawson M. White, Peter R. Nelson, Richard H. Ree, and José Eduardo Meireles for their hard work on the *Dryas* manuscript (Chapter 1) which is now published in *New Phytologist*. I would especially like to thank Dawson M. White for completing the genomic analyses presented in this manuscript. I would also like to thank Catherine Chan and Kevaughan Smith for helping with the data collection in Alaska, Shan Kothari and Laura Williams for a supportive review of the *Dryas* manuscript, Brian J. McGill for his statistical advice regarding MANOVAs and linear mixed-effects models, and Daniel Stanton for sharing the lichen spectral dataset that was used in Chapter 2. Finally, I extend my gratitude to my co-advisers José Eduardo Meireles and Peter R. Nelson for their continued support and advice throughout the development of this thesis, and I would like to especially thank José Eduardo Meireles for the countless meetings, late night paper edits, and continued mentorship during my time at the University of Maine.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................ iv

LIST OF TABLES .................................................................................................................... vii

LIST OF FIGURES .................................................................................................................. viii

LIST OF ABBREVIATIONS ........................................................................................................ x

Chapter

1. READING LIGHT: LEAF SPECTRA CAPTURE FINE-SCALE DIVERSITY

   OF CLOSELY RELATED, HYBRIDIZING ARCTIC SHRUBS ............................................ 1

   Introduction ......................................................................................................................... 1

   Materials and Methods ....................................................................................................... 4

      Taxon Sampling ............................................................................................................... 4

      Genetic Structure and Ancestry ....................................................................................... 6

   Spectral Data ....................................................................................................................... 7

   Spectral Analyses ................................................................................................................ 8

   Morphological Analyses ..................................................................................................... 11

   Results ................................................................................................................................. 12

      Genetic Analyses ............................................................................................................ 12

      Spectral Analyses ............................................................................................................ 15

      Morphological Analyses ............................................................................................... 21

   Discussion ............................................................................................................................ 22
2. HERBARIUM SPECIMENS: A SOURCE FOR BUILDING SPECTRAL LIBRARIES FOR LICHENS

Introduction.............................................................................................................28

Materials and Methods..........................................................................................31
  Data Collection....................................................................................................31
  Response to Aging...............................................................................................35
  Classification........................................................................................................37

Results....................................................................................................................38
  Response to Aging...............................................................................................38
  Classification........................................................................................................42

Discussion...............................................................................................................47
  Response to Aging...............................................................................................47
  Classification........................................................................................................50

BIBLIOGRAPHY.....................................................................................................53

APPENDIX A. DRYAS............................................................................................63

APPENDIX B. LICHEN............................................................................................77

BIOGRAPHY OF THE AUTHOR..........................................................................94
**LIST OF TABLES**

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 2.1.</td>
<td>Lichen species taxonomic classification and morphology</td>
<td>32</td>
</tr>
<tr>
<td>Table 2.2.</td>
<td>Linear mixed-effects model types</td>
<td>36</td>
</tr>
<tr>
<td>Table 2.3.</td>
<td>Brightness vs age model comparison</td>
<td>42</td>
</tr>
<tr>
<td>Table 2.4.</td>
<td>PLS-DA results</td>
<td>43</td>
</tr>
<tr>
<td>Table 2.5.</td>
<td>T-test: Comparing model accuracy with and without age</td>
<td>44</td>
</tr>
<tr>
<td>Table A.1.</td>
<td>Distances between sampling sites (km)</td>
<td>72</td>
</tr>
<tr>
<td>Table A.2.</td>
<td>Number of individuals and scans per taxon and site</td>
<td>72</td>
</tr>
<tr>
<td>Table A.3.</td>
<td>Accuracy and number of components for each classification unit</td>
<td>73</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1.1. Map of study area in Alaska, United States..............................................5
Figure 1.2. Dryas species and population genetic structure.................................13
Figure 1.3. Species and population classification from leaf reflectance spectroscopy....16
Figure 1.4. Confusion matrix from the PLS-DA model discriminating samples by
species and collection site.................................................................................18
Figure 1.5. Predicted Dryas alaskensis ancestry from spectral reflectance versus
genomic ancestry..............................................................................................20
Figure 1.6. Pairwise correlations between taxonomically informative leaf traits........22
Figure 2.1. Delta AIC for all wavelengths...............................................................39
Figure 2.2. Results of the fixed slope model for every wavelength......................40
Figure 2.3. Confusion matrix for classifying species without age as a covariate....45
Figure A.1. Proportion of ancestry from Dryas alaskensis.....................................63
Figure A.2. STRUCTURE plots delimiting D. ajaneneis (DAJ) and
D. alaskensis (DAK)..........................................................................................64
Figure A.3. Variable importance for classifying species........................................65
Figure A.4. Classification matrix for collection site.............................................66
Figure A.5. Classification matrix for species and site without hybrids.................67
Figure A.6. Classification of site with the visible spectrum and water absorption
features removed from the reflectance spectra.................................................68
Figure A.7. Variable importance for classifying site of collection........................69
Figure A.8. Phylogenetic signal by 1nm wavelength.............................................70
Figure A.9. Leaf morphology by species...............................................................71
LIST OF ABBREVIATIONS

Alaska Locations
BG - Bison Gulch
ES - Eagle Summit
MD - Murphy Dome
TM - Twelve Mile Summit
WDA - Wickersham Dome A
WDB - Wickersham Dome B

Dryas Species
DAJ - *Dryas ajanensis*
DAK - *Dryas alaskensis*
DX - *Dryas* hybrids

Spectra
VIS - Visible wavelengths of light between 400 and 700 nm.
NIR - Near infrared wavelengths of light between 700 and 1100 nm.
SWIR - Short-wave infrared wavelengths of light between 1400 and 2400 nm.
nm - nanometers

Statistics
AIC - Akaike information criterion
MANOVA - Multivariate analysis of variance
PCA - Principal components analysis
PLS-DA - Partial least squares discriminant analysis
PLS-R - Partial least squares regression
READING LIGHT: LEAF SPECTRA CAPTURE FINE-SCALE DIVERSITY OF CLOSELY RELATED, HYBRIDIZING ARCTIC SHRUBS

Introduction

As biodiversity faces significant threats worldwide (Bellard et al., 2012; Isbell et al., 2015; Pimm et al., 2014), it is critical that we improve our ability to assess diversity in order to understand and conserve ecological and evolutionary processes (Atwater & Callaway, 2015; Cavender-Bares et al., 2017). Although biodiversity is typically assessed at the species level, it is now clear that fine-scale diversity – the genetic and phenotypic diversity present within species and taxonomic complexes – also warrants evaluation and monitoring. Fine-scale diversity facilitates taxon-specific adaptive potential (Hoffmann et al., 2017) and it can have significant positive effects on community structure and ecosystem productivity (Atwater & Callaway, 2015; Crawford & Rudgers, 2013; Raffard et al., 2019). The influences and consequences of fine-scale diversity are particularly important to consider among Arctic-alpine plants because they frequently exhibit high within-species genetic diversity (Grundt et al., 2006).

In addition to fine-scale diversity, hybrid taxa are another dimension of biodiversity requiring attention because hybrids may exhibit ecologically or evolutionarily significant capabilities, such as the effects of transgression on adaptive potential and invasiveness (Abasolo et al., 2012; Dittrich-Reed & Fitzpatrick, 2013; Gallego-Tévar et al., 2018; Rieseberg et al., 2003, 2007). Alternatively, phenotypically intermediate hybrids can enable the genetic swamping and extinction of parent taxa (Todesco et al., 2016). These low levels of biological organization are most frequently, and appropriately, defined using genomic data, but it is imperative that we
advance our methodologies for estimating and monitoring this fine-scale diversity to maximize biodiversity data in a rapidly changing world.

Leaf reflectance spectroscopy is emerging as a powerful tool to assess plant phylogenetic and functional diversity and monitor how it changes over space and time (Cavender-Bares et al., 2017; Jetz et al., 2016; Meireles et al., 2020; Schweiger et al., 2018; Serbin et al., 2019; Turner, 2014). This approach is based on the principle that the electromagnetic radiation reflected off leaves (400–2500 nm) carries information about their structural and chemical traits (Cavender-Bares et al., 2017; Curran, 1989; Ustin et al., 2009a). We know that pigments absorb in the visible region (VIS; 400–700 nm) whereas light in the near infrared region (NIR; 700–1100 nm) is scattered by leaf anatomical, tissue, water, and surface features, and light in the short-wave infrared region (SWIR; 1400–2500 nm) is scattered and absorbed by anatomical features and biochemicals like cellulose, phenolics, and water (Carter, 1991; Fang et al., 2017; Gates et al., 1965; Kokaly & Skidmore, 2015; Ustin et al., 2009a). Thus, leaf spectra are dense, complex, and dynamic phenomic datasets influenced by both environmental and genetic factors. They can be effectively leveraged using supervised statistical models to estimate phenotypic traits (Ustin & Gamon, 2010) or estimate diversity (Durgante et al., 2013; Lang et al., 2017; Meireles et al., 2020; Schweiger et al., 2018).

Leaf spectra have mostly been used to differentiate groups at or above the species level, and these taxa are usually classified with high accuracy (Durgante et al., 2013; Lang et al., 2017; Meireles et al., 2020; R. Wang & Gamon, 2019). The few studies that have assessed fine-scale diversity have shown variable success. Cavender-Bares and colleagues (2016) classified leaves from four Quercus oleoides populations with low accuracy, but Madritch and colleagues (2014) were able to classify 79 genotypes of Populus tremuloides with moderate to high accuracy. Thus,
it is important to continue to test the limits of spectral detection using genomic data and small spatial scales for closely related taxa, populations, and hybrids.

While we expect traits to be able to separate populations and species, hybrid individuals will not necessarily form a phenotypically cohesive group because individuals will inherit and express different sets of parental alleles resulting in variable phenotypes and weakly correlated traits (Cheng et al., 2011; Guo et al., 2004). While such individuals are difficult to statistically classify using macro-morphological data like organ sizes and shapes (Abasolo et al., 2012; Field et al., 2009), reflectance spectra are proving to be detailed enough to identify such challenging hybrid taxa. For instance, an analysis of bloodwoods (*Corymbia*), showed their weedy, morphologically complex hybrids could not be accurately classified using 30 morphological traits, but classifications of these same hybrids using spectral data was 72–100% accurate (Abasolo et al. 2012; Abasolo et al. 2013). Reflectance spectra have also been used to accurately detect *Citrus* (Páscoa et al., 2018) and *Populus* (Deacon et al., 2017) hybrids.

In this study, we estimated the ability of leaf spectra to detect phylogenetic divergence and low taxonomic levels in a complicated, yet frequently encountered, biological scenario of closely related, co-occurring, hybridizing plants. *Dryas octopetala* L. *sensu lato* has been described as one to nine species and numerous subspecific taxa representing geographically and/or ecologically distinct lineages, many of which form hybrids with intergrading morphologies (Elven et al., 2011; Hultén, 1959; Skrede et al., 2006). Two noteworthy taxa (currently described as species) in this complex are *D. ajanensis* Juzepczuk ssp. *beringensis* Jurtzev (hereafter *D. ajanensis*) and *D. alaskensis* A.E. Porsild (in accordance with the *Flora of North America* treatment; Springer and Paritt, 2014), which were established as ecotypes through the exemplary experiments of (McGraw & Antonovics, 1983). These species are very close
relatives, possibly sister species, that are believed to have diverged during the Pleistocene (Hultén 1959). Individuals expressing intermediate leaf morphologies, long presumed to be hybrids, are found in tight contact zones between habitats (Hultén, 1968; Max et al., 1999).

We generated spectral and genomic datasets for Dryas individuals, collected across six alpine sites in the interior of Alaska, and used them to determine: (1) if we can use spectra to accurately classify species, hybrids, and genetically defined populations, and (2) to what degree spectral variation correlates with genetic variation at these low taxonomic levels. These closely related, co-occurring, hybridizing taxa provide a challenging system to explore the boundaries of biodiversity detection via spectroscopy.

**Materials and Methods**

**Taxon Sampling**

We sampled *D. ajanensis*, *D. alaskensis*, and putative hybrids from six alpine sites in the interior of Alaska, United States (Fig. 1). *Dryas ajanensis* is an abundant species found in isolated patches on dry, rocky fellfields across boreal North America. It is readily identified by small tomentose leaves (5–15 mm) with rust-colored ‘scales’ ( multicellular, feathered hairs) occurring on the midvein on the bottom (abaxial) side of the leaf. *Dryas alaskensis* occurs in wet tundra microhabitats in Alaska and the Yukon Territories. It has larger leaves (15–50 mm) with less pubescence and stipitate glandular trichomes on the abaxial midvein, and sometimes with adaxial wax secretions (Hultén, 1959; Hultén, 1968; McGraw & Antonovics, 1983). Individual plants displaying both kinds of adaxial midvein pubescence (rarely was this on the same leaf) were identified in the field as hybrids.
Figure 1.1. Map of study area in Alaska, United States. Sampling sites: BG = Bison Gulch, ES = Eagle Summit, MD = Murphy Dome, TM = Twelve Mile Summit, WDA = Wickersham Dome Site A, WDB = Wickersham Dome Site B. The maximum distance between sites was 252 km, the minimum distance was 3 km, and the average distance was 118 km. The red rectangle on the inset map of North America represents the study area.
We collected approximately 10 to 15 leaves from evenly spaced Dryas individuals along a 100 meter transect (except no transect was used at Wickersham Dome B) that traversed wet and dry habitats, and two voucher specimens per taxon were collected at each site and deposited at the Field Museum herbarium. We collected leaves from the following number of individual plants from each site; Bison Gulch (BG: 20 D. ajanensis), Eagle Summit (ES: 19 D. alaskensis, 21 D. ajanensis, 2 hybrids), Murphy Dome (MD: 20 D. ajanensis), Twelve Mile Summit (TM: 22 D. alaskensis, 20 D. ajanensis, 2 hybrids), Wickersham Dome A (WDA: 20 D. ajanensis) and Wickersham Dome B (WDB: 11 D. alaskensis, 16 D. ajanensis, 5 hybrids; Fig. 1; Table A.1).

**Genetic Structure and Ancestry**

We isolated DNA from silica-dried leaf samples and prepared GBS libraries using the ApeKI restriction endonuclease (Elshire et al., 2011). Libraries were prepared and sequenced at the University of Wisconsin DNA Sequencing Center on an Illumina NovaSeq 6000 (2 x 150 bp reads). We aligned reads to the D. drummondii genome (GCA_003254865.1) and called single nucleotide polymorphisms (SNPs) using ipyrad (Eaton & Overcast, 2020). A description of steps from DNA isolation to SNP calling is available in Methods A.1.

To establish the population genetic structure of this system, we first conducted principal components analysis (PCA) in R adegenet v2.1.3 using all SNPs to summarize genetic variation and discontinuities among individuals and sites (Jombart, 2008; R Core Team, 2020). Next, we used the model-based clustering method STRUCTURE to estimate the number of populations in the dataset via each sample’s proportional assignment to a set number of inferred ancestral groups (Pritchard et al., 2000). For one to ten groups (K), we completed ten replicate runs of 1 million generations plus 500,000 generations of burn-in. STRUCTURE’s Q value is the proportion of an individual’s ancestry from ancestral group K, therefore K=2 represents an index
of genomic ancestry from *D. ajanensis* or *D. alaskensis*. We then completed separate
STRUCTURE runs for each species to delineate populations. To gauge further support for the
separation of lineages by populations or sites, we dropped the hybrid samples and reconstructed
a maximum-likelihood phylogeny from the full, concatenated SNP dataset using IQ-TREE
v1.6.12 with the GTR+ASC model of nucleotide evolution and 1000 ultrafast bootstrap
pseudoreplicates (Hoang et al., 2018; Minh et al., 2020).

**Spectral Data**

We collected leaves in labeled tea bags, placed them within plastic bags containing ample
silica gel desiccant, and left them to dry for 36 to 60 hours before scanning. Water and water
vapor have strong effects on spectral reflectance, especially in the NIR and SWIR regions
(Carter, 1991), so thoroughly drying the leaves with silica gel reduces the effect of this
environmental variable as well as reveals distinctive anatomical and chemical features in these
spectral regions (Costa et al., 2018; Páscoa et al., 2018). While the limitations of field work
imposed variability in the duration of drying times, we are confident that these small leaves with
thin cuticles were fully desiccated in less than 24 hours of storage in silica gel. This is based on
the empirical work of Carrió & Rosselló (2014) described below and in Methods A.2.

Reflectance measurements (scans hereafter) were taken after a warm-up period of at least 15
minutes using a PSR+ portable spectroradiometer (Spectral Evolution, Haverhill, MA, USA)
with reflectance contact probe (with tungsten halogen light source) and leaf probe clip. A
Spectralon ® white reference was recorded every five samples to recalibrate, then the leaf clip
was reversed to the black background during spectral readings.

We scanned the leaves in three stages due to the small size of the leaves. First, we
arranged two to five leaves across the leaf clip to cover as much of the entire field of view as
possible without overlap and with the adaxial leaf surface facing the probe. We then scanned spectral reflectance from 350–2500 nm two times to include possible variability from the instrument or slight variations in the output from the light source. For the second stage, we added one to three leaves to cover more of the exposed black background and took two more scans. We added one to three more leaves for the third stage, resulting in a total of six scans per specimen. The addition of leaves was implemented to explore the two possible shortcomings of measuring leaves that were smaller than the spot size of the reflectance contact probe; those being exposed background that generates noise in the spectra and partially overlapping leaves that alter the shape and magnitude of the reflected spectrum (Neuwirthová et al., 2017).

**Spectral Analyses**

We processed the reflectance spectra by removing erroneous scans that had reflectance values greater than 1.0. The resulting dataset consisted of 1045 scans at 1 nm resolution representing 178 individual plants (Table A.2). We then trimmed the spectra to a length of 400 to 2400 nm to remove regions of higher noise (350–399 nm, 2401–2500 nm; Cavender-Bares et al., 2016). All manipulations of the spectra were conducted using the Spectrolab R package (Meireles et al., 2017). These processed spectra were used in the following analyses, and all six scans (or fewer if scans had reflectance values greater than 1.0) were used to represent the respective specimen.

We classified species (with and without hybrids), populations, collection sites, and a combination of species identity and collection site (with and without hybrids) using partial least squares discriminant analysis (PLS-DA), a multivariate analysis that classifies observations from PLS regression on indicator variables (Chevallier et al., 2006). This method works well with high dimensional multicollinear datasets such as those acquired via leaf reflectance spectroscopy.
(Barker & Rayens, 2003; Cavender-Bares et al., 2016). We executed the following PLS-DA procedure independently for each classification unit (species, populations, sites, and species plus site). For classifying species, populations, and sites, we partitioned the data such that a random 80% of the spectral data was used as a training set, and the remaining 20% was used for a testing set. However, for classifying the combination of species and location, we split the data so that 50% was used for testing and 50% was used for training because low sample sizes in the in some of the classes, such as hybrids belonging to a particular location, prevented the use of smaller allocations of data for testing the models. Further, the spectra exhibited large variation at the scale of the individual, so we chose to treat each scan as an independent measurement in these models.

We used a two-step procedure to account for class imbalance. In the first step, we ran 100 iterations of PLS-DA with ten-fold cross validation repeated three times. Each iteration used an independent partition of the data (see above for how the partition was made) in which the training set was downsampled, and the final model of the respective iteration was tested against the testing set to determine model accuracy. The number of components for which the average overall accuracy was maximized was used as the optimal number of components in the second step. This second step was the same as the first except that the training data was upsampled in each iteration and the number of components used was equal to the number chosen from the first step to prevent overfitting. Finally, the overall classification accuracy was calculated as the mean accuracy extracted from confusion matrices across all the 100 iterations used with the upsampling procedure (second step).

We evaluated the potential effect of differences in drying times on our classifications by using the same PLS-DA procedure, with 10 iterations instead of 100, to classify the site of
collection using spectra with water absorption features and the visible spectrum removed. We assessed site-based classification because it covaried with drying time, and we removed the water absorption features and the visible spectrum because pigments and water content are most likely to be affected by drying (Carter, 1991; Chen et al., 2012). The specific wavelengths removed from the spectra for this analysis were the visible spectrum 400–749 nm and water absorption features at 960–980 nm, 1170–1190 nm, 1235–1255 nm, 1300–1460 nm, 1750–2030 nm, 2040–2060 nm, 2135–2155 nm, and 2153–2173 nm (Thenkabail & Lyon, 2016; Z. Wang et al., 2019).

Additionally, we created a PLS beta regression model, implemented in the R package plsRbeta, to predict the proportion of *D. alaskensis* ancestry as quantified by the genetic analyses (Fig. A.1a), for each sample using the spectra as predictor variables (Bertrand et al., 2013). The PLS beta regression restricts the response variable to a beta distribution which is necessary to predict the continuous ancestry variable inherently restricted between 0 and 1. We used *Dryas alaskensis* ancestry and spectral data from the three sites in which both *D. alaskensis* and *D. ajanensis* occurred. The model was assessed after ten-fold cross validation repeated five times. We chose the optimal number of components (52) as the lowest number of components within two Akaike information criteria (AIC) of the number of components with the lowest AIC. Model RMSE and r-squared statistics were determined by comparing the mean predicted ancestry values from the 5 repeats to the ancestry values estimated by the genetic analyses.

We determined the phylogenetic signal, or the degree to which closely related individuals resemble each other more than any two randomly drawn individuals from the same phylogenetic tree (Blomberg & Garland, 2002), for each wavelength in the reflected spectrum. We calculated Blomberg’s K, a measure of phylogenetic signal that compares the similarity of any two
individuals to that expected by Brownian motion (Blomberg et al., 2003; Meireles et al., 2020), using the phylosig function in the R Phytools package with an ultrametric tree, the mean spectra per each individual, the standard errors associated with those means, and 500 tip-swap simulations per trait (1 nm wavelength; Paradis & Schliep, 2019; Revell, 2012). The result was then compared to the maximum Blomberg’s K value estimated from data simulated on a tree with no phylogenetic covariance.

**Morphological Analyses**

We analyzed leaf morphologies and trait correlations among parent species and hybrids to test our hypothesis that hybrids have variable and inconsistent morphologies compared to the parent taxa. We measured four taxonomically informative, though not necessarily spectrally informative, traits from 135 total leaves (5 leaves x 9 individuals x 3 taxa total; (Elven et al., 2011; E. Hultén, 1959; Springer et al., 2014): leaf length (millimeters), abaxial midvein scales (presence or absence), abaxial midvein stipitate glandular trichomes (presence or absence), and degree of adaxial tomentum (glabrous–or nearly so, intermediate, or dense). Pubescence and midvein morphology were observed with a digital microscope. We used a multivariate analysis of variance (MANOVA) to test for differences between the traits of parent species and the traits of the hybrids (R Core Team, 2020). We then constructed trait correlation matrices for parent species and hybrids using Pearson product-moment correlations. The correlation structure of the parents was compared to that of the hybrids using a Mantel test with 23 permutations (complete enumeration) in the Vegan R package (Oksanen et al., 2019).
Results

Genetic Analyses

We generated a median of 3.08 M reads per sample with a median of 99.8% of these passing quality filters. We then mapped a median of 22,345 loci per sample to the reference genome with an average depth of coverage of 38.7 reads, resulting in 52,325 filtered SNPs from 12,961 GBS loci mapped to 367 *D. drummondii* scaffolds. We subsampled SNPs at least 50,000 bp apart to get a final dataset of 3,042 unlinked SNPs.

The STRUCTURE results corroborated this clear separation of *D. alaskensis* and *D. ajanensis* individuals. When two clusters were defined (*K* = 2), the putative morphologically intermediate individuals were inferred to be 31.5–48.1% admixed hybrids (Fig. 2a). When we dropped the hybrids and analyzed the two species independently across several *K* (Fig. A.2), we found *D. alaskensis* was best modeled as two populations, *D. alaskensis*-WDB and *D. alaskensis*-ESTM (Eagle Summit plus Twelve Mile), which is sensible because ES and TM are connected by contiguous habitat (Fig. 1). *Dryas ajanensis* also showed lack of gene flow among sites, being best modeled as four populations: *D. ajanensis*-BG, *D. ajanensis*-ESTM, *D. ajanensis*-MD, and *D. ajanensis*-WD (Wickersham Dome A plus Wickersham Dome B; Fig. 2a, Fig. A.2).
Figure 1.2. *Dryas* species and population genetic structure. (a) STRUCTURE bar plots showing the proportion of each individuals’ genome assigned to *D. alaskensis* (DAK) and *D. ajanensis* (DAJ) species (*K*=2, top), and proportional assignment to the optimal six groups (*K*=6, bottom). (b) Plot of first two principal components, explaining 13.9% of the total genetic variation among individuals with 95% quantile ellipses circumscribing each site for each taxon. (c) Maximum likelihood cladogram without hybrids; branches with <70% bootstrap support have been collapsed, backbone nodes with >95% bootstrap support are indicated by circles. Colors for site and population assignments are provided in legend. Hybrids are abbreviated as DX. Site abbreviations are as follows: BG = Bison Gulch, ES = Eagle Summit, ESTM = Eagle Summit and Twelve Mile, MD = Murphy Dome, TM = Twelve Mile, WDA = Wickersham Dome A, WDB = Wickersham Dome B, and WD = Wickersham Dome A and B.
The PCA of SNPs also revealed a significant genetic separation of species and sites (Fig. 2b). The first principal component (PC) explained 11.1% of the total genetic variation and clearly separated *D. ajanensis* and *D. alaskensis* individuals, with hybrids falling in the middle, and the second PC separated individuals by site (2.8% of variance explained). In both species, individuals from ES and TM overlapped on both PC axes, supporting those as a single population. *Dryas ajanensis*-WDA and *D. ajanensis*-WDB individuals had proximal values on these PC axes but overlapped only on a single axis.

Phylogenetic inference corroborated a robust separation of *D. alaskensis* and *D. ajanensis* as well as the six total populations inferred from STRUCTURE (Fig. 2c). Within *D. alaskensis*, all WDB individuals, except a single individual, form a clade with 96% bootstrap support that is set within several ES and/or TM lineages (Fig. 2c). Within the *D. ajanensis* clade, BG, ESTM, and MD all form well-supported monophyletic groups. Individuals from WDB form a nested series of lineages subtending a nearly monophyletic clade of WDA individuals (Fig. 2c).

Overall, the results of our genetic analyses support the delineation of six reproductively isolated groups: *D. alaskensis* is split into *D. alaskensis*-ESTM and *D. alaskensis*-WDB, and *D. ajanensis* is split into *D. ajanensis*-BG, *D. ajanensis*-ESTM, *D. ajanensis*-MD, and *D. ajanensis*-WD. The PCA and STRUCTURE show some genetic differentiation among the *D. ajanensis*-WDA vs WDB sites, but due to their overall similarity and proximity (3 km), as well as the fact that our main conclusions were robust to splitting these sites into two populations, we chose to treat them as a single population.
Spectral Analyses

The mean reflectance for *D. ajanensis*, *D. alaskensis*, and the hybrids appeared superficially similar (Fig. 3a), yet we were able to train PLS-DA models to classify spectral scans of the three taxa with 92.9 ±1.8% accuracy (Fig. 3b; the ‘±’ indicates one standard deviation). *Dryas alaskensis* was predicted with 95.9 ±2.7% accuracy and *D. ajanensis* with 92.4 ±2.6% accuracy. Hybrids alone were predicted with 80.3 ±13.9% accuracy (Fig. 3b). All regions of the reflected spectrum were useful for separating these taxa with the VIS and SWIR appearing to be relatively more informative than the NIR according to variable importance calculations - the contribution of coefficients associated with each wavelength weighted proportionally to the reduction in the sums of squares (Fig. A.3; Kuhn, 2021). This may indicate that these species are best separated by their leaf pigments, such as chlorophyll, carotenoids, and anthocyanins, as well as their lignin, cellulose, and phenolic compound contents (Kokaly & Skidmore, 2015; Thenkabail & Lyon, 2016; Ustin et al., 2009a). When PLS-DA models were trained without hybrids, *D. alaskensis* and *D. ajanensis* were classified with an overall accuracy of 99.7 ±0.4%.

In contrast to the minor spectral reflectance differences at the species level, the mean spectral reflectance values of populations noticeably varied (Fig. 3c). The PLS-DA models correctly classified scans to their genetically determined population with 97.7 ±1.9% to 100% accuracy depending on the population (98.9 ±0.7% overall accuracy; Fig. 3d). See Table A.3 for complete overall accuracy statistics and the number of components used for each set of classification models.
Figure 1.3. Species and population classification from leaf reflectance spectroscopy. a) Mean reflectance values for each species. b) Confusion matrix from the PLS-DA model discriminating species. c) Mean reflectance values for each population. d) Confusion matrix from the PLS-DA model discriminating populations. For the spectra plots, DAJ indicates *D. ajanensis*, DAK indicates *D. alaskensis*, and DX indicates hybrids. In the confusion matrices, the number in each cell represents the proportion of scans from the reference class (row) classified into the predicted class (column). Correctly classified scans fall into the diagonal and misclassifications are off-diagonal. White cells represent zeroes. The proportion of correct identifications are indicated.
by the size and shade of the orange squares. Site names are abbreviated as follows: ESTM = Eagle Summit and Twelve Mile, WDB = Wickersham Dome B, BG = Bison Gulch, MD = Murphy Dome, and WD = Wickersham Dome A and Wickersham Dome B.

We were able to classify both the species identity and collection site, the finest classification resolution, with an overall accuracy of 92.0 ±1.2% (Fig. 4). Site alone was predicted with an accuracy of 99.8 ±0.3 % (Fig. A.4). Leaf scans from all *D. alaskensis* sites as well as *D. ajanensis* from BG and MD were correctly assigned with over 94% accuracy when hybrids were included in the model. Occasionally, one species was misclassified as the other species belonging to the same site with the main contributions to this phenomenon coming from 8.5 ± 5.9% of *D. ajanensis* from WDB being misclassified as *D. alaskensis* from WDB, which could be due to the higher levels of admixture at this site (Fig. 2a). Most of the misclassifications were attributed to *D. alaskensis* or *D. ajanensis* being classified as a hybrid from the same site. For example, 6.2 ±4.0% of *D. ajanensis* from ES, 9.0 ±3.1% of *D. ajanensis* from TM, and 9.9 ±6.3% of *D. ajanensis* from WDB were misclassified as hybrids from the respective sites. When we modeled species and site without hybrids, individuals were classified with 98.9 ±0.8% accuracy (Fig. A.5). Most of the error was due to *D. ajanensis* from ES and WDB being classified as *D. alaskensis* from the respective site (2.5±2.9% for ES, 5.0 ±6.7% for WDB).

When we removed water absorption features and the visible wavelengths from the spectra, our PLS-DA models classified location with 99.8 ±0.3% accuracy (51 components, Fig. A.6) which indicates that our drying procedure did not create an unintended drought artifact. Further, our analysis of residual leaf water content across sites showed weak spatial autocorrelation (Methods A.2, Fig. A.7), so drying the leaves for 36–60 hours in silica gel prior to scanning was sufficient to control for this environmental factor.
**Figure 1.4.** Confusion matrix from the PLS-DA model discriminating samples by species and collection site. Row and column names are abbreviations as the site of collection. The numbers in each cell correspond to the proportion of classifications per row averaged over 100 iterations. Larger squares and darker shades of orange indicate larger proportions of scans classified as the corresponding cell. White cells correspond to true zeros, and zeros indicate proportions less than 0.001. Correctly identified species and collection site combinations are represented in the
diagonal and misclassifications are found on the off-diagonal. Site abbreviations are as follows:

BG = Bison Gulch, ES = Eagle Summit, MD = Murphy Dome, TM = Twelve Mile, WDA = Wickersham Dome A, and WDB = Wickersham Dome B.

The PLS beta regression predicted the proportion of *D. alaskensis* ancestry of the samples with a mean RMSE of 0.13 and an $r$-squared of 0.91 (Fig. 5). Also, our test for phylogenetic signal indicated that all 1 nm wavelengths carried phylogenetic information with the VIS and SWIR being most similar between closely related individuals (Fig. A.8). Blomberg’s K was greater than the maximum Blomberg’s K estimated from the null model and significant for all wavelengths ($p = 0.002$). These results corroborate the evidence from our discriminant analyses that leaf reflectance spectra carry useful genetic information and reflect phylogenetic relationships at these fine levels of biological organization.
Figure 1.5. Predicted *Dryas alaskensis* ancestry from spectral reflectance versus genomic ancestry. The points indicate the mean predicted ancestry for each individual plant, and the bars represent the full range of ancestry predicted from individual scans per plant. The diagonal line represents the 1:1 perfect relationship between predicted and actual values. Note: The *D. ajanensis* individual in the top left of the plot had only one scan after the spectral cleanup procedure.
Morphological Analyses

We evaluated the hypothesis that hybrids exhibit inconsistent morphologies by quantifying a few taxonomically (though not necessarily spectrally) informative morphological traits and their correlations. *Dryas ajanensis* and *D. alaskensis* were clearly separated by abaxial midvein pubescence and by adaxial tomentum, in agreement with previous descriptions (Eric Hultén, 1968), but hybrid individuals are not uniformly intermediate across all traits (Fig. A.9). Although both midvein morphologies (glands and scales) were present on hybrid plants observed in the field, individual leaves were not consistent in these traits.

Our analysis of four taxonomically informative leaf traits (leaf length, tomentum, glands, scales) showed high trait correlation coefficients between leaves of parent *D. ajanensis* or *D. alaskensis* individuals, indicating consistent trait expression (absolute values from 0.46–0.87; Fig. 6a). Unsurprisingly, the traits of hybrids were significantly different from the parent species (MANOVA: $F = 3.7, p = 0.007$), but the correlation coefficients between traits in hybrid individuals were much weaker (absolute values from 0.08–0.31, and 0.50 for scales and glands; Fig. 6b). Although we do not expect these few traits to drive spectral reflectance values, this supports our hypothesis that traits have differentially segregated among hybrid individuals and thus their overall morphology is not consistently intermediate between parents. The Mantel test confirmed the correlation structure between traits within hybrids was not the same as the parent species ($r = 0.77, p = 0.083$).
Figure 1.6. Pairwise correlations between taxonomically informative leaf traits. a) Trait correlations among leaves of the parent species, *D. ajanensis* and *D. alaskensis*. b) Trait correlations among leaves of hybrids. More detailed descriptions of traits are found in the Materials and Methods section. Negative correlations are blue, positive correlations are red, and the shading corresponds to the magnitude of the correlations.

**Discussion**

Biodiversity detection from reflectance spectra has mostly focused on taxa at or above the species level (Durgante et al., 2013; Féret & Asner, 2012; Ustin & Gamon, 2010) despite the importance fine-scale genetic and phenotypic diversity has for ecological and evolutionary processes (Crawford & Rudgers, 2013; Hoffmann et al., 2017). Thus, our goal was to determine if fine-scale diversity characterized from genomic sequence data over a small geographic scale could be detected via leaf spectral reflectance. We demonstrated that reflectance spectroscopy captures genetic information that can be used to accurately classify leaves to species, hybrids, and populations in a taxonomically challenged group of arctic dwarf shrubs.
Genomic sequencing proved an effective method for establishing species and population-level structure in *Dryas*. Max and colleagues (1999) hypothesized that *Dryas* plants at alpine sites in interior Alaska would have distinct allozyme profiles due to the effects of drift on isolated mountaintop populations. We were initially unsure if we would find population genetic structure in this study region because Max and colleagues (1999) found no allozyme differences among sites and very few differences between species, and we expected that widely distributed species like *D. ajanensis* would be genetically cohesive at this spatial scale (Fig. 1). However, our dense sampling and deep sequencing strategy readily revealed the genetic differentiation of *Dryas* species and mountaintop population structure (Fig. 2). Together, our studies confirm there is negligible gene flow across unsuitable habitat at lower, forested elevations. Also, the genomic data showed *D. ajanensis* and *D. alaskensis* were clearly differentiated, and while it would be useful to sample other parts of the range, our results indicated these are valid (plant) species exhibiting morphological and phylogenetic distinction despite recurrent hybrid formation (Baum, 2009).

Several pieces of evidence indicate that our models can recover this fine-scale genetic structure from leaf reflectance spectra. First, we found significant phylogenetic signal in the spectra (Fig. A.8), which adds to the body of evidence that spectra convey evolutionary relatedness (Cavender-Bares et al., 2016; Madritch et al., 2014; Meireles et al., 2020; Schweiger et al., 2018). We also demonstrate that the full reflected spectrum is useful for determining phylogenetic relationships below larger taxonomic units, such as families or orders, as is presented by Meireles and colleagues (2020). Second, our PLS beta regression model accurately estimated the proportion of *D. alaskensis* ancestry from the spectra, and the models explained 91% of the variation in the ancestry of the individuals. This is a novel and alternative approach to
spectral classification of taxa that is a promising direction for dynamic analyses of hybrid zones and broader analyses of population structure and admixture. This analysis was also successful in revealing spectral similarities within species sampled from multiple sites and the intermediate spectral values of hybrid individuals (Fig. 5).

Third, the genomic resolution of species and population-level structure was captured very well by the spectral reflectance data, which were used to classify leaves to their populations with 98.9% average accuracy (Fig. 3d). The PLS-DA models successfully classified leaves from populations ESTM and WD to their correct species with 97.8% (D. ajanensis-WD) to 99.7% (Dryas alaskensis-ESTM) accuracy despite these populations spanning two sites (Fig. 3d). Our classification of these six Dryas populations was more accurate than the discrimination of four populations of Quercus oleoides grown in a common garden (Cavender-Bares et al., 2016). Conversely, our ability to classify these populations is comparable to the classification of Populus tremuloides genotypes from aerial imaging spectroscopy (Madritch et al., 2014), which further demonstrates the ability of spectra to detect fine-scale genetic diversity in-situ.

Lastly, we accurately classified species identities from reflectance spectra of D. alaskensis and D. ajanensis that co-occurred (Fig. 3b), and we observed near-perfect classification accuracy (99.7%) when we trained the models without hybrids. We classified these two species of Dryas from leaf reflectance spectra more accurately than reported for classifying Quercus species (Cavender-Bares et al., 2016); however, the difference in the number of species included in the models (two species of Dryas versus 28 species of Quercus) and the differences in sample preparation (dry Dryas leaves versus fresh Quercus leaves) may account for the differences in model accuracy. The accuracy of our models that classified both of the Dryas
species and the hybrids were less accurate than previous studies that classified aspen (*Populus*; Deacon et al., 2017) and bloodwood (*Corymbia*; Abasolo et al., 2013) species and hybrids.

The most accurate classification model of our study (99.8% overall accuracy) discriminated samples by their collection sites alone (Fig. A.4). An initial reaction to this could be that we were classifying the environment as opposed to genetics, but this model actually added the signal from genetics with any signal from site-based environmental factors. The confounding effects of environmental factors on phenotype was only directly observed in the few instances where taxa were misclassified as a different taxon belonging to the same site, which was rare and mostly involved hybrids (Fig. 4). While it would be desirable to understand the environmental versus genetic components of the observed variability, the unanticipated covariance of genetic structure with mountaintop sites in this *Dryas* system is not conducive to such an analysis. Nonetheless, we provide multiple lines of evidence that fine-scale genetic diversity is an identifiable signal in the spectra.

Our analysis of trait correlations was an effort to understand why hybrids failed to achieve classification accuracies as high as parent taxa (80% vs. 92 to 96%; Fig. 3b). This analysis indeed showed that hybrid leaves overall had very weak correlations among several taxonomically informative traits compared to leaves of parent species, as would be expected from the differential inheritance and expression of alleles in hybrid individuals (Cheng et al., 2011; Guo et al., 2004). However, we did not classify hybrids poorly at every site, and the average classification accuracy of 80.3% is most likely pulled down by 41.3% of hybrids from Eagle Summit (ES) being classified as *D. alaskensis* from ES (Fig. 4). This may be because the genetic admixture of the hybrid sample ES03_DX was 31.5% (Fig. A.1b) – the least balanced hybrid we sampled – and the leaf morphology of both ES hybrids leaned toward *D. alaskensis*.
(glandular trichomes, sparse tomentum, little to no midvein scales). In contrast to the ES hybrids, the Twelve Mile (TM) and Wickersham Dome B (WDB) hybrids were classified quite accurately (88.4–99.5% accuracy; Fig. 4). The two samples from TM had genetic admixture close to 50% (47.2% and 47.7%), and the five samples from WDB had admixtures ranging from 36.2–48.1%; WDB hybrids showed a range of admixture but benefitted from a larger sample size.

In a similar study, three hybrid aspen individuals with intermediate leaf morphology between parent species were classified using spectra with 94–99% accuracy (Deacon et al., 2017). On the other hand, over 900 individuals representing several Corymbia hybrid taxa could be classified using 30 morphological traits with only 9.5–76% accuracy, but these morphologically challenging and inconsistent hybrid taxa could be classified using spectra with 72–100% accuracy (Abasolo et al., 2012, 2013). In summary, we should expect classification accuracy of genetically or morphologically inconsistent hybrids to improve when more individuals are scanned – a standard effect of sample size in statistical classification (Foody, 2009). Our study design differs from the others because we were not primarily focused on hybrid taxon classification, so our hybrid sample sizes were determined by their relative abundance along a transect.

We have shown that closely related, co-occurring plant species, their hybrids, and their populations can be distinguished by the way light is reflected from their leaves. Our study extends the body of evidence on the utility of leaf-level spectral profiles by showing that they can successfully detect fine-scale genetic variation and thus can be applied to all levels of biological diversity above the level of the individual (Bickford et al., 2007; Cavender-Bares et al., 2016; Madritch et al., 2014; Meireles et al., 2020). Scaling biodiversity detection from leaf-level spectra to remotely-sensed imagery comes with challenges that should not be
understated, such as the presence of pixels that contain multiple individuals of different species (Rocchini, 2007; R. Wang & Gamon, 2019). However, the fact that fine-scale evolutionary diversity is captured by the spectrum of a leaf – a fundamental biological unit – suggests that assessing genetic variation using remote sensing should be possible if scaling methods improve and higher spatial resolution spectral images become more widespread. Similarly, the fact that hybrids can be accurately classified from spectra indicates that spectroscopy could be leveraged for studying plants along hybrid zones and accelerate the study of speciation, functional ecology, and community interactions (Abbott, 2017; Campbell et al., 2018; Evans et al., 2012).
CHAPTER 2

HERBARIUM SPECIMENS: A SOURCE FOR BUILDING SPECTRAL LIBRARIES FOR LICHENS

Introduction

Biodiversity detection and monitoring via the use of spectroscopy has shown great success (Cavender-Bares et al., 2016; Deacon et al., 2017; Durgante et al., 2013; Madritch et al., 2014; Meireles et al., 2020; Stasinski et al., 2021) due to the fact that the light reflected from a biological surface captures information about that organism’s structural components and chemistry (Cavender-Bares et al., 2017). For instance, the visible spectrum (VIS; 400–700 nm) carries information about plant pigments (Ustin et al., 2009a), and the near infrared (NIR; 700–1100 nm) and short-wave infrared (SWIR; 1400–2500 nm) regions capture structural information, such as lignin and cellulose concentrations (Curran, 1989) and water content (Carter, 1991), and information about phenolic compounds (Kokaly & Skidmore, 2015). Despite the utility of spectroscopy, this technology has mostly been applied to the detection of vascular plant diversity (Asner et al., 2017; Cavender-Bares et al., 2017; Jetz et al., 2016; Ustin & Gamon, 2010; R. Wang & Gamon, 2019). Other ecologically important autotrophs that inhabit most terrestrial ecosystems have received comparatively less attention (Guzmán Q. et al., 2020; Nelson et al., 2013; Neta et al., 2010). Lichens in particular have been often overlooked despite their existence in nearly all terrestrial ecosystems and their contribution to ecological processes such as nitrogen fixation, productivity, and nutrient cycling (Asplund & Wardle, 2017; Bokhorst et al., 2015; Nash, 2010; Rundel, 1978; Zedda et al., 2014).
Bridging the gap between plant and lichen spectral libraries could take a substantial amount of effort because lichens will have to be collected and identified – a lengthy process that often involves chemical spot tests and microscopy – before a useful spectral dataset can be established. One way to reduce the time needed to build such a spectral library could be to use herbarium specimens because they have already been identified, and they typically include additional useful metadata such as lichen chemistry. Lichen spectral libraries built from the spectra of herbarium specimens may assist in training models that can be used to identify historical specimens, capture misidentifications, identify functional traits such as thallus mass per area and the identity and concentration of secondary compounds (Asplund et al., 2012; Asplund & Wardle, 2017; Ellis & Coppins, 2006), and they may even serve to classify contemporary lichen diversity (Zomer et al., 2009). The process of specimen aging, however, is an obstacle that stands in the way of using herbarium specimens to build useful spectral libraries. It remains unclear whether lichen specimens change as they are subjected to the long-term desiccation that is inherent with herbarium storage, and, if lichens do change with time spent in storage (referred to as age hereafter), we do not know how much it would affect applications of such a spectral library such as classifying herbarium specimens.

The aging process of herbarium samples may be complicated by the fact that lichens are a diverse group of symbioses that have separately evolved several times (Nash, 2010), and they respond differently to the long-term desiccation (Kranner et al., 2008). For example, the rate at which chlorophyll concentrations decrease during desiccation is dependent upon the identity of the photosynthetic partner in the symbiosis (Kranner et al., 2003). Moreover, lichens also display a diversity of morphologies and secondary metabolites that may change differently with
Changes in photosynthetic and cortical pigments as well as thallus structure are captured by different patterns of reflectance changes throughout the spectrum (Kokaly & Skidmore, 2015; Neta et al., 2010; Ustin et al., 2009b). The VIS contains information about cortical and photosynthetic pigments such as chlorophyll and carotenoids (Curran, 1989; Ustin et al., 2009a); therefore, chlorophyll degradation and the breakdown of accessory pigments should be captured in this region. In contrast, changes in the reflectance in the NIR and shortwave-infrared SWIR would be related to changes in thallus mass per area, structural proteins and polysaccharides, water content, and phenolic compounds (Carter, 1991; Curran, 1989; Kokaly & Skidmore, 2015; Thenkabail & Lyon, 2016).

I had three expectations for how long term aging in herbaria would be reflected in the spectra and how it would influence classification models. First, thallus reflectance spectra of different species would respond differently to aging due to the diversity of secondary cortical compounds and morphologies exhibited by diverse lichen species. Second, although I expected reflectance to vary at different rates between different species, I generally expected the following patterns to occur: 1) Reflectance in the VIS would increase for blue (400–500 nm) and red (600–700 nm) wavelengths while it would decrease for green wavelengths (500–550 nm) due to the decomposition of chlorophyll which absorbs red and blue light (Curran, 1989). 2) Reflectance in the NIR and SWIR would decrease due to the slow decomposition of thallus structure via the decomposition of cellular components caused by enzymatic activity from microbes found on herbarium specimens (Bieker et al., 2020) or by chemical interactions with the substrate or the atmosphere (Molak & Ho, 2011). Third, classification models would be
influenced by specimen age due to the expected correlation between reflectance and specimen age.

My overarching goals were to determine how lichen specimens change with age and if those changes influenced the ability to classify herbarium specimens. I used a dataset of reflectance spectra from 30 lichen species spanning a specimen age range of 126 years to determine: (1) if the reflectance spectra change with specimen age, (2) if the rate of change is similar between different species, and (3) if age influenced accuracy when classifying species and higher taxonomic ranks.

**Materials and Methods**

**Data Collection**

The Stanton Lab at the University of Minnesota collected four to six 400–2400nm reflectance measurements (scans hereafter) from 301 individual thalli of lichen herbarium specimens, stored in the Bell Museum in St. Paul Minnesota. Samples represented 30 species, 19 families, 16 orders, and 6 classes (Table 2.1). Each scan per individual was taken from a random and different region of the lichen thallus using a leaf clip attachment of an SVC HR1024i spectroradiometer (Spectra Vista Corp., Poughkeepsie, NY, USA) in a dark room. Samples were chosen with the goal of maximizing phylogenetic coverage while also selecting species that could be easily measured (a flat surface large enough to fill the area of the leaf clip aperture); thus, most specimens have either crustose or foliose morphologies, and one species, *Ephebe ocellata*, has an appressed filamentous fruticose morphology. Most of the specimens were originally collected in Minnesota, United States, but some collections were from the greater Great Lakes region of the United States. The age of specimens ranged from 7 years to 133 years;
however, most of the specimens are less than 60 years old due to a general lack of collectors from 1910 to 1960 (Figure B.1). The number of individual thalli and length of time covered was not equal between all species due to the limited number of specimens available in the herbarium.

I cleaned the spectra by matching overlap regions between detector arrays (wavelengths 900 nm and 1900 nm), trimming the spectra to wavelengths between 400 and 2400 nm to remove regions of higher noise (Schweiger, 2020), and by resampling the spectra to 1nm resolution. All cleanup procedures were conducted using functions in the Spectrolab R package (Meireles et al., 2017).

Table 2.1 Lichen species taxonomic classification and morphology.

<table>
<thead>
<tr>
<th>Species</th>
<th>Family</th>
<th>Order</th>
<th>Class</th>
<th>Morphology</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Acarospora americana</em></td>
<td>Acarosporaceae</td>
<td>Acarosporales</td>
<td>Lecanoromycetes</td>
<td>crustose</td>
</tr>
<tr>
<td><em>Baeomyces rufus</em></td>
<td>Baeomycetaceae</td>
<td>Baeomycetales</td>
<td>Lecanoromycetes</td>
<td>crustose</td>
</tr>
<tr>
<td><em>Caloplaca flavovirens</em></td>
<td>Teloschistaceae</td>
<td>Teloschistales</td>
<td>Lecanoromycetes</td>
<td>crustose</td>
</tr>
<tr>
<td><em>Candelaria concolor</em></td>
<td>Candelariaceae</td>
<td>Candelariales</td>
<td>Candelariomycetes</td>
<td>foliose</td>
</tr>
<tr>
<td><em>Chysothrix candelaris</em></td>
<td>Chrysotrichaceae</td>
<td>Arthoniales</td>
<td>Arthoniomycetes</td>
<td>crustose</td>
</tr>
<tr>
<td><em>Dimelaena oreina</em></td>
<td>Caliciaceae</td>
<td>Teloschistales</td>
<td>Lecanoromycetes</td>
<td>crustose</td>
</tr>
<tr>
<td><em>Ephebe ocellata</em></td>
<td>Lichinaceae</td>
<td>Lichinales</td>
<td>Lichinomycetes</td>
<td>fruticose</td>
</tr>
<tr>
<td><em>Flavoparmelia baltimoresis</em></td>
<td>Parmeliaceae</td>
<td>Lecanorales</td>
<td>Lecanoromycetes</td>
<td>foliose</td>
</tr>
<tr>
<td><em>Flavoparmelia caperata</em></td>
<td>Parmeliaceae</td>
<td>Lecanorales</td>
<td>Lecanoromycetes</td>
<td>foliose</td>
</tr>
<tr>
<td><em>Flavoparmelia euplecta</em></td>
<td>Parmeliaceae</td>
<td>Lecanorales</td>
<td>Lecanoromycetes</td>
<td>foliose</td>
</tr>
<tr>
<td>Species</td>
<td>Family</td>
<td>Order</td>
<td>Class</td>
<td>Life Form</td>
</tr>
<tr>
<td>---------</td>
<td>--------</td>
<td>-------</td>
<td>-------</td>
<td>-----------</td>
</tr>
<tr>
<td>Flavoparmelia haysomii</td>
<td>Parmeliaceae</td>
<td>Lecanorales</td>
<td>Lecanoromycetes</td>
<td>foliose</td>
</tr>
<tr>
<td>Flavoparmelia rutidota</td>
<td>Parmeliaceae</td>
<td>Lecanorales</td>
<td>Lecanoromycetes</td>
<td>foliose</td>
</tr>
<tr>
<td>Flavoparmelia soredians</td>
<td>Parmeliaceae</td>
<td>Lecanorales</td>
<td>Lecanoromycetes</td>
<td>foliose</td>
</tr>
<tr>
<td>Flavoparmelia subcapitata</td>
<td>Parmeliaceae</td>
<td>Lecanorales</td>
<td>Lecanoromycetes</td>
<td>foliose</td>
</tr>
<tr>
<td>Flavopunctelia darrowi</td>
<td>Parmeliaceae</td>
<td>Lecanorales</td>
<td>Lecanoromycetes</td>
<td>foliose</td>
</tr>
<tr>
<td>Flavopunctelia flaventior</td>
<td>Parmeliaceae</td>
<td>Lecanorales</td>
<td>Lecanoromycetes</td>
<td>foliose</td>
</tr>
<tr>
<td>Flavopunctelia praesignis</td>
<td>Parmeliaceae</td>
<td>Lecanorales</td>
<td>Lecanoromycetes</td>
<td>foliose</td>
</tr>
<tr>
<td>Flavopunctelia soredica</td>
<td>Parmeliaceae</td>
<td>Lecanorales</td>
<td>Lecanoromycetes</td>
<td>foliose</td>
</tr>
<tr>
<td>Graphis scripta</td>
<td>Graphidaceae</td>
<td>Ostropales</td>
<td>Lecanoromycetes</td>
<td>crustose</td>
</tr>
<tr>
<td>Ionaspis lacustris</td>
<td>Hymenieliaeae</td>
<td>Lecanorales</td>
<td>Lecanoromycetes</td>
<td>crustose</td>
</tr>
<tr>
<td>Lecidea tessellata</td>
<td>Lecideaeae</td>
<td>Lecideales</td>
<td>Lecanoromycetes</td>
<td>crustose</td>
</tr>
<tr>
<td>Loxospora elatina</td>
<td>Sarrameanaceae</td>
<td>Lecanorales</td>
<td>Lecanoromycetes</td>
<td>crustose</td>
</tr>
<tr>
<td>Neofuscelia verruculifera</td>
<td>Parmeliaceae</td>
<td>Lecanorales</td>
<td>Lecanoromycetes</td>
<td>foliose</td>
</tr>
<tr>
<td>Peltigera elisabethae</td>
<td>Peltigeraceae</td>
<td>Peltigerales</td>
<td>Lecanoromycetes</td>
<td>foliose</td>
</tr>
<tr>
<td>Pertusaria opthalmiza</td>
<td>Variolariaeae</td>
<td>Pertusariales</td>
<td>Lecanoromycetes</td>
<td>crustose</td>
</tr>
<tr>
<td>Rhizocarpon grande</td>
<td>Rhizocarpaceae</td>
<td>Rhizocarpales</td>
<td>Lecanoromycetes</td>
<td>crustose</td>
</tr>
<tr>
<td>Strigula submuriformis</td>
<td>Strigulaceae</td>
<td>Strigulales</td>
<td>Dothideomycetes</td>
<td>crustose</td>
</tr>
<tr>
<td>Lichen Name</td>
<td>Family</td>
<td>Order</td>
<td>Class</td>
<td>Life Form</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------------------</td>
<td>----------------</td>
<td>--------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Trypethelium virens</td>
<td>Trypetheliaceae</td>
<td>Trypetheliales</td>
<td>Dothideomycetes</td>
<td>crustose</td>
</tr>
<tr>
<td>Umbilicaria muehlenbergii</td>
<td>Umbilicariaceae</td>
<td>Umbilicariales</td>
<td>Lecanoromycetes</td>
<td>foliose</td>
</tr>
<tr>
<td>Verrucaria fuscella</td>
<td>Verrucariaceae</td>
<td>Verrucariales</td>
<td>Chaetothyriomycetes</td>
<td>crustose</td>
</tr>
<tr>
<td>Xanthoparmelia darrowii</td>
<td>Parmeliaceae</td>
<td>Lecanorales</td>
<td>Lecanoromycetes</td>
<td>foliose</td>
</tr>
</tbody>
</table>

**Response to Aging**

Before analyzing the relationship between reflectance and specimen age, I restricted the age range to between 0 and 60 years old to prevent poor estimates driven by the lack of samples beyond 60 years old (Fig. B.1). I also chose to use the average reflectance per individual to remove clustering at the level of the individual. In addition, I multiplied reflectance values by 100 to convert reflectance values from proportions to percentages to increase interpretability of model outputs. I used linear mixed-effects models, implemented in the lme4 R packages (Bates et al., 2015), to determine the effect of thallus age on the reflected spectrum. Linear mixed-effects models are a suitable choice for the unbalanced and taxonomically clustered lichen spectral dataset because the data fit the model assumptions, and linear mixed-effects models are robust to imbalanced datasets and they can model non-independent (clustered) samples (Schielzeth et al., 2020).

First, I ran three linear mixed-effects models on the reflectance values versus lichen age for each individual wavelength in the spectrum, and I treated species as a grouping factor (random effect) in each model (Table 2.2). The first model was an intercept only model in which the intercept (the expected mean reflectance when age is equal to zero) was allowed to vary...
between species. The second model was a fixed slope - variable intercept (fixed slope hereafter) model in which reflectance was a function of specimen age and intercepts were allowed to vary between species. The third model was a variable slope - variable intercept (variable slope hereafter) model that was the same as the second model (fixed slope) except that both the slope (relationship between reflectance and age) and the intercept were allowed to vary between species.

I used model comparison, using the Akaike information criterion (AIC), for each wavelength to determine which model was the best fit. Accepting the intercept only model would indicate that there was no effect of age on spectral reflectance. The fixed slope model would indicate that different species have different starting reflectance values but reflectance for each wavelength changes consistently between species. The variable slope model would suggest that different species have different starting reflectance values and those values change at different rates for different species. I used a delta AIC of two or more as the standard to decide if one model was a better fit to the data than another model. I calculated Marginal r-squared, the variation explained by the fixed effects, and the conditional r-squared, the variation explained by both the fixed and random effects, using the partR2 R package for the model that was best for most wavelengths (Nakagawa & Schielzeth, 2013; Stoffel et al., 2020). I then partitioned the random effects variance, the variance not explained by the fixed effects, by dividing the variance of a particular random effect, intercept for example, by the sum of the random effects variance.
I also used linear mixed-effects models to determine if there was a relationship between overall brightness, or the magnitude of the reflectance values, of the spectrum and specimen age. I used vector magnitude, the length of a vector (a spectrum) calculated as the square root of the sum of all squared elements (reflectance values), as a single value that summarizes the overall brightness of each spectrum. Larger magnitudes correspond to spectra with higher reflectance values. I then created intercept only, fixed slope, and variable slope linear mixed-effects models where vector magnitude was a function of age, and I used AIC to compare the models. I also wanted to know if brightness in a particular region of the spectrum varied with age, so I used the same procedure to evaluate the relationship between brightness and age for subsections of the spectrum which included the VIS, NIR, SWIR and a combination of the NIR and SWIR for wavelengths 700 to 1900 nm.

**Classification**

I used partial least squares discriminatory analysis (PLS-DA) to classify spectra to taxonomic ranks. Partial least squares discriminatory analysis is a multivariate analysis that classifies samples from a regression on underlying indicator variables (Chevallier et al., 2006), and it has been shown to work well with the high dimensional datasets such as full-range spectra (Barker & Rayens, 2003; Chevallier et al., 2006; Gavaghan et al., 2002). I conducted eight
I conducted two separate analyses to determine the influence of age on classification accuracy. The first four of the analyses were conducted to classify spectra to species, family, order, and class, respectively. The second four analyses classified spectra to species, family, order, and class, and specimen age was included as a covariate in each respective model. I did not classify specimens to genus because genera covaried with species for all but two genera. I treated scans as independent samples due to the large amount of spectral variation at the scale of the individual, and I used the full 124 year age range for these analyses.

I implemented a two-step procedure for each of the eight classifications. For the first step, I ran 100 iterations of PLS-DA in which the data was downsampled. For each iteration, the data was parsed such that a random 70% was used to train a PLS-DA model and the remaining 30% was used to test it. The model for each iteration was created using tenfold cross-validation repeated three times. Finally, the optimal number of components to use in the second step was chosen by selecting the number of components for which the average accuracy of the 100 iterations was maximized.

The second step was the same as the first step except that an upsampling procedure was implemented instead of the downsampling procedure, and the number of components used was equal to the optimal number of components chosen in the first step. The overall classification accuracy was calculated as the mean accuracy extracted from confusion matrices across all the 100 iterations used with the upsampling procedure.

I then conducted two-sample t-tests, using the R base package, to compare the accuracy between the model that did not include age and the model that did include age for each of the four taxonomic ranks (R Core Team, 2020). I conducted the t-tests using the classification accuracy values produced from 100 iterations of the respective models.
Results

Response to Aging

The fixed slope models fit the data best for all wavelengths in the NIR and about half of the SWIR (Fig. 2.1). The intercept only model was a better or equal fit than the fixed slope model for the VIS region and for SWIR wavelengths above about 1900 nm. The variable slope model was worse or equal to the fixed slope model for all wavelengths except for VIS wavelengths between about 400 and 450 nm, but the variable slope model still performed worse than the intercept only model at these wavelengths. These results indicate that reflectance in the NIR and part of the SWIR changes at about the same rate for each species.
Figure 2.1 Delta AIC for all wavelengths. Delta AIC was calculated as the difference between the AIC of the fixed slope model and the intercept only model (blue) and the difference between the fixed slope model and the variable slope model (red). Negative ΔAIC indicates that the fixed slope model was a better fit than the other models. The gray rectangle corresponds to an absolute ΔAIC of 2 or less which indicates that the fixed slope model is not statistically better or worse than the model being compared to.

The results of the fixed slope model indicated that reflectance generally decreased with age (Fig 2.2a) and, on average, it explained over 58% of the variation for all wavelengths and up to 81% of the variation for wavelengths in the NIR (Fig 2.2c). The effect of age, however, explained less than 2.5% of the variation for all wavelengths. Intercept estimates varied greatly between species (Fig. 2.2b), and the differences between species accounted for 58 to 80% of the variation not explained by the fixed effects (Figs. 2.2d,e).
**Figure 2.2** Results of the fixed slope model for every wavelength. a) Fixed effect of age. The shaded area indicates the 95% confidence interval. b) Fixed intercept compared to the intercepts for each species. The shaded blue area indicates the 95% confidence interval for the fixed intercept estimate. c) Conditional (blue) and marginal (red) r-squared values. The shaded areas correspond to 95% confidence intervals. d) Random effects variance accounted for by different species.
intercepts for each species (blue) and residual variance. e) Proportion of random effects variance explained by different intercepts for each species.

The intercept only model was the best fit to the data when evaluating the relationship between brightness and age for the VIS region (Table 2.2). The fixed slope model had the lowest AIC for the full spectrum and the NIR and SWIR, but the difference between the intercept only AIC and fixed slope AIC was not greater than 2 for these regions. Thus, by parsimony, the intercept only model should be accepted. However, the fixed slope model was the best fit when I compared models for a combination of NIR and SWIR wavelengths (700 to 1900 nm; Fig. B.2). This model had a fixed effect of age of -0.042 magnitude per year, and it had a conditional r-squared of 0.733 (95% CI [0.611, 0.778]) and a marginal r-squared of 0.015 (95% CI [0.005, 0.053]. These results further show that the reflectance in the VIS and tail end of the SWIR (>1900 nm) is not strongly tied to specimen age while the reflectance of wavelengths between 700 and 1900 nm generally decreases with age.
Table 2.3 Brightness vs age model comparison. AIC and conditional r-squared (Cond. R2 in the table) for each model type per region of the spectrum for the relationship between vector magnitude and age. The intercept only model does not have any explanatory variables so there are no associated conditional r-square values.

<table>
<thead>
<tr>
<th>Spectrum</th>
<th>Intercept only</th>
<th>Fixed slope</th>
<th>Variable slope</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AIC</td>
<td>Cond. R2</td>
<td>AIC</td>
</tr>
<tr>
<td></td>
<td>(95% CI)</td>
<td>(95% CI)</td>
<td>(95% CI)</td>
</tr>
<tr>
<td>VIS</td>
<td>639.90 **</td>
<td>646.84</td>
<td>650.83</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>0.71</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.65 - 0.81)</td>
<td>(0.62 - 0.78)</td>
</tr>
<tr>
<td>NIR</td>
<td>908.87</td>
<td>907.46 *</td>
<td>911.42</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>0.78</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.66 - 0.87)</td>
<td>(0.68 - 0.87)</td>
</tr>
<tr>
<td>SWIR</td>
<td>1019.65 *</td>
<td>1019.93</td>
<td>1023.93</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>0.66</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.52 - 0.74)</td>
<td>(0.44 - 0.73)</td>
</tr>
<tr>
<td>Full Spectrum</td>
<td>1196.97</td>
<td>1194.98 *</td>
<td>1198.97</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>0.72</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.68 - 0.85)</td>
<td>(0.62 - 0.81)</td>
</tr>
<tr>
<td>700-1900nm</td>
<td>1143.35</td>
<td>1140.54 **</td>
<td>1144.53</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>0.73</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.58 - 0.80)</td>
<td>(0.64 - 0.81)</td>
</tr>
</tbody>
</table>

* Lowest AIC per region of the spectrum
** Lowest AIC by more than 2 AIC

Classification

Average overall classification accuracy ranged from 77.0 ±0.02% (species) to 94.5 ±0.01% (family; the ‘±’ indicates one standard deviation) when classifying each of the taxonomic ranks (species, family, order, class) without adding age as a covariate (Table 2.3). Adding age as covariate in the models did significantly (p < 0.05) improve classification accuracy when classifying species, order, and class; however, it did not significantly improve
accuracy when classifying family (Table 2.3). Further, the age covariate had the lowest variable import ance value - the contribution of each model coefficient weighted proportionally to its reduction in the sum of the squares of variation - for classifying any taxonomic unit in any model when compared to the variable importance values for the individual wavelengths in the spectrum (Figs B.3, B.4, B.5, and B.6; Kuhn, 2021). Therefore, the age of a specimen does inform a PLS-DA model (Table 2.3); however, the gain in classification accuracy is marginal (Table 2.3) with effect sizes of 3.5%, 0.5%, and 1.4% for classifying species, orders, and classes, respectively. This suggests that lichen specimens do age or decompose over time within an herbarium setting, but an older specimen resembles a younger specimen much more than it resembles a different species of any age.

**Table 2.4 PLS-DA results.** Accuracy indicates average overall model accuracy after 100 iterations of PLS-DA. SD indicates one standard deviation.

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
<th>SD</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Species</td>
<td>0.770</td>
<td>0.020</td>
<td>26</td>
</tr>
<tr>
<td>Species + age</td>
<td>0.805</td>
<td>0.019</td>
<td>32</td>
</tr>
<tr>
<td>Family</td>
<td>0.945</td>
<td>0.011</td>
<td>60</td>
</tr>
<tr>
<td>Family + age</td>
<td>0.944</td>
<td>0.013</td>
<td>60</td>
</tr>
<tr>
<td>Order</td>
<td>0.890</td>
<td>0.013</td>
<td>53</td>
</tr>
<tr>
<td>Order + age</td>
<td>0.895</td>
<td>0.014</td>
<td>57</td>
</tr>
<tr>
<td>Class</td>
<td>0.825</td>
<td>0.021</td>
<td>24</td>
</tr>
<tr>
<td>Class + age</td>
<td>0.839</td>
<td>0.020</td>
<td>28</td>
</tr>
</tbody>
</table>
Table 2.5 T-test: Comparing model accuracy with and without age.

Difference of means indicates the difference between the mean accuracy of the model that does not include age as a covariate and the model that does include age as a covariate.

<table>
<thead>
<tr>
<th>Taxonomic rank</th>
<th>Difference of means (%)</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Species</td>
<td>-3.54</td>
<td>-13.9</td>
<td>2.2E-16 *</td>
</tr>
<tr>
<td>Family</td>
<td>0.07</td>
<td>0.43</td>
<td>0.668</td>
</tr>
<tr>
<td>Order</td>
<td>-0.49</td>
<td>-2.59</td>
<td>0.010 *</td>
</tr>
<tr>
<td>Class</td>
<td>-1.40</td>
<td>-4.77</td>
<td>3.52E-06 *</td>
</tr>
</tbody>
</table>

* p < 0.05

Classifying taxonomic units without including specimen age was generally successful. Species were classified with an overall accuracy of 77.0 ±2.0%. Eighteen of the species (60% of species) were classified with over 80% accuracy, 8 species (27% of species) were classified with 50 to 79% accuracy, and only 4 species (13% of species) were classified with less than 50% accuracy (Fig. 2.3). *Flavoparmelia haysomii* was poorly classified with only 10.5 ±10.5% accuracy, and it was misclassified as the congeneric *F. euplecta* 47.3 ±21.0% of the time. *Acorospora americana* was the most poorly classified species (2.1 ±4.7% accuracy).

Approximately 52% of *Acorospora americana* misclassifications were attributed to species within the same class, Lecanoromycetes, but about 46% of the misclassification were attributed to the much more distantly related species *Ephebe ocellata* and *Verrucaria fuscella.*
Figure 2.3 Confusion matrix for classifying species without age as a covariate. The rows correspond to the reference and columns correspond to the prediction. Correct classifications are shown by squares in the diagonal, and misclassifications are shown in the off diagonal. Larger and darker orange squares indicate higher proportions of classifications belonging to the respective cell. Numbers correspond to the proportion of classifications per species.
The classification of lichen families had the greatest overall classification accuracy (94.5 ±1.1%; Fig. B.7). Only four families (21% of families) were classified with less than 90% accuracy, and three of which were classified with over 80% accuracy. Acarosporaceae, the family that contains *Acarospora americana*, was classified with only 52.8 ±16.1% accuracy. Classifying lichen families may have been more accurate than classifying species because the variation within the 13 species in *Parmeliaceae* is captured better by a single unit than 13 separate units.

Lichen orders were classified with an overall accuracy of 89.0 ±1.3%, and all but 3 orders were classified with over 90% accuracy (Fig. B.8). Specimens belonging to Teloschistales were classified with 84.4 ±6.9% accuracy and all misclassifications were attributed to specimens in the same class, Lecanormycetes. The order Lecanorales was classified with 77.1 ±3.2% and most of the misclassifications were attributed to other orders in the same class, Lecanoromycetes. Only approximately 2.5% of Lecanorales classification were attributed to lichens belonging in different classes. Acarosporales was the least accurately classified order at 62.1 ±15.4% accuracy, and only about 47.5% of misclassifications were attributed to lichens within the same class (Lecanoromycetes).

Finally, five of the six classes of lichens were classified with over 95% overall accuracy (Fig. B.9). However, Lecanoromycetes, the most well represented class in this study, was classified with 77.2 ±2.9% accuracy. This class may have the lowest classification accuracy due to its large variation that may overlap with the variation of other lichen classes.
**Discussion**

Biodiversity assessment via the use of spectroscopy has largely focused on vascular plants (Cavender-Bares et al., 2016; Deacon et al., 2017; Jetz et al., 2016; Schweiger et al., 2018) despite the ecological importance of other terrestrial autotrophs such as lichens (Asplund & Wardle, 2017; Bokhorst et al., 2015; Zedda et al., 2014). The gap between plant and lichen spectral libraries may be narrowed by using widely available herbarium specimens to build spectral libraries that are useful both in and out of the herbarium setting. However, it has been unclear whether lichen specimens change as they age. This aging process could hinder or complicate the use of such a spectral library for various applications such as classification and estimating historical functional diversity from herbarium specimens (Heberling et al., 2019; Kothari et al., 2021; Meineke et al., 2018). As such, my goals were to elucidate the relationship between spectral reflectance and specimen age and to determine if specimen age influenced the results of classification models. Here, I demonstrated that lichens did change as they aged, but this change had only a marginal influence on the classification of herbarium specimens.

**Response to Aging**

The raw spectra showed a general decrease in reflectance (Figs. B.10 and B.11), and this was corroborated by the results of the linear mixed-effects models. However, my expectation that reflectance would change at different rates for different species was not supported. This suggests that photosynthetic and structural compounds that make up a lichen thallus decompose or change at the same rates irrespective of species identity. The differences between species, then, are driven by differences in morphology and chemistry that the lichens had to begin with.

Not all regions of the spectrum had a clear response to age. The VIS did not show significant change with age demonstrated by the superior AIC of the intercept only linear
mixed-effects model for this region (Fig. 2.1) and by the 95% confidence interval and near-zero slope estimates of the fixed slope linear mixed-effects models (Fig. 2.2a). Therefore, my expectation that this region would change in response to changes in chlorophyll content did not hold (Ustin et al., 2009a). However, there are at least three possibilities that explain the lack of change with age in the VIS. First, chlorophyll content may decrease quickly within the first several days of dessication, as has been shown for *Lobaria pulmonaria*, and *Peltigera polydactyla* (Kranner et al., 2003). The current study only involves lichens that have been desiccated for several years, so it could be the case chlorophyll content degrades to a minimal amount, potentially due to reactive oxygen species (Smirnoff, 1993), well before time window investigated here. Alternatively, chlorophyll content may not significantly change in more desiccation-tolerant species such as *Pseudevernia furfuracea* (Kranner et al., 2003). Second, the expected increase in blue (400 – 500 nm) and red (600 – 700 nm) VIS wavelengths could be complicated by changes in thallus structure as is implicated by the changes in the NIR and SWIR (Curran, 1989). The physical changes to thallus structure may modify the way light interacts (scatters, absorbs, refracts, and reflects) with photosynthetic pigments in the photobiont layer of lichen thallus (Vogelmann, 1993; Vogelmann et al., 1996). Third, there are over 700 known secondary metabolites produced by lichens, much of which are deposited in the lichen cortex and correspond to lichen color, such as usnic acid (Huneck, 1999; Nash, 2010). As such, it may be the case that these pigments mask photosynthetic pigments or that the decomposition of some of these secondary metabolites counteracts the expected changes in reflectance from chlorophyll decomposition.
Unlike the VIS, reflectance for wavelengths in the NIR and SWIR, at least between 700 and 1900 nm, appeared to decrease with age (Table 2.3); however, this decrease was generally less than 0.2% reflectance per year (Fig. 2.2a). Further, my expectation that different species would show different rates of reflectance change with age was not supported considering the fixed slope model was the best fit for this spectral region. Since the NIR and SWIR capture structural characteristics and components, such as thallus mass per area, proteins, polysaccharides, sugars, and nitrogen (Curran, 1989; Thenkabail & Lyon, 2016), these results indicate that the structure of a lichen generally changes at about the same rate irrespective of the species (Fig. 2.2a), but different species have different initial structural conditions (Fig. 2.2b,d,e). This suggests that the differences between lichens should persist as lichens age. The general decrease in reflectance for this region is also in agreement with findings that show reflectance across the spectrum to decrease for decomposing monocots (Proctor et al., 2017) and pine needles (Goward et al., 1994).

Although the changes in the NIR and SWIR indicate a structural response to age, the exact structural changes are unclear. The cytoplasm in mycobiont cells is known to form a glassy state from non-reducing sugars, a process called vitrification, that protects proteins and membranes during desiccation (Kranner et al., 2008). However, vitrification does not necessarily prevent protein denaturing and membrane damage indefinitely (Crowe et al., 1998), especially for the cortical cells that are directly exposed to oxidation and hydrolysis from the atmosphere (Molak & Ho, 2011). I speculate that lichen cells do decompose over time, but vitrification does contribute to the rather low decreases in reflectance with specimen age (less than 0.2% reflectance per year). Further, lichens may be slowly decomposed by bacterial and fungal microbes that have been found to contaminate herbarium specimens (Bieker et al., 2020).
Although their damage to lichen thalli is not extensive, it is not unreasonable to presume that herbarium specimens act as suitable environments for microbial communities much like their living counterparts (Pankratov et al., 2017). Lastly, Scheidegger et al. (1995) demonstrated that both mycobiont hyphae and algal cells cavitate during desiccation. It could be the case that lichens stored in herbaria undergo continuous cavitation at the cellular level as the desiccated state persists; however, cavitation rates may be at or near zero once the humidity in the interior of the lichen thallus is at equilibrium with that of the surrounding herbarium environment. Regardless of the mechanism that underlies the changes in reflectance, it seems that more light between 700 and 1900 nm is absorbed by a lichen thallus as it ages (Gates et al., 1965).

Interestingly, however, reflectance beyond 1900 nm seems to be rather stable, as is indicated by the linear mixed-effects model comparisons (Fig. 2.1). Since water absorbs light in the SWIR (Carter, 1991; Thenkabail & Lyon, 2016), the lack of a clear decrease in reflectance beyond 1900 nm may be attributed to a consistent thallus water content.

**Classification**

The classifications of lichen species, orders, and classes were improved by including specimen age as a covariate within the models; however, the increase in accuracy was marginal (Tables 2.3 and 2.4). Further, the classification of lichen families was not improved by including specimen age, and age was the least important variable for classifying any member of any taxonomic group (Figs. B.3, B.4, B.5, B.6). This indicates the variation driven by the aging process was overshadowed by the variation between species.

The classification of species without including age was comparable to previous classifications of plants and lichens using reflectance spectra as explanatory variables. The overall classification accuracy of 77% was lower than that reported by Stasinski et al. (2021) and
by Durgante et al. (2013) who classified species of *Dryas* and Amazonian tree species, respectively, with over 90% accuracy. However, lichen species were classified with greater accuracy than that presented by Cavender-Bares et al. (2016) who classified 28 species of *Quercus*. It is important to note that *Quercus* species were more closely related, thus more phenotypically similar, than the lichen species presented here, so differences in classification accuracies may be driven by the available variability within each study system.

The classification of lichen families without age (94.5% accuracy) was comparable to the results presented by Blanco-Sacristán et al. (2019) who used visible-near infrared spectra to classify lichen genera and moss coexisting within biocrusts with 97.8% overall accuracy. Furthermore, the classifications of lichen families and orders presented here (Table 2.3) were more accurate than classifying plant families (67.6%) and orders (66.4%) presented by Meireles et al. (2020); however, the differences between classification accuracies may be driven by the great disparity between the sizes of spectral datasets (1,209 spectra for 30 lichen species versus 16,000 spectra for 544 plant species).

I used a spectral dataset with a large phylogenetic breadth in order to make generalizations about how lichen respond to long-term desiccation in herbarium storage. Overall, spectral reflectance indicates that lichen specimens appear to change over time; however, this change seems to be most evident for wavelengths between 700 and 1900 nm, and it does not strongly influence the accuracy of classification models. These are promising results for employing herbarium specimens to build spectral libraries for applications within herbarium settings and potentially for contemporary biodiversity assessments (Price, 1994; Zomer et al., 2009). Going forward, it would be beneficial to determine the relationship between spectra, lichen functional traits, and the aging process. Models built from such information could be used
to study contemporary and historical functional diversity which would give us further insight into the ecological function of lichens.


APPENDIX A: DRYAS

Figure A.1 Proportion of ancestry from *Dryas alaskensis*. a) Ancestry by species. b) Ancestry for hybrids by site. In the box plots, the bold lines indicate the median, the top and bottom edges of the boxes indicate the first and third quartiles, the lines at the ends of the whiskers indicate the minimum (quartile 1 - 1.5 * interquartile range (IQR)) and maximum (quartile 3 + 1.5* IQR), and the dots indicate outliers.
Figure A.2 STRUCTURE plots delimiting *D. ajaneneis* (DAJ) and *D. alaskensis* (DAK). a) DAK proportional genome assignments to two (*K*=2) or three (*K*=3) ancestral populations. Plots of posterior probability (red lines) and Evanno’s Delta *K* (blue lines) indicate the most information content at *K*=3 for DAK (b) and *K*=5 for DO (c). d) DAJ proportional genome assignments to four (*K*=4) and five (*K*=5) ancestral populations.
**Figure A.3** Variable importance for classifying species. Red corresponds to *D. ajanensis*, blue corresponds to *D. alaskensis*, and black corresponds to hybrids. Variable importance is the contribution of each coefficient (wavelength) weighted proportionally to its reduction in the sum of squares. The y-axis shows the relative contribution of each wavelength scaled from 0-100. The dark lines indicate the mean variable importance and the shaded regions indicate the 95% confidence intervals.
Figure A.4 Classification matrix for collection site. The diagonal represents correct classifications. The size and shade of the orange square in each cell corresponds to the proportion of classifications within each row. White cells represent true zeros and zeros indicate proportions less than 0.001.

<table>
<thead>
<tr>
<th></th>
<th>BG</th>
<th>ES</th>
<th>MD</th>
<th>TM</th>
<th>WDA</th>
<th>WDB</th>
</tr>
</thead>
<tbody>
<tr>
<td>BG</td>
<td>0.999</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ES</td>
<td></td>
<td>0.995</td>
<td></td>
<td>0</td>
<td>0.005</td>
<td>0</td>
</tr>
<tr>
<td>MD</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TM</td>
<td>0.001</td>
<td></td>
<td></td>
<td>0.999</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WDA</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>WDB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.998</td>
</tr>
</tbody>
</table>
**Figure A.5** Classification matrix for species and site without hybrids. The diagonal represents correct classifications. The size and shade of the orange square in each cell corresponds to the proportion of classifications within each row. White cells represent true zeros and zeros indicate proportions less than 0.01.
Figure A.6 Classification of site with the visible spectrum and water absorption features removed from the reflectance spectra. The diagonal represents correct classifications. The size and shade of the orange square in each cell corresponds to the proportion of classifications within each row. White cells represent true zeros and zeros indicate proportions less than 0.01.
Figure A.7 Variable importance for classifying site of collection. Variable importance is the contribution of each coefficient (wavelength) weighted proportionally to its reduction in the sum of squares. The y-axis shows the relative contribution of each wavelength scaled from 0-100. The dashed vertical lines indicate the primary water absorption features at 1450nm and 1940 nm (Carter, 1991). The dark lines indicate the mean variable importance and the shaded regions indicate the 95% confidence intervals.
Figure A.8 Phylogenetic signal by 1nm wavelength. a) Blomberg’s K per individual 1 nm wavelength. The dashed gray line indicates the maximum K estimated from a tree with no phylogenetic covariance. b) p-values for the K estimates per wavelength.
**Figure A.9.** Leaf morphology by species. a) Leaf length in millimeters. b) Presence of abaxial midvein glandular trichomes, feathery scales, both, or neither. c) Adaxial tomentum ranging from sparse to dense. For the box plot (a), the bold lines indicate the median, the top and bottom edges of the boxes indicate the first and third quartiles, the ends of the whiskers indicate the minimum (quartile 1 - 1.5 * interquartile range (IQR)) and maximum (quartile 3 + 1.5* IQR), and the dots indicate outliers.
Table A.1 Distances between sampling sites (km)

<table>
<thead>
<tr>
<th></th>
<th>BG</th>
<th>ES</th>
<th>MD</th>
<th>TM</th>
<th>WDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>ES</td>
<td>252</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MD</td>
<td>130</td>
<td>153</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TM</td>
<td>228</td>
<td>28</td>
<td>126</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>WDA</td>
<td>160</td>
<td>128</td>
<td>32</td>
<td>100</td>
<td>-</td>
</tr>
<tr>
<td>WDB</td>
<td>163</td>
<td>127</td>
<td>35</td>
<td>100</td>
<td>3</td>
</tr>
</tbody>
</table>

Abbreviations: BG = Bison Gulch, ES = Eagle Summit, MD = Murphy Dome, TM = Twelve Mile, WDA = Wickersham Dome A, WDB = Wickersham Dome B.

Table A.2 Number of individuals and scans (in parentheses) per taxon and site

<table>
<thead>
<tr>
<th>Site</th>
<th>D. ajanensis</th>
<th>D. alaskensis</th>
<th>Hybrids</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eagle Summit</td>
<td>21 (124)</td>
<td>19 (114)</td>
<td>2 (12)</td>
<td>42 (250)</td>
</tr>
<tr>
<td>Twelve Mile</td>
<td>20 (124)</td>
<td>22 (124)</td>
<td>2 (12)</td>
<td>44 (260)</td>
</tr>
<tr>
<td>Wickersham Dome B</td>
<td>16 (79)</td>
<td>11 (66)</td>
<td>5 (30)</td>
<td>32 (175)</td>
</tr>
<tr>
<td>Bison Gulch</td>
<td>20 (120)</td>
<td>-</td>
<td>-</td>
<td>20 (120)</td>
</tr>
</tbody>
</table>
Table A.2 Continued.

<table>
<thead>
<tr>
<th>Classification Unit</th>
<th>Murphy Dome</th>
<th>Wickersham Dome A</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20 (120)</td>
<td>-</td>
<td>20 (120)</td>
</tr>
<tr>
<td></td>
<td>20 (120)</td>
<td>-</td>
<td>20 (120)</td>
</tr>
<tr>
<td>Total</td>
<td>117 (687)</td>
<td>51 (304)</td>
<td>9 (54)</td>
</tr>
</tbody>
</table>

Table A.3 Accuracy and number of components for each classification unit

<table>
<thead>
<tr>
<th>Classification Unit</th>
<th>Accuracy (%)*</th>
<th>Number of Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Species with hybrids</td>
<td>92.85 ±1.83</td>
<td>27</td>
</tr>
<tr>
<td>Species without hybrids</td>
<td>99.65 ±0.44</td>
<td>33</td>
</tr>
<tr>
<td>Populations</td>
<td>98.91 ±0.68</td>
<td>45</td>
</tr>
<tr>
<td>Site</td>
<td>99.80 ±0.31</td>
<td>59</td>
</tr>
<tr>
<td>Species + site</td>
<td>92.05±1.25</td>
<td>30</td>
</tr>
<tr>
<td>Species + site without hybrids</td>
<td>98.92 ±0.82</td>
<td>55</td>
</tr>
</tbody>
</table>

*The ‘±’ symbol indicates one standard deviation.
Methods A.1 DNA sequencing and Genetic Analysis.

For DNA isolation, we first homogenized 30–40 mg of silica-dried leaves with a single 3.2 mm chrome-coated steel bead for 1 minute at 25 Hz in a TissueLyser® II (QIAGEN, Hilden, Germany). We then extracted with the Invisorb® Spin Plant Mini Kit (STRATEC Molecular, Birkenfeld, Germany) with the following modifications: 500 µl Lysis Buffer P, 25 µl Proteinase S centrifuged at 11,000 g for two minutes, 250 µl Binding Buffer P, 20 µl RNase, and eluting in 55 µl prewarmed (50º C) nuclease-free water with 10 mM Tris-HCl (pH 8.0), incubating elution buffer for 3 minutes, centrifuging, and eluting again with the flow-through. Genomic DNA (350 ng – 1 ug per sample) was plated onto one of two 96-well plates and sent to the University of Wisconsin Biotechnology Center DNA Sequencing Facility for genotyping-by-sequencing library preparation and sequencing (GBS; ApeKI endonuclease and size selected for 200–500 bp fragments). Each plate of 95 samples was pooled and paired-end sequenced (150 bp) on the NovaSeq 6000 (Illumina Inc., San Diego, CA, USA).

To generate our genetic dataset, we demultiplexed (perfect match to the barcode sequence), trimmed and filtered (removed reads less than 35 bp long or with 10 or more bp having a q-score <20; strict adapter trimming), and mapped our reads to the D. drummondii genome (GCA_003254865.1) using the ipyrad analysis toolkit (Eaton & Overcast, 2020). We removed GBS loci with >10% polymorphic sites (SNPs), more than five indels, or if they were missing any samples, and, finally, removed all multiallelic SNPs. Since Dryas are long-lived, mat-forming shrubs that can be connected underground, we identified clones as individuals with over 99% pairwise similarity and the individual with more missing data was dropped (vcf_clone_detect.py; github.com/pimbongaerts). Lastly, we removed SNPs with a minor allele
count less than two and then removed all SNPs violating Hardy-Weinberg equilibrium at the species level using VCFtools (Danecek et al., 2011).

Methods A.2 Determining the effect of differing drying times on our classification accuracy.

We suspected that measuring the reflectance of our specimens after differing periods of drying (36–60 hours) may have increased the differences between plants collected at different sites (plants collected from the same site were measured at relatively the same time), thus increasing the accuracy of our models when classifying collection sites and populations. We estimated the influence of water from the reflectance at 1450 nm and 1940 nm, which are known water absorption features (Carter, 1991), the Normalized Difference Water Index (NDWI; (Gao, 1996), and variable importance values from PLS-DA for sites. We compared samples collected from different sites using pairwise t-tests with these variables. *Dryas ajanensis* and *D. alaskensis* samples were compared separately.

The pairwise t-tests of the water absorption features and NDWI, compared between the two species, only agreed that there was a significant difference between the water content of leaves collected from Eagle Summit and Wickersham Dome B. However, 2.8% of Wickersham Dome B *D. ajanensis* scans were still misclassified as Eagle Summit *D. ajanensis* in our PLS-DA (fig. 4) which indicates that factors other than water content may have influenced our classification accuracy. Furthermore, when we only accounted for differences between *D. ajanensis* (excluding *D. alaskensis*), our pairwise t-tests of water absorption features and NDWI only agreed on differences between Murphy Dome and Wickersham Dome B and between Wickersham Dome A and Wickersham Dome B. Despite differing water content, 2.0% and 2.4% Wickersham Dome B *D. ajanensis* scans were misclassified as belonging to Murphy Dome and
Wickersham Dome A, respectively (fig. 4). These observations led us to conclude that differing
drying times did not strongly influence accuracy of our classification models.

Finally, the variable importance values output by the PLS-DA for classifying sites did not
show strong peaks at or near the primary water absorption features (1450 nm and 1940 nm) that
were substantially different from peaks in other regions of the spectra. There was a peak near
1940 nm for classifying Twelve Mile specimens (Fig. A.5); however, we cannot conclude that
this peak indicates an importance of water content or a morphological feature that is more
apparent when water is removed.
Figure B.1. Histogram of specimen age.
Figure B.2 Relationship between vector magnitude and age for wavelengths 700 to 1900 nm. Different colors correspond to different species. The black line corresponds to the fixed effect estimate.
Figure B.3 The most and least important variables for predicting species. The green bars correspond to the five most important variables and brown bars correspond to the five least important variables for predicting species. Numbers on the y-axis correspond to individual wavelengths and age indicates the age covariate. The x-axis indicates the variable importance values.
Figure B.4 The most and least important variables for predicting family. The green bars correspond to the five most important variables and brown bars correspond to the five least important variables for predicting family. Numbers on the y-axis correspond to individual wavelengths and age indicates the age covariate. The x-axis indicates the variable importance values.
Figure B.5 The most and least important variables for predicting order. The green bars correspond to the five most important variables and brown bars correspond to the five least important variables for predicting order. Numbers on the y-axis correspond to individual wavelengths and age indicates the age covariate. The x-axis indicates the variable importance values.
Figure B.6 The most and least important variables for predicting class. The green bars correspond to the five most important variables and brown bars correspond to the five least important variables for predicting class. Numbers on the y-axis correspond to individual wavelengths and age indicates the age covariate. The x-axis indicates the variable importance values.
**Figure B.7** Confusion matrix for the classification of family without age as a covariate. The rows correspond to the reference and columns correspond to the prediction. Correct classifications are shown by squares in the diagonal, and misclassifications are shown in the off diagonal. Larger and darker orange squares indicate higher proportions of classifications belonging to the respective cell. Zeros correspond to values less than 0.005 and white spaces indicate true zeros.
**Figure B.8** Confusion matrix for the classification of order without age as a covariate. The rows correspond to the reference and columns correspond to the prediction. Correct classifications are shown by squares in the diagonal, and misclassifications are shown in the off-diagonal. Larger and darker orange squares indicate higher proportions of classifications belonging to the respective cell. Zeros correspond to values less than 0.005 and white spaces indicate true zeros.
**Figure B.9** Confusion matrix for the classification of class without age as a covariate. The rows correspond to the reference and columns correspond to the prediction. Correct classifications are shown by squares in the diagonal, and misclassifications are shown in the off diagonal. Larger and darker orange squares indicate higher proportions of classifications belonging to the respective cell. Zeros correspond to values less than 0.005 and white spaces indicate true zeros.

<table>
<thead>
<tr>
<th></th>
<th>Arthoniomyces</th>
<th>Candelariomyces</th>
<th>Chaetothyriomyces</th>
<th>Dothideomycetes</th>
<th>Lecanoromycetes</th>
<th>Lichinomycetes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arthoniomyces</td>
<td>0.96</td>
<td>0</td>
<td>0.03</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Candelariomyces</td>
<td>0</td>
<td>0.97</td>
<td>0</td>
<td>0</td>
<td>0.02</td>
<td>0</td>
</tr>
<tr>
<td>Chaetothyriomyces</td>
<td></td>
<td>0.96</td>
<td>0</td>
<td>0</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Dothideomycetes</td>
<td>0</td>
<td>0</td>
<td>0.99</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Lecanoromycetes</td>
<td>0</td>
<td>0.03</td>
<td>0.12</td>
<td>0.05</td>
<td>0.77</td>
<td>0.03</td>
</tr>
<tr>
<td>Lichinomycetes</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.99</td>
</tr>
</tbody>
</table>
Figure B.10 Mean reflectance for all individuals within each age cohort. Dark green indicates specimens younger than 15 years old, light green indicates specimens between 16 and 30 years old, light brown indicates specimens between 31 and 45 years old, and dark brown indicates specimens 46 to 60 years old.
Figure B.11 Mean spectra per individual per age cohort. Dark green indicates specimens younger than 15 years old, light green indicates specimens between 16 and 30 years old, light brown indicates specimens between 31 and 45 years old, and dark brown indicates specimens 46 to 60 years old.
Figure B.11 Continued.
Figure B.11 Continued.
Figure B.11 Continued.
Figure B.11 Continued.
BIOGRAPHY OF THE AUTHOR

Lance Stasinski was born in Denver, Colorado on December 18, 1995. He was raised in Greeley, Colorado and graduated from Frontier Academy in 2014. He attended Chadron State College in Chadron, Nebraska and graduated in 2019 with a Bachelor’s degree in Science. He moved to Maine and entered the Biology and Ecology graduate program at the University of Maine in the fall of 2019. After receiving his degree, Lance will pursue a career in web development with a hope to build web-based platforms that assist with biological research or conservation. Lance is a candidate for the Master of Science degree in Botany and Plant Pathology from the University of Maine in December 2021.