

¹ Metagenomic Barcoding of Pollen Loads Offers Insights on the
² Foraging Patterns of Queen Bumble Bees

³ Reed Clark Benkendorf^{*1,2}, Jane E. Ogilvie ³, Emily J. Woodworth ^{1,2}, Sophie Taddeo ^{1,2}, Paul J. CaraL

⁴ **Abstract**

⁵ .
⁶ 1) DNA Barcoding has been remarkably successful in nearly all kingdoms of life and has allowed for the
⁷ rapid analysis of ecological assemblies. Successful DNA barcoding in the plant kingdom has been more
⁸ difficult than other kingdoms. Due to this understanding plants in ecological contexts and understandings
⁹ of their synecology in some instances may begin to lag behind other kingdoms. 2) Here we utilize hyb-seq,
¹⁰ museum studies, and species distribution modelling, to detect the plant species present in pollen loads
¹¹ collected from Queen Bumble Bees. 3) By utilizing Species distribution modelling we allow for one
¹² to process hyb-seq data, create user specified sequence databases which may use MORE ACCURATE
¹³ alignment algorithms on personal computers over realistic time periods. 4) We show that hyb-seq using
¹⁴ the Angiosperms 353 probes, which are currently being used in the largest ever plant systematic endeavor,
¹⁵ offers significant promise to metagenomic approaches in real world scenarios. 5) We conclude that these
¹⁶ probes offer promise for the identification of plant tissue in both single sample, and metasample contexts.

¹⁷ **1 | INTRODUCTION**

¹⁸ The inability to reliably identify plants to the level of species often leaves our understanding of ecosystem
¹⁹ function and interactions wanting. Current methods to ameliorate this situation include: ignoring these
²⁰ ecologically relevant levels of detail, revisiting plots as diagnostic material becomes temporally available,
²¹ assistance from taxonomic specialists, or the use of barcoding or other molecular techniques. These approaches
²² are untenable in light of the benefits offered by: species in several morphologically difficult genera which
²³ serve as bioindicators, preferred partners in ecological interactions, as well as an increasing lack of taxonomic
²⁴ experts (Hebert *et al.* (2003)). Many genera, especially with the formalized advent of integrative taxonomy,

*Correspondence: reedbenkendorf2021@u.northwestern.edu

25 have species which are well defined based upon ecological and behavioral rather than morphological properties,
26 the identification of these taxa in degraded areas or without their mutualistic partners is fraught with difficulty.
27 Hindering an understanding of the breadth of habitat which some species occupy, and the interactions they
28 have with other species.

29 The identification of many plant species to terminal taxon is an essential component of nearly all land
30 management programs, where many species in the same genus (e.g. Sagebrush - *Artemisia* L., Willows -
31 *Salix* L., and Sedges - *Carex* L.) serve as bioindicators (respectively for ‘rangelands’, streams, and wetlands),
32 as well as in academic research (Gage & Cooper (2013), AIM). This endeavour is often mired by lack of
33 diagnostic characters (e.g. flowers, fruits, roots or combinations thereof), and increasingly the description of
34 cryptic species (Janzen *et al.* (2017), Oliver *et al.* (2009)). Solutions to this problem are wanting, certain
35 programmes have relied increasingly upon revisiting field sites to identify material using morphological or
36 chemical approaches, whereas academic research has often used high copy number plastid genes as barcodes
37 (Rosentreter *et al.* 2021, MORE MORE). However, both approaches have significant downsides, the former
38 resource intensive at the landscape scale - and often does not work, while the latter seldom works due to a
39 lack of variability in the currently available barcodes (Liu *et al.* (2014)).

40 Recently barcoding (the identification of a sample from a single organism *e.g.* a piece of leaf), and metabarcod-
41 ing (the identification of a sample containing a mix of organisms *e.g.* soil), have shown considerable promise
42 in all Kingdoms of life (Ruppert *et al.* (2019)). With plants the identification of members of certain clades
43 using barcoding has been quite successful (REF), whereas with other clades results have been elusive (Liu *et*
44 *al.* (2014), Group *et al.* (2011), Coissac *et al.* (2012)), however metabarcoding incurs additional challenges to
45 those which exist for the currently implemented barcodes (Li *et al.* (2015), Kress & Erickson (2007), Group
46 *et al.* (2009), Coissac *et al.* (2012)). Particular challenges with barcodes include the utilization of high-copy
47 number sequences are associated with their rates of divergence, gene tree conflict, and hybridization (Coissac
48 *et al.* (2016), Fazekas *et al.* (2009)). Particular challenges with the utilization of high-copy number sequences
49 are associated with their rates of divergence, gene tree conflict, and hybridization (Coissac *et al.* (2016)).

50 Currently the largest plant systematic endeavor ever undertaken, the Kew Plant and Fungal Tree of Life
51 (PAFTOL), is approaching completion (Baker *et al.* (2021)). This dataset will contain Hyb-Seq data from
52 at least one species representing each genus in the plant kingdom using the popular A353 probes (Baker *et*
53 *al.* (2021)), resulting in over 14,000 represented species. These publicly available data serve to provide a
54 taxonomically comprehensive backbone for plant metabarcoding. Data from the 10kP project, which seeks to
55 develop reference genomes from a phylogenetically diverse suite of plants will contribute many more records
56 upon its intended completion, now slated to be by 2030, similar projects which seek to sequence high amounts

57 of genomes in regions e.g. the ‘Darwin Tree of Life’ are being undertaken which will contribute data applicable
58 to enormous spatial domains (Cheng *et al.* (2018), Life Project Consortium *et al.* (2022), Lewin *et al.* (2022)).

59 These data will promote the ability to apply metabarcoding to resolve a diversity of questions relevant to
60 theoretical and applied ecology (cite). However, the application of metabarcoding still face challenges relating
61 to the enormity of the genomic datasets and the computational power required to process sequence data.
62 Herein we have resolved major components of the problems of identifying plant material without diagnostic
63 morphological character states using the Angiosperms353 (A353) Hyb-Seq probes (Johnson *et al.* (2019)),
64 and custom species sequence databases derived via species distribution modelling, and temporal filtering.

65 To increase the quality of metabarcoding results in plants, we suggest reducing the number of possible plant
66 species candidates by generating user selected sequence databases relevant to the study region and its
67 ecological characteristics (CITE !?). To achieve this goal, we first create a list of candidate species using digital
68 collections gleaned from herbaria, survey work, and citizen science (e.g. iNaturalist), from a region exceeding
69 the study area. To these candidate species, modelling approaches - such as logistic regression, may be used
70 to identify taxa which warrant further exploration e.g. modelling to determine their possibility of presence
71 in metabarcoding samples. We then use species distribution models to create potential distribution maps
72 for the candidate species to limit the impact of spatial and taxonomic biases in the species list and account
73 for spatial variations in niche availability throughout the study area. Species distribution models (SDM’s)
74 examine the ecological conditions associated with known occurrence of a species to identify where else in the
75 study area might suitable habitats be found. This approach has the additional benefit of greatly reducing the
76 size of a sequence database, which allows for the usage of genomic size data on personal computers. This can
77 also significantly reduce processing time, particularly as most next-generation sequence data is deposited
78 as raw-sequence reads.

79 Considerable amounts of species interactions vary along time (CaraDonna *et al.* (2021)). For the tropics
80 and subtropics, contrasts in the flowering periods of many plant species, can provide an additional filter for
81 identifying material in many metagenomic samples (Janzen (1967), Newstrom *et al.* (1994)). In temperate
82 regions, pollination interactions also vary temporally (CaraDonna *et al.* (2017)), however the overall shorter
83 extent of the active growing season in these systems results in the presence of few to any natural breaks
84 in these systems which reduces the utility of these to operate as filters in the post-processing of sequence
85 matches. Nonetheless, we work through a process which seems applicable to the tropics and subtropics to
86 utilize the temporal dimension for classifying sequencing results.

87 We apply these metagenomic and informatics approaches to determine whether the foraging record of Queen
88 Bumble Bee’s is consistent across direct observations and the pollen record, an incongruity in several

89 floral visitation networks involving smaller bodied fauna (Barker & Arceo-Gomez (2021), Zhao *et al.* (2019),
90 Alarcón (2010)). The two foraging phases of the Queen Bumble Bee life cycle is essential to 1) increase
91 their weight before diapause, 2) increase their ovary weights while establishing their recently found nests,
92 both of these time periods represent potential demographic bottlenecks in bumble bee populations (Sarro *et*
93 *al.* (2022)). Bumblebees are one of the only groups of insects with unequivocal quantitative evidence for
94 numerous populations declines, while simultaneously serving as the most effective pollinators in temperate
95 montane ecosystems (Cameron & Sadd (2020), Goulson *et al.* (2008), Williams (1982), Colla *et al.* (2012),
96 Bergman *et al.* (1996), Bingham & Orthner (1998)). *Heterogeneous montane ecosystems simultaneously*
97 *represent one of the most ecologically resilient areas in the temperate and oftentimes offer the sole potential*
98 *refugia for multiple dimensions of biodiversity under climate change, whilst experiencing the most abrupt*
99 *changes in mean annual temperature (CITE). An immediate understanding of how to manage previously*
100 *overlooked keystone insect species, such as bumble bees, is essential if these refugial ecosystem services are to*
101 *be utilized while maintaining their current species richness.*

102 2 | METHODS

103 Study System & Field Work

104 Observations and sample collection was conducted at The Rocky Mountain Biological Laboratory (RMBL;
105 38°57.5" N, 106°59.3" W (WGS 84), 2900 m.a.s.l.), Colorado, USA (APPENDIX 1 for site information).
106 Pollinator observations of Bombus Latreille spp. (Apidae Latreille) were conducted from June - August of 2015
107 in six study sites characterized by high-montane/subalpine Parkland vegetation communities. Observations
108 of Bombus foraging took place for one hour at each field site in three 100m transects, where all flowers were
109 also counted and placed into abundance bins. Corbiculae loads were, non-lethally, collected once from all
110 Queen individuals encountered.

111 2.1 | Spatial Analyses

112 **2.1.1 Candidate Species** We downloaded from the Botanical Information and Ecology Network ‘BIEN’
113 (Maitner (2022)) all records adjacent to the field sites to develop an ecologically relevant list of vascular plant
114 species, with expected biotic pollination, which may be present at the study area . We then generated Species
115 Distribution Models (SDMs) to predict their distribution throughout the study area. These maps of potential
116 distribution served as a reference to reduce the list of species to include in the genomic sequence databases.
117 In order to minimise the number of species for which SDM’s were to be generated, BIEN was queried at

118 a distance of up to 100km from our study area and all plant species records were downloaded. ***In order***
119 ***to emulate the stochasticity of botanical collecting, this dataset was bootstrap re-sampled 250***
120 ***times, with 90% of samples selected, to create a testing dataset.*** The median of the logistic
121 regression assessing the probability of occurrence of a species record as a function of distance from the study
122 area was used as a threshold distance, under which, to include species as candidates for distribution modelling.

123 **2.1.2 Distribution Modelling** We used all occurrence records from BIEN (n = 23,919) within a 50km
124 border of the Omernik level 3 ecoregion, which includes the study area (*No. 21 “Southern Rockies”*) to
125 construct the species distribution model (Omernik (1987)). These records were copied into two, initially
126 identical, sets, one for generating machine learning models (ML; Random Forest, and Boosted Regression
127 Tree’s), and the other for Generalised Linear (GLM) and Generalized Additive Models (GAM) (Barbet-
128 Massin *et al.* (2012)). **Ensembled predictions have been shown to outperform their constituent**
129 **models, on average, and to reduce the ecological signal to the analytical noise of individual runs**
130 **(Araujo & New (2007)). No single method of producing SDMs has been shown to universally**
131 **outperform others when faced with a large and diverse number of applications, in our case a**
132 **great number of species with differing biologies and ecologies (Elith* *et al.* (2006), Qiao *et al.***
133 **(2015)). In the spirit of these findings, multiple families of models, which can be generated**
134 **together as they have similar requirements regarding the number and ratios of Presence to**
135 **Absence records were ensembled together (Barbet-Massin *et al.* (2012)).**

136 We then generated 4,029 absence points , locations where the focal taxon is anticipated missing, through a
137 random stratification of 19% of the land cover in the area and included them in (BLM CITATION - need
138 appropriate format for journal). To achieve a larger absence dataset, we generated 1,000 pseudo-absence
139 records for each taxon by randomly selecting coordinates located at least 10km away from an occurrence
140 record. For ML models, these pseudo-absences were reduced so that the ratio of presence to absence records
141 were balanced (Barbet-Massin *et al.* (2012)). To achieve this, we removed absence records inside of 10% of
142 the mean sample value of the presence records; the required number of absence records were then randomly
143 sampled.

144 We used 26 environmental variables at 30m resolution to predict the potential distribution of each species,
145 six related to climate, five soil, four topographic, four related to cloud cover, with the remaining reflecting
146 assorted abiotic parameters (Wilson & Jetz (2016), Wang *et al.* (2016), Hengl *et al.* (2017), Robinson *et al.*
147 (2014)) (*APPENDIX 6*). **These publicly available datasets, were selected as they For linear**
148 regression models these predictors underwent both *vifstep* (theta = 10, max observations = 12,500) and *vifcor*

¹⁴⁹ (theta = 0.7, max observations = 12,500) to detect highly correlated variables, and collinear features were
¹⁵⁰ removed leaving 16 variables (Naimi *et al.* (2014)).

¹⁵¹ Modelling: Random Forest and Boosted Regression Trees, were sub sampled with 30% test and two replicates
¹⁵² each before weighted ensemble based on True Skill Statistics (tss) (Naimi & Araujo (2016)). Generalised linear
¹⁵³ models (GLM) and Generalised additive models (GAM) with 30% sub sampling and three replicates each
¹⁵⁴ were also ensembled using the tss (Naimi & Araujo (2016), @). TSS was chosen as the ensemble criterion as it
¹⁵⁵ has been shown to work across a range of species occurrences prevalences (Allouche *et al.* (2006)). The results
¹⁵⁶ of these models were extracted on a cell-by-cell basis to a polygon feature derived from a minimum-spanning
¹⁵⁷ tree which encompasses the study sites, and species from either ensemble with greater than 50% mean habitat
¹⁵⁸ suitability across all cells were considered present for further purposes (Prim (1957)).

¹⁵⁹ 535 species were modelled using Generalized Linear Models and Generalized Additive Models. 534 species
¹⁶⁰ were modelled using Random Forest and Boosted Regression Trees. To evaluate the accuracy of the species
¹⁶¹ distribution models, additional presence records from GBIF (n = 61,789), and AIM (n = 12,730) were used
¹⁶² as test and training sets (n = 74,519) for logistic regression (CITE AIM AND Occdownload Gbif.Org (2021)).
¹⁶³ Additional novel absence records were generated from the AIM dataset to create a dataset where each species
¹⁶⁴ has balanced presence and absences. 11 or more paired presence and absence records were required for this
¹⁶⁵ testing, resulting in 334 species being included in the logistic regression (Mdn = 110.0, \bar{x} = 223.1, max =
¹⁶⁶ 1568 record pairs used) with a 70% test split (Kuhn (2022)).

¹⁶⁷ 2.2 | Molecular Lab Work

¹⁶⁸ All lab work was carried out at The Daniel F. and Ada L. Rice Plant Conservation Science Center at the
¹⁶⁹ Chicago Botanic Garden, Glencoe, Illinois, U.S.A.

¹⁷⁰ **2.2.1 | Reference Plant Library Generation** Using five years of observational data on *Bombus* Queen
¹⁷¹ Bee foraging at these studies sites, we identified the plant taxa most frequently visited by Queens across
¹⁷² all years. We sequenced the 12 most commonly visited taxa twice using samples from one site within the
¹⁷³ Gunnison River Drainage and one individual from another population. In addition, for any of these 12 focal
¹⁷⁴ species which did not have a congener pair in this filtered sample, we included a congener - or a species from
¹⁷⁵ a closely related genus to serve as an outgroup. We also sequenced another 15 abundant taxa commonly
¹⁷⁶ visited by *Bombus* workers, based on the aforementioned data set (*APPENDIX 4*).

177 **2.2.2 | Plant Genomic DNA Extraction** Plant genomic DNA was isolated from ~ 1 cm² of leaf tissue
178 from silica-gel dried or herbarium material using a modified cetyltrimethylammonium (CTAB) protocol
179 (Doyle & Doyle (1987)) that included two chloroform washes. DNA was quantified using a Nanodrop 2000
180 (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and Qubit fluorometer (Thermo Fisher Scientific).

181 **2.2.3 | Pollen Genomic DNA Extraction** Pollen genomic DNA was extracted from corbiculae using a
182 CTAB based protocol modified from Lahlamgahi et al. and Guertler et al. (2014, 2014). A SDS extraction
183 buffer (350 µL , 100mM Tris-HCl, 50 mM EDTA, 50 mM NaCl, 10% SDS v/v., pH 7.5) was added followed
184 by vortexing to allow dissolution of corbiculae. Pollen grains were then macerated with Kontes Pellet Pestles,
185 and the tip of these washed with 130 µL of the SDS extraction buffer, samples were then incubated for 1 hour
186 at 30°C. This was followed by the addition of 10% CTAB solution (450ul, of 20 mM Tris-Cl pH. 8.0, 1.4 M
187 NaCl, 10 mM EDTA pH 7.5, 10% CTAB, 5% PVP, ~85% Deionized water) and RNase (10 uL of 10 mg/mL)
188 and samples were incubated for 40 minutes at 37°C, on heat block (Multi-Blok, Thermo Fisher Scientific,
189 Waltham Massachusetts) set to 40°C. After 20 minutes incubation, Proteinase K (15 µL of 20mg/ml) and
190 DTT (12.5 µL of 1M in water) were added, and the samples were further incubated at 60°C for 1 hour.
191 Samples were then incubated overnight at 40°C. 500 µL of Phenol-Chloroform-Isoamyl alcohol (25:24:1) were
192 added, vortexed, and centrifuged at 10,000 rpm for 10 minutes and the aqueous phase was pipetted to a 1.5
193 ml centrifuge tube.

194 To precipitate the DNA, chilled Isopropyl alcohol & 3 mM Sodium acetate (5:1) equivalent to 2/3 of the
195 volume of sample were added, with 1 hour of chilling at -20°C, followed by 10 minutes of centrifuging at
196 13,000 rpm. The supernatant was pipetted to a new 1.5 ml centrifuge tube, and 70% EtOH (400 µL) were
197 added before chilling at -20°C for 20 minutes followed by centrifugation at 13,000 rpm for 10 minutes. Both
198 tubes were then washed with 75% EtOH (400 µL), inverted, centrifuged at 13,000 rpm for 4 minutes, and the
199 solution discarded, then washed with 95% EtOH (400 µL) , inverted, centrifuged at 13,000 rpm for 4 minutes,
200 and the solution discarded. Pellets were dried at room temperature overnight before resuspension in Nuclease
201 free H₂O. Extractions were assessed using a Nanodrop 2000 (Thermo Fisher Scientific) and Qubit fluorometer
202 (Thermo Fisher Scientific). DNA extracts were then cleaned using 2:1 v./v. Sera-Mag beads (Cytiva, Little
203 Chalfont, UK) to solute following the manufacturer's protocol, eluted in 0.5x TE, and the eluent allowed to
204 reduce by half volume in ambient conditions. DNA was quantified using a Qubit fluorometer.

205 **2.2.4 | Fragmentation, Library Preparation & Target Enrichment** Library preparation was per-
206 formed using the NEBNext Ultra II FS-DNA Library Prep Kit for Illumina (New England BioLabs, Ipswich,
207 Massachusetts, USA) using slightly modified manufacturers recommendation. Fragmentation was performed

208 at $\frac{1}{2}$ volume of reagents and $\frac{1}{4}$ enzyme mix for 40 minutes at 37°C, with an input of 500 ng cleaned DNA.
209 Adapter Ligation and PCR enrichment were performed with $\frac{1}{2}$ volumes, while cleanup of products was
210 performed with $\frac{1}{2}$ volume of SPRI beads (Beckman Coulter, Indianapolis, Indiana, USA) and recommended
211 volumes of 80% v./v. ethanol washes. The exception was the herbarium specimens which were not fragmented
212 and only end repaired, with similar library preparation of all samples. Products were analysed on 4% agarose
213 gels, and a Qubit fluorometer. Libraries were pooled and enriched with the Angiosperms 353 probe kit V.4
214 (Arbor Biosciences myBaits Target Sequence Capture Kit) by following the manufacturer's protocol and
215 Brewer et al. 2019. Sequencing was performed using an Illumina mi-Seq with 150-bp end reads, (NUSeq Core,
216 Chicago, Illinois).

217 **2.2.5 | Computational Processes and Analyses.**

218 **2.2.5.1 | Reference Library Data Processing** Sequences were processed using Trimmomatic, which
219 removed sequence adapters, clipped the first 3 bp, discarding reads less than 36 bp, and removing reads
220 if their average PHRED score dropped beneath 20 over a window of 5 bp (Bolger & Giorgi (2014), Tange
221 (2021)). Contigs were generated using HybPiper using target files created by M353 (Johnson *et al.* (2016),
222 McLay *et al.* (2021)).

223 **2.2.5.2 | Sequence Identification** A custom Kraken2 database was created by downloading representative
224 species of each genus indicated as being present in the study area by the spatial analyses from the Sequence
225 Read Archive (SRA) NCBI (Wood *et al.* (2019)). These sequences were processed in the same manner as
226 our novel sequences before being placed into the database. The Kraken2 database was built using default
227 parameters. Kraken2 was run on sequences using default parameters (*APPENDIX 5*). Following Kraken2,
228 Bracken was used to classify sequences to terminal taxa (Lu *et al.* (2017)). Results from both Kraken2 and
229 Bracken, results were reclassified manually to identify terminal taxa. For example, when only a single species
230 of a genus was known in the study area, but our database used a representative of another taxon in the genus,
231 this species was coded as the result. The re-coding of sequences from another representative species for the
232 genus to the sole RMBL representative allowed the identification of XX & % more species.

233 **2.2.5.3 | Identification of Sequence Matching Loci** A local NCBI database was built using the same
234 processed novel and downloaded sequences (Camacho *et al.* (2009)).

²³⁵ **2.2.5.4 | Morphological Pollen identification**

²³⁶ To develop a reference library of pollen grains which may be present in corbiculae loads, an image reference
²³⁷ collection of fuchsin-jelly stained (Beattie (1971)) slides was assembled from slides previously prepared by the
²³⁸ authors (n = 21), and other researchers (n = 38) (Brosi & Briggs (2013)). Using five years of observational
²³⁹ data on *Bombus* Queen Bee foraging at these studies sites (Ogilvie unpublished), as well as the Vascular
²⁴⁰ Plant Checklist (Frase & Buck (2007)), an additional 62 voucher slides for species were prepared and imaged
²⁴¹ at 400x (Leica DMLB, Leica MC170 HD Camera, Leica Application Suite V. 4.13.0) from non accessioned
²⁴² herbarium collections to supplement the number of species and clades covered (Appendix 3).
²⁴³ We used Divisive Hierarchical Clustering techniques to determine which plant taxa were distinguishable via
²⁴⁴ light microscopy, and to develop a dichotomous key to pollen morphotypes. Ten readily discernible categorical
²⁴⁵ traits were collected from each specimen in the image collection. These traits were transformed using Gower
²⁴⁶ distances, and clustered using Divisive Hierarchical clustering techniques (Maechler *et al.* (2022)). Using
²⁴⁷ the cluster dendrogram, elbow plot, and heatmaps (Hennig (2020)), of these results morphological groups of
²⁴⁸ pollen which could not be resolved via microscopy were delineated, and a dichotomous key was prepared
²⁴⁹ (APPENDIX NO.). This key was then used to identify the pollen grains sampled from corbiculae loads to
²⁵⁰ morphotypes in a consistent manner. To prepare the pollen slides from corbiculae, all corbiculae loads were
²⁵¹ broken apart and rolled using dissection needlepoints to increase heterogeneity of samples. *Cerca* 0.5mm² of
²⁵² pollen was placed onto a ~4mm² fuchsin jelly cube (Beattie (1971)) atop a graticulated microscope slide,
²⁵³ with 20 transects and 20 rows (400 quadrants) (EMS, Hartfield, PA). The jelly was melted, with stirring, until
²⁵⁴ pollen grains were homogeneously spread across the microscope slide. Slides were sealed with Canada Balsam
²⁵⁵ (Rublev Colours, Willits, CA) followed by sealing with nail polish; all samples are noted in APPENDIX 3.
²⁵⁶ To identify the pollen present in corbiculae loads, light microscopy at 400x (Zeiss Axioscope A1) was used. In
²⁵⁷ initial sampling in three transects, each pollen grain was identified to morphotype and counted; an additional
²⁵⁸ two transects were scanned for morphotypes unique to that slide, if either transect contained an unique
²⁵⁹ morphotype than all grains in that transect were also identified and counted. Subsequent to the first round
²⁶⁰ of sampling, non-parametric species richness rarefaction curves (Oksanen *et al.* (2022)), and non-parametric
²⁶¹ species diversity rarefaction curves were used to assess the completeness of sampling (Chao *et al.* (2014),
²⁶² Hsieh *et al.* (2020)). Slides not approaching the asymptote of the rarefaction curve were then re-sampled,
²⁶³ and analysed iteratively for up to a total of seven transects APPENDIX 2.

264 **2.3 | Temporal Analyses**

265 To estimate the duration of dates in which plant species were flowering weibull estimates of several phenological
266 parameters all spatially modelled taxa were developed (Belitz *et al.* (2020), Pearse *et al.* (2017)). Only
267 BIEN records which occurred in the Omernik Level 4 Ecoregions within 15km of the study area ($n = 5$
268 Level 4 Ecoregions, or conditionally 6 if enough records not be found in the nearest 5), and which were from
269 herbarium records were included. To remove temporally irrelevant herbarium records, i.e. material collected
270 during times which flowering is impossible at the study area due to snow cover, we used the SnowUS dataset
271 (Iler *et al.* (2021), Tran *et al.* (2019)) from 2000-2017 was analyzed for the first three days of contiguous snow
272 absence, and the first three days of contiguous snow cover in Fall. Herbarium records after the 3rd quantile
273 for melt, and the 1st quantile for snow cover of these metrics were removed. Species with > 10 records had
274 their weibull distributions generated for the date when 10% of individuals had begun flowering, when 50%
275 were flowering, and when 90% of individuals had flowered.

276 **2.4 | Floral Observations**

277 **3 | RESULTS**

278 **3.1 | Spatial Analyses**

279 [Table 1 about here.]

280 [Table 2 about here.]

281 The median (25.009 km) of the logistic regression assessing the probability of occurrence of a species record as
282 a function of distance from the study area was used as a threshold distance to include species for distribution
283 modelling. A 2-sample test for equality of proportions with continuity correction ($X^2 = 13.254$, df
284 = 1, p-value = 0.000136, 95% CI 0.04-1.00) was used to test whether more of the records located in the
285 broad ecological sites present at the field station, between the distance of the median (25.009 km) to the
286 third quantile (ca 43.830 km) of the regression distance, where true presences at the field station. Including
287 these records would have resulted in modelling an additional 222 species distributions of which 30 are true
288 presences these taxa were not modelled.

289 Across the entire spatial domain of modelling all ensembled models ($n = 968$) had an accuracy of 0.84 (95%
290 CI 0.8356 - 0.8443), kappa 0.68, p-value < 0.001, sensitivity = 0.80, specificity = 0.87, AUC = 0.92.

291 At the field site, of the 554 vascular plants with biotic pollination syndromes, the 493 ML ensembles accurately
292 predicted the presence of 362 (65.3%), incorrectly predicted the presence of 64 (11.6%), incorrectly predicted

293 34 true presences (6.1%) as being absent, and correctly predicted the true absence of 33 (6.0%). The balanced
294 accuracy of the ensembled models is 0.627 (Sensitivity = 0.340, Specificity 0.914). Of the 554 vascular plants
295 with biotic pollination syndromes, the 475 LM ensembles accurately predicted the presence of 286 (51.6%),
296 incorrectly predicted the presence of 41 (14.3%), incorrectly predicted 93 true presences (16.8%) as being
297 absent, and correctly predicted the true absence of 55 (9.9%). The balanced accuracy of the ensembled
298 models is 0.664 (Sensitivity = 0.573, Specificity 0.754). Of the 554 vascular plants with biotic pollination
299 syndromes in the flora 13 (2.3%) were in the Orchid family and 41 (7.4%) are non-natives, both of which are
300 restricted from the database, and can only reduce the number of true predicted presences by roughly 10%.

301 At the six study plots, of the 117 plant species identified to the species level across the spatial extents of all
302 plots and duration of queen bee activity, the ML ensembles predicted the presence of 105 (89.7%) of them,
303 and LM ensembles 102 (87.2%). Of the missing species two (1.7%) are Orchids, six (5.1%) are non-native,
304 and one (0.85%) is of contested taxonomic standing, all of which (7.65%) are restricted from the initial query
305 database.

306 3.2 | Microscopic Pollen identification

307 Using the fuchsin jelly preparation and light microscopic analyses of grains and scoring of 10 character
308 states resulted in the establishment of 28 morphotypes which grains could be classified into. APPENDIX
309 7. 60 samples were counted and based on rarefaction **had over % of expected morphotypes found**
310 (morphotype richness, $\bar{x} = 4.5$, Mdn = 4, min = 1, max = 9), all samples had expected morphotype diversity
311 reach the asymptote APPENDIX 8. The number of counted pollen grains in each sample range from (MIN -
312 16,293, $\bar{x} = 2788.685$, Mdn = 1453).

313 [Figure 1 about here.]

314 note this figure is draft mode, i reached out to C.H. Cole to get the official APG colors so we are gonna colour
315 edges with that, I have also drawn phylo pics for almost all the labelled order and need to add them in !

316 3.3 | Metabarcoding Pollen identification

317 54 corbiculae loads had DNA extracted and underwent various steps towards hyb-seq, in the end a total of 44
318 corbiculae samples were sequenced, 7,752,353 reads were recovered from sequencing. The number of reads
319 per sequence varied widely (range = 76 - 508,795, $\bar{x} = 176,189.8$, Mdn = 138,395). Of the possible 353 loci,
320 the number which were recovered from each sample, and informative to BLAST were range = 24 - 353, $\bar{x} =$
321 305.5, Mdn = 331. The number of reads per loci from across all samples had a range of 178 - 506,653, $\bar{x} =$

³²² 20,688, Mdn = 12,616.

³²³ APPENDIX X.

³²⁴ After trimming 7,865,680 sequences remained. 10,682,538 reads were matched using Kraken, of the reads
³²⁵ classified by Kraken 10,160,768 reads were matched using Bracken, of the reads classified by Kraken 7,302,876
³²⁶ reads were matched using BLAST.

³²⁷ Based upon subjective review of the three classifiers, BLAST was chosen as the classification method which
³²⁸ yielded the most probable results, and it's values were used for all subsequent analyses.

³²⁹ **3.4 | Temporal Analyses**

³³⁰ The first date of modeled snow melt in the Gothic area ($n = 17$, $\bar{x} = 137.9$, Mdn = 135, 3rd quantile =
³³¹ 151), and the first date of a consistent winter snow base ($n = 17$, $\bar{x} = 299.9$, Mdn = 300, 1st quantile = 291)
³³² from 2000-2017, were used as delimiters for the inclusions of herbarium records in modelling. Of the **500**
³³³ species predicted likely present in the area via logistic regression, 332 species (64.4%) with more than 10
³³⁴ records in the focal level 4 ecoregions ($\bar{x} = 35.01657$, Mdn = 35, max = 96) had weibull estimates calculated,
³³⁵ an additional 56 species (11.2%) with enough contributing records from the “Sedimentary Mid-Elevation
³³⁶ Forests”, a large ecoregion in general just beneath the elevation bands occupied by the five ecoregions around
³³⁷ the study area had weibull estimates also calculated ($\bar{x} = 13.86885$, Mdn = 13, max = 24).

³³⁸ Only 58 of these 388 species ($n = 34.56897$, Mdn = 31) were able to be compared to plot based observational
³³⁹ data from the long term (1974–2012) dataset. Of these species relatively high accord was observed between
³⁴⁰ the long-term ground truthed dataset, and the modelled species. There was very strong evidence that the
³⁴¹ weibull estimates were positively associated with the observed onset ($r^2 = 0.72$, $p < 0.0001$, $\tau = 0.61$) and
³⁴² peak ($r^2 = 0.70$, $p < 0.0001$, $\tau = 0.65$) of flowering, and that the number of herbarium samples had a
³⁴³ moderate effect on the estimates ($p = 0.004$ and $p = 0.034$ respectively). There was very strong evidence
³⁴⁴ that the weibull estimates had a positive association with the observed cessation of flowering ($r^2 = 0.4339$, p
³⁴⁵ < 0.0001 , $\tau = 0.489$), however their was no evidence that sample size had an effect ($p = 0.349$). There was
³⁴⁶ moderate evidence that the weibull estimates, with an effect of sample size, had a weak positive association
³⁴⁷ with the observed duration of flowering ($p = 0.0401$, $r^2 = 0.07$, $\tau = 0.17$).

³⁴⁸

[Figure 2 about here.]

³⁴⁹ **3.5 | Floral Observations**

³⁵⁰ The six sites were surveyed for a total of 52 hours from May 27-July 27. A total of 723 queen-pollen foraging
³⁵¹ interactions were observed (range per bee species by week range = 1 - 18, $\bar{x} = 3.46$, Mdn = 2), with a range
³⁵² of total observed interactions per bee species across this time period (min = 1, $\bar{x} = 59.08$, Mdn = 19, max =
³⁵³ 184). Plants varied widely in the number of interactions which they partook in with each species of bee (range
³⁵⁴ per plant species by week min = 1 - 20, $\bar{x} = 3.51$, Mdn = 2), with a range of total observed interactions per
³⁵⁵ plant species over this time period (min = 1, $\bar{x} = 20.26$, Mdn = 4, max = 141). The number of plant species
³⁵⁶ which bees were observed interacting with varied more narrowly (range = 1 - 18, $\bar{x} = 8$, Mdn = 6).

³⁵⁷ A total of 66 corbiculae loads were collected from Bees, 64 of them from Queens.

³⁵⁸ **4 | DISCUSSION**

³⁵⁹ Although we were able to use an actually fine scale flora to determine the species present at the field site,
³⁶⁰ we suspect a similar approach may be accomplished via quick species richness inventories at sites, and then
³⁶¹ utilizing a bootstrap approach akin to ours, to the taxa returned from databases to derive these estimates.

³⁶² ** Mention in here that the general effect of sample size on predictions means that larger samples sizes are
³⁶³ required for this to work!!!! ** Although our temporal results were lackluster, we note that our study area
³⁶⁴ has an incredibly brief growing period. and we suspect these temporal results would be useful in sub-tropical
³⁶⁵ and tropical ecosystems. FURTHER, the sites used for ground truthing the temporal flowering periods
³⁶⁶ were not randomly selected across the study area, and cannot be used to make inference to the population
³⁶⁷ across the entirety of the study area as we did here. Regardless both show good agreement on flower onset,
³⁶⁸ peak flowering, and moderate agreement with flowering cessation. The disagreement in flowering cessation
³⁶⁹ is perhaps due to more microclimates which retain water, rather than microclimates which allow the early
³⁷⁰ accumulation of heat.

³⁷¹ Fewer modelling runs for SDM's likely to be effective for determining inclusion, elastic inclusion criteria. The
³⁷² actual dataset which was used for training and testing all of the models incorporated into SDM's represented
³⁷³ only roughly one quarter of the records available for such purposes. We consciously chose to do this in order
³⁷⁴ to showcase the possibility of this approach working in less data rich areas.

³⁷⁵ Bayesian framework

³⁷⁶ Future Directions:

³⁷⁷ While at the time of writing this there are limited A353 sequence data, the Plant and Fungal Trees of Life

378 (PAFTOL) project, which is sequencing at least a species of each genera in the plant Kingdom will produce
379 sequence data from over 14,000 species. Given the extant publicly available genomic data, we conservatively
380 estimate that upon completion of PAFTOL there will be no fewer than 15,500 species (4.4% of all ca. 350,000
381 plant species) for which sequence data of a majority of these loci exist (Govaerts *et al.* (2021)). Accordingly,
382 projects in the near future may increase the number of metagenomics samples while decreasing the need to
383 create their own plant sequence reference libraries. As a result of PAFTOL the first ever comprehensive
384 phylogenetic hypotheses of all plant genera will be presented. In tandem with an increased number of
385 digitised and geo-referenced herbarium specimens, and monitoring programs in natural areas, we believe that
386 geo-informatics, and phylogenetic inference will increase the ability of researchers applying this technique
387 to identifying sequence reads. While our approach emphasises the use of this metagenomic technique for
388 the purpose of identifying pollen, I argue the template and resources we provide here make this approach a
389 suitable candidate for many plant metagenomic tasks. While we did not have the resources to explore the
390 possibility of characterising infraspecific characteristics, preliminary results from others (Wenzell *et al.* (2021),
391 Loke et al. in prep) indicate a possibility for these probes to also collect data at the level of populations and
392 individuals. **

393 In regards to better understanding the foraging preferences of *Bombus* feeding in subalpine ecosystems.

394 **JANE AND PAUL SET UP FOR NEAR FUTURE RESULTS?**

395 **AUTHOR CONTRIBUTIONS:** R.C.B conducted botanical collections, conducted all molecular lab work,
396 lead all analyses, and writing. J.E.O conceived, designed, and conducted all ecological fieldwork, assisted
397 with analyses, and writing. E.J.W. prepared, imaged, and collected trait data on pollen reference slides, and
398 assisted with analysis of trait data and writing a dichotomous key. S.T. assisted with spatial analyses and
399 writing. P.J.C assisted with ecological analyses and writing. J.B.F. conceived, and designed all lab work,
400 analyses, assisted with writing, and secured funding for molecular work.

401 **ACKNOWLEDGMENTS:** Nyree Zerega for assistance obtaining herbaria loans and accessioning our
402 collections at CHIC. Pat Herendeen for assistance with virtually all aspects of preparing pollen vouchers
403 and the identification process. Ian Breckheimer for sharing the SDM predictor variables. Hilary Noble, Zoe
404 Diaz-Martinez, Angela McDonnell, & Elena Loke for assistance with genomic library preparation. We thank
405 the curators at the following herbaria for supplying tissue: Ben Legler at Stillinger (ID), Charles (Rick)
406 Williams at Ray J. Davis (IDS), (B)Ernie Nelson at Rocky Mountain (RM). We thank the original collectors
407 of these specimens: D. Knoke, L. Brummer, J. Boyd, C. Davidson, I. Gilman, M. Kirkpatrick, S. McCauley,
408 J. Smith, K. Taylor, & C. Williams. David Giblin for sharing relevant sections of an advanced draft of FNA
409 V. 15. The Bureau of Land Management is thanked as many plant specimens were collected by R.C.B as

⁴¹⁰ a partner or contractor to the agency, Sarah Burnett and Lauren Price are thanked for sharing AIM data.
⁴¹¹ Sandra and NEB are gratefully acknowledged for technical support and generously sharing samples.

⁴¹² **CONFLICT OF INTERESTS** The authors declare no conflicts of interest.

⁴¹³ **PEER REVIEW** The peer review history for this document is available at ...

⁴¹⁴ **DATA AVAILABILITY STATEMENT** The queries required to download all data used in this project
⁴¹⁵ are located in... All novel sequencing data are located at NCBI...

⁴¹⁶ **ORCID**

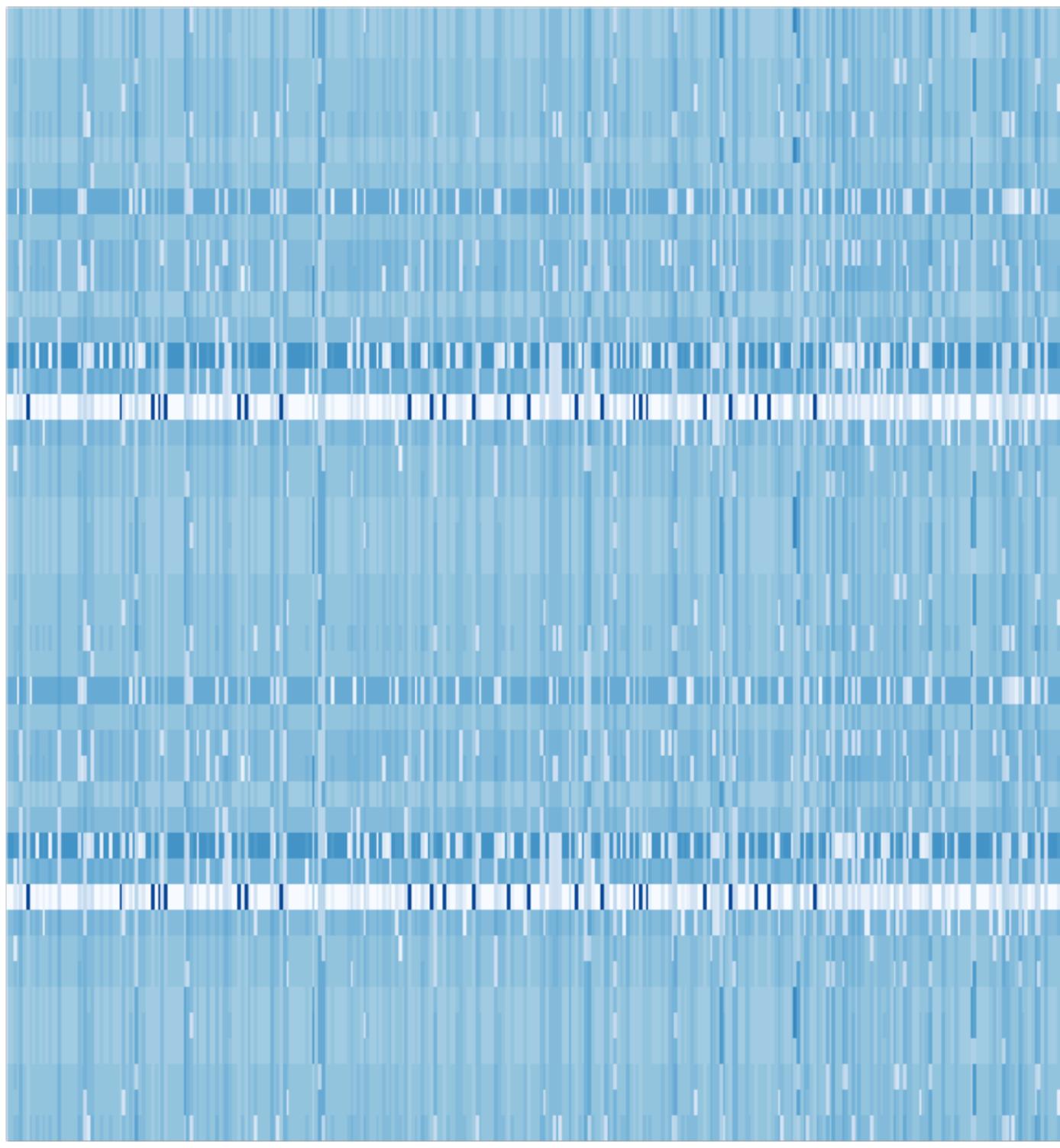
⁴¹⁷ Paul CaraDonna <https://orcid.org/0000-0003-3517-9090>
⁴¹⁸ Jeremie Fant <https://orcid.org/0000-0001-9276-1111>
⁴¹⁹ Jane Ogilvie <https://orcid.org/0000-0001-8546-0417>
⁴²⁰ Sophie Taddeo <https://orcid.org/0000-0002-7789-1417>

⁴²¹ **References**

⁴²² **Supporting**

⁴²³ Appendix XX - Reads Per Loci

Percent matched reads per locus by sample

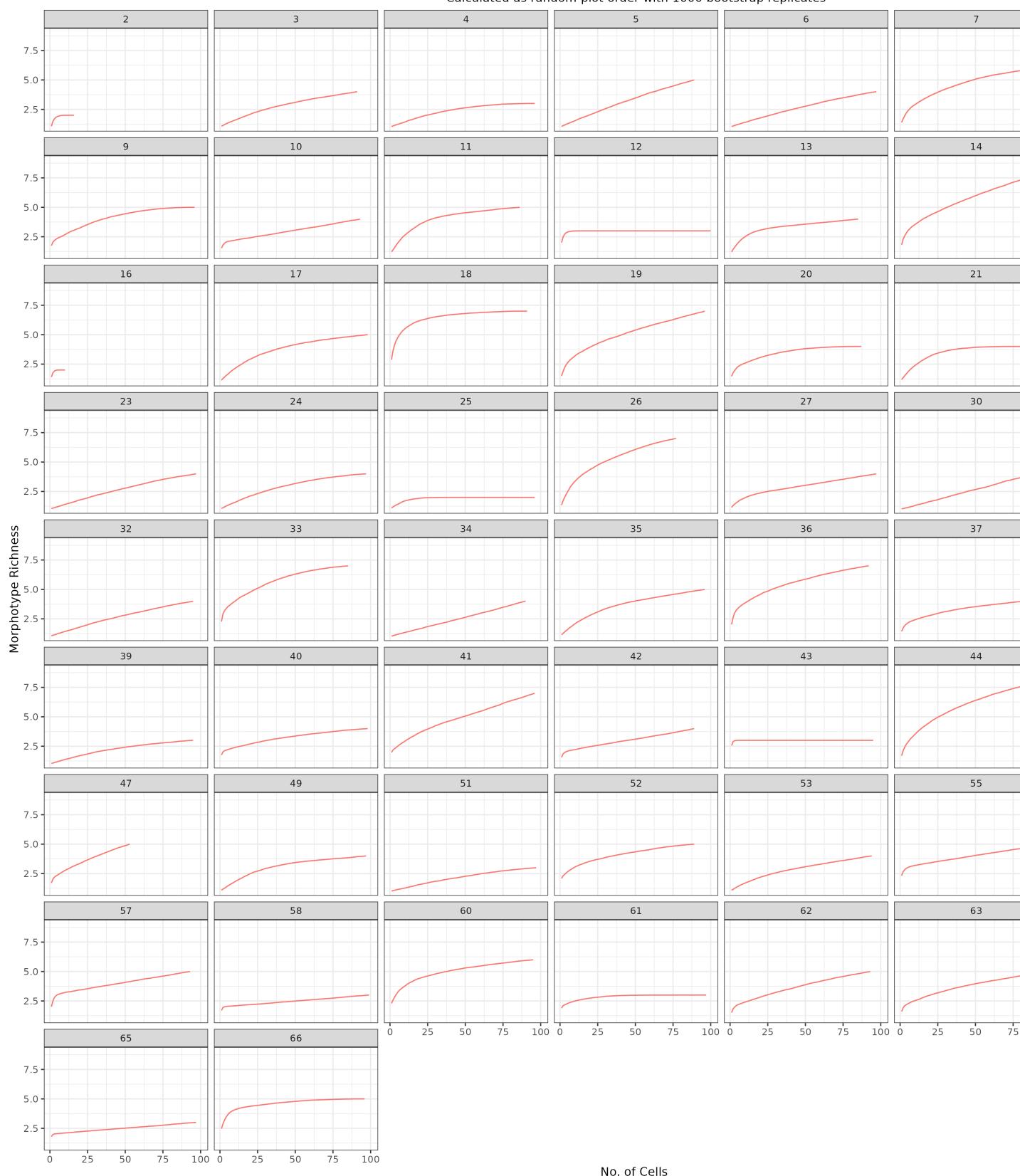


Locus

425 Appendix XX - Species Distribution Models Predictors

Layer	Description	Source
1.	Mean annual cloudiness - MODIS	Wilson et al. 2016
2.	Cloudiness seasonality 1 - MODIS	Wilson et al. 2016
3.	Cloudiness seasonality 2 - MODIS	Wilson et al. 2016
4.	Cloudiness seasonality 3 - MODIS	Wilson et al. 2016
5.	Beginning of the frost-free period	Wang et al.
6.	Climatic moisture deficit	Wang et al.
7.	Degree-days above 5C from	Wang et al.
8.	Mean annual precipitation	Wang et al.
9.	Mean annual precipitation as snow	Wang et al.
10.	Temperature seasonality	Wang et al.
11.	2015 Percent Grass/Herbaceous cover - MODIS	(MOD44B)
12.	2015 Percent Tree cover from Landsat 7/8	(GLCF)
13.	Soil probability of bedrock (R Horizon)	SoilGrids
14.	Soil organic carbon (Tonnes / ha)	SoilGrids
15.	Surface soil pH in H ₂ O	SoilGrids
16.	Surface soil percent sand	SoilGrids
17.	Soil USDA class	SoilGrids
18.	Topographic elevation	EarthEnv DEM
19.	Topographic elevation, moving window.	EarthEnv DEM
20.	Topographic percent slope	EarthEnv DEM
21.	Topographic wetness index	EarthEnv DEM
22.	Topographic aspect from	EarthEnv DEM
23.	Annual potential solar radiation computed	r.sun
24.	Estimated actual (w/-cloud) solar radiation r	.sun / Wilson et al. 2016
25.	Log-transformed distance to surface water Gl	obal Surface Water Explorer
26.	Percent surface water Gl	obal Surface Water Explorer

Rarefaction Curves of Species Richness
Calculated as random plot order with 1000 bootstrap replicates



⁴²⁸ Appendix XX - Pollen Morphotype Abundance Rarefaction Curves



⁴²⁹

Table 1: samples used in creating the Reference Library

Taxon	Family	Accession	Pres.	Locality	Date Col.	GenBank	Dist. (km)
<i>Cirsium parryi</i> (A. Gray) Petr.	Asteraceae	CHIC tba	S	Colorado, Gunnison	20.IX.2020	tba	3.6
<i>Cirsium parryi</i> (A. Gray) Petr.	Asteraceae	CHIC tba	S	Colorado, Gunnison	20.IX.2020	tba	3.6
<i>Ericameria parryi</i> (A. Gray) G.L. Nesom & Baird	Asteraceae	CHIC tba	S	Colorado, Gunnison	20.IX.2020	tba	3.6
<i>Erigeron speciosus</i> (Lindley) De Candolle	Asteraceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Erigeron subtrinervis</i> Rydb. Ex Porter & Britton	Asteraceae	CHIC tba	S	Colorado, Gunnison	20.VII.2020	tba	3.6
<i>Helianthella quinquenervis</i> (Hook.) A. Gray	Asteraceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Helianthus multiflora</i> Nutt.	Asteraceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Heterotheca villosa</i> (Pursh) Shinners	Asteraceae	CHIC tba	S	Colorado, Gunnison	20.IX.2020	tba	3.6
<i>Senecio sera</i> Hook.	Asteraceae	CHIC tba	P	Idaho, Idaho	26.VII.2020	tba	105.0
<i>Symplytrichum foliacum</i> (Lindl. Ex D.C.) G.L. Nesom	Asteraceae	CHIC tba	S	Illinois, McHenry	28.VII.2020	tba	1624.6
<i>Taraxacum officinale</i> F.H. Wigg.	Asteraceae	CHIC tba	S	Illinois, McHenry	28.VII.2020	tba	1624.6
<i>Mertensia ciliata</i> (James ex Torr.) G. Don	Boraginaceae	ID 1754185	S	Idaho, Valley	18.VI.2018	tba	979.3
<i>Mertensia ciliata</i> (James ex Torr.) G. Don	Boraginaceae	ID 1698387	P	Idaho, Adams	10.VII.2014	tba	991.5
<i>Mertensia fusiformis</i> Greene	Boraginaceae	RMH 720522	P	Colorado, Gunnison	7.VI.1997	tba	44.8
<i>Mertensia fusiformis</i> Greene	Boraginaceae	RMH 720600	P	Colorado, Gunnison	9.VII.1997	tba	38.9
<i>Campanula rotundifolia</i> L.	Campanulaceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Lathyrus lanszwertii</i> Kellogg var. leucanthus (Ryd.) Dorn	Fabaceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Lathyrus lanszwertii</i> Kellogg var. leucanthus (Ryd.) Dorn	Fabaceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Lupinus argenteus</i> Pursh	Fabaceae	CHIC tba	P	Nevada, Pershing	29.V.2018	tba	971.2
<i>Lupinus argenteus</i> Pursh	Fabaceae	ISU 10387	P	Colorado, Gunnison	29.VI.2010	tba	0.2
<i>Lupinus bakeri</i> Greene	Fabaceae	ISU 10142	P	Colorado, Gunnison	15.VIII.2010	tba	2.6
<i>Vicia americana</i> Muhl. ex Willd.	Fabaceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Vicia americana</i> Muhl. ex Willd. var. minor Hook.	Fabaceae	CHIC tba	S	Montana, Carbon	4.VII.2019	tba	10020.8
<i>Frasera speciosa</i> Douglas ex Griseb	Gentianaceae	RMH 721930	P	Colorado, Gunnison	20.VI.1997	tba	66.2
<i>Frasera speciosa</i> Douglas ex Griseb	Gentianaceae	RMH 719305	P	Colorado, Gunnison	7.VII.1997	tba	19.8
<i>Hydrophyllum capitatum</i> Douglas ex. Benth	Hydrophyllaceae	RMH tba	P	Colorado, Mesa	30.VI.2011	tba	64.6
<i>Hydrophyllum capitatum</i> Douglas ex. Benth	Hydrophyllaceae	RMH tba	P	Colorado, Delta	8.VI.2011	tba	65.3
<i>Hydrophyllum fendleri</i> (Gray) Heller	Hydrophyllaceae	ID 161100	P	Washington, Yakima	9.VI.2008	tba	1429.7
<i>Hydrophyllum fendleri</i> (Gray) Heller	Hydrophyllaceae	ID 164040	P	Idaho, Idaho	27.V.2009	tba	1014.4
<i>Agastache pallidiflora</i> (Heller) Rydberg	Lamiaceae	CHIC tba	S	Arizona, Coconino	17.VII.2020	tba	617.7
<i>Chamerion angustifolium</i> (L.) Holub	Lamiaceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Delphinium barbeyi</i> (Huth) Huth	Ranunculaceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Delphinium nuttallianum</i> Pritz.	Ranunculaceae	ID 166162	P	Idaho, Gem	15.VI.2011	tba	9825.5
<i>Delphinium nuttallianum</i> Pritz.	Ranunculaceae	ID 179376	P	Idaho, Gooding	29.IV.2017	tba	733.7
<i>Potentilla fruticosa</i> Pursh	Rosaceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Potentilla fruticosa</i> Pursh	Rosaceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Potentilla hippiana</i> Lehman.	Rosaceae	CHIC tba	S	New Mexico, Catron	15.VIII.2020	tba	573.8

(Continued on Next Page)

Table 1: samples used in creating the Reference Library (*continued*)

Taxon	Family	Accession	Pres.	Locality	Date Col.	GenBank	Dist. (km)
Potentilla pulcherrima Lehman.	Rosaceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6

^a Accession includes both Herbarium and Accession number^b Pres. refers to Preservation method. 'S' denotes silica-gel dried, 'P' denotes pressed^c All Localities are in the United States of America

Table 1: All species present in the Reference Sequence Databases
(Kraken and BLAST)

Order	Family	Taxon
Alismatales	Potamogetonaceae	<i>Potamogeton wrightii</i>
Apiales	Apiaceae	<i>Osmorhiza aristata</i>
Asparagales	Amaryllidaceae	<i>Allium stamineum</i>
	Asparagaceae	<i>Streptopus amplexifolius</i>
Asterales	Asteraceae	<i>Anaphalis margaritacea</i> <i>Antennaria carpatica</i> <i>Antennaria dioica</i> <i>Artemisia sibirica</i> <i>Brickellia dentata</i> <i>Chrysanthemus greenei</i> <i>Cirsium pannonicum</i> <i>Cirsium parryi</i> <i>Cirsium vulgare</i> <i>Crepis pygmaea</i> <i>Ericameria parryi</i> <i>Erigeron ecuadoriensis</i> <i>Erigeron grandiflorus</i> <i>Erigeron rosulatus</i> <i>Erigeron uniflorus</i> <i>Helianthella quinquenervis</i> <i>Heterotheca villosa</i> <i>Hieracium avilae</i> <i>Hieracium jubatum</i> <i>Hymenoxys hoopesii</i> <i>Leucanthemum graminifolium</i> <i>Microseris lindleyi</i> <i>Omalotheca supina</i> <i>Packera quercetorum</i> <i>Pseudognaphalium attenuatum</i> <i>Pseudognaphalium frigidum</i> <i>Pseudognaphalium lacteum</i> <i>Pseudognaphalium oxyphyllum</i> <i>Rudbeckia hirta</i> <i>Scabrethia scabra</i> <i>Senecio adenophyllus</i> <i>Senecio algens</i> <i>Senecio apolobambensis</i> <i>Senecio candollei</i> <i>Senecio chionogeton</i> <i>Senecio formosus</i> <i>Senecio funcii</i> <i>Senecio gilliesii</i> <i>Senecio humillimus</i> <i>Senecio nutans</i> <i>Senecio puchei</i> <i>Senecio rufescens</i> <i>Senecio spinosus</i> <i>Senecio tephrosioides</i>

(Continued on Next Page)

Table 1: All species present in the Reference Sequence Databases (Kraken and BLAST) (*continued*)

Order	Family	Taxon
Boraginales	Campanulaceae	<i>Solidago chilensis</i> <i>Stilpnolepis intricata</i> <i>Symphytum foliacium</i> <i>Taraxacum cucullatum</i> <i>Taraxacum officinale</i>
		<i>Tonestus lyallii</i> <i>Townsendia formosa</i>
		<i>Campanula argaea</i> <i>Campanula rotundifolia</i>
		<i>Cynoglossum amplifolium</i> <i>Cynoglossum anachusoides</i>
		<i>Cynoglossum pringlei</i> <i>Mertensia ciliata</i> <i>Mertensia fusiformis</i>
	Hydrophyllaceae	<i>Hydrophyllum canadense</i> <i>Hydrophyllum capitatum</i> <i>Hydrophyllum fendleri</i> <i>Nemophila menziesii</i>
		<i>Arenaria globiflora</i> <i>Arenaria serpyllifolia</i> <i>Cerastium arvense</i> <i>Cerastium lanceolatum</i>
		<i>Minuartia recurva</i> <i>Odontostemma leucasterium</i> <i>Pseudostellaria heterophylla</i> <i>Sagina procumbens</i>
		<i>Schizotechium monospermum</i> <i>Shivparvania glanduligera</i> <i>Stellaria graminea</i> <i>Stellaria holostea</i>
		<i>Stellaria obtusa</i> <i>Rumex induratus</i> <i>Rumex spinosus</i>
Celastrales	Celastraceae	<i>Parnassia faberi</i> <i>Parnassia palustris</i> <i>Paxistima canbyi</i>
		<i>Gaultheria procumbens</i> <i>Moneses uniflora</i> <i>Orthilia secunda</i>
		<i>Vaccinium vitis-idaea</i> <i>Collomia grandiflora</i> <i>Ipomopsis aggregata</i>
		<i>Phlox douglasii</i>
Fabales	Primulaceae	<i>Androsace studiosorum</i> <i>Androsace vitaliana</i>
		<i>Astragalus pelecinus</i>
		<i>Lupinus argenteus</i>
		<i>Lupinus sericeus</i>

(Continued on Next Page)

Table 1: All species present in the Reference Sequence Databases (Kraken and BLAST) (*continued*)

Order	Family	Taxon
Gentianales	Gentianaceae	<i>Vicia americana</i> <i>Frasera speciosa</i> <i>Gentiana cruciata</i>
Hyphomicrobiales	Xanthobacteraceae	<i>Azorhizobium caulinodans</i>
Lamiales	Lamiaceae	<i>Agastache pallidiflora</i>
Liliales	Colchicaceae	<i>Prosartes smithii</i>
	Liliaceae	<i>Erythronium dens-canis</i>
	Melanthiaceae	<i>Anticlea elegans</i> <i>Veratrum viride</i>
Malpighiales	Hypericaceae	<i>Hypericum perforatum</i>
	Salicaceae	<i>Populus alba</i>
	Violaceae	<i>Viola odorata</i>
Myrtales	Onagraceae	<i>Chamaenerion angustifolium</i> <i>Epilobium canum</i> <i>Epilobium parviflorum</i>
Ranunculales	Berberidaceae	<i>Berberis sibirica</i>
	Papaveraceae	<i>Corydalis aitchisonii</i>
	Ranunculaceae	<i>Actaea heracleifolia</i> <i>Anemone anemonoides</i> <i>Anemone obtusiloba</i> <i>Aquilegia ecalcarata</i> <i>Caltha palustris</i> <i>Delphinium barbeyi</i> <i>Delphinium gracile</i> <i>Delphinium nuttallianum</i> <i>Pulsatilla chinensis</i> <i>Thalictrum thalictroides</i> <i>Thalictrum tuberosum</i> <i>Trollius europaeus</i>
Rosales	Elaeagnaceae	<i>Shepherdia argentea</i>
	Rosaceae	<i>Crataegus bipinnatifida</i> <i>Dasiphora fruticosa</i> <i>Geum ternatum</i> <i>Hedlundia austriaca</i> <i>Holodiscus argenteus</i> <i>Karpatiosorbus devoniensis</i> <i>Micromeles japonica</i> <i>Potentilla anserina</i> <i>Potentilla pulcherrima</i> <i>Potentilla tetrandra</i> <i>Rubus chingii</i>
Sapindales	Sapindaceae	<i>Acer campestre</i>
Saxifragales	Crassulaceae	<i>Rhodiola rosea</i> <i>Sedum nudum</i>
	Grossulariaceae	<i>Ribes rubrum</i>
	Saxifragaceae	<i>Lithophragma parviflorum</i> <i>Saxifraga biflora</i> <i>Saxifraga fortunei</i>

(Continued on Next Page)

Table 1: All species present in the Reference Sequence Databases
(Kraken and BLAST) (*continued*)

Order	Family	Taxon
		Saxifraga maderensis
		Saxifraga oppositifolia
		Saxifraga portosanctana
		Saxifraga x geum

x geum* \end{longtable}

442 Appendix XX - All Pollen Reference Slides Used to Establish Morphotypes

Table 1: All Pollen Voucher Slides Consulted

Taxon	Family	Locality	Accession	Type	Prepared by	Date
<i>Cymopterus lemnoides</i> (J.M. Coulter & Rose) Dorn	Apiaceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Heracleum sphondylium</i> L.	Apiaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Ligusticum porteri</i> J.M. Coulter & Rose	Apiaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Osmorhiza depauperata</i> Phil.	Apiaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Maianthemum stellatum</i> (L.) Link	Asparagaceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Achillea millefolium</i> L.	Asteraceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Achillea millefolium</i> L.	Asteraceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Acourtia wrightii</i> (A. Gray) Reveal & King	Asteraceae	NV, Clark, Gold Butte	tba	Novo	E.J.W.	2021
<i>Antennaria racemosissima</i> Hook.	Asteraceae	WY, Park, Ishawooa Mesa	tba	Novo	E.J.W.	2021
<i>Arnica latifolia</i> Bong.	Asteraceae	ID, Blaine, Wildhorse Canyon	tba	Novo	E.J.W.	2021
<i>Artemisia scopulorum</i> A. Gray	Asteraceae	CO, Hinsdale, Uncompahgre Peak	tba	Novo	E.J.W.	2021
<i>Canadanthus modestus</i> (Lindl.) G.L. Nesom	Asteraceae	ID, Idaho, Whiskey Creek	tba	Novo	E.J.W.	2021
<i>Chaenactis douglasii</i> (Hook.) Hook. & Arn.	Asteraceae	MT, Carbon, Pryor Mtn. Rd.	tba	Novo	E.J.W.	2021
<i>Erigeron corymbosus</i> Nutt.	Asteraceae	MT, Carbon, Pryor Mtn. Rd.	tba	Novo	E.J.W.	2021
<i>Erigeron flagellaris</i> A. Gray	Asteraceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Erigeron speciosus</i> (Lind.) DC.	Asteraceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Erigeron speciosus</i> (Lindl.) DC.	Asteraceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Helianthella quinquenervis</i> (Hook.) A. Gray	Asteraceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Heliotropium multiflorum</i> Nutt.	Asteraceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Heliotropium multiflorum</i> Nutt.	Asteraceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Heterotheca villosa</i> (Pursh) Shinners	Asteraceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Heterotheca villosa</i> (Pursh) Shinners	Asteraceae	AZ, Coconino, Lake Mary Rd. & 209	tba	Novo	E.J.W.	2021
<i>Hymenoxys hoopesii</i> (A. Gray) Bierner	Asteraceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Hymenoxys rusbyi</i> (A. Gray) Cockerell	Asteraceae	AZ, Coconino, Lake Mary Rd. & 209	tba	Novo	E.J.W.	2021
<i>Ionactis stenomeria</i> (A. Gray) Greene	Asteraceae	ID, Idaho, Marshall Mountains	tba	Novo	E.J.W.	2021
<i>Senecio hydrophilus</i> Nutt.	Asteraceae	ID, Custer, E. fl. Salmon River	tba	Novo	E.J.W.	2021
<i>Senecio integrerrimus</i> Nutt.	Asteraceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Senecio serrula</i> Hook.	Asteraceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Senecio wootonii</i> Greene	Asteraceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Solidago lepida</i> DC.	Asteraceae	ID, Idaho, American River	tba	Novo	E.J.W.	2021
<i>Sympotrichum foliacum</i> (Lindl. ex DC.) G.L. Nesom	Asteraceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Sympotrichum subspicatum</i> (Nees) G.L. Nesom	Asteraceae	ID, Custer, E. fl. Salmon River	tba	Novo	E.J.W.	2021
<i>Taraxacum officinale</i> F.H. Wigg	Asteraceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Taraxacum officinale</i> F.H. Wigg	Asteraceae	IL, McHenry, Barrington	tba	Novo	E.J.W.	2021
<i>Lappula squarrosa</i> (Retz.) Dumort.	Boraginaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Mertensia ciliata</i> (James ex Torr.) G. Don	Boraginaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Mertensia fusiformis</i> Greene	Boraginaceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Boechera</i>	Brassicaceae	NV, Washoe, Mt. Rose	tba	Novo	E.J.W.	2021
<i>Boechera stricta</i> (Graham) Al-Shehbaz	Brassicaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Cardamine cordifolia</i> A. Gray	Brassicaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Draba aurea</i> Vahl. Ex Hornem	Brassicaceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014

(Continued on Next Page)

444 Appendix XX - All Pollen Reference Slides Used to Establish Morphotypes (con't)

Table 1: All Pollen Voucher Slides Consulted (*continued*)

Taxon	Family	Locality	Accession	Type	Prepared by	Date
<i>Draba spectabilis</i> Greene	Brassicaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Erysimum capitatum</i> (Douglas ex Hook.) Greene	Brassicaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Lepidium montanum</i> Nutt.	Brassicaceae	NM, Catron, Zuni Dry Lake	tba	Novo	E.J.W.	2021
<i>Smelowskia americana</i> Rydb.	Brassicaceae	ID, Blaine, Pioneer Mtns Crest	tba	Novo	E.J.W.	2021
<i>Thlaspi arvense</i> L.	Brassicaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Campanula rotundifolia</i> L.	Campanulaceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Campanula rotundifolia</i> L.	Campanulaceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
Downingia	Campanulaceae	CA, Nevada, Truckee Meadows	tba	Novo	E.J.W.	2021
<i>Lonicera involucrata</i> (Richardson) Banks ex Spreng.	Caprifoliaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Minuartia nuttallii</i> (Pax.) Briq.	Caryophyllaceae	ID, Blaine, Wildhorse Canyon	tba	Novo	E.J.W.	2021
<i>Stellaria longifolia</i> Muhl. Ex Willd.	Caryophyllaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Parnassia kotzebuei</i> Cham. ex Spreng	Celastraceae	ID, Lemhi, Terrace Lakes	tba	Novo	E.J.W.	2021
<i>Sedum lanceolatum</i> Torr.	Crassulaceae	ID, Lemhi, Terrace Lakes	tba	Novo	E.J.W.	2021
<i>Astragalus robbinsii</i> (Oakes) A. Gray	Fabaceae	ID, Custer, Lake Creek	tba	Novo	E.J.W.	2021
<i>Glycyrrhiza lepidota</i> Nutt.	Fabaceae	ID, Butte, Warm Springs Creek	tba	Novo	E.J.W.	2021
<i>Lathyrus eucosmus</i> Butters & H. St. John	Fabaceae	NM, Catron, Zuni Salt Lake	tba	Novo	E.J.W.	2021
<i>Lathyrus lanszwertii</i> var. <i>leucanthus</i> (Rydb.) Dorn	Fabaceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Lathyrus lanszwertii</i> var. <i>leucanthus</i> (Rydb.) Dorn	Fabaceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Lupinus argenteus</i> Pursh	Fabaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Lupinus argenteus</i> Pursh	Fabaceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Lupinus crassus</i> Payson	Fabaceae	NV, Pershing, Star Peak Canyon	tba	Novo	E.J.W.	2021
<i>Lupinus sericeus</i> Pursh	Fabaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Melilotus albus</i> Medik.	Fabaceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Trifolium hybridum</i> L.	Fabaceae	NM, Catron, Hwy 159	tba	Novo	E.J.W.	2021
<i>Trifolium pratense</i> L.	Fabaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Vicia americana</i> Muhl. Ex Willd.	Fabaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Vicia americana</i> Muhl. Ex Willd.	Fabaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Vicia americana</i> Muhl. Ex Willd.	Fabaceae	UT, Cache, Spawn Creek	tba	Novo	E.J.W.	2021
<i>Geranium</i>	Fabaceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Geranium richardsonii</i> Fisch. Trautv.	Geraniaceae	NM, Catron, Jim Smith TH. Rd.	tba	Novo	E.J.W.	2021
<i>Phacelia</i> sp.	Geraniaceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Iris missouriensis</i> Nutt.	Hydrophyllaceae	NV, Nye, Toiyabe Crest	tba	Novo	E.J.W.	2021
<i>Agastache palliflora</i> (A. Heller) Rydb.	Iridaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Erythronium grandiflorum</i> Pursh	Lamiaceae	AZ, Coconino, Lake Mary Rd. & 209	tba	Novo	E.J.W.	2021
<i>Linum lewisii</i> Pursh	Liliaceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Zigadenus elegans</i> Pursh	Linaceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Anticlea elegans</i> (A. Gray) Zomlefer & Judd	Melanthiaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Cistanthe/Calyptidium</i>	Melianthiaceae	ID, Blaine, Wildhorse Canyon	tba	Novo	E.J.W.	2021
<i>Chamerion angustifolium</i> (L.) Holub	Montiaceae	NV, Nye, Toiyabe Crest	tba	Novo	E.J.W.	2021
<i>Epilobium obcordatum</i> A. Gray	Onagraceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Castilleja miniata</i> Douglas ex Hook.	Orobanchaceae	ID, Lemhi, Bighorn Crags	tba	Novo	E.J.W.	2021
		CO, Gunnison, Gothic	None	Image	H.M.B.	2010

(Continued on Next Page)

446 Appendix XX - All Pollen Reference Slides Used to Establish Morphotypes (con't)

Table 1: All Pollen Voucher Slides Consulted (*continued*)

Taxon	Family	Locality	Accession	Type	Prepared by	Date
<i>Castilleja sulphurea</i> Rydb.	Orobanchaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Pedicularis groenlandica</i> Retz.	Orobanchaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Pedicularis racemosa</i> Douglas ex Benth.	Orobanchaceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Corydalis aurea</i> Willd.	Papaveraceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Erythranthe guttata</i> (DC.) G.L. Nesom	Phrymaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Penstemon wilcoxii</i> Rydb.	Plantaginaceae	MT, Missoula, Mission Mtns vic.	tba	Novo	E.J.W.	2021
<i>Collomia linearis</i> Nutt.	Polemoniaceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Phlox condensata</i> (A. Gray) E.E. Nelson	Polemoniaceae	CO, Hinsdale, Uncompahgre Peak	tba	Novo	E.J.W.	2021
<i>Polemonium foliosissimum</i> A. Gray	Polemoniaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Polemonium occidentale</i> Greene	Polemoniaceae	ID, Custer, Lake Creek	tba	Novo	E.J.W.	2021
<i>Polemonium viscosum</i> Nutt.	Polemoniaceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2010
<i>Bistorta bistortoides</i> (Pursh) Small	Polygonaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Eriogonum</i> spp.	Polygonaceae	NV, Washoe, Hwy 445	tba	Novo	E.J.W.	2021
<i>Polygala barbeyana</i> Chodat	Polygonaceae	NM, Eddy, Yeso Hills	tba	Novo	E.J.W.	2021
<i>Polygonum polygaloides</i> L.	Polygonaceae	MT, Missoula, Mission Mtns	tba	Novo	E.J.W.	2021
<i>Androsace filiformis</i> Retz.	Primulaceae	ID, Custer, Bradshaw Creek	tba	Novo	E.J.W.	2021
<i>Aquilegia coerulea</i> E. James	Ranunculaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Aquilegia coerulea</i> E. James	Ranunculaceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Aquilegia elegantula</i> Greene	Ranunculaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Delphinium barbeyi</i> (Huth) Huth	Ranunculaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Delphinium nuttallianum</i> Pritz. Ex Walp.	Ranunculaceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Ranunculus alismifolius</i> Geyer ex Benth.	Ranunculaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Ranunculus glaberrimus</i> Hook.	Ranunculaceae	ID, Lemhi, Agency Creek	tba	Novo	E.J.W.	2021
<i>Ranunculus inamoenus</i> Greene	Ranunculaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Ranunculus</i> spp.	Ranunculaceae	NV, Washoe, Mt. Rose	tba	Novo	E.J.W.	2021
<i>Thalictrum sparsiflorum</i> Tucz. Ex Fisch. & C.A. Mey.	Ranunculaceae	ID, Custer, E. flk. Salmon River	tba	Novo	E.J.W.	2021
<i>Dasiphora fruticosa</i> (L.) Rydb.	Rosaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Dasiphora fruticosa</i> (L.) Rydb.	Rosaceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Fragaria virginiana</i> Duchesne	Rosaceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Geum triflorum</i> Pursh	Rosaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Potentilla biennis</i> Greene	Rosaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Potentilla hippiana</i> Lehm.	Rosaceae	NM, Catron, Jim Smith TH.Rd.	tba	Novo	E.J.W.	2021
<i>Potentilla pulcherrima</i> Lehm.	Rosaceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Potentilla pulcherrima</i> Lehm.	Rosaceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Salix bebbiana</i> Sarg.	Salicaceae	ID, Custer, Lake Creek	tba	Novo	E.J.W.	2021
<i>Salix geyeriana</i> Andersson	Salicaceae	ID, Butte, Clyde	tba	Novo	E.J.W.	2021
<i>Mitella stauropetala</i> Piper	Saxifragaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Valeriana occidentalis</i> A. Heller	Valerianaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Viola canadensis</i> L.	Violaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010

* All Localities are in the United States of America

† Accession refers to whole-plant vouchers, all specimens are deposited at CHIC.

‡ Type Refers to whether both a physical and digital copy exist; 'Image' denotes only digital

§ Date refers to the Date of preparation.

- 1a: Pollen shed in clumps (tetrads/polyads); grains generally triangular, with an annulus subtending the porate apertures (go 34)
- 1b: Pollen generally dispersed as single units (monads); grains seldom if ever with annulus.
- 2a: Apertures porate, always lacking colpi
 - 3a: grain outline from equatorial view circular
 - 4a: Pores distributed along the equator.
 - 5a: Pores > 5 (stephanoporate)
 - 6a: Ornamentation homobrochate (~ *MENTZELIA*)
 - 6b: Ornamentation otherwise (~ *POLYGALA*)
 - 5b: Pores < 5 (*CURRENTLY OPEN*)
 - 4b: Pores +/- distributed across grain (pantoporate)
 - 7a: Ornamentation with striate ornamentation (~ *POLEMONIUM*)
 - 7b: Ornamentation otherwise
 - 8a: Ornamentation, slightly irregular - without regularly repeating features (scabrate) (~ *STELLARIA*)
 - 8b: Ornamentation forming regularly repeating (reticulate) cells of varying shapes.
 - 9a: spacing between the grid cells large (lophate), the walls of the cells with another set of projecting ornamentation (~ *OPUNTIA*)
 - 9b: spacing between cells small, the wall of the cells without projecting features.
 - 10a: Pores extending beyond the reticulate grids (~ *ARENARIA*)
 - 10b: Pores extending beyond the reticulate grids (~ *PHLOX*)
 - 3b: Outline from equatorial view otherwise (usually slightly triangular)
 - 11a: Outline elliptic (*CURRENTLY EMPTY*)
 - 11b: Outline not elliptic, grains often with acute, if rounded, angles along sides (e.g., triangular, polygonal) (*EMPTY*)
 - 2b: Apertures with colpi, occasionally also with pores in addition (colporate)
 - 12a: Grains with bristles tapering to points (echinate), and tri-colporate.
 - 13a: Grains uniformly echinate, less the apertures. (Asteraceae 1)
 - 13b: Grains with echinate bristles on ridges of lophae (Asteraceae 2)
 - 12b: Grains without echinate ornamentation - this lead includes projections with ornamentation with round tips.
 - 14a: Grains with either less than 3 apertures, or with two distinct ornamentation types (generally $\frac{1}{2}$ psilate, $\frac{1}{2}$ reticulate).
 - 15a. Grains apparently lacking any apertures. (~ *IRIS*)
 - 15b. Grains aperturate
 - 16a. Ornamentation on one face of grain psilate, the other homobrochate (~ *ZIGADENUS + ANTICLEA*)
 - 16b. Ornamentation psilate across both faces of grain (~*ERYTHONIUM*)
 - 14b. Grains with either 3 or more apertures, or with an elongated spiral like aperture
 - 17a. Grain with spiral like colpi
 - 18a. Spiral with deep well-defined furrows (~ *ERYTHRANTHE GUTTATA*, syn. obsolete. *MIMULUS*)
 - 18b. Spirals without well-defined grooves, ornamentation evidently perforate (~ *RANUNCULUS ALISMIFOLIUS*)
 - 17b. Grains with colpi these not forming irregular spiral motifs.
 - 19a. Grains elliptic, essentially perfectly cylindrical along longest axis, except for minor inundations along equatorial region. Apertures, of two types (heteroaperturate). (~ *BORAGINACEAE*)
 - 19b Grains shaped similar or not, but never heteroaperturate.
 - 20a. From a polar view, grains notably polygonal (hexagonal), also evident when seldom seen from a equatorial view. (~ *PHACELIA*/ maybe *Hydrophyllaceae*, *Hydrophyllum* not sampled)
 - 20b. From a polar view, grains not with 6 convex apices

- 21a. Grains elliptic, with a short colporate aperture on each psilate face, the edges of each face and the apices with a distinct (homobrochate) textured ornamentation. (~ POLYGONUM)
- 21b. Grains otherwise, not featuring a mix of ornamentations independent of the apertures.
- 22a. Ornamentation perforate, the three colpi very short, their longest axis parallel to the equator rather than perpendicular. These colpi often times almost appearing to be slightly raised on an annulus like feature (~ LONICERA)
- 22b. Grains not as described in all aspects of the above.
- 23a. Apertures colporate
- 24a. Outline of grain in equatorial view circular, ornamentation smooth. (~MORPHOTYPE A).
- 24b. Grains otherwise
 - 25a Grains distinctly triangular from polar view (go 26)
 - 25b Grains elliptic (go 27)
 - 26a Grains very large, clearly strongly triangular in cross section. (~ GERANIUM)
 - 26b Grains smaller (SIZE), weakly triangular in cross section (~ POTENTILLA/DASIPHORA in part)
 - 27a Grains elliptic to weakly circular (~MORPHOTYPE B)
 - 27b Grains elliptic, much longer pole to pole than across equator.
 - 28a Grains with evident protrusions of the pore, colpi short, scarcely noticeable (~ APIACEAE)
- 23b. Apertures colpate
 - 30a Ornamentation psilate (~MORPHOTYPE C).
 - 30b Ornamentation otherwise
 - 31a Ornamentation homobrochate (~ MORPHOTYPE D)
 - 31b Ornamentation otherwise
 - 32a. Ornamentation bacculate, grains large, (~LINUM)
 - 32b. Ornamentation otherwise
 - 33a. Ornamentation of minor cross-corrugated grooves (fossulate) (~CORYDALIS)
 - 33b. Ornamentation of scarcely perceptible irregular features (scabrate) (~RANUNCULUS IN PART)

34a: Annula subtending the apertures – making grains appear more or less triangular; Pollen often with viscin threads (ONAGRACEAE)

34b: Apertures not annulate – grains appear more or less circular (~ERICACEAE)

Morphotype A: Trifolium, Lupinus, Glycrrhiza, Mitella, Geum

Morphotype B: Lupinus, Lathyrus, Potentilla, Androsace, Bistorta, Vicia

Morphotype C: Jeffersonia, Micranthes, Prunus, Delphinium, Androsace, Penstemon, Orthocarpus, Scutellaria, Aquilegia, Castilleja, Draba)

Morphotype D: Salix, Boechera

451 Appendix XX - Models used for Species Distribution Model Ensembles

452 *Generalised Linear Models (GLM)*

453 *Generalised Additive Models (GAM)*

454 Ensemble learning utilizes many sets of trees, each composed of many decisions, to create a single model.

455 Each independent variable (- or *feature*, may) become a node on the tree, a location on the tree where a
456 binary decision will move towards a predicted outcome. Each of the decision tree models which ensemble

457 learning utilizes is a weak models, each of which may suffer due to high variance or bias, but which produce

458 better outcomes than would be expected via random chance. When ensembled these models generate a strong

459 model, a model which should have more appropriately balanced variance and bias and predicts outcomes

460 which are more strongly correlated with the expected values than the individual weak models.

461 *Random Forest (RF)* the training data are continually bootstrap re-sampled, in combination with random

462 subsets of features, to create nodes which attempt to optimally predict a known outcome. A large number of

463 trees are then aggregated, via the most common predictions, to generate a final classification prediction tree.

464 Each individual prediction tree is generated independently of the others.

465 *Boosted Regression Tree (BRT)* An initial tree is grown, and all other trees are derived sequentially from it,

466 as each new tree is grown the errors in responses from the last tree are weighed more heavily so that the

467 model focuses on selecting dependent variables which refine predictions. All response data and predictor

468 variables are kept available to all trees.

- 469 Alarcón, R. (2010). Congruence between visitation and pollen-transport networks in a California plant–
470 pollinator community. *Oikos*, **119**, 35–44. Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0706.2009.17694.x>
- 471
- 472 Allouche, O., Tsoar, A. & Kadmon, R. (2006). Assessing the accuracy of species distribution models:
473 Prevalence, kappa and the true skill statistic (TSS). *Journal of applied ecology*, **43**, 1223–1232.
- 474 Araujo, M.B. & New, M. (2007). Ensemble forecasting of species distributions. *Trends in ecology & evolution*,
475 **22**, 42–47.
- 476 Baker, W.J., Bailey, P., Barber, V., Barker, A., Bellot, S., Bishop, D., Botigué, L.R., Brewer, G., Carruthers,
477 Clarkson, J.J., Cook, J., Cowan, R.S., Dodsworth, S., Epitawalage, N., Françoso, E., Gallego, B.,
478 Johnson, M.G., Kim, J.T., Leempoel, K., Maurin, O., McGinnies, C., Pokorný, L., Roy, S., Stone, M.,
479 Toledo, E., Wickett, N.J., Zuntini, A.R., Eiserhardt, W.L., Kersey, P.J., Leitch, I.J. & Forest, F. (2021).
480 A Comprehensive Phylogenomic Platform for Exploring the Angiosperm Tree of Life. *Systematic Biology*,
481 **71**, 301–319. Retrieved from <https://doi.org/10.1093/sysbio/syab035>
- 482 Barbet-Massin, M., Jiguet, F., Albert, C.H. & Thuiller, W. (2012). Selecting pseudo-absences for species
483 distribution models: How, where and how many? *Methods in ecology and evolution*, **3**, 327–338.
- 484 Barker, D.A. & Arceo-Gómez, G. (2021). Pollen transport networks reveal highly diverse and temporally
485 stable plant–pollinator interactions in an Appalachian floral community. *AoB PLANTS*, **13**. Retrieved
486 from <https://doi.org/10.1093/aobpla/plab062>
- 487 Beattie, A. (1971). A technique for the study of insect-borne pollen. *The Pan-Pacific Entomologist*, **47**, 82.
- 488 Belitz, M.W., Larsen, E.A., Ries, L. & Guralnick, R.P. (2020). The accuracy of phenology estimators for use
489 with sparsely sampled presence-only observations. *Methods in Ecology and Evolution*, **11**, 1273–1285.
- 490 Bergman, P., Molau, U. & Holmgren, B. (1996). Micrometeorological impacts on insect activity and plant
491 reproductive success in an alpine environment, Swedish Lapland. *Arctic and Alpine research*, **28**, 196–202.
- 492 Bingham, R.A. & Orthner, A.R. (1998). Efficient pollination of alpine plants. *Nature*, **391**, 238–239.
- 493 Bolger, A. & Giorgi, F. (2014). Trimmomatic: A flexible read trimming tool for Illumina NGS data.
494 *Bioinformatics*, **30**, 2114–2120.
- 495 Brosi, B.J. & Briggs, H.M. (2013). Single pollinator species losses reduce floral fidelity and plant reproductive
496 function. *Proceedings of the National Academy of Sciences*, **110**, 13044–13048.
- 497 Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. & Madden, T.L. (2009).
498 BLAST+: Architecture and applications. *BMC bioinformatics*, **10**, 1–9.
- 499 Cameron, S.A. & Sadd, B.M. (2020). Global trends in bumble bee health. *Annual review of entomology*, **65**,
500 209–232.
- 501 CaraDonna, P.J., Burkle, L.A., Schwarz, B., Resasco, J., Knight, T.M., Benadi, G., Blüthgen, N., Dormann,

- 502 C.F., Fang, Q., Fründ, J. & others. (2021). Seeing through the static: The temporal dimension of
503 plant–animal mutualistic interactions. *Ecology Letters*, **24**, 149–161.
- 504 CaraDonna, P.J., Petry, W.K., Brennan, R.M., Cunningham, J.L., Bronstein, J.L., Waser, N.M. & Sanders,
505 N.J. (2017). Interaction rewiring and the rapid turnover of plant–pollinator networks. *Ecology letters*, **20**,
506 385–394.
- 507 Chao, A., Gotelli, N.J., Hsieh, T.C., Sande, E.L., Ma, K.H., Colwell, R.K. & Ellison, A.M. (2014). Rarefaction
508 and extrapolation with hill numbers: A framework for sampling and estimation in species diversity studies.
509 *Ecological Monographs*, **84**, 45–67.
- 510 Cheng, S., Melkonian, M., Smith, S.A., Brockington, S., Archibald, J.M., Delaux, P.-M., Li, F.-W., Melkonian,
511 B., Mavrodiev, E.V., Sun, W., Fu, Y., Yang, H., Soltis, D.E., Graham, S.W., Soltis, P.S., Liu, X., Xu, X.
512 & Wong, G.K.-S. (2018). 10KP: A phylogenetic genome sequencing plan. *GigaScience*, **7**. Retrieved from
513 <https://doi.org/10.1093/gigascience/giy013>
- 514 Coissac, E., Hollingsworth, P.M., Lavergne, S. & Taberlet, P. (2016). From barcodes to genomes: Extending
515 the concept of DNA barcoding.
- 516 Coissac, E., Riaz, T. & Puillandre, N. (2012). Bioinformatic challenges for DNA metabarcoding of plants and
517 animals. *Molecular ecology*, **21**, 1834–1847.
- 518 Colla, S.R., Gadallah, F., Richardson, L., Wagner, D. & Gall, L. (2012). Assessing declines of north american
519 bumble bees (*bombus* spp.) Using museum specimens. *Biodiversity and Conservation*, **21**, 3585–3595.
- 520 Doyle, J.J. & Doyle, J.L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue.
521 *Phytochemical Bulletin*, **19**, 11–15.
- 522 Elith*, J., H. Graham*, C., P. Anderson, R., Dudik, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann,
523 F., R. Leathwick, J., Lehmann, A. & others. (2006). Novel methods improve prediction of species'
524 distributions from occurrence data. *Ecography*, **29**, 129–151.
- 525 Fazekas, A.J., Kesanakurti, P.R., Burgess, K.S., Percy, D.M., Graham, S.W., Barrett, S.C., Newmaster, S.G.,
526 Hajibabaei, M. & Husband, B.C. (2009). Are plant species inherently harder to discriminate than animal
527 species using DNA barcoding markers? *Molecular Ecology Resources*, **9**, 130–139.
- 528 Frase, Barbara A. & Buck, P. (2007). Vascular Plants of the Gothic Area. Retrieved from https://www.digitallibrarystatus.org/wp-content/uploads/2016/05/vascularplantlist_20071.pdf
- 529 Gage, E. & Cooper, D.J. (2013). Historical range of variation assessment for wetland and riparian ecosystems,
530 u.s. Forest service rocky mountain region
- 531 Goulson, D., Lye, G. & Darvill, B. (2008). The decline and conservation of bumblebees. *Annual review of
532 entomology*, **53**, 191–208.
- 533 Govaerts, R., Nic Lughadha, E., Black, N., Turner, R. & Paton, A. (2021). The world checklist of vascular

- 535 plants, a continuously updated resource for exploring global plant diversity. *Scientific Data*, **8**, 1–10.
- 536 Group, C.P.W., Hollingsworth, P.M., Forrest, L.L., Spouge, J.L., Hajibabaei, M., Ratnasingham, S., Bank,
537 M. van der, Chase, M.W., Cowan, R.S., Erickson, D.L. & others. (2009). A DNA barcode for land plants.
538 *Proceedings of the National Academy of Sciences*, **106**, 12794–12797.
- 539 Group, C.P.B., Li, D.-Z., Gao, L.-M., Li, H.-T., Wang, H., Ge, X.-J., Liu, J.-Q., Chen, Z.-D., Zhou, S.-L.,
540 Chen, S.-L. & others. (2011). Comparative analysis of a large dataset indicates that internal transcribed
541 spacer (ITS) should be incorporated into the core barcode for seed plants. *Proceedings of the National
542 Academy of Sciences*, **108**, 19641–19646.
- 543 Hebert, P.D., Cywinski, A., Ball, S.L. & DeWaard, J.R. (2003). Biological identifications through DNA
544 barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **270**, 313–321.
- 545 Hengl, T., Mendes de Jesus, J., Heuvelink, G.B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A.,
546 Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B. & others. (2017). SoilGrids250m:
547 Global gridded soil information based on machine learning. *PLoS one*, **12**, e0169748.
- 548 Hennig, C. (2020). *Fpc: Flexible procedures for clustering*. Retrieved from <https://CRAN.R-project.org/package=fpc>
- 549 Hsieh, T.C., Ma, K.H. & Chao, A. (2020). *iNEXT: Interpolation and extrapolation for species diversity*.
550 Retrieved from http://chao.stat.nthu.edu.tw/wordpress/software_download/
- 551 Iler, A.M., Humphrey, P.T., Ogilvie, J.E. & CaraDonna, P.J. (2021). Conceptual and practical issues limit
552 the utility of statistical estimators of phenological events. *Ecosphere*, **12**, e03828.
- 553 Janzen, D.H. (1967). Synchronization of sexual reproduction of trees within the dry season in central america.
554 *Evolution*, **21**, 620–637.
- 555 Janzen, D.H., Burns, J.M., Cong, Q., Hallwachs, W., Dapkey, T., Manjunath, R., Hajibabaei, M., Hebert, P.D.
556 & Grishin, N.V. (2017). Nuclear genomes distinguish cryptic species suggested by their DNA barcodes
557 and ecology. *Proceedings of the National Academy of Sciences*, **114**, 8313–8318.
- 558 Johnson, M.G., Gardner, E.M., Liu, Y., Medina, R., Goffinet, B., Shaw, A.J., Zerega, N.J. & Wickett,
559 N.J. (2016). HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput
560 sequencing reads using target enrichment. *Applications in plant sciences*, **4**, 1600016.
- 561 Johnson, M.G., Pokorny, L., Dodsworth, S., Botigue, L.R., Cowan, R.S., Devault, A., Eiserhardt, W.L.,
562 Epitawalage, N., Forest, F., Kim, J.T. & others. (2019). A universal probe set for targeted sequencing of
563 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Systematic biology*, **68**,
564 594–606.
- 565 Kress, W.J. & Erickson, D.L. (2007). A two-locus global DNA barcode for land plants: The coding rbcL
566 gene complements the non-coding trnH-psbA spacer region. *PLoS one*, **2**, e508.
- 567

- 568 Kuhn, M. (2022). *Caret: Classification and regression training*. Retrieved from <https://CRAN.R-project.org/package=caret>
- 569
- 570 Lewin, H.A., Richards, S., Aiden, E.L., Allende, M.L., Archibald, J.M., Bálint, M., Barker, K.B., Baumgartner, B., Belov, K., Bertorelle, G., Blaxter, M.L., Cai, J., Caperello, N.D., Carlson, K., Castilla-Rubio, J.C., Chaw, S.-M., Chen, L., Childers, A.K., Coddington, J.A., Conde, D.A., Corominas, M., Crandall, K.A., Crawford, A.J., DiPalma, F., Durbin, R., Ebenezer, T.E., Edwards, S.V., Fedrigo, O., Flieck, P., Formenti, G., Gibbs, R.A., Gilbert, M.T.P., Goldstein, M.M., Graves, J.M., Greely, H.T., Grigoriev, I.V., Hackett, K.J., Hall, N., Haussler, D., Helgen, K.M., Hogg, C.J., Isobe, S., Jakobsen, K.S., Janke, A., Jarvis, E.D., Johnson, W.E., Jones, S.J.M., Karlsson, E.K., Kersey, P.J., Kim, J.-H., Kress, W.J., Kuraku, S., Lawniczak, M.K.N., Leebens-Mack, J.H., Li, X., Lindblad-Toh, K., Liu, X., Lopez, J.V., Marques-Bonet, T., Mazard, S., Mazet, J.A.K., Mazzoni, C.J., Myers, E.W., O'Neill, R.J., Paez, S., Park, H., Robinson, G.E., Roquet, C., Ryder, O.A., Sabir, J.S.M., Shaffer, H.B., Shank, T.M., Sherkow, J.S., Soltis, P.S., Tang, B., Tedersoo, L., Uliano-Silva, M., Wang, K., Wei, X., Wetzer, R., Wilson, J.L., Xu, X., Yang, H., Yoder, A.D. & Zhang, G. (2022). The earth BioGenome project 2020: Starting the clock. *Proceedings of the National Academy of Sciences*, **119**, e2115635118. Retrieved from <https://www.pnas.org/doi/abs/10.1073/pnas.2115635118>
- 571
- 572
- 573
- 574
- 575
- 576
- 577
- 578
- 579
- 580
- 581
- 582
- 583
- 584
- 585
- 586
- 587
- 588
- 589
- 590
- 591
- 592
- 593
- 594
- 595
- 596
- 597
- 598
- 599
- 600
- Life Project Consortium, D.T. of, Blaxter, M., Mieszkowska, N., Palma, F.D., Holland, P., Durbin, R., Richards, T., Berriman, M., Kersey, P., Hollingsworth, P., Wilson, W., Twyford, A., Gaya, E., Lawniczak, M., Lewis, O., Broad, G., Howe, K., Hart, M., Flieck, P. & Barnes, I. (2022). Sequence locally, think globally: The darwin tree of life project. *Proceedings of the National Academy of Sciences*, **119**, e2115642118. Retrieved from <https://www.pnas.org/doi/abs/10.1073/pnas.2115642118>
- Liu, J., Shi, L., Han, J., Li, G., Lu, H., Hou, J., Zhou, X., Meng, F. & Downie, S.R. (2014). Identification of species in the angiosperm family apiaceae using DNA barcodes. *Molecular ecology resources*, **14**, 1231–1238.
- Li, X., Yang, Y., Henry, R.J., Rossetto, M., Wang, Y. & Chen, S. (2015). Plant DNA barcoding: From gene to genome. *Biological Reviews*, **90**, 157–166.
- Lu, J., Breitwieser, F.P., Thielen, P. & Salzberg, S.L. (2017). Bracken: Estimating species abundance in metagenomics data. *PeerJ Computer Science*, **3**, e104.
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. & Hornik, K. (2022). *Cluster: Cluster analysis basics and extensions*. Retrieved from <https://CRAN.R-project.org/package=cluster>
- Maitner, B. (2022). *BIEN: Tools for accessing the botanical information and ecology network database*. Retrieved from <https://CRAN.R-project.org/package=BIEN>
- McLay, T.G., Birch, J.L., Gunn, B.F., Ning, W., Tate, J.A., Nauheimer, L., Joyce, E.M., Simpson, L., Schmidt-Lebuhn, A.N., William J & others. (2021). New targets acquired: Improving locus recovery from

- 601 the Angiosperms353 probe set. *Applications in plant sciences*, **9**.
- 602 Naimi, B. & Araujo, M.B. (2016). Sdm: A reproducible and extensible r platform for species distribution
603 modelling. *Ecography*, **39**, 368–375.
- 604 Naimi, B., Hamm, N. a.s., Groen, T.A., Skidmore, A.K. & Toxopeus, A.G. (2014). Where is positional
605 uncertainty a problem for species distribution modelling. *Ecography*, **37**, 191–203.
- 606 Newstrom, L.E., Frankie, G.W. & Baker, H.G. (1994). A new classification for plant phenology based on
607 flowering patterns in lowland tropical rain forest trees at la selva, costa rica. *Biotropica*, **26**, 141–159.
- 608 Occdownload Gbif.Org. (2021). Occurrence download. Retrieved from <https://www.gbif.org/occurrence/download/0206948-200613084148143>
- 609 Oksanen, J., Simpson, G.L., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Solymos,
610 P., Stevens, M.H.H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho,
611 G., Chirico, M., De Caceres, M., Durand, S., Evangelista, H.B.A., FitzJohn, R., Friendly, M., Furneaux,
612 B., Hannigan, G., Hill, M.O., Lahti, L., McGlinn, D., Ouellette, M.-H., Ribeiro Cunha, E., Smith, T.,
613 Stier, A., Ter Braak, C.J.F. & Weedon, J. (2022). *Vegan: Community ecology package*. Retrieved from
614 <https://CRAN.R-project.org/package=vegan>
- 615 Oliver, P.M., Adams, M., Lee, M.S., Hutchinson, M.N. & Doughty, P. (2009). Cryptic diversity in vertebrates:
616 Molecular data double estimates of species diversity in a radiation of australian lizards (diplodactylus,
617 gekkota). *Proceedings of the Royal Society B: Biological Sciences*, **276**, 2001–2007.
- 618 Omernik, J.M. (1987). Ecoregions of the conterminous united states. *Annals of the Association of American
619 geographers*, **77**, 118–125.
- 620 Pearse, W.D., Davis, C.C., Inouye, D.W., Primack, R.B. & Davies, T.J. (2017). A statistical estimator for
621 determining the limits of contemporary and historic phenology. *Nature Ecology & Evolution*, **1**, 1876–1882.
- 622 Prim, R.C. (1957). Shortest connection networks and some generalisations. *Bell System Technical Journal*,
623 **36**, 1389–1401.
- 624 Qiao, H., Soberon, J. & Peterson, A.T. (2015). No silver bullets in correlative ecological niche modelling:
625 Insights from testing among many potential algorithms for niche estimation. *Methods in Ecology and
626 Evolution*, **6**, 1126–1136.
- 627 Robinson, N., Regetz, J. & Guralnick, R.P. (2014). EarthEnv-DEM90: A nearly-global, void-free, multi-
628 scale smoothed, 90m digital elevation model from fused ASTER and SRTM data. *ISPRS Journal of
629 Photogrammetry and Remote Sensing*, **87**, 57–67.
- 630 Ruppert, K.M., Kline, R.J. & Rahman, M.S. (2019). Past, present, and future perspectives of environmental
631 DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global
632 eDNA. *Global Ecology and Conservation*, **17**, e00547.

- 634 Sarro, E., Tripodi, A. & Woodard, S.H. (2022). Bumble bee (*bombus vosnesenskii*) queen nest searching
635 occurs independent of ovary developmental status. *Integrative Organismal Biology*, **4**, obac007.
- 636 Tange, O. (2021). GNU parallel 20220322 (savannah). Retrieved from <https://doi.org/10.5281/zenodo.6377950>
- 637 Tran, H., Nguyen, P., Ombadi, M., Hsu, K., Sorooshian, S. & Qing, X. (2019). A cloud-free MODIS snow
638 cover dataset for the contiguous united states from 2000 to 2017. *Scientific data*, **6**, 1–13.
- 639 Wang, T., Hamann, A., Spittlehouse, D. & Carroll, C. (2016). Locally downscaled and spatially customizable
640 climate data for historical and future periods for north america. *PloS one*, **11**, e0156720.
- 641 Wenzell, K.E., McDonnell, A.J., Wickett, N.J., Fant, J.B. & Skogen, K.A. (2021). Incomplete reproductive
642 isolation and low genetic differentiation despite floral divergence across varying geographic scales in
643 *castilleja*. *American Journal of Botany*, **108**, 1270–1288.
- 644 Williams, P.H. (1982). The distribution and decline of british bumble bees (*bombus latr.*). *Journal of
645 Apicultural Research*, **21**, 236–245. Retrieved from <https://doi.org/10.1080/00218839.1982.11100549>
- 646 Wilson, A.M. & Jetz, W. (2016). Remotely sensed high-resolution global cloud dynamics for predicting
647 ecosystem and biodiversity distributions. *PLoS biology*, **14**, e1002415.
- 648 Wood, D.E., Lu, J. & Langmead, B. (2019). Improved metagenomic analysis with kraken 2. *Genome biology*,
649 **20**, 1–13.
- 650 Zhao, Y.-H., Lázaro, A., Ren, Z.-X., Zhou, W., Li, H.-D., Tao, Z.-B., Xu, K., Wu, Z.-K., Wolfe, L.M., Li,
651 D.-Z. & Wang, H. (2019). The topological differences between visitation and pollen transport networks:
652 A comparison in species rich communities of the himalaya–hengduan mountains. *Oikos*, **128**, 551–562.
653 Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1111/oik.05262>

654 **List of Figures**

| | | |
|------------|---------------------|----|
| 655 1 | A caption | 41 |
| 656 2 | A caption | 42 |

Biotically pollinated plant genera
with morphological or molecular data

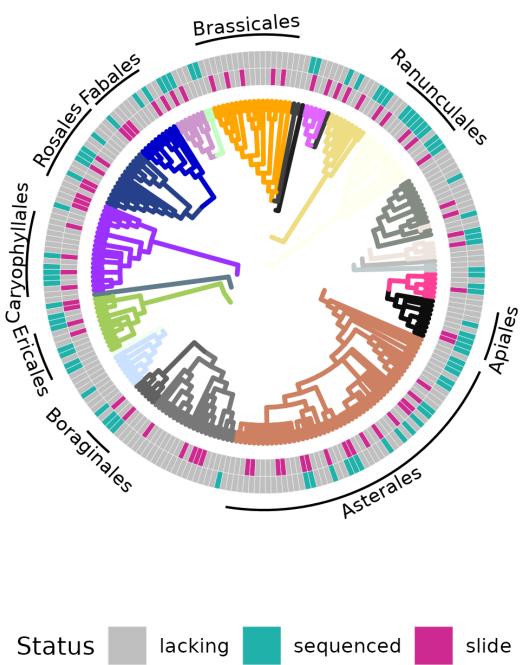


Figure 1: A caption

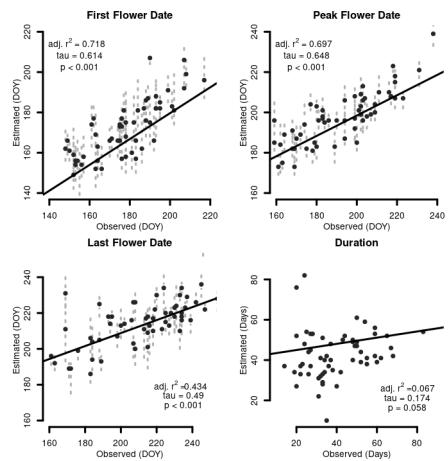


Figure 2: A caption

657 **List of Tables**

| | | |
|------------|--|----|
| 658 2 | Logistic regression assessing accuracy of SDMs | 44 |
| 659 3 | SDM evaluation contingency table | 45 |

Table 2: Logistic regression assessing accuracy of SDMs

| Metric | Value | Metric | Value |
|---------------------|-------|-------------|-------|
| Accuracy (Training) | 83.75 | F-Score | 0.84 |
| Accuracy (Test) | 84.00 | AUC | 0.92 |
| Recall | 81.03 | Concordance | 0.92 |
| True Neg. Rate | 86.97 | Discordance | 0.08 |
| Precision | 88.04 | Tied | 0.00 |

Table 3: SDM evaluation contingency table

| | | Training | | Testing | |
|---------|----------|----------|----------|---------|----------|
| | | Absence | Presence | Absence | Presence |
| Absence | Absence | 25620 | 3838 | 11130 | 1653 |
| | Presence | 6614 | 28248 | 2758 | 12024 |