

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^{3,4}, Emily J. Woodworth^{1,2}
Sophie Taddeo^{1,2}, Paul J. CaraDonna^{1,2,3}, Jeremie B. Fant^{1,2}

¹Chicago Botanic Garden, 1000 Lake Cook Road, Glencoe, Illinois 60022, USA

²Plant Biology and Conservation, Northwestern University, Evanston, Illinois 60208, USA

³ Rocky Mountain Biological Laboratory, P.O. Box 519, Crested Butte, Colorado 81224, USA

⁴ Australia post fire project, —, —, —, AU

⁴ **Abstract**

⁵ 1) DNA Barcoding has been remarkably successful in nearly all kingdoms of life and has allowed
⁶ for the rapid analysis of ecological assemblages. Successful DNA barcoding in the plant kingdom
⁷ has been more difficult than other kingdoms. Due to this understanding plants in ecological
⁸ contexts and understandings of their synecology in some instances may begin to lag behind other
⁹ kingdoms.

¹⁰ 2) Here we utilize hyb-seq, museum studies, and species distribution modelling, to detect the plant
¹¹ species present in pollen loads collected from Queen Bumble Bees.

¹² 3) By utilizing Species distribution modelling we allow for one to process hyb-seq data, create user
¹³ specified sequence databases which may use MORE ACCURATE alignment algorithms on personal
¹⁴ computers over realistic time periods.

¹⁵ 4) We show that hyb-seq using the Angiosperms 353 probes, which are currently being used in the
¹⁶ largest ever plant systematic endeavor, offers significant promise to metagenomic approaches in
¹⁷ real world scenarios.

²⁰ *Correspondence: reedbenkendorf2021@u.northwestern.edu

21 5) We conclude that these probes offer promise for the identification of plant tissue in both single
22 sample, and metasample contexts.

23 **1 | INTRODUCTION**

24 The inability to reliably identify plants down to species can limit our understanding of ecosystem function
25 and interactions (Bortolus (2008)). This is especially true for genera where species are defined based upon
26 ecological and behavioral rather than morphological properties, and hence can serve as key habitat bioindi-
27 cators (e.g. different species of Sagebrush- *Artemisia* L., Willows - *Salix* L., and Sedges - *Carex* L.) (Gage &
28 Cooper (2013)). The lack of species level data can hinder our understanding of the breadth of habitat which
29 some species occupy, and the interactions they have with other species. Current methods to ameliorate this
30 situation include: ignoring these ecologically relevant levels of detail, revisiting plots as diagnostic mate-
31 rial becomes temporally available, assistance from taxonomic specialists, or the use of barcoding or other
32 molecular techniques.

33 The identification to species is often mired by lack of diagnostic characters (e.g. flowers, fruits, roots or
34 combinations thereof), an increasing lack of taxonomic experts (Hebert *et al.* (2003)) and increasingly the
35 description of cryptic species (Janzen *et al.* (2017), Oliver *et al.* (2009)). And revisiting field sites to identify
36 material using morphological or chemical approaches, can be resource intensive and often does not work.

37 Recently barcoding (the identification of a sample from a single organism *e.g.* a piece of leaf), and metabar-
38 coding (the identification of a sample containing a mix of organisms *e.g.* soil), have shown considerable
39 promise in all Kingdoms of life (Ruppert *et al.* (2019)). With plants the identification of members of cer-
40 tain clades using barcoding has been quite successful (Kress (2017)), whereas with other clades results have
41 been elusive (Liu *et al.* (2014), Group *et al.* (2011), Coissac *et al.* (2012)), however metabarcoding incurs
42 additional challenges to those which exist for the currently implemented barcodes (Li *et al.* (2015), Kress &
43 Erickson (2007), Group *et al.* (2009), Coissac *et al.* (2012)). Particular challenges with barcodes include the
44 utilization of high-copy number sequences are associated with their rates of divergence, gene tree conflict,
45 and hybridization (Coissac *et al.* (2016), Fazekas *et al.* (2009)).

46 Currently the largest plant systematic endeavor ever undertaken, the Kew Plant and Fungal Tree of Life
47 (PAFTOL), is approaching completion (Baker *et al.* (2021a)). This data set will contain hybridization
48 capture (Hyb-Seq) data from at least one species representing each genus in the plant kingdom using the
49 popular A353 probes (Baker *et al.* (2021a)), resulting in over 14,000 represented species. These publicly
50 available data serve to provide a taxonomically comprehensive backbone for plant metabarcoding, and the

51 A353 probes are currently being used in many other plant phylogenetic issues increasing the sampling depth
52 of many clades (Baker *et al.* (2021b)). Data from the 10kP project, which seeks to develop reference
53 genomes from a phylogenetically diverse suite of plants will contribute many more records upon it's intended
54 completion, now slated to be by 2030, similar projects which seek to sequence high amounts of genomes in
55 regions e.g. the 'Darwin Tree of Life' are being undertaken which will contribute data applicable to enormous
56 spatial domains (Cheng *et al.* (2018), Life Project Consortium *et al.* (2022), Lewin *et al.* (2022)). These
57 data will promote the ability to apply metabarcoding to resolve a diversity of questions relevant to theoretical
58 and applied ecology (Kress (2017)). However, the application of metabarcoding still face challenges relating
59 to the enormity of the genomic data sets and the computational power required to process sequence data.
60 Herein we have resolved major components of the problems of identifying plant material without diagnostic
61 morphological character states using the Angiosperms353 (A353) Hyb-Seq probes (Johnson *et al.* (2019)),
62 and custom species sequence databases derived via species distribution modelling, and temporal filtering.

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

77 Considerable amounts of species interactions vary along time (CaraDonna *et al.* (2021)). For the tropics
78 and subtropics, contrasts in the flowering periods of many plant species, can provide an additional filter for
79 identifying material in many metagenomic samples (Janzen (1967), Newstrom *et al.* (1994)). In temperate
80 regions, pollination interactions also vary temporally (CaraDonna *et al.* (2017)), however the overall shorter
81 extent of the active growing season in these systems results in the presence of few to any natural breaks
82 in these systems which reduces the utility of these to operate as filters in the post-processing of sequence

83 matches. Nonetheless, we work through a process which seems applicable to the tropics and subtropics to
84 utilize the temporal dimension for classifying sequencing results.

85 To test these metagenomic and informatics approaches to determine whether the foraging record of Queen
86 Bumble Bee's is consistent across direct observations and the pollen record, an incongruency in several floral
87 visitation networks involving smaller bodied fauna (Barker & Arceo-Gomez (2021), Zhao *et al.* (2019),
88 Alarcón (2010)). The assessment of the plant species compositions of pollen is a desired results, with several
89 applications, and numerous complications (Poron *et al.* (2017), Bell *et al.* (2017), Sickel *et al.* (2015),
90 Bell *et al.* (2019), Suchan *et al.* (2019), Johnson *et al.* (2021)). The two foraging phases of the Queen
91 Bumble Bee life cycle is essential to 1) increase their weight before diapause, 2) increase their ovary weights
92 while establishing their recently found nests, both of these time periods represent potential demographic
93 bottlenecks in bumble bee populations (Sarro *et al.* (2022)). Bumblebees are one of the only groups of insects
94 with unequivocal quantitative evidence for numerous populations declines, while simultaneously serving as
95 the most effective pollinators in temperate montane ecosystems (Cameron & Sadd (2020), Goulson *et al.*
96 (2008), Williams (1982), Colla *et al.* (2012), Bergman *et al.* (1996), Bingham & Orthner (1998)). Montane
97 areas often represent the most diverse areas in the temperate and oftentimes offer the sole potential refugia
98 for multiple dimensions of biodiversity under climate change, whilst simultaneously experiencing the greatest
99 proportional changes in mean annual temperature (Brito-Morales *et al.* (2018), Pepin *et al.* (2022)). An
100 immediate understanding of how to manage previously overlooked keystone insect species, such as bumble
101 bees, is essential if the refugial potential of the temperate mountains are to be utilized while maintaining
102 their current diversity (Loarie *et al.* (2009), Dobrowski & Parks (2016))

103 2 | METHODS

104 Study System & Field Work

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

112 **2.1 | Spatial Analyses**

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

119 In order to minimise the number of species for which SDM’s were to be generated, BIEN was queried at
120 a distance of up to 100km from our study area and all plant species records were downloaded. In order
121 to emulate the stochasticity of botanical collecting, this data set was bootstrap re-sampled 250 times, with
122 90% of samples selected, to create a testing data set. The median of the logistic regression assessing the
123 probability of occurrence of a species record as a function of distance from the study area was used as a
124 threshold distance, under which, to include species as candidates for distribution modelling.

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

138 We then generated 4,029 absence points , locations where the focal taxon is anticipated missing, through a
139 random stratification of 19% of the land cover in the area and included them in (BLM CITATION - need
140 appropriate format for journal). To achieve a larger absence data set, we generated 1,000 pseudo-absence
141 records for each taxon by randomly selecting coordinates located at least 10km away from an occurrence
142 record. For ML models, these pseudo-absences were reduced so that the ratio of presence to absence records

¹⁴³ were balanced (Barbet-Massin *et al.* (2012)). To achieve this, we removed absence records inside of 10% of
¹⁴⁴ the mean sample value of the presence records; the required number of absence records were then randomly
¹⁴⁵ sampled.

¹⁴⁶ We used 26 environmental variables at 30m resolution to predict the potential distribution of each species,
¹⁴⁷ six related to climate, five soil, four topographic, four related to cloud cover, with the remaining reflecting
¹⁴⁸ assorted abiotic parameters (Wilson & Jetz (2016), Wang *et al.* (2016), Hengl *et al.* (2017), Robinson *et al.*
¹⁴⁹ (2014)) (**APPENDIX 6**). **These publicly available data sets, were selected as they . . .** For linear
¹⁵⁰ regression models these predictors underwent both *vifstep* (theta = 10, max observations = 12,500) and
¹⁵¹ *vifcor* (theta = 0.7, max observations = 12,500) to detect highly correlated variables, and collinear features
¹⁵² were removed leaving 16 variables (Naimi *et al.* (2014)).

¹⁵³ Modelling: Random Forest and Boosted Regression Trees, were sub sampled with 30% test and two replicates
¹⁵⁴ each before weighted ensemble based on True Skill Statistics (tss) (Naimi & Araujo (2016)). Generalised
¹⁵⁵ linear models (GLM) and Generalised additive models (GAM) with 30% sub sampling and three replicates
¹⁵⁶ each were also ensembled using the tss (Naimi & Araujo (2016), @). TSS was chosen as the ensemble
¹⁵⁷ criterion as it has been shown to work across a range of species occurrences prevalence (Allouche *et al.*
¹⁵⁸ (2006)). The results of these models were extracted on a cell-by-cell basis to a polygon feature derived from
¹⁵⁹ a minimum-spanning tree which encompasses the study sites, and species from either ensemble with greater
¹⁶⁰ than 50% mean habitat suitability across all cells were considered present for further purposes (Prim (1957)).

¹⁶¹ 535 species were modelled using Generalized Linear Models and Generalized Additive Models. 534 species
¹⁶² were modelled using Random Forest and Boosted Regression Trees. To evaluate the accuracy of the species
¹⁶³ distribution models, additional presence records from GBIF (n = 61,789), and AIM (n = 12,730) were used as
¹⁶⁴ test and training sets (n = 74,519) for logistic regression (Ocdownload Gbif.Org (2021), Land Management
¹⁶⁵ (2020)). Additional novel absence records were generated from the AIM data set to create a data set where
¹⁶⁶ each species has balanced presence and absences. 11 or more paired presence and absence records were
¹⁶⁷ required for this testing, resulting in 334 species being included in the logistic regression (Mdn = 110.0, \bar{x} =
¹⁶⁸ 223.1, max = 1568 record pairs used) with a 70% test split (Kuhn (2022)).

¹⁶⁹ 2.2 | Molecular Lab Work

¹⁷⁰ All lab work was carried out at The Daniel F. and Ada L. Rice Plant Conservation Science Center at the
¹⁷¹ Chicago Botanic Garden, Glencoe, Illinois, U.S.A.

172 **2.2.1 | Reference Plant Library Generation** Using five years of observational data on *Bombus* Queen
173 Bee foraging at these studies sites, we identified the plant taxa most frequently visited by Queens across
174 all years. We sequenced the 12 most commonly visited taxa twice using samples from one site within the
175 Gunnison River Drainage and one individual from another population. In addition, for any of these 12 focal
176 species which did not have a congener pair in this filtered sample, we included a congener - or a species from a
177 closely related genus to serve as an outgroup. We also sequenced another 15 abundant taxa commonly visited
178 by *Bombus* workers, based on the aforementioned data set (*APPENDIX 4*). Plant collections were identified
179 via a variety, and typically a combination, of dichotomous keys and primary literature as required (Flora of
180 North America Editorial Committee (1993+), Hitchcock & Cronquist (2018), Ackerfield (2015), Lesica *et al.*
181 (2012), Cronquist *et al.* (1977+), Allred & Ivey (2012), *Jepson flora project* (2020), Mohlenbrock (2002)).

182 **2.2.2 | Plant Genomic DNA Extraction** Plant genomic DNA was isolated from ~ 1 cm² of leaf tissue
183 from silica-gel dried or herbarium material using a modified cetyltrimethylammonium (CTAB) protocol
184 (Doyle & Doyle (1987)) that included two chloroform washes. DNA was quantified using a Nanodrop 2000
185 (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and Qubit fluorometer (Thermo Fisher Scientific).

186 **2.2.3 | Pollen Genomic DNA Extraction** Pollen genomic DNA was extracted from corbiculae using a
187 CTAB based protocol modified from Lahlamgiah et al. and Guertler et al. (2014, 2014). A SDS extraction
188 buffer (350µL , 100mM Tris-HCl, 50 mM EDTA, 50 mM NaCl, 10% SDS v/v., pH 7.5) was added followed by
189 vortexing to allow dissolution of corbiculae. Pollen grains were then macerated with Kontes Pellet Pestles,
190 and the tip of these washed with 130 µL of the SDS extraction buffer, samples were then incubated for
191 1 hour at 30°C. This was followed by the addition of 10% CTAB solution (450ul, of 20 mM Tris-Cl pH.
192 8.0, 1.4 M NaCl, 10 mM EDTA pH 7.5, 10% CTAB, 5% PVP, ~85% Deionized water) and RNase (10
193 uL of 10 mg/mL) and samples were incubated for 40 minutes at 37°C, on heat block (Multi-Blok, Thermo
194 Fisher Scientific, Waltham Massachusetts) set to 40°C. After 20 minutes incubation, Proteinase K (15 µL of
195 20mg/ml) and DTT (12.5 µL of 1M in water) were added, and the samples were further incubated at 60°C
196 for 1 hour. Samples were then incubated overnight at 40°C. 500 µL of Phenol-Chloroform-Isoamyl alcohol
197 (25:24:1) were added, vortexed, and centrifuged at 10,000 rpm for 10 minutes and the aqueous phase was
198 pipetted to a 1.5 ml centrifuge tube.

199 To precipitate the DNA, chilled Isopropyl alcohol & 3 mM Sodium acetate (5:1) equivalent to 2/3 of the
200 volume of sample were added, with 1 hour of chilling at -20°C, followed by 10 minutes of centrifuging at
201 13,000 rpm. The supernatant was pipetted to a new 1.5 ml centrifuge tube, and 70% EtOH (400 µL) were

202 added before chilling at -20°C for 20 minutes followed by centrifugation at 13,000 rpm for 10 minutes. Both
203 tubes were then washed with 75% EtOH (400 µL), inverted, centrifuged at 13,000 rpm for 4 minutes, and
204 the solution discarded, then washed with 95% EtOH (400 µL) , inverted, centrifuged at 13,000 rpm for 4
205 minutes, and the solution discarded. Pellets were dried at room temperature overnight before resuspension
206 in Nuclease free H₂O. Extractions were assessed using a Nanodrop 2000 (Thermo Fisher Scientific) and
207 Qubit fluorometer (Thermo Fisher Scientific). DNA extracts were then cleaned using 2:1 v./v. Sera-Mag
208 beads (Cytiva, Little Chalfont, UK) to solute following the manufacturer's protocol, eluted in 0.5x TE, and
209 the eluent allowed to reduce by half volume in ambient conditions. DNA was quantified using a Qubit
210 fluorometer.

211 **2.2.4 | Fragmentation, Library Preparation & Target Enrichment** Library preparation was per-
212 formed using the NEBNext Ultra II FS-DNA Library Prep Kit for Illumina (New England BioLabs, Ipswich,
213 Massachusetts, USA) using slightly modified manufacturers recommendation. Fragmentation was performed
214 at ½ volume of reagents and ¼ enzyme mix for 40 minutes at 37°C, with an input of 500 ng cleaned DNA.
215 Adapter Ligation and PCR enrichment were performed with ½ volumes, while cleanup of products was
216 performed with ½ volume of SPRI beads (Beckman Coulter, Indianapolis, Indiana, USA) and recommended
217 volumes of 80% v./v. ethanol washes. The exception was the herbarium specimens which were not frag-
218 mented and only end repaired, with similar library preparation of all samples. Products were analysed on
219 4% agarose gels, and a Qubit fluorometer. Libraries were pooled and enriched with the Angiosperms 353
220 probe kit V.4 (Arbor Biosciences myBaits Target Sequence Capture Kit) by following the manufacturer's
221 protocol and Brewer et al. 2019. Sequencing was performed using an Illumina mi-Seq with 150-bp end reads,
222 (NUSeq Core, Chicago, Illinois).

223 **2.2.5 | Computational Processes and Analyses.**

224 **2.2.5.1 | Reference Library Data Processing** Sequences were processed using Trimmomatic, which
225 removed sequence adapters, clipped the first 3 bp, discarding reads less than 36 bp, and removing reads
226 if their average PHRED score dropped beneath 20 over a window of 5 bp (Bolger & Giorgi (2014), Tange
227 (2021)). Contigs were generated using HybPiper using target files created by M353 (Johnson *et al.* (2016),
228 McLay *et al.* (2021)).

229 **2.2.5.2 | Sequence Identification** A custom Kraken2 database was created by downloading represen-
230 tative species of each genus indicated as being present in the study area by the spatial analyses from the

231 Sequence Read Archive (SRA) NCBI (Wood *et al.* (2019)). These sequences were processed in the same
232 manner as our novel sequences . The Kraken2 database was built using default parameters. Kraken2 was
233 run on sequences using default parameters (*APPENDIX 5*). Following Kraken2, Bracken was used to clas-
234 sify sequences to terminal taxa (Lu *et al.* (2017)). Results from both Kraken2 and Bracken, results were
235 reclassified manually to identify terminal taxa. For example, when only a single species of a genus was known
236 in the study area, but our database used a representative of another taxon in the genus, this species was
237 coded as the result. The re-coding of sequences from another representative species for the genus to the sole
238 RMBL representative allowed the identification of XX & % more species.

239 **2.2.5.3 | Identification of Sequence Matching Loci** A local NCBI database was built using the same
240 processed novel and downloaded sequences as the previous databases (Camacho *et al.* (2009)).

241 **2.2.5.4 | Morphological Pollen identification**

242 To develop a reference library of pollen grains which may be present in corbiculae loads, an image reference
243 collection of fuchsin-jelly stained (Beattie (1971)) slides was assembled from slides previously prepared by the
244 authors (n = 21), and other researchers (n = 38) (Brosi & Briggs (2013)). Using five years of observational
245 data on *Bombus* Queen Bee foraging at these studies sites (Ogilvie unpublished), as well as the Vascular
246 Plant Checklist (Frase & Buck (2007)), an additional 62 voucher slides for species were prepared and imaged
247 at 400x (Leica DMLB, Leica MC170 HD Camera, Leica Application Suite V. 4.13.0) from non accessioned
248 herbarium collections to supplement the number of species and clades covered (Appendix 3).

249 We used Divisive Hierarchical Clustering techniques to determine which plant taxa were distinguishable via
250 light microscopy, and to develop a dichotomous key to pollen morphotypes. Ten readily discernible categorical
251 traits were collected from each specimen in the image collection. These traits were transformed using Gower
252 distances, and clustered using Divisive Hierarchical clustering techniques (Maechler *et al.* (2022)). Using
253 the cluster dendrogram, elbow plot, and heatmaps (Hennig (2020)), of these results morphological groups
254 of pollen which could not be resolved via microscopy were delineated, and a dichotomous key was prepared
255 (*APPENDIX NO.*). This key was then used to identify the pollen grains sampled from corbiculae loads to
256 morphotypes in a consistent manner. To prepare the pollen slides from corbiculae, all corbiculae loads were
257 broken apart and rolled using dissection needlepoints to increase heterogeneity of samples. *Cerca* 0.5mm²
258 of pollen was placed onto a ~4mm² fuchsin jelly cube (Beattie (1971)) atop a graticulated microscope
259 slide, with 20 transects and 20 rows (400 quadrants) (EMS, Hartfield, PA). The jelly was melted, with
260 stirring, until pollen grains were homogeneously spread across the microscope slide. Slides were sealed with

261 Canada Balsam (Rublev Colours, Willits, CA) followed by sealing with nail polish; all samples are noted in
262 *APPENDIX 3*. To identify the pollen present in corbiculae loads, light microscopy at 400x (Zeiss Axioscope
263 A1) was used. In initial sampling in three transects, each pollen grain was identified to morphotype and
264 counted; an additional two transects were scanned for morphotypes unique to that slide, if either transect
265 contained an unique morphotype than all grains in that transect were also identified and counted. Subsequent
266 to the first round osf sampling, non-parametric species richness rarefaction curves (Oksanen *et al.* (2022)),
267 and non-parametric species diversity rarefaction curves were used to assess the completeness of sampling
268 (Chao *et al.* (2014), Hsieh *et al.* (2020)). Slides not approaching the asymptote of the rarefaction curve
269 were then re-sampled, and analysed iteratively for up to a total of seven transects *APPENDIX 2*.

270 **2.3 | Temporal Analyses**

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

282 **2.4 | Floral Observations**

283 **3 | RESULTS**

284 **3.1 | Spatial Analyses**

285 [Table 1 about here.]

286 [Table 2 about here.]

287 The median (25.009 km) of the logistic regression assessing the probability of occurrence of a species record as
288 a function of distance from the study area was used as a threshold distance to include species for distribution
289 modelling. A 2-sample test for equality of proportions with continuity correction (X-squared = 13.254, df
290 = 1, p-value = 0.000136, 95% CI 0.04-1.00) was used to test whether more of the records located in the
291 broad ecological sites present at the field station, between the distance of the median (25.009 km) to the
292 third quantile (ca 43.830 km) of the regression distance, where true presences at the field station. Including
293 these records would have resulted in modelling an additional 222 species distributions of which 30 are true
294 presences these taxa were not modelled.

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

297 At the field site, of the 554 vascular plants with biotic pollination syndromes, the 493 ML ensembles accu-
298 rately predicted the presence of 362 (65.3%), incorrectly predicted the presence of 64 (11.6%), incorrectly
299 predicted 34 true presences (6.1%) as being absent, and correctly predicted the true absence of 33 (6.0%).
300 The balanced accuracy of the ensembled models is 0.627 (Sensitivity = 0.340, Specificity 0.914). Of the 554
301 vascular plants with biotic pollination syndromes, the 475 LM ensembles accurately predicted the presence
302 of 286 (51.6%), incorrectly predicted the presence of 41 (14.3%), incorrectly predicted 93 true presences
303 (16.8%) as being absent, and correctly predicted the true absence of 55 (9.9%). The balanced accuracy of
304 the ensembled models is 0.664 (Sensitivity = 0.573, Specificity 0.754). Of the 554 vascular plants with biotic
305 pollination syndromes in the flora 13 (2.3%) were in the Orchid family and 41 (7.4%) are non-natives, both
306 of which are restricted from the database, and can only reduce the number of true predicted presences by
307 roughly 10%.

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

313 3.2 | Microscopic Pollen identification

314 Using the fuchsin jelly preparation and light microscopic analyses of grains and scoring of 10 character
315 states resulted in the establishment of 28 morphotypes which grains could be classified into. APPENDIX
316 7. 60 samples were counted and based on rarefaction **had over % of expected morphotypes found**

317 (morphotype richness, $\bar{x} = 4.5$, Mdn = 4, min = 1, max = 9), all samples had expected morphotype diversity
318 reach the asymptote APPENDIX 8. The number of counted pollen grains in each sample range from (MIN
319 - 16,293, $\bar{x} = 2788.685$, Mdn = 1453).

320 [Figure 1 about here.]

321 3.3 | Metabarcoding Pollen identification

322 54 corbiculae loads had DNA extracted and underwent various steps towards hyb-seq, in the end a total of
323 44 corbiculae samples were sequenced, 7,752,353 reads were recovered from sequencing. The number of reads
324 per sequence varied widely (range = 76 - 508,795, $\bar{x} = 176,189.8$, Mdn = 138,395). Of the possible 353 loci,
325 the number which were recovered from each sample, and informative to BLAST were range = 24 - 353, $\bar{x} =$
326 305.5, Mdn = 331. The number of reads per loci from across all samples had a range of 178 - 506,653, $\bar{x} =$
327 20,688, Mdn = 12,616.

328 APPENDIX X Reads Per Loci.

329 After trimming 7,865,680 sequences remained. 10,682,538 reads were matched using Kraken, of the reads
330 classified by Kraken 10,160,768 reads were matched using Bracken, of the reads classified by Kraken 7,302,876
331 reads were matched using BLAST. Based upon subjective review of the three classifiers **APPENDIX X**
332 **MOLECULAR NETWORKS - 3 DIFFERENT ONES**, BLAST was chosen as the classification
333 method which yielded the most probable results, and it's values were used for all subsequent analyses.

334 To determine at which level species in pollen loads could be detected the results of light microscopy were
335 compared to the molecular results. The pollen samples contained three morphotypes which could readily
336 be identified via microscopy. Two of these mapped to the clades (Boraginaceae & Heliantheae Alliance),
337 and one to a Asteraceae less Heliantheae. Boraginaceae grains were detected in 85.7% of samples where the
338 proportion of target grains were between 0.01-1 ($n = 14$ Mdn = 0.572). Asteraceae type 1, non-helianthoids,
339 were detected in 50% of samples where the proportion of target grains were between 0.001-0.01 ($n = 6$ Mdn
340 = 0.002) Asteraceae type 2, Helianthoids, were detected in 62.5% of samples where the proportion of target
341 grains were between 0.001-0.01 ($n = 8$ Mdn = 0.003). Both morphotypes of Asteraceae pollen were detected
342 in 100% of samples where the proportion of target grains were between 0.01-1 ($n = 3$ Mdn = 0.011), and
343 Ericaceae were detected in 50% of samples where the proportion of target grains were between 0.001-0.1 (n
344 = 2 Mdn = 0.01).

345 **3.4 | Temporal Analyses**

346 The first date of modeled snow melt in the Gothic area ($n = 17$, $\bar{x} = 137.9$, Mdn = 135, 3rd quantile = 151),
347 and the first date of a consistent winter snow base ($n = 17$, $\bar{x} = 299.9$, Mdn = 300, 1st quantile = 291) from
348 2000-2017, were used as delimiters for the inclusions of herbarium records in modelling. Of the **500** species
349 predicted likely present in the area via logistic regression, 332 species (64.4%) with more than 10 records in the
350 focal level 4 ecoregions ($\bar{x} = 35.01657$, Mdn = 35, max = 96) had weibull estimates calculated, an additional
351 56 species (11.2%) with enough contributing records from the “Sedimentary Mid-Elevation Forests”, a large
352 ecoregion in general just beneath the elevation bands occupied by the five ecoregions around the study area
353 had weibull estimates also calculated ($\bar{x} = 13.86885$, Mdn = 13, max = 24).

354 Only 58 of these 388 species ($n = 34.56897$, Mdn = 31) were able to be compared to plot based observational
355 data from the long term (1974–2012) data set. Of these species relatively high accord was observed between
356 the long-term ground truthed data set, and the modelled species. There was very strong evidence that the
357 weibull estimates were positively associated with the observed onset ($r^2 = 0.72$, $p < 0.0001$, $\tau = 0.61$)
358 and peak ($r^2 = 0.70$, $p < 0.0001$, $\tau = 0.65$) of flowering, and that the number of herbarium samples had
359 a moderate effect on the estimates ($p = 0.004$ and $p = 0.034$ respectively). There was very strong evidence
360 that the weibull estimates had a positive association with the observed cessation of flowering ($r^2 = 0.4339$, p
361 < 0.0001 , $\tau = 0.489$), however there was no evidence that sample size had an effect ($p = 0.349$). There was
362 moderate evidence that the weibull estimates, with an effect of sample size, had a weak positive association
363 with the observed duration of flowering ($p = 0.0401$, $r^2 = 0.07$, $\tau = 0.17$).

364 [Figure 2 about here.]

365 **3.5 | Floral Observations**

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

³⁷⁴ A total of 66 corbiculae loads were collected from Bees, 64 of them from Queens.

³⁷⁵ [Figure 3 about here.]

³⁷⁶ 3.6 | Integrated Observational, Molecular, and Palynological Network

³⁷⁷ For example a common UNKNOWN sequence mapped to the Asteraceae family, but which was flagged by
³⁷⁸ temporal filters and is present in both *B. nevadensis* and *B. rufocinctus* pollen is most likely *Frasera*, failed
³⁷⁹ extraction. A similar likely mismatch could be between what was fide molecular evidence as *Agastache*
³⁸⁰ *pallidiflora* but where feeding was infrequently observed on *Pedicularis*, likely due to this entire order being
³⁸¹ represented by only a single molecular reference species.

³⁸² It is not unlikely that much of the difference in the results between the observational and molecular work
³⁸³ are attributable to the challenges in detecting rare events in these smaller sizes. For example, no more than
³⁸⁴ 10 bee corbiculae loads per species were sequenced with the Mdn = 7 . . . , and the median of interactions
³⁸⁵ with the top 5 plant sizes constituted 0.8142857 of the top.

³⁸⁶ . . . many of our results indicate foraging on *Viola* spp, zygomorphic flowers with architecture which would
³⁸⁷ require subtle handling and strength to reach the pollen and nectar loads. . . . Or the *Epilobium* sp.
³⁸⁸ results indicating that a species such as *Chamerion angustifolium* or *latifolium* is occasionally utilized, as it
³⁸⁹ supported by limited paylnological data.

³⁹⁰ 4 | DISCUSSION

³⁹¹ We have demonstrated how Angiosperms533 hyb-seq probes may be used for plant barcoding in a metage-
³⁹² nomic context. This was exemplified in an ecologically relevant scenario, where the results have immediate
³⁹³ implications for natural history driven fundamental science and the applied science of land management.
³⁹⁴ The test pollen loads contained a number of closely related taxa, some in notoriously taxonomically difficult
³⁹⁵ clades (e.g. *Mertensia*, *Lupinus*), at naturally occurring proportions. We incorporated spatial and tempo-
³⁹⁶ ral approaches for creating custom sequence databases an approach which is readily applicable to any lab
³⁹⁷ group with the capacity to perform next-generation sequencing across the entirety of multiple continents,
³⁹⁸ and which we expect to be highly beneficial in many study areas. By combining insights from these novel
³⁹⁹ approaches with an extensive observational field based study we show how these methods may be applied to
⁴⁰⁰ test a variety of hypotheses related to ecological interactions.

401 We anticipate that many of the complications which we faced, using opportunistically collected pollen loads
402 and the first implementation of this method may readily be overcome. It seems apparent that we had issues
403 detecting pollen from several genera of plants, based upon these and other observational studies most likely
404 *Vicia*, *Lathyrus*, and *Frasera* (Inouye (1980), Pleasants (1980)), this is most likely related to user error in ob-
405 taining high quality DNA during the plant reference library generation period. (**REED SHOULD HUNT**
406 **FOR LOCI RETURNS FROM BLAST LIKE HE DID POLLEN**). Additional complications seem
407 to relate to the presence of closely related false positives, e.g. frequent classifications of sequences as *Trollius*,
408 *Caltha*, and *Thalictrum* alongside a more common species in the family, e.g. *Delphinium*.. Many of our
409 errors are known to us and multiple mnemonics are in *APPENDIX XX* to assist others in future attempts
410 to achieve better results. However, the line between false and negative positives may be blurred in some
411 of these instances and warrant further work, for example **Ericaceae pollen grains were observed in a**
412 **number of samples in trace quantities....**

413 These results show that the overall results between **Bumble Bee ecology** observational and barcoding are
414 largely congruent. But that ... We analyzed pollen loads from all of the most common bumble bee species
415 in the area(Pyke (1982)) Future analyses of the long term data set...

416 **Spatial ... & Timing** (filters!) We have concerns regarding the number of persons training to become and
417 practice botany, and grave concerns regarding the funding mechanisms for floristic and field based botanical
418 research and for centralized authorities to produce consensus opinions on alpha taxonomy (Prather *et al.*
419 (2004b), Kramer & Havens (2015), Prather *et al.* (2004a), Crisci *et al.* (2020), Manzano (2021), Stroud
420 *et al.* (2022)). To reduce the effects of a low population density of botanists on the maintenance of and
421 production of flora's and to foster metagenomics across landscapes without field stations we utilized Species
422 Distribution Modelling to generate predictive species lists. In this proof of concept example we performed
423 several iterations of modelling runs, and several approaches (i.e. the ‘linear models’, and the ‘machine
424 learning’), which took notable amounts of compute power. We suspect the possible deleterious nature of
425 this endeavor may be reduced by: 1) more field surveying by crews will reduce the need to generate as many
426 species 2) fewer runs of models, 3) only running machine learning models which do not require an explicitly
427 process to reduce spatial autocorrelation. However, given the time required to perform all aspects of a study,
428 even our amount of computation was negligible. Further, we are very optimistic about the possibility for
429 persons to perform these tasks, as mentioned we utilized roughly only one quarter of the records which
430 were digitally available for presence, and we suspect others will have enough records to perform this process
431 nearly anywhere else in the temperate. Tandem to the lack of continued expertise required to generate
432 and maintain species lists, is the expertise required to continue tracking when major phenological events

433 occur in many plant species at relatively fine scales or under novel climates. Knowledge of these events is
434 currently limited to general time periods of only a handful of phenological events and groups of organisms
435 (e.g. flowering initiation, or trees) (Prather *et al.* (2004a), Li *et al.* (2016)). While many programs and
436 initiatives exist to collect phenological information on subsets of easily identifiable charismatic species to
437 detect major trends in phenology, by design these capture only a subset of the extent diversity (Betancourt
438 *et al.* (2005), Havens *et al.* (2007)). In many instances it appears that while landscapes respond similarly
439 to environmental variables which predict phenological responses, that individual species vary widely in their
440 responses to similar environmental cues, or respond to different cues (Augspurger & Zaya (2020), Xie *et al.*
441 (2015), Xie *et al.* (2018), CaraDonna *et al.* (2014)). As can be seen here, predictions of when a single, major
442 phenological event occurs is already data limited, with sample size having an effect on the subset of species
443 which we could even generate weibull estimates for.

444 **Molecular . . .** The nearly complete Plant and Fungal Tree of Life (PAFTOL) will provide a comprehensive
445 phylogenetic backbone of the entire plant kingdom, and the inclusion of A353 probes with lineage specific
446 probe sets is common in producing massive genetic datasets (Baker *et al.* (2021b)). We predict that the
447 A353 probes which it is utilizing to work nearly immediately for DNA barcoding of whole plant material, and
448 that more elaborate validation studies in controlled metabarcoding settings, utilizing existing experimental
449 designs, will have favorable results (Bell *et al.* (2017), Bell *et al.* (2019), Bell *et al.* (2021), Lamb *et al.*
450 (2019)). In particular the harvesting of loci with more variation in certain lineages, and or with more variable
451 flanking regions, will prove promising for identifying closely related plant material (CITE). We suspect that
452 conserved reaches of genes resulted in the high amounts of reads in somewhat obscure species. Given that
453 the A353 loci are nuclear, single copy, and a variety are present the possibility of identifying target loci for
454 quantitative purposes is high, without continual PCR enrichment is possible; this would align with relatively
455 high efficacy of WGS (Lang *et al.* (2019), Peel *et al.* (2019), Bell *et al.* (2021)). Recent evidence indicates
456 that the potential for identifying nearly cryptic taxa and even infra-specific inference, of either whole plant
457 material, and perhaps in metagenomic context are possible (Ottenlips *et al.* (2021), Wenzell *et al.* (2021),
458 Loke *et al.* in prep, Slimp *et al.* (2021), Beck *et al.* (2021)).

459 4 | CONCLUSION

460 We believe that the combination of spatial and temporal models, united and guided by localized natural
461 history knowledge, provides the essential components of a bayesian framework for approaching the coarse
462 elucidation of ecological interactions using DNA Barcoding. Herein we crudely utilized this thinking via

463 binary outcomes, should a species predicted be predicted present or not? Is it unequivocally flowering
464 or not? Myriad data show biological systems and ecological interactions have more variance than can be
465 reasonably discretely parsed. We expect that within a bayesian framework studies of pollinator behavior
466 may be enacted via this approach at a landscape level, e.g. the scale of an entire drainage basin such as the
467 Gunnison which is quickly becoming one of the worlds few model ecosystems. We hope that the promise of
468 A353 probes as tools for metabarcoding play a role in these endeavors.

469 **AUTHOR CONTRIBUTIONS:** R.C.B conducted botanical collections, conducted all molecular lab
470 work, lead all analyses, and writing. J.E.O conceived, designed, and conducted all ecological fieldwork,
471 assisted with analyses, and writing. E.J.W. prepared, imaged, and collected trait data on pollen reference
472 slides, and assisted with analysis of trait data and writing a dichotomous key. S.T. assisted with spatial
473 analyses and writing. P.J.C assisted with ecological analyses and writing. J.B.F. conceived, and designed all
474 lab work, analyses, and integration of approaches, assisted with writing, and secured funding for molecular
475 work.

476 **ACKNOWLEDGMENTS:** Nyree Zerega for assistance obtaining herbaria loans and accessioning our
477 collections at CHIC. Pat Herendeen for assistance with virtually all aspects of preparing pollen vouchers
478 and the identification process. Ian Breckheimer for sharing the SDM predictor variables. Hilary Noble, Zoe
479 Diaz-Martinez, Angela McDonnell, & Elena Loke for assistance with genomic library preparation. We thank
480 the curators at the following herbaria for supplying tissue: Ben Legler at Stillinger (ID), Charles (Rick)
481 Williams at Ray J. Davis (IDS), (B)Ernie Nelson at Rocky Mountain (RM); and the collectors: D. Knoke,
482 L. Brummer, J. Boyd, C. Davidson, I. Gilman, M. Kirkpatrick, S. McCauley, J. Smith, K. Taylor, & C.
483 Williams. David Giblin & for sharing relevant sections of an advanced draft of FNA V. 15. The Bureau of
484 Land Management is thanked as many plant specimens were collected by R.C.B as a partner or contractor
485 to the agency, Sarah Burnett and Lauren Price are thanked for sharing AIM data. Sanda and NEB are
486 gratefully acknowledged for technical support and generously sharing samples. T.C.H. Cole for sharing the
487 Angiosperm Phylogeny 4 colour palette.

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

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

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

492 **ORCID**

493 Paul CaraDonna <https://orcid.org/0000-0003-3517-9090>

⁴⁹⁴ Jeremie Fant <https://orcid.org/0000-0001-9276-1111>

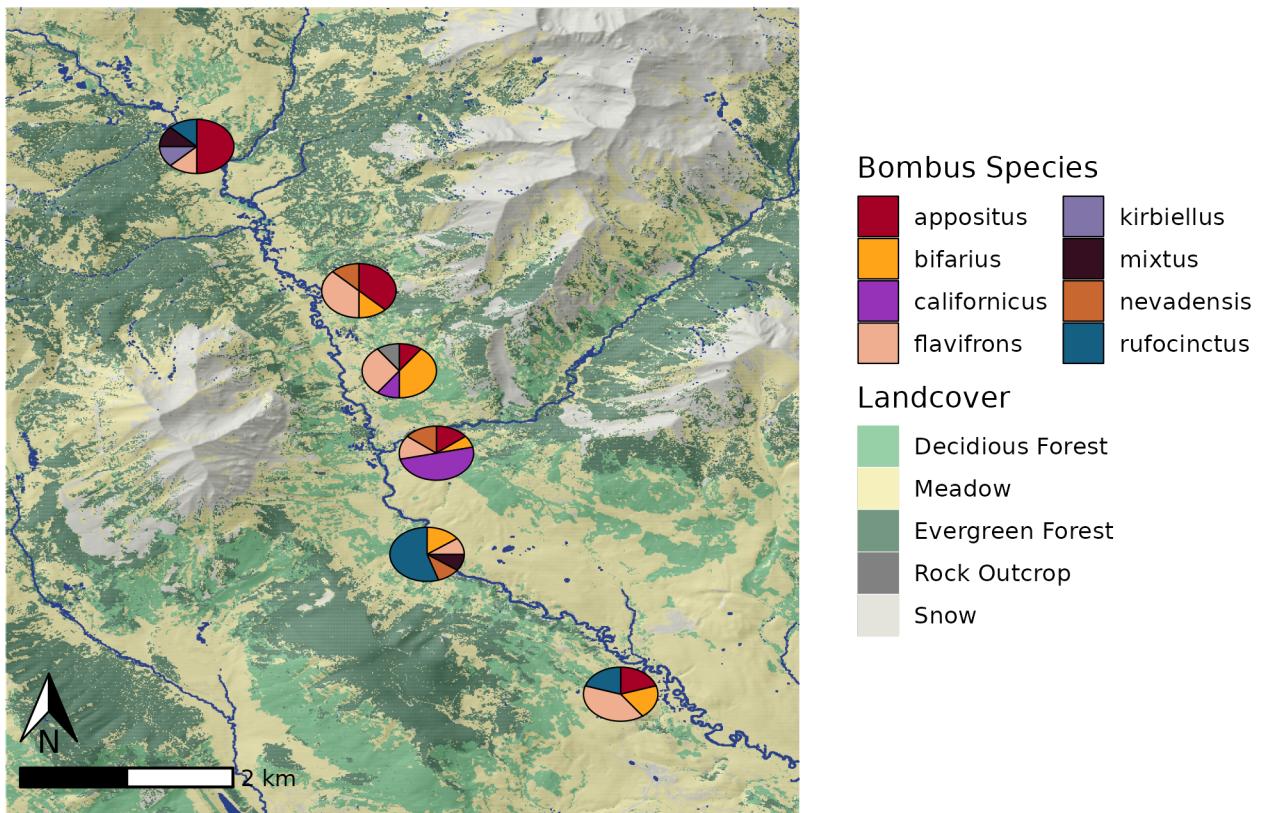
⁴⁹⁵ Jane Ogilvie <https://orcid.org/0000-0001-8546-0417>

⁴⁹⁶ Sophie Taddeo <https://orcid.org/0000-0002-7789-1417>

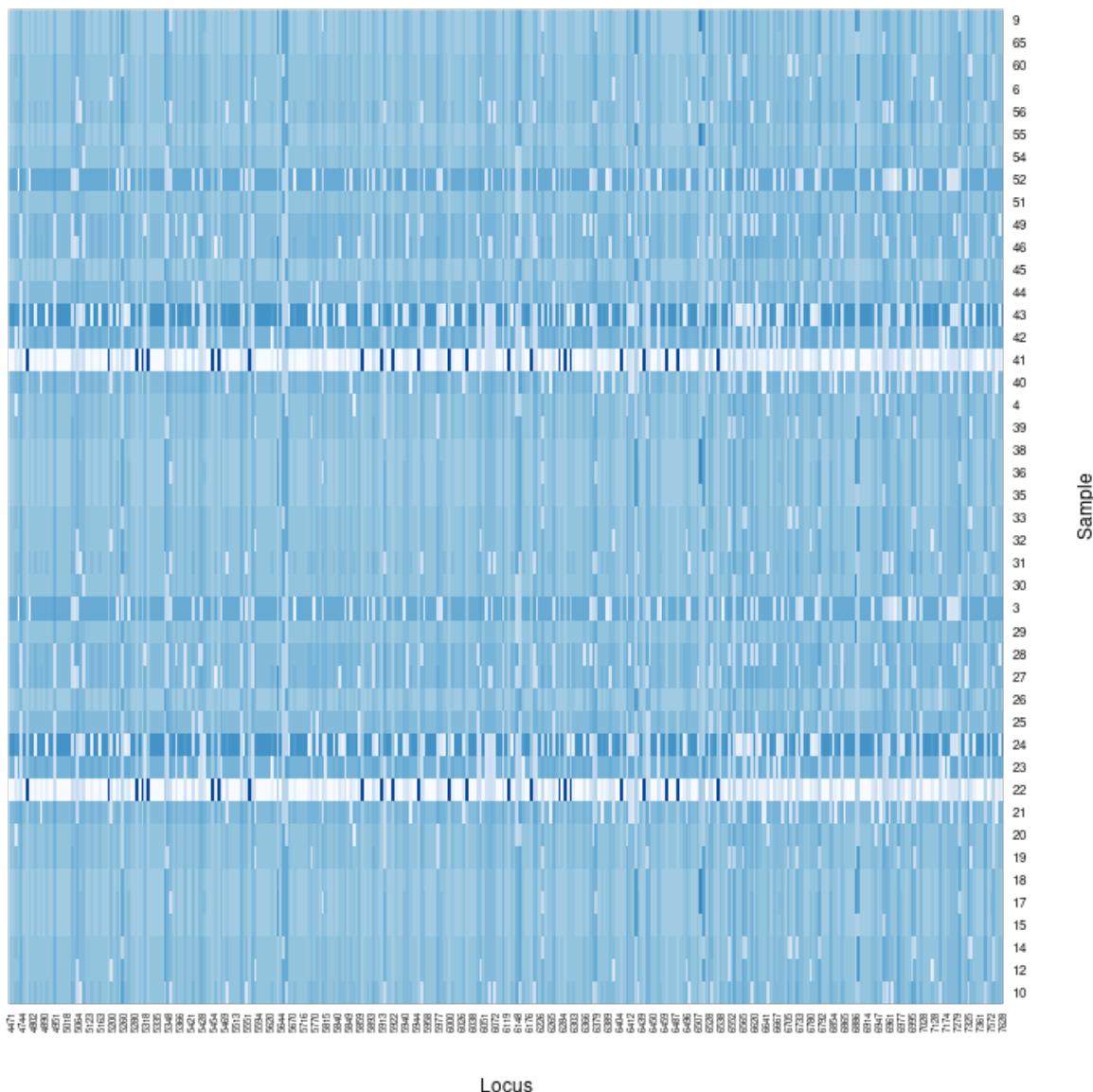
⁴⁹⁷ **References**

⁴⁹⁸ **Supporting**

Origins of Corbiculae Loads

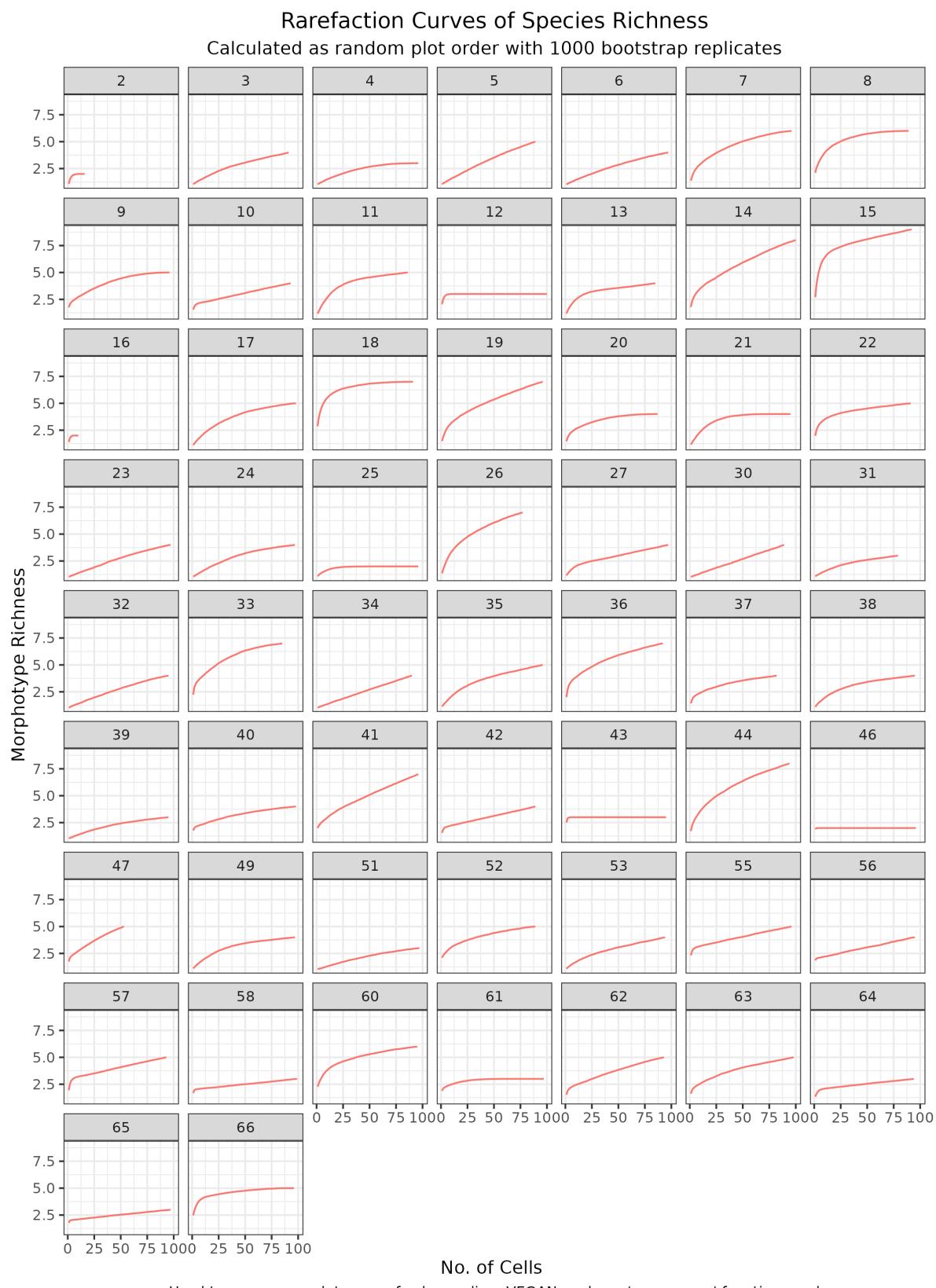


Percent matched reads per locus by sample



503 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



Species Richness Abundance Estimating via Hill Numbers ($q = 0$)

Confidence Interval of 99% with 1000 Bootstrap replicates

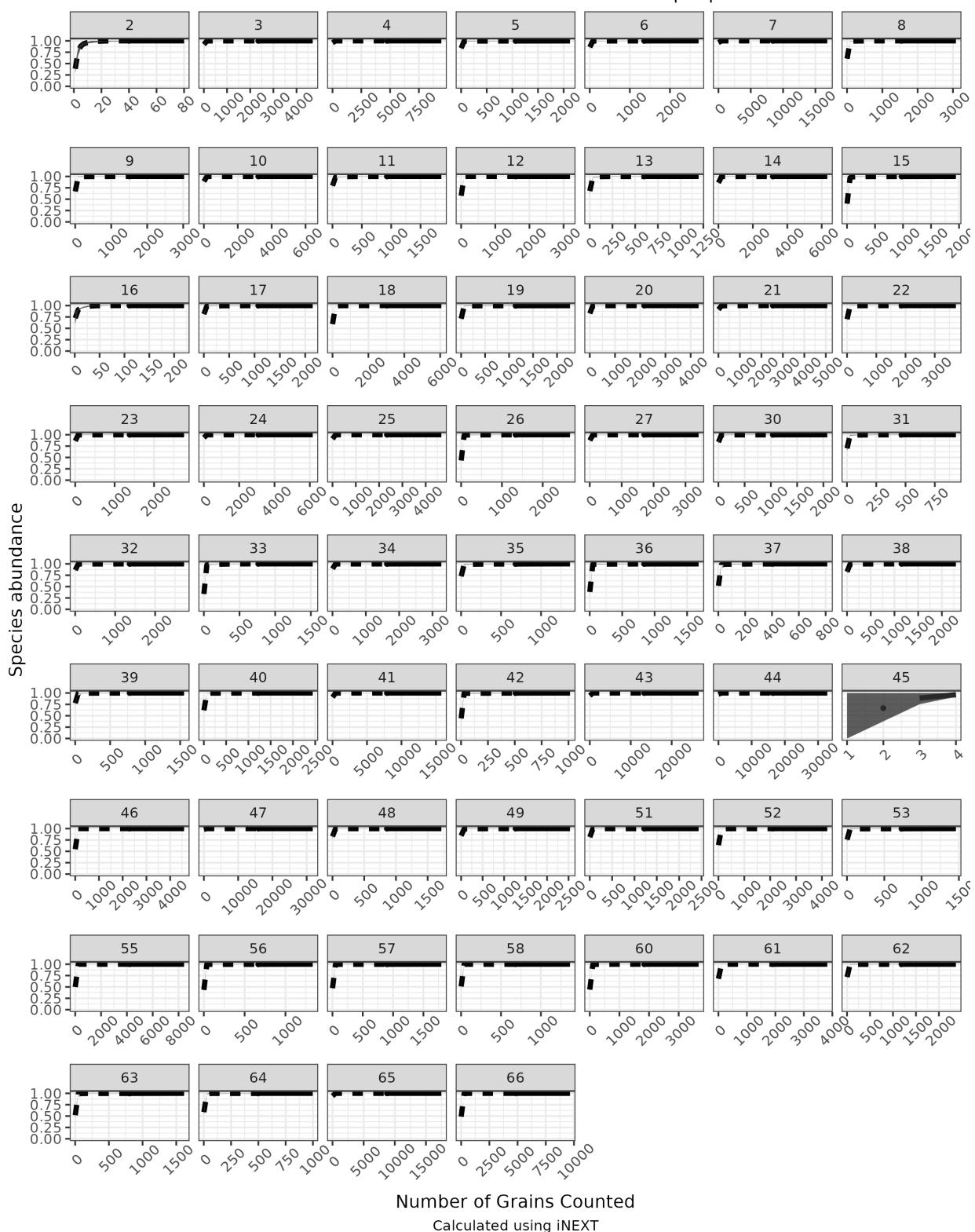


Table 1: samples used in creating the Reference Library

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 169837	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 ecuadorensis</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 foliaceum</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>
	Hydrophyllaceae	<i>Cynoglossum amplifolium</i> <i>Cynoglossum anchusoides</i> <i>Cynoglossum pringlei</i> <i>Mertensia ciliata</i> <i>Mertensia fusiformis</i>
		<i>Hydrophyllum canadense</i>
		<i>Hydrophyllum capitatum</i>
		<i>Hydrophyllum fendleri</i>
		<i>Nemophila menziesii</i>
Caryophyllales	Caryophyllaceae	<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>Shivparvatia glanduligera</i>
		<i>Stellaria graminea</i>
		<i>Stellaria holostea</i>
		<i>Stellaria obtusa</i>
		<i>Rumex induratus</i>
		<i>Rumex spinosus</i>
		<i>Parnassia faberi</i>
		<i>Parnassia palustris</i>
		<i>Paxistima canbyi</i>
		<i>Gaultheria prostrata</i>
Celastrales	Ericaceae	<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>
		<i>Primulaceae</i>
		<i>Androsace studiosorum</i>
		<i>Androsace vitaliana</i>
	Fabaceae	<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)

516 Appendix XX - All Species in the Sequence Databases (con't)

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}

518 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 lemmmonii</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 multiflora</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 stenomeres</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 serra</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)

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
<i>Downingia</i>	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> Muh. 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 eucomus</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 argenteus</i> Pursh	Fabaceae	NV, Pershing, Star Peak Canyon	tba	Novo	E.J.W.	2021
<i>Lupinus crassus</i> Payson	Fabaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Lupinus sericeus</i> Pursh	Fabaceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Melilotus albus</i> Medik.	Fabaceae	NM, Catron, Hwy 159	tba	Novo	E.J.W.	2021
<i>Trifolium hybridum</i> L.	Fabaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<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	UT, Cache, Spawn Creek	tba	Novo	E.J.W.	2021
<i>Vicia americana</i> Muhl. Ex Willd.	Fabaceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Geranium</i>	Geraniaceae	NM, Catron, Jim Smith TH. Rd.	tba	Novo	E.J.W.	2021
<i>Geranium richardsonii</i> Fisch. Trautv.	Geraniaceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Phacelia</i> sp.	Hydrophyllaceae	NV, Nye, Toiyabe Crest	tba	Novo	E.J.W.	2021
<i>Iris missouriensis</i> Nutt.	Iridaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Agastache palliflora</i> (A. Heller) Rydb.	Lamiaceae	AZ, Coconino, Lake Mary Rd. & 209	tba	Novo	E.J.W.	2021
<i>Erythronium grandiflorum</i> Pursh	Liliaceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Linum lewisii</i> Pursh	Linaceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Zigadenus elegans</i> Pursh	Melanthiaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Anticlea elegans</i> (A. Gray) Zomlefer & Judd	Melanthiaceae	ID, Blaine, Wildhorse Canyon	tba	Novo	E.J.W.	2021
<i>Cistanthe/Calyptidium</i>	Montiaceae	NV, Nye, Toiyabe Crest	tba	Novo	E.J.W.	2021
<i>Chamerion angustifolium</i> (L.) Holub	Onagraceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Epilobium obcordatum</i> A. Gray	Onagraceae	ID, Lemhi, Bighorn Crags	tba	Novo	E.J.W.	2021
<i>Castilleja miniata</i> Douglas ex Hook.	Orobanchaceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010

(Continued on Next Page)

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> Tuzc. Ex Fisch. & C.A. Mey.	Ranunculaceae	ID, Custer, E. fk. 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 (coporate)
- 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)

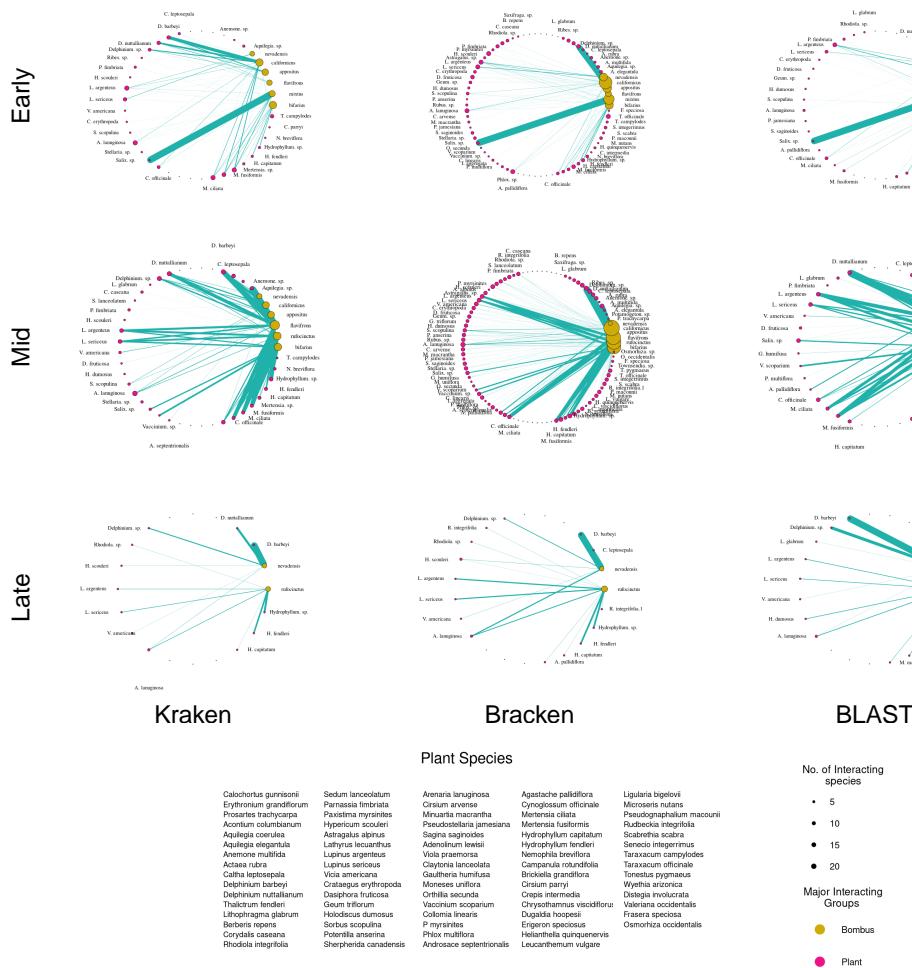
Morphotype A: Trifolium, Lupinus, Glycrrhiza, Mitella, Geum

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

Comparision of Foraging Patterns from Three Sequence Alignment Algorithms



526 Appendix XX - Models used for Species Distribution Model Ensembles

527 *Generalised Linear Models (GLM)*

528 *Generalised Additive Models (GAM)*

529 The two machine learning models utilize Ensemble learning.

530 Decision trees, ...

531 Ensemble learning utilizes many sets of trees, each tree being composed of many binary decisions, to create
532 a single model. Each independent variable (- or *feature*) may become a node on the tree - i.e. a location on
533 the tree where a binary decision will move towards a predicted outcome. Each of the decision tree models
534 which ensemble learning utilizes is a weak model, each of which may suffer due to high variance or bias, but
535 which produce better outcomes than would be expected via chance. When ensembled these models generate
536 a strong model, a model which should have more appropriately balanced variance and bias and predicts
537 outcomes which are more strongly correlated with the expected values than the individual weak models.

538 *Random Forest (RF)* the training data are continually bootstrap re-sampled, in combination with random
539 subsets of features, to create nodes which attempt to optimally predict a known outcome. A large number
540 of trees are then aggregated, via the most common predictions, to generate a final classification prediction
541 tree. Each individual prediction tree is generated independently of the others.

542 *Boosted Regression Tree (BRT)* (or Gradient Boosted tree) An initial tree is grown, and all other trees
543 are derived sequentially from it, as each new tree is grown the errors in responses from the last tree are
544 weighed more heavily so that the model focuses on selecting dependent variables which refine predictions.
545 All response data and predictor variables are kept available to all trees.

546 Random Forest have high bias and low variance, where boosted regressions trees have low bias and high
547 variances.

- 548 Ackerfield, J. (2015). *Flora of colorado*. BRIT Press Fort Worth.
- 549 Alarcón, R. (2010). Congruence between visitation and pollen-transport networks in a California plant–
550 pollinator community. *Oikos*, **119**, 35–44. Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0706.2009.17694.x>
- 551
- 552 Allouche, O., Tsoar, A. & Kadmon, R. (2006). Assessing the accuracy of species distribution models:
553 Prevalence, kappa and the true skill statistic (TSS). *Journal of applied ecology*, **43**, 1223–1232.
- 554 Allred, K.W. & Ivey, R. (2012). Flora neomexicana III: An illustrated identification manual. *Lulu. com*.
- 555 Araujo, M.B. & New, M. (2007). Ensemble forecasting of species distributions. *Trends in ecology & evolution*,
556 **22**, 42–47.
- 557 Augspurger, C.K. & Zaya, D.N. (2020). Concordance of long-term shifts with climate warming varies among
558 phenological events and herbaceous species. *Ecological Monographs*, **90**, e01421.
- 559 Baker, W.J., Bailey, P., Barber, V., Barker, A., Bellot, S., Bishop, D., Botigué, L.R., Brewer, G., Carruthers,
560 T., Clarkson, J.J., Cook, J., Cowan, R.S., Dodsworth, S., Epitawalage, N., Françoso, E., Gallego, B.,
561 Johnson, M.G., Kim, J.T., Leempoel, K., Maurin, O., McGinnie, C., Pokorny, L., Roy, S., Stone, M.,
562 Toledo, E., Wickett, N.J., Zuntini, A.R., Eiserhardt, W.L., Kersey, P.J., Leitch, I.J. & Forest, F. (2021a).
563 A Comprehensive Phylogenomic Platform for Exploring the Angiosperm Tree of Life. *Systematic Biology*,
564 **71**, 301–319. Retrieved from <https://doi.org/10.1093/sysbio/syab035>
- 565 Baker, W., Dodsworth, S., Forest, F., Graham, S., Johnson, M., McDonnell, A., Pokorny, L., Tate, J., Wicke,
566 S. & Wickett, N. (2021b). Exploring Angiosperms353: An open, community toolkit for collaborative
567 phylogenomic research on flowering plants. *American Journal of Botany*, **108**.
- 568 Barbet-Massin, M., Jiguet, F., Albert, C.H. & Thuiller, W. (2012). Selecting pseudo-absences for species
569 distribution models: How, where and how many? *Methods in ecology and evolution*, **3**, 327–338.
- 570 Barker, D.A. & Arceo-Gomez, G. (2021). Pollen transport networks reveal highly diverse and temporally
571 stable plant–pollinator interactions in an Appalachian floral community. *AoB PLANTS*, **13**. Retrieved
572 from <https://doi.org/10.1093/aobpla/plab062>
- 573 Beattie, A. (1971). A technique for the study of insect-borne pollen. *The Pan-Pacific Entomologist*, **47**, 82.
- 574 Beck, J.B., Markley, M.L., Zielke, M.G., Thomas, J.R., Hale, H.J., Williams, L.D. & Johnson, M.G. (2021).
575 Are palmer's elm-leaf goldenrod and the smooth elm-leaf goldenrod real? The Angiosperms353 kit
576 provides within-species signal in solidago ulmifolia sl. *Systematic Botany*, **46**, 1107–1113.
- 577 Belitz, M.W., Larsen, E.A., Ries, L. & Guralnick, R.P. (2020). The accuracy of phenology estimators for use
578 with sparsely sampled presence-only observations. *Methods in Ecology and Evolution*, **11**, 1273–1285.
- 579 Bell, K.L., Burgess, K.S., Botsch, J.C., Dobbs, E.K., Read, T.D. & Brosi, B.J. (2019). Quantitative and
580 qualitative assessment of pollen DNA metabarcoding using constructed species mixtures. *Molecular*

- 581 *Ecology*, **28**, 431–455.
- 582 Bell, K.L., Fowler, J., Burgess, K.S., Dobbs, E.K., Gruenewald, D., Lawley, B., Morozumi, C. & Brosi, B.J.
583 (2017). Applying pollen DNA metabarcoding to the study of plant–pollinator interactions. *Applications*
584 *in plant sciences*, **5**, 1600124.
- 585 Bell, K.L., Petit III, R.A., Cutler, A., Dobbs, E.K., Macpherson, J.M., Read, T.D., Burgess, K.S. & Brosi,
586 B.J. (2021). Comparing whole-genome shotgun sequencing and DNA metabarcoding approaches for
587 species identification and quantification of pollen species mixtures. *Ecology and evolution*, **11**, 16082–
588 16098.
- 589 Bergman, P., Molau, U. & Holmgren, B. (1996). Micrometeorological impacts on insect activity and plant
590 reproductive success in an alpine environment, swedish lapland. *Arctic and alpine research*, **28**, 196–202.
- 591 Betancourt, J.L., Schwartz, M.D., Breshears, D.D., Cayan, D.R., Dettinger, M.D., Inouye, D.W., Post, E.
592 & Reed, B.C. (2005). Implementing a US national phenology network.
- 593 Bingham, R.A. & Orthner, A.R. (1998). Efficient pollination of alpine plants. *Nature*, **391**, 238–239.
- 594 Bolger, A. & Giorgi, F. (2014). Trimmomatic: A flexible read trimming tool for illumina NGS data. *Bioin-*
595 *formatics*, **30**, 2114–2120.
- 596 Bortolus, A. (2008). Error cascades in the biological sciences: The unwanted consequences of using bad
597 taxonomy in ecology. *AMBIO: A journal of the human environment*, **37**, 114–118.
- 598 Brito-Morales, I., Molinos, J.G., Schoeman, D.S., Burrows, M.T., Poloczanska, E.S., Brown, C.J., Ferrier,
599 S., Harwood, T.D., Klein, C.J., McDonald-Madden, E. & others. (2018). Climate velocity can inform
600 conservation in a warming world. *Trends in ecology & evolution*, **33**, 441–457.
- 601 Brosi, B.J. & Briggs, H.M. (2013). Single pollinator species losses reduce floral fidelity and plant reproductive
602 function. *Proceedings of the National Academy of Sciences*, **110**, 13044–13048.
- 603 Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. & Madden, T.L. (2009).
604 BLAST+: Architecture and applications. *BMC bioinformatics*, **10**, 1–9.
- 605 Cameron, S.A. & Sadd, B.M. (2020). Global trends in bumble bee health. *Annual review of entomology*, **65**,
606 209–232.
- 607 CaraDonna, P.J., Burkle, L.A., Schwarz, B., Resasco, J., Knight, T.M., Benadi, G., Blüthgen, N., Dormann,
608 C.F., Fang, Q., Fründ, J. & others. (2021). Seeing through the static: The temporal dimension of
609 plant–animal mutualistic interactions. *Ecology Letters*, **24**, 149–161.
- 610 CaraDonna, P.J., Iler, A.M. & Inouye, D.W. (2014). Shifts in flowering phenology reshape a subalpine plant
611 community. *Proceedings of the National Academy of Sciences*, **111**, 4916–4921.
- 612 CaraDonna, P.J., Petry, W.K., Brennan, R.M., Cunningham, J.L., Bronstein, J.L., Waser, N.M. & Sanders,
613 N.J. (2017). Interaction rewiring and the rapid turnover of plant–pollinator networks. *Ecology letters*,

- 614 **20**, 385–394.
- 615 Chao, A., Gotelli, N.J., Hsieh, T.C., Sande, E.L., Ma, K.H., Colwell, R.K. & Ellison, A.M. (2014). Rarefac-
616 tion and extrapolation with hill numbers: A framework for sampling and estimation in species diversity
617 studies. *Ecological Monographs*, **84**, 45–67.
- 618 Cheng, S., Melkonian, M., Smith, S.A., Brockington, S., Archibald, J.M., Delaux, P.-M., Li, F.-W., Melko-
619 nian, B., Mavrodiev, E.V., Sun, W., Fu, Y., Yang, H., Soltis, D.E., Graham, S.W., Soltis, P.S., Liu,
620 X., Xu, X. & Wong, G.K.-S. (2018). 10KP: A phylogenetic genome sequencing plan. *GigaScience*, **7**.
621 Retrieved from <https://doi.org/10.1093/gigascience/giy013>
- 622 Coissac, E., Hollingsworth, P.M., Lavergne, S. & Taberlet, P. (2016). From barcodes to genomes: Extending
623 the concept of DNA barcoding.
- 624 Coissac, E., Riaz, T. & Puillandre, N. (2012). Bioinformatic challenges for DNA metabarcoding of plants
625 and animals. *Molecular ecology*, **21**, 1834–1847.
- 626 Colla, S.R., Gadallah, F., Richardson, L., Wagner, D. & Gall, L. (2012). Assessing declines of north american
627 bumble bees (*bombus* spp.) Using museum specimens. *Biodiversity and Conservation*, **21**, 3585–3595.
- 628 Crisci, J.V., Katinas, L., Apodaca, M.J. & Hoch, P.C. (2020). The end of botany. *Trends in Plant Science*,
629 **25**, 1173–1176.
- 630 Cronquist, A., Holmgren, A.H., Holmgren, N.H., Reveal, J.L., Holmgren, P.K., Barneby, R & others.
631 (1977+). *Intermountain flora. Vascular plants of the intermountain west, USA volume six. The mono-*
632 *cotyledons*. Columbia University.
- 633 Dobrowski, S.Z. & Parks, S.A. (2016). Climate change velocity underestimates climate change exposure in
634 mountainous regions. *Nature Communications*, **7**, 1–8.
- 635 Doyle, J.J. & Doyle, J.L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue.
636 *Phytochemical Bulletin*, **19**, 11–15.
- 637 Elith*, J., H. Graham*, C., P. Anderson, R., Dudik, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann,
638 F., R. Leathwick, J., Lehmann, A. & others. (2006). Novel methods improve prediction of species'
639 distributions from occurrence data. *Ecography*, **29**, 129–151.
- 640 Fazekas, A.J., Kesanakurti, P.R., Burgess, K.S., Percy, D.M., Graham, S.W., Barrett, S.C., Newmaster,
641 S.G., Hajibabaei, M. & Husband, B.C. (2009). Are plant species inherently harder to discriminate than
642 animal species using DNA barcoding markers? *Molecular Ecology Resources*, **9**, 130–139.
- 643 Flora of North America Editorial Committee, eds. (1993+). *Flora of north america north of mexico [online]*.
644 Oxford University Press on Demand.
- 645 Fraser, Barbara A. & Buck, P. (2007). Vascular Plants of the Gothic Area. Retrieved from https://www.digitalrmbi.org/wp-content/uploads/2016/05/vascularplantlist_20071.pdf

- 647 Gage, E. & Cooper, D.J. (2013). Historical range of variation assessment for wetland and riparian ecosystems,
648 u.s. Forest service rocky mountain region
- 649 Goulson, D., Lye, G. & Darvill, B. (2008). The decline and conservation of bumblebees. *Annual review of
650 entomology*, **53**, 191–208.
- 651 Group, C.P.W., Hollingsworth, P.M., Forrest, L.L., Spouge, J.L., Hajibabaei, M., Ratnasingham, S., Bank,
652 M. van der, Chase, M.W., Cowan, R.S., Erickson, D.L. & others. (2009). A DNA barcode for land
653 plants. *Proceedings of the National Academy of Sciences*, **106**, 12794–12797.
- 654 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.,
655 Chen, S.-L. & others. (2011). Comparative analysis of a large dataset indicates that internal transcribed
656 spacer (ITS) should be incorporated into the core barcode for seed plants. *Proceedings of the National
657 Academy of Sciences*, **108**, 19641–19646.
- 658 Havens, K., Vitt, P., Schwarz, J., Orr, B. & Crimmins, T. (2007). Chicago botanic garden's conservation
659 and outreach efforts on climate change. *BGjournal*, **4**, 13–16.
- 660 Hebert, P.D., Cywinska, A., Ball, S.L. & DeWaard, J.R. (2003). Biological identifications through DNA
661 barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **270**, 313–321.
- 662 Hengl, T., Mendes de Jesus, J., Heuvelink, G.B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A.,
663 Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B. & others. (2017). SoilGrids250m:
664 Global gridded soil information based on machine learning. *PLoS one*, **12**, e0169748.
- 665 Hennig, C. (2020). *Fpc: Flexible procedures for clustering*. Retrieved from <https://CRAN.R-project.org/>
666 package=fpc
- 667 Hitchcock, C.L. & Cronquist, A. (2018). *Flora of the pacific northwest: An illustrated manual*. University
668 of Washington Press.
- 669 Hsieh, T.C., Ma, K.H. & Chao, A. (2020). *iNEXT: Interpolation and extrapolation for species diversity*.
670 Retrieved from http://chao.stat.nthu.edu.tw/wordpress/software_download/
- 671 Iler, A.M., Humphrey, P.T., Ogilvie, J.E. & CaraDonna, P.J. (2021). Conceptual and practical issues limit
672 the utility of statistical estimators of phenological events. *Ecosphere*, **12**, e03828.
- 673 Inouye, D.W. (1980). The effect of proboscis and corolla tube lengths on patterns and rates of flower
674 visitation by bumblebees. *Oecologia*, **45**, 197–201.
- 675 Janzen, D.H. (1967). Synchronization of sexual reproduction of trees within the dry season in central america.
676 *Evolution*, **21**, 620–637.
- 677 Janzen, D.H., Burns, J.M., Cong, Q., Hallwachs, W., Dapkey, T., Manjunath, R., Hajibabaei, M., Hebert,
678 P.D. & Grishin, N.V. (2017). Nuclear genomes distinguish cryptic species suggested by their DNA
679 barcodes and ecology. *Proceedings of the National Academy of Sciences*, **114**, 8313–8318.

- 680 *Jepson flora project*. (2020).
- 681 Johnson, M.D., Fokar, M., Cox, R.D. & Barnes, M.A. (2021). Airborne environmental DNA metabarcoding
682 detects more diversity, with less sampling effort, than a traditional plant community survey. *BMC Ecology
683 and Evolution*, **21**, 1–15.
- 684 Johnson, M.G., Gardner, E.M., Liu, Y., Medina, R., Goffinet, B., Shaw, A.J., Zerega, N.J. & Wickett,
685 N.J. (2016). HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput
686 sequencing reads using target enrichment. *Applications in plant sciences*, **4**, 1600016.
- 687 Johnson, M.G., Pokorny, L., Dodsworth, S., Botigue, L.R., Cowan, R.S., Devault, A., Eiserhardt, W.L.,
688 Epitawalage, N., Forest, F., Kim, J.T. & others. (2019). A universal probe set for targeted sequencing
689 of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Systematic biology*,
690 **68**, 594–606.
- 691 Kramer, A.T. & Havens, K. (2015). Report in brief: Assessing botanical capacity to address grand challenges
692 in the united states. *Natural Areas Journal*, **35**, 83–89.
- 693 Kress, W.J. (2017). Plant DNA barcodes: Applications today and in the future. *Journal of systematics and
694 evolution*, **55**, 291–307.
- 695 Kress, W.J. & Erickson, D.L. (2007). A two-locus global DNA barcode for land plants: The coding rbcL
696 gene complements the non-coding trnH-psbA spacer region. *PLoS one*, **2**, e508.
- 697 Kuhn, M. (2022). *Caret: Classification and regression training*. Retrieved from <https://CRAN.R-project.org/package=caret>
- 698 Lamb, P.D., Hunter, E., Pinnegar, J.K., Creer, S., Davies, R.G. & Taylor, M.I. (2019). How quantitative is
699 metabarcoding: A meta-analytical approach. *Molecular ecology*, **28**, 420–430.
- 700 Land Management, B. of. (2020). U.s. Department of interior bureau of land management, BLM - assess-
701 ment, inventory, and monitoring (AIM) terrestrial indicators raw dataset. Retrieved from [https://gbp-
703 blm-egis.hub.arcgis.com/pages/aim](https://gbp-
702 blm-egis.hub.arcgis.com/pages/aim)
- 704 Lang, D., Tang, M., Hu, J. & Zhou, X. (2019). Genome-skimming provides accurate quantification for pollen
705 mixtures. *Molecular Ecology Resources*, **19**, 1433–1446.
- 706 Lesica, P., Lavin, M. & Stickney, P.F. (2012). *Manual of montana vascular plants*. Brit Press.
- 707 Lewin, H.A., Richards, S., Aiden, E.L., Allende, M.L., Archibald, J.M., Bálint, M., Barker, K.B., Baumgart-
708 ner, B., Belov, K., Bertorelle, G., Blaxter, M.L., Cai, J., Caperello, N.D., Carlson, K., Castilla-Rubio,
709 J.C., Chaw, S.-M., Chen, L., Childers, A.K., Coddington, J.A., Conde, D.A., Corominas, M., Crandall,
710 K.A., Crawford, A.J., DiPalma, F., Durbin, R., Ebenezer, T.E., Edwards, S.V., Fedrigo, O., Flicek, P.,
711 Formenti, G., Gibbs, R.A., Gilbert, M.T.P., Goldstein, M.M., Graves, J.M., Greely, H.T., Grigoriev,
712 I.V., Hackett, K.J., Hall, N., Haussler, D., Helgen, K.M., Hogg, C.J., Isobe, S., Jakobsen, K.S., Janke,

- 713 A., Jarvis, E.D., Johnson, W.E., Jones, S.J.M., Karlsson, E.K., Kersey, P.J., Kim, J.-H., Kress, W.J.,
714 Kuraku, S., Lawniczak, M.K.N., Leebens-Mack, J.H., Li, X., Lindblad-Toh, K., Liu, X., Lopez, J.V.,
715 Marques-Bonet, T., Mazard, S., Mazet, J.A.K., Mazzoni, C.J., Myers, E.W., O'Neill, R.J., Paez, S.,
716 Park, H., Robinson, G.E., Roquet, C., Ryder, O.A., Sabir, J.S.M., Shaffer, H.B., Shank, T.M., Sherkow,
717 J.S., Soltis, P.S., Tang, B., Tedersoo, L., Uliano-Silva, M., Wang, K., Wei, X., Wetzer, R., Wilson,
718 J.L., Xu, X., Yang, H., Yoder, A.D. & Zhang, G. (2022). The earth BioGenome project 2020: Starting
719 the clock. *Proceedings of the National Academy of Sciences*, **119**, e2115635118. Retrieved from
720 <https://www.pnas.org/doi/abs/10.1073/pnas.2115635118>
- 721 Life Project Consortium, D.T. of, Blaxter, M., Mieszkowska, N., Palma, F.D., Holland, P., Durbin, R.,
722 Richards, T., Berriman, M., Kersey, P., Hollingsworth, P., Wilson, W., Twyford, A., Gaya, E., Lawniczak,
723 M., Lewis, O., Broad, G., Howe, K., Hart, M., Flieck, P. & Barnes, I. (2022). Sequence locally, think globally:
724 The darwin tree of life project. *Proceedings of the National Academy of Sciences*, **119**, e2115642118.
725 Retrieved from <https://www.pnas.org/doi/abs/10.1073/pnas.2115642118>
- 726 Li, X., Jiang, L., Meng, F., Wang, S., Niu, H., Iler, A.M., Duan, J., Zhang, Z., Luo, C., Cui, S. & others.
727 (2016). Responses of sequential and hierarchical phenological events to warming and cooling in alpine
728 meadows. *Nature Communications*, **7**, 1–8.
- 729 Liu, J., Shi, L., Han, J., Li, G., Lu, H., Hou, J., Zhou, X., Meng, F. & Downie, S.R. (2014). Identification
730 of species in the angiosperm family apiaceae using DNA barcodes. *Molecular ecology resources*, **14**,
731 1231–1238.
- 732 Li, X., Yang, Y., Henry, R.J., Rossetto, M., Wang, Y. & Chen, S. (2015). Plant DNA barcoding: From gene
733 to genome. *Biological Reviews*, **90**, 157–166.
- 734 Loarie, S.R., Duffy, P.B., Hamilton, H., Asner, G.P., Field, C.B. & Ackerly, D.D. (2009). The velocity of
735 climate change. *Nature*, **462**, 1052–1055.
- 736 Lu, J., Breitwieser, F.P., Thielen, P. & Salzberg, S.L. (2017). Bracken: Estimating species abundance in
737 metagenomics data. *PeerJ Computer Science*, **3**, e104.
- 738 Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. & Hornik, K. (2022). *Cluster: Cluster analysis basics
739 and extensions*. Retrieved from <https://CRAN.R-project.org/package=cluster>
- 740 Maitner, B. (2022). *BIEN: Tools for accessing the botanical information and ecology network database*.
741 Retrieved from <https://CRAN.R-project.org/package=BIEN>
- 742 Manzano, S. (2021). Flippant attitudes towards plant identification jeopardize early career botanists. *Trends
743 in Plant Science*, **26**, 987–988.
- 744 McLay, T.G., Birch, J.L., Gunn, B.F., Ning, W., Tate, J.A., Nauheimer, L., Joyce, E.M., Simpson, L.,
745 Schmidt-Lebuhn, A.N., William J & others. (2021). New targets acquired: Improving locus recovery

- 746 from the Angiosperms353 probe set. *Applications in plant sciences*, **9**.
- 747 Mohlenbrock, R.H. (2002). *Vascular flora of illinois*. SIU Press.
- 748 Naimi, B. & Araujo, M.B. (2016). Sdm: A reproducible and extensible r platform for species distribution
749 modelling. *Ecography*, **39**, 368–375.
- 750 Naimi, B., Hamm, N. a.s., Groen, T.A., Skidmore, A.K. & Toxopeus, A.G. (2014). Where is positional
751 uncertainty a problem for species distribution modelling. *Ecography*, **37**, 191–203.
- 752 Newstrom, L.E., Frankie, G.W. & Baker, H.G. (1994). A new classification for plant phenology based on
753 flowering patterns in lowland tropical rain forest trees at la selva, costa rica. *Biotropica*, **26**, 141–159.
- 754 Occownload Gbif.Org. (2021). Occurrence download. Retrieved from <https://www.gbif.org/occurrence/>
755 download/0206948-200613084148143
- 756 Oksanen, J., Simpson, G.L., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Solymos,
757 P., Stevens, M.H.H., Szoechs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D.,
758 Carvalho, G., Chirico, M., De Caceres, M., Durand, S., Evangelista, H.B.A., FitzJohn, R., Friendly,
759 M., Furneaux, B., Hannigan, G., Hill, M.O., Lahti, L., McGlinn, D., Ouellette, M.-H., Ribeiro Cunha,
760 E., Smith, T., Stier, A., Ter Braak, C.J.F. & Weedon, J. (2022). *Vegan: Community ecology package*.
761 Retrieved from <https://CRAN.R-project.org/package=vegan>
- 762 Oliver, P.M., Adams, M., Lee, M.S., Hutchinson, M.N. & Doughty, P. (2009). Cryptic diversity in vertebrates:
763 Molecular data double estimates of species diversity in a radiation of australian lizards (diplodactylus,
764 gekkota). *Proceedings of the Royal Society B: Biological Sciences*, **276**, 2001–2007.
- 765 Omernik, J.M. (1987). Ecoregions of the conterminous united states. *Annals of the Association of American
766 geographers*, **77**, 118–125.
- 767 Ottenlips, M.V., Mansfield, D.H., Buerki, S., Feist, M.A.E., Downie, S.R., Dodsworth, S., Forest, F., Plun-
768 kett, G.M. & Smith, J.F. (2021). Resolving species boundaries in a recent radiation with the An-
769 giosperms353 probe set: The lomatium packardiae/l. anomalam clade of the l. triternatum (apiaceae)
770 complex. *American journal of botany*, **108**, 1217–1233.
- 771 Pearse, W.D., Davis, C.C., Inouye, D.W., Primack, R.B. & Davies, T.J. (2017). A statistical estimator
772 for determining the limits of contemporary and historic phenology. *Nature Ecology & Evolution*, **1**,
773 1876–1882.
- 774 Peel, N., Dicks, L.V., Clark, M.D., Heavens, D., Percival-Alwyn, L., Cooper, C., Davies, R.G., Leggett,
775 R.M. & Yu, D.W. (2019). Semi-quantitative characterisation of mixed pollen samples using MinION
776 sequencing and reverse metagenomics (RevMet). *Methods in Ecology and Evolution*, **10**, 1690–1701.
- 777 Pepin, N., Arnone, E., Gobiet, A., Haslinger, K., Kotlarski, S., Notarnicola, C., Palazzi, E., Seibert, P.,
778 Serafin, S., Schöner, W. & others. (2022). Climate changes and their elevational patterns in the mountains

- 779 of the world. *Reviews of geophysics*, **60**, e2020RG000730.
- 780 Pleasants, J.M. (1980). Competition for bumblebee pollinators in rocky mountain plant communities. *Ecol-*
781 *ogy*, **61**, 1446–1459.
- 782 Pernon, A., Andalo, C., Burrus, M. & Escaravage, N. (2017). DNA metabarcoding data unveils invisible
783 pollination networks. *Scientific Reports*, **7**, 1–11.
- 784 Prather, L.A., Alvarez-Fuentes, O., Mayfield, M.H. & Ferguson, C.J. (2004a). Implications of the decline in
785 plant collecting for systematic and floristic research. *Systematic Botany*, **29**, 216–220.
- 786 Prather, L.A., Alvarez-Fuentes, O., Mayfield, M.H. & Ferguson, C.J. (2004b). The decline of plant collecting
787 in the united states: A threat to the infrastructure of biodiversity studies. *Systematic Botany*, **29**, 15–28.
- 788 Prim, R.C. (1957). Shortest connection networks and some generalisations. *Bell System Technical Journal*,
789 **36**, 1389–1401.
- 790 Pyke, G.H. (1982). Local geographic distributions of bumblebees near crested butte, colorado: Competition
791 and community structure. *Ecology*, **63**, 555–573.
- 792 Qiao, H., Soberon, J. & Peterson, A.T. (2015). No silver bullets in correlative ecological niche modelling:
793 Insights from testing among many potential algorithms for niche estimation. *Methods in Ecology and*
794 *Evolution*, **6**, 1126–1136.
- 795 Robinson, N., Regetz, J. & Guralnick, R.P. (2014). EarthEnv-DEM90: A nearly-global, void-free, multi-
796 scale smoothed, 90m digital elevation model from fused ASTER and SRTM data. *ISPRS Journal of*
797 *Photogrammetry and Remote Sensing*, **87**, 57–67.
- 798 Ruppert, K.M., Kline, R.J. & Rahman, M.S. (2019). Past, present, and future perspectives of environmental
799 DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global
800 eDNA. *Global Ecology and Conservation*, **17**, e00547.
- 801 Sarro, E., Tripodi, A. & Woodard, S.H. (2022). Bumble bee (*bombus vosnesenskii*) queen nest searching
802 occurs independent of ovary developmental status. *Integrative Organismal Biology*, **4**, obac007.
- 803 Sickel, W., Ankenbrand, M.J., Grimmer, G., Holzschuh, A., Hartel, S., Lanzen, J., Steffan-Dewenter, I. &
804 Keller, A. (2015). Increased efficiency in identifying mixed pollen samples by meta-barcodeing with a
805 dual-indexing approach. *BMC ecology*, **15**, 1–9.
- 806 Slimp, M., Williams, L.D., Hale, H. & Johnson, M.G. (2021). On the potential of Angiosperms353 for
807 population genomic studies. *Applications in Plant Sciences*, **9**.
- 808 Stroud, S., Fennell, M., Mitchley, J., Lydon, S., Peacock, J. & Bacon, K.L. (2022). The botanical education
809 extinction and the fall of plant awareness. *Ecology and Evolution*, **12**, e9019.
- 810 Suchan, T., Talavera, G., Saez, L., Ronikier, M. & Vila, R. (2019). Pollen metabarcoding as a tool for
811 tracking long-distance insect migrations. *Molecular Ecology Resources*, **19**, 149–162.

- 812 Tange, O. (2021). GNU parallel 20220322 (savannah). Retrieved from <https://doi.org/10.5281/zenodo.6377950>
- 814 Tran, H., Nguyen, P., Ombadi, M., Hsu, K., Sorooshian, S. & Qing, X. (2019). A cloud-free MODIS snow
815 cover dataset for the contiguous united states from 2000 to 2017. *Scientific data*, **6**, 1–13.
- 816 Wang, T., Hamann, A., Spittlehouse, D. & Carroll, C. (2016). Locally downscaled and spatially customizable
817 climate data for historical and future periods for north america. *PloS one*, **11**, e0156720.
- 818 Wenzell, K.E., McDonnell, A.J., Wickett, N.J., Fant, J.B. & Skogen, K.A. (2021). Incomplete reproductive
819 isolation and low genetic differentiation despite floral divergence across varying geographic scales in
820 *castilleja*. *American Journal of Botany*, **108**, 1270–1288.
- 821 Williams, P.H. (1982). The distribution and decline of british bumble bees (*bombus latr.*). *Journal of
822 Apicultural Research*, **21**, 236–245. Retrieved from <https://doi.org/10.1080/00218839.1982.11100549>
- 823 Wilson, A.M. & Jetz, W. (2016). Remotely sensed high-resolution global cloud dynamics for predicting
824 ecosystem and biodiversity distributions. *PLoS biology*, **14**, e1002415.
- 825 Wood, D.E., Lu, J. & Langmead, B. (2019). Improved metagenomic analysis with kraken 2. *Genome biology*,
826 **20**, 1–13.
- 827 Xie, Y., Wang, X. & Silander Jr, J.A. (2015). Deciduous forest responses to temperature, precipitation, and
828 drought imply complex climate change impacts. *Proceedings of the National Academy of Sciences*, **112**,
829 13585–13590.
- 830 Xie, Y., Wang, X., Wilson, A.M. & Silander Jr, J.A. (2018). Predicting autumn phenology: How deciduous
831 tree species respond to weather stressors. *Agricultural and Forest Meteorology*, **250**, 127–137.
- 832 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,
833 D.-Z. & Wang, H. (2019). The topological differences between visitation and pollen transport networks:
834 A comparison in species rich communities of the himalaya–hengduan mountains. *Oikos*, **128**, 551–562.
835 Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1111/oik.05262>

836 List of Figures

837 1	Phylogenetic tree of all biotically pollinated plant genera in the study area. The innermost	47
838 ring indicates every genus which Queen Bee's were observed to visit. The intermediate ring		
839 indicates that at least a single morphological pollen voucher slide was prepared for a member		
840 of the genus. The outermost ring indicates that sequence data were available for at least a		
841 member of that genus. Branch colors follow APG 4.		
842 2	Modelled dates of when major flowering events occurred	48
843 3	Number of the ten most commonly visited plants which are also in the top ten most common	
844 sequences	49	

Biotically pollinated plant genera with morphological or molecular data



Figure 1: Phylogenetic tree of all biotically pollinated plant genera in the study area. The innermost ring indicates every genus which Queen Bee's were observed to visit. The intermediate ring indicates that at least a single morphological pollen voucher slide was prepared for a member of the genus. The outermost ring indicates that sequence data were available for at least a member of that genus. Branch colors follow APG 4.

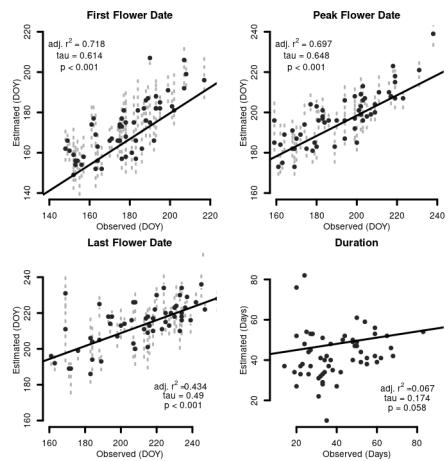


Figure 2: Modelled dates of when major flowering events occurred

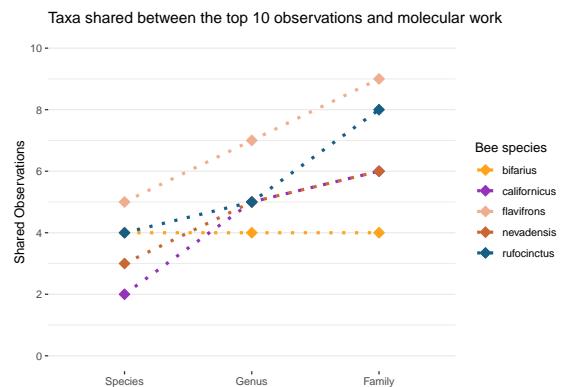


Figure 3: Number of the ten most commonly visited plants which are also in the top ten most common sequences

⁸⁴⁵ **List of Tables**

⁸⁴⁶	2	Logistic regression assessing accuracy of SDMs	51
⁸⁴⁷	3	Species Distribution Modeling evaluation contingency table	52

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: Species Distribution Modeling evaluation contingency table

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